A Review on Deep Learning Based Image Classification of Plant Diseases

Praveen S. Muthukumarana and Achala C. Aponso

Abstract—**A large portion of agricultural crop yield is lost due to plant diseases. The impact of this is more severe in developing countries that do not have sufficient trained professionals to identify and treat diseases. Deep learning has shown promising results in the field of image classification and is adopted in fields such as medicine. However, its' adoption in the field of agriculture has been slow in comparison. There are many examples in literature that had trained deep learning models to detect plant diseases by images. However, there is still no successful application developed that works in the real world. In this paper, the authors review the research efforts that have been done in the area of image-based plant disease detection with deep learning and try to analyze the challenges faced in adopting it in the agricultural sector. The authors examine datasets used, image pre-processing conducted and deep learning technologies utilized.**

Index Terms—**CNN, computer vision, disease classification, plant diseases, visible symptoms, automatic identification.**

I. INTRODUCTION

According to Strange *et al*. [1] plant diseases cause an estimated loss of 16% of agricultural annual crop yield globally and leads to famines and food crisis worldwide. Swift [2] states that spread of plant diseases is also one of the few reasons accepted by the World Trade Organization for blocking importation of agricultural produce. According to the Federation of American Scientists [3] this causes significant loss of revenue to nations. In developed countries there are more professionals in the agriculture sector to correctly diagnose and treat crop diseases before it becomes an epidemic, there are safety nets to support farmers who are affected and food reserves are maintained to avoid famines if a major food crop is affected by a disease. Vurro *et al*. [4] notes that most developing countries do not have such resources and therefore are severely affected by outbreaks of plant diseases.

Sladojevic *et al.* [5] states that most plant diseases show symptoms in the visible spectrum. Therefore, majority of diseases can be diagnosed by visual examinations by professionals in the field of agriculture. As stated by Vurro *et al*. [4] developing countries lack sufficient trained professionals to meet the demand. Due to recent advancements of deep learning in the field of computer vision, it is a prime candidate to democratize the tools required to accurately diagnose plant diseases.

Smartphones have become as powerful as computers in their processing power and memory according to Boulos *et al.*

[6] and has high quality cameras. Boissin *et al*. [7] states that modern smartphone cameras are powerful enough to be used for imaged-based teleconsultations in medical practice instead of digital cameras. As observed by the Economist [8] price of smartphones have also significantly dropped making them accessible to a wider demographic of society even in developing countries. Due to the state of processing, memory, camera and affordability of modern smartphones it is a good hardware choice for applications that diagnose plant diseases.

II. DEEP LEARNING ARCHITECTURES USED IN LITERATURE

A. Convolutional Neural Networks

According to Kamilaris *et al*. [9] Convolutional Neural Networks (CNN) are a type of deep, feed-forward Artificial Neural Networks (ANN) which are widely used in the literature for computer vision based tasks with high accuracies. According to Amara *et al*. [10] and Ramcharan *et al*. [11], one of the main advantages of deep learning methods such as CNNs over traditional machine learning methods is the lack of need for extracting feature manually which is a time consuming and labour intensive process. CNNs are able to learn the features automatically by convolving multiple filters across the image pixels.

B. Imagenet Dataset

As mentioned by Russakovsky *et al*. [12] the improvements CNN architectures within the past decade can be mainly credited to the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC). ImageNet is a public dataset that contains millions of annotated images belonging to around thousand object classes. This dataset is now used as a benchmark to evaluate the effectiveness of different computer vision algorithms. The annual challenge sets challenges based on either image classification, object detection or single-object localization.

C. Architectures Used in the Literature

Some studies in the literature have repeated plant disease classification experiments with multiple deep learning architectures to compare their performance. The findings are consistent with how these algorithms have performed in the ILSVRC. AlexNet described by Krizhevsky [13] won the ILSVRC challenge in 2012. GoogleNet described by Szegedy [14] won ILSVRC in 2014 by improving upon AlexNet. Mohanty *et al*. [15], Brahimi *et al*. [16], Fuentes *et al*. [17] compared the CNN architectures AlexNet and GoogleNet using the PlantVillage dataset and found GoogleNet to perform better than AlexNet.

