IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Agricultural Crop Disease Protection And Leaf Disease Prediction Using Machine Learning

Spoorthi.S 1 , V.Bindushree 2, Anusha.P.R 3,Wasim Yasin
[2]UG Student, [3]UG Student, [4]UG Student, [5]UG Student,
[1]Department of Computer Science & Design
[1] K S Institute of Technology, Bengaluru, India

Abstract: Precision agriculture is an emerging area that applies modern information technology and machine learning to create news ways of identifying and diagnosing plant diseases to promote sustainable farming practices. This paper aims to review the application of machine learning and deep learning techniques in plant disease detection and classification in precision agriculture. It also proposes a different approach in classifying relevant literature which is based on the employed methodology - classification or object detection, and reviews the literature on datasets available for plant disease detection and classification. This work comprises a comprehensive analysis within the scope of object detection and classification of plant diseases utilising the PlantDoc dataset. The conclusion reached in this research is that YOLOV5 is the best object detection algorithm and that ResNet50 and MobileNetv2 models are the best image classifier models relative to the time cost of training the models and the accuracy of produced images.

***Index Terms* -** Classification,deeplearning,disease detection,machine learning,object detection,precision agriculture

1.Introduction

Almost every country heavily depends on agriculture for sustaining its economy. Furthermore, the ever-increasing population growth, resource constraints, and changing climatic conditions make satisfying food demands even more challenging. Plant diseases are also becoming more virulent and widespread, causing devastation to agriculture. The earlier these diseases are diagnosed, the better the chance the farmers will have to minimize losses. There is a great opportunity to improve farming practices through the use of ML, a part of AI that deals with the identification of patterns in data. It is now possible to give a foreign device the ability to learn all by itself without having to directly program it. Coupled with IoT gadgets, ML is transforming farming as we know it. Various works have been conducted on the use of ML algorithms for analyzing images of plants or leaves for the purpose of plant disease detection. The majority of the works classify the state of the plants' leaves as healthy or diseased and use methods such as Random Forest and Deep Learning.

Nonetheless, there is a scarcity of research on identifying the particular illness or measuring affected regions, which is more advanced object recognition. This is especially beneficial for drones that photograph crops over great distances. It is more complicated than classification because it has to work in non-stable environments with a great deal of noise competition. This document summarizes the analysis of ML and DL techniques for classification and detection of plant diseases. Earlier reviews concentrated on ML methods in farming and the role of AI and IoT technology in agriculture. Other reviews have concentrated on the application of DL techniques to the recognition of diseases on plant leaves, the fast and real time detection problems, and the inadequacy of available image datasets and needed transfer learning approaches. Although there are numerous reviews of the application of ML and DL in precision agriculture, very few deal with plant disease detection and classification, or both classification and object recognition. Due to the considerable increase in publications in this area, an effort should be undertaken to study the state of the art in these methods.

SECTION I

I.Literature Review

This segment contains a review of disease detection methods laid out in published works before October, 2022. I performed a thorough search using the following databases: Scopus, ScienceDirect, Scholar, Web of Science using the following keywords: “machine learning”, “deep learning”, “classification,” “disease detection,” “healthy plant,” and “diseased plant.” The relevant studies’ references were also screened for further studies. Each paper’s abstract was read to filter inclusion eligibility, and papers concentrating on non ML algorithms were removed from consideration. The final set consisted of 79 papers with 65 focusing on classification and 14 on object detection.

A. Classification

Various researchers have attempted to use CNN models for the detection of diseases in plants. Mohanty et al. trained both [1]AlexNet and GoogLeNet on the Plant Village dataset for 26 diseases across 14 crop species, achieving high accuracy with GoogLeNet obtaining 99.35%. Sladojevic et al. implemented a CNN approach with the Caffe DL framework and computed an average accuracy of 96.3%. DeChant et al. used a three-stage approach with CNN models to identify NLB infected maize plants, achieving an accuracy of 96.7%. Liu et al. used the AlexNet architecture to identify apple leaf diseases and achieved an accuracy of 97.62%. Kathole and Munot performed a[2] comparison of various CNN models for diseases of the tomato crop and recorded the highest accuracy of 97.94% with GoogLeNet.