Brahimi *et al*. [16] repeated the experiment with AlexNet, GoogleNet and the shallow neural network SVM. GoogleNet

and AlexNet both performed significantly better than the shallower networks. Ramcharan *et al*. [11] created a model with Inception v3 architecture (an improved version of GoogleNet) and repeated the experiment with the final output SoftMax layer of the model replaced by a Support Vector Machine (SVM) and K-nearest neighbor (KNN). Both SVM and original SoftMax had average classification accuracies above 90% while KNN performed the worst with 71% average classification accuracy.

Ferentinos [18] trained a model with another CNN architecture VGG [19] and found it to perform better than both GoogleNet and AlexNet.

Too *et al*. [20] repeated the experiment for ResNet, DenseNet and VGG 16 and found The former two to perform better than VGG 16. However, Fuentes *et al*. [17] who trained models with tomato plant disease dataset reports that VGG-16 outperforms different versions of ResNet architecture such as ResNet-50, ResNet-101, ResNet-152 and ResNeXt-50.

From studies done in the literature it is seen that modern deep learning model architectures outperform shallow networks. Also, performance of the models in classification of plant diseases is consistent with how the performance of these models are compared generally. For example, their performance in ILSVRC.

D. Transfer Learning

It is expensive and time consuming to create large image datasets for plant diseases. It is more difficult for diseases that are rare. Kaya *et al*. [21] and Too *et al*. [20] states that with a limited dataset, training a network from scratch is not efficient and leads to overfitting. Many authors such as Sladojevic *et al*. [5], Mohanty *et al*. [15], Brahimi *et al*. [16], Ferentinos *et al*. [18], Too *et al*. [20], Kaya *et al*. [21] use transfer leaning to overcome this problem.

When using a re-trained network, the first layers use weights pre-trained with a large dataset such as ImageNet dataset. This will be used to extract useful general features. The final layers are modified to detect the specific features inherent to the image class. The model is then retrained with a dataset to update the weight values. Transfer learning could reduce training time according to Ramcharan [11] and overfitting according to Barbedo [22]. Mohanty *et al*. [15], Brahimi *et al*. [16] reports significant improvement in accuracy when the model is pre-trained compared to training it from scratch. Kaya *et al*. [21] states that specially when having smaller datasets, classification accuracy is higher with transfer learning than without.

III. FACTORS AFFECTING CLASSIFICATION ACCURACY

The literature shows deep learning models can be trained to classify plant diseases using image datasets with high level of accuracy. Many of the reviewed literature on plant disease detection use the PlantVillage dataset for training. Mohanty *et al*. [15], Brahimi *et al*. [16], Ferentinos [18], Barbedo [23], Too *et al*. [20] are some examples. According to Too *et al*. [20] PlantVillage is a free dataset created by Penn State University containing 54,306 images of 26 different plant diseases of 14 crops. Other methods discussed in the literature uses datasets created by the authors themselves.

Ramcharan *et al*. [11], Amara *et al*. [10], Fuentes *et al*. [17], Habaragamuwa *et al*. [24], Oppenheim [25] are some who have done this.

In all the studies, they divide the dataset into a training and testing sets. After training the model, the accuracy of the model is found by using the test dataset and testing is not repeated with different datasets. However, Mohanty *et al*. [15] and Ferentinos [18] observe that even though deep neural networks models trained using the PlantVillage dataset achieve classification accuracies exceeding 90%, the accuracy drops significantly when tested on images outside the PlantVillage dataset and taken in different conditions. For Mohanty *et al*. [15], accuracies dropped from being above 90% to just above 31%. Barbedo *et al*. [23] notes that this is likely due to several conditions other than the symptom regions which get picked up by the training process. So, the model works fine when used on images from the same dataset with the same conditions. But when tested on images taken on different days, different locations and different capture conditions, the model performs poorly. Therefore, to make advancements in the field it is important to discover what these conditions are and study on how to mitigate them.

A. Images Taken in Field Conditions vs Laboratory Conditions

Majority of the images in the PlantVillage dataset which is used in most studies are taken in laboratory conditions. This means many of the factors such as angle of capture, background, size of symptom region and light conditions are controlled. (Shown in Fig. 1).