Turkoglu et al. implemented hybrid CNN-SVM models and performed deep feature extraction based classification with transfer learning on six pre-trained DL networks. They achieved an accuracy of 97.56% on the Turkey-PlantDataset. Vallabhajosyula et al. implemented data augmentation techniques using Deep Ensemble Neural Networks (DENN) and outperformed all other state of the art models by achieving an accuracy of 99.99% on the Plant Village dataset.

B. Object Detection

Fuenteset al. combined three object detection models (Faster RCNN, R-FCN, and SSD with five feature extractors (VGG16, ResNet50, ResNet101, ResNet152, ResNeXt-50) for the recognition of tomato plant disease and pests. They applied the CNN models to a self-collected dataset of around 5, 000 images using various methods for dataset expansion. The results show that R-FCN with ResNet50 provided the highest mAP of 86%. [4]

Jiang et alutilized a DL approach based on the GoogLeNet Inception model and the rainbow concatenation to detect apple leaf diseases. They generated a special dataset of 2, 029 images of apple leaf infections and trained their program to recognize five common apple leaf diseases: Alternaria leaf spot, brown spot, mosaic, grey spot, and rust. Various augmentation operations have been implemented in an effort to develop a more generalizable dataset. They achieved a detection performance of 78.80% mAP.

The detection of Northern maize leaf blight[5] disease in maize under complex field conditions was proposed by Sun et al. using multi-scale feature fusion and the SSD algorithm. To this end, a dataset of 18,000 images captured by a camera on a five-meter boom or a UAV was used. The new approach consists of three steps: dataset preprocessing, fine-tuning of the model, and disease detection. The NLB dataset was used because it is calibrated by human plant pathologists and has high accuracy. Preparing data was mainly done to reduce the effect of intense light on image identification and improve detection accuracy. The new model was 91.83% accurate overall.

SECTION II.

III.Classification Scheme

In this section, we introduce a classification scheme and classify all the relevant works in the corresponding classes. Moreover, we provide[3] statistics of the most frequently utilizedcrops/data types/methods/algorithms/datasets/metrics of the surveyed papers. The classes that we employ in our classification scheme are the following:

•Crop: the crop type that every study employs as a case study

•Input data: the nature of the input data utilized in the ML algorithms

•Dataset: the name of the dataset utilized if such information is presented in the paper

•Models/Algorithms: the ML algorithm utilized

•Method: whether classification (C) or object detection is addressed (C)

•Metrics: the performance metrics utilized to quantify the performance of the ML algorithms

•Results: a concise description of the results achieved by the ML algorithms

FIGURE 1. Visualization of the PlantDoc dataset with image annotations

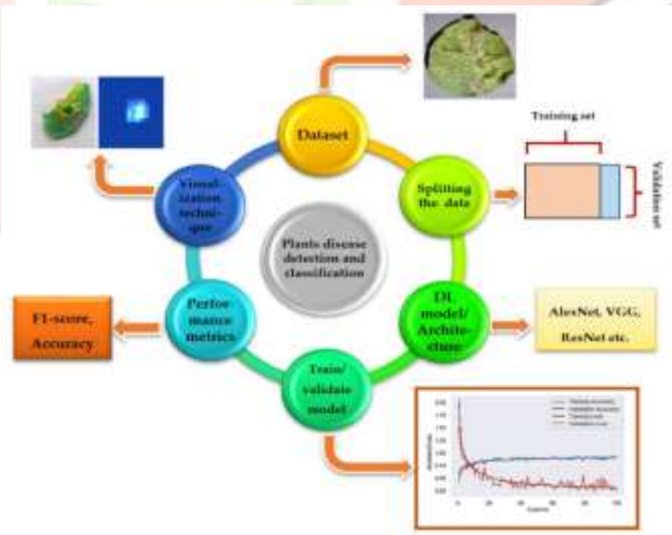

FIGURE 2. Stages of plant disease recognition

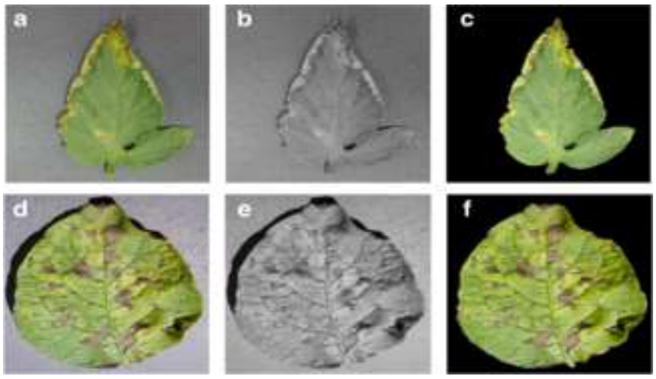

FIGURE 3. plant disease recognition and classification

SECTION III.

I.Discussion

In computer vision, the choice process of appropriate methodologies typically considers trade-offs among classification and object detection approaches. Classification as an approach is a fundamental machine learning concept with the benefits of simplicity and interpretability. It is simple to classify inputs into well-defined classes. It is also computationally efficient, hence very effective for limited resource applications.[6] Nonetheless, the classification approach has its own shortcomings. It is missing fine-grained information, including spatial information or multiple object handling capability. It is thus less ideal for applications that demand accurate localization or detection of multiple objects. Object detection approaches, however, not only return class labels but also include accurate spatial details, which solve the localization issue. Practically, some of these projects are fine illustrations of the application[7] of such advanced algorithms in precision agriculture:

• PlantVillage uses computer vision and machine learning to diagnose crops for disease. Farmers are able to photograph the crops using PlantVillage's simple mobile phone app and obtain robotic diagnoses for disease. This project shows an example of a classification algorithm implemented in real-time disease identification benefitting small farmers.

•Object detection is depicted by John Deere's See & Spray technology. It applies[8] computer vision and AI to detect and precisely identify specific weeds, reducing the use of blanket herbicide spraying. This project exemplifies how object detection algorithms optimize the use of resources.

• Blue River Technology uses machine vision and computer learning to drive their See & Spray system, self-detecting and selectively spraying herbicides only where they're needed. The project demonstrates[9] real-world use cases of object detection algorithms in pursuing more sustainable, cost-effective crop management.

• IBM's Smart Agriculture solution uses AI and machine learning to provide crop health insights and disease detection. With the use of data from[15] diverse sources such as satellite imaging and IoT sensors, it presents farmers with actionable recommendations. The project demonstrates how innovative algorithms can be combined with a data-centric strategy in precision agriculture.

SECTION VI.
II.Challenges in Plant Disease Detection

After a thorough examination of ML and DL algorithms for plant disease classification and[10] detection and the exhaustive computational effort on five state-of-the-art plant disease detection object detection algorithms and eighteen state-of-the-art plant disease classification algorithms on a widely used dataset, we have identified a few issues in real-world applications of plant disease detection:

1. There are not many models that handle non-image data. Most of the current object detection and classification algorithms handle only image data without taking into account other useful information such as temperature and humidity.[11] Methods to handle non-image data are required for more precise predictions.
2.There are few completely annotated open datasets. Many studies are founded on the PlantVillage dataset, which was obtained under a controlled lab environment. It is crucial to create larger datasets under conditions reflecting the real world. Collective effort is required to create representative datasets.
3. Most works treat the disease detection task as a classification task, either multi-class or binary classification. Even though most works[12] treat disease detection as a classification task, object detection must be treated with greater importance to detect both the disease[13] type and region of interest in the image.
4. A majority of the papers employ one dataset that is used for testing and training the model. Single-dataset trained models perform very badly on other datasets. Care should be taken to investigate other datasets in order to increase the robustness of the model.
5. Too much reliance on CNN architectures: While CNNs yield excellent performance, other neural network architectures like recurrent neural networks should be explored in order to make disease detection methods richer.
6. Early-stage disease detection and small leaf identification: Current databases are mostly filled with images having large leaves. Early-stage disease detection and small leaf identification need to be annotated in current datasets.
7. Occlusion and illumination issues: Current algorithms cannot tackle images with varying illumination and occlusion. More robust approaches are needed to address these issues.
8.Computational efficiency: Most models are computationally demanding, limiting real-time processing. Computational efficiency of their models should be improved by the researchers.

SECTION V.

III.Future Directions in Plant Disease Detection

Aside from the above challenges, there are a number of potential avenues for future research in plant disease detection:
1. Integration of non-image data: Design models that can properly incorporate non-image data, e.g., environmental conditions, into disease detection models to enhance [14]prediction accuracy.
2. Development of diverse and real-world datasets:[15] Work with experts to create large, representative datasets under actual agricultural conditions to improve model generalizability.
3.\tFocus on object detection: Investigate the prospect of object detection techniques in estimating plant diseases that can offer greater localization detail for disease.
4.\tDatasets robustness: Create models that work steadily well on any dataset to warrant their real-world applicability.
5.\tInvestigation into other neural[16] network architectures: Conduct experiments using alternate neural network structures like recurrent neural networks to unveil their promise of disease detection.
6. Early-stage and small leaf identification: Label datasets specifically for early-stage disease identification and the detection of diseases on plants or leaves of small sizes.
7. Handling illumination and[17] occlusion issues: Implement methods to improve the resilience of algorithms under varying lighting conditions and occluded images.
8. Enhanced computational efficiency: Prioritize optimizing model structures and algorithms to make them appropriate for real-time use.

SECTION VI.

I.Conclusion

The purpose[18] of this research is to review existing literature which employs ML and DL methods in precision farming, specifically plant disease detection and classification techniques. Moreover, a new classification system is proposed, dividing all applicable works into their corresponding classes. We distinguish between the studies into two broad categories based on the[19] methodology that they adopt (i.e. classification and object detection). In addition, we introduce the existing datasets for plant disease detection and classification, and give information regarding their classes and data, and whether the particular dataset can be used for classification or object detection.

In future research, we hope to investigate more algorithms for object detection and classification on more datasets and observe if the results are robust across various datasets.[20] We hope to investigate some image preprocessing and data augmentation strategies to observe if the accuracy of the algorithms can be enhanced with these methods.

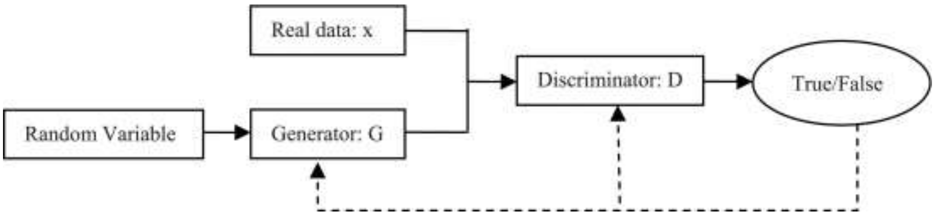

FIGURE 4. Flowchart of plant disease recognition

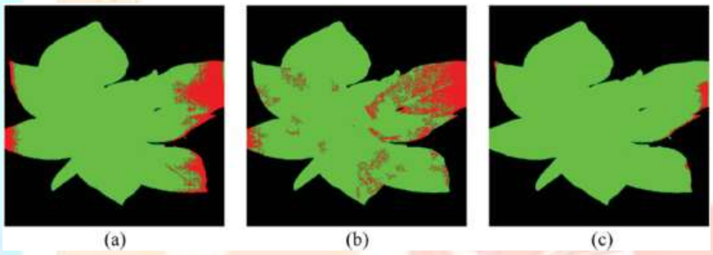

FIGURE 5. Output of plant disease recognition

REFERENCES

1.New Standards to Curb the Global Spread of Plant Pests and Diseases, Nov. 2019, [online] Available: https://www.fao.org/news/story/en/item/1187738/icode/.
Show in Context Google Scholar
2.Y. Mekonnen, S. Namuduri, L. Burton, A. Sarwat and S. Bhansali, "Machine learning techniques in wireless sensor network based precision agriculture", J. Electrochem. Soc., vol. 167, no. 3, Jan. 2020.
Show in Context CrossRef Google Scholar
3.L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review", Sensors, vol. 21, no. 11, pp. 3758, May 2021.
Show in Context CrossRef Google Scholar
4.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.Show in Context View Article Google Scholar
5.J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: A review", Plant Methods, vol. 17, no. 1, pp. 1-18, Dec. 2021. Show in Context CrossRef Google Scholar
6.Y. Yuan, L. Chen, H. Wu and L. Li, "Advanced agricultural disease image recognition technologies: A review", Inf. Process. Agricult., vol. 9, no. 1, pp. 48-59, Mar. 2022. Show in Context CrossRef Google Scholar
7.S. P. Mohanty, D. P. Hughes and M. Salathé, "Using deep learning for image-based plant disease detection", Frontiers Plant Sci., vol. 7, pp. 1419, Sep. 2016. Show in Context CrossRef Google Scholar
8.A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Proc. Adv. Neural Inf. Process. Syst., vol. 25, pp. 1097-1105, 2012. Show in Context Google Scholar
9.C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., "Going deeper with convolutions", Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 1-9, Jun. 2015.Show in Context View Article Google Scholar

10.Plantvillage Dataset, Nov. 2018, [online] Available: https://www.kaggle.com/datasets/emmarex/plantdisease.
Show in Context Google Scholar

11.Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al., "Caffe: Convolutional architecture for fast feature bedding", Proc. 22nd ACM Int. Conf. Multimedia, pp. 675-678, Nov. 2014. Show in Context CrossRef Google Scholar

12.S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification", Comput. Intell. Neurosci., vol. 2016, pp. 1-11, May 2016. Show in Context CrossRef Google Scholar

13.J. Amara, B. Bouaziz and A. Algergawy, "A deep learning-based approach for banana leaf diseases classification", Proc. Datenbanksys. Für Bus. Technol. Web (BTW) Workshopband, pp. 1-24, Jul. 2017.Show in Context Google Scholar

14.Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, et al., "Backpropagation applied to handwritten zip code recognition", Neural Comput., vol. 1, no. 4, pp. 541-551, Dec. 1989.Show in Context View Article Google Scholar

15.M. Brahimi, K. Boukhalfa and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization", Appl. Artif. Intell., vol. 31, no. 4, pp. 299-315, Apr. 2017.Show in Context CrossRef Google Scholar

16.M. H. Saleem, J. Potgieter and K. M. Arif, "Plant disease detection and classification by deep learning", Plants, vol. 8, no. 11, pp. 468-489, Oct. 2019.Show in Context CrossRef Google Scholar

17.L. C. Ngugi, M. Abelwahab and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition—A review", Inf. Process. Agricult., vol. 180, pp. 26-50, Apr. 2020.Show in Context CrossRef Google Scholar

18.W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", Bull. Math. Biophys., vol. 5, no. 4, pp. 115-133, Dec. 1943.Show in Context CrossRef Google Scholar

19.A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Proc. Adv. Neural Inf. Process. Syst., vol. 25, pp. 1097-1105, Sep. 2012.Show in Context Google Scholar

20.K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition", Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 770-778, Jun. 2016.Show in Context View Article Google Scholar.