Ferantinos [18] trained one model with images taken in controlled laboratory conditions and another model with images taken in field conditions (such as Fig. 2) for the same disease. When tested on images taken in the field, the model achieved better performance when it was trained on images taken in field conditions. Since users will be taking images in field conditions, it is important to capture the training images in the field itself. This should be taken into consideration when creating datasets in the future.

B. Impact of the Image Background

When images are captured in laboratory conditions such as in Fig. 1, a uniform background could be maintained. However, as discussed in the previous chapter this is not practical.

According to Barbedo [23] studies conducted with traditional machine learning algorithms always removed the background of the images in the training dataset before training the model. Amara [10] states that in these approaches' features used for classification were hand-crafted. Therefore, additional elements in the image were removed to prevent them interfering. In contrast deep learning techniques such as CNN's automatically creates the features necessary for classification. Ferantinos *et al*. [18] states that since deep learning architectures such as Convolutional Neural Networks (CNNs) can identify important and non-important features from images, there is little risk of the model learning unnecessary background features.

Mohanty *et al*. [15] segmented the images to remove the extra background from the images with the assumption that it would improve classification accuracy by removing

distracting background features. But it resulted in no significant difference. However, it should be noted that most PlantVillage dataset images were taken in laboratory conditions such as Fig. 1 and therefore, will not have very complex backgrounds to begin with. Ramcharan *et al*. [11] used images taken in field conditions with complex backgrounds containing items such as soil, sky, other vegetation, feet and hands but the model produced high classification accuracies despite this. Therefore, Ramcharan notes that it is not necessary to remove the background when using deep learning. Barbedo [23] found that removing the background improved classification accuracy from 76% to 79%. The author notes that in the dataset he used, the background had elements that mimicked the plant symptoms. For example, Fig. 4 shows soil having similar colour to the disease symptom on the leaf. This suggest that removing the background can be useful in specific scenarios and it should be reviewed in case by case basis.

There are several things that can be done to reduce the background busyness when taking images for the training dataset. One is to capture the images such that the symptom region encompasses majority of the image. Another simple solution is to hold a piece of single coloured card behind the plant part when capturing the image.

To test the impact on classification accuracy of these changes, the model should be tested on images with busy backgrounds. End user can also be prompted to crop the region of interest to minimize the impact of the background. With the prevalence of touch controlled mobile devices this would not be a difficult task. However, if they crop out too much of the image including parts of the symptom region, this would have a negative effect by reducing accuracy due to the model not having enough information.

C. Impact of Dataset Size

It is expensive and time consuming to make large databases. However, Kamilaris *et al*. [9] states that at least a few hundred images are required per disease class for an accurate diagnosis.

Many methods in the literature perform image augmentation to increase the dataset size and reduce overfitting. The augmentation methods used in the literature are rotation performed by Zeng *et al*. [26], Fuentes *et al*. [17], Barbedo [23] Barbe *et al*. [5], cropping performed by Zeng *et al*. [26], Habaragamuwa *et al*. [24], mirroring by Habaragamuwa *et al*. [24], Barbedo [23], contrast and brightness adjustment by Fuentes [17] and Barbedo [23], affine transformations and perspective transformations by Barbe *et al*. [5]. Fuentes *et al*. [17] found the accuracy to increase from 0.5564 to 0.8306 by increasing the dataset size by data augmentation. Barbedo [23] divided each image in the dataset into to multiple images containing individual symptom regions to increase dataset size and see how the CNN would perform with more localized information (shown in Fig. 5). This significantly improved classification accuracy from 76% to 87% compared to original dataset. To see if the improvement is solely due to increase in dataset size or also due to the effect of cropping the image to only contain the Region of Interest (ROI), experiment was repeated with the cropped image dataset being the same size as the original dataset. Accuracy fell to 81% but was still higher than the

original dataset. This showed that both cropping the symptom region as well as increasing the dataset size had an impact in increasing accuracy. Ramcharan *et al*. [11] manually cropped images of each cassava leaf into individual leaflets as shown in Fig. 3 to increase the dataset size almost 7 times. However, this did not cause significant improvement in classification accuracy which was already high for original dataset. Since this is a time-consuming process, it should not be pursued unless the model is performing poorly due to lack of data in the dataset.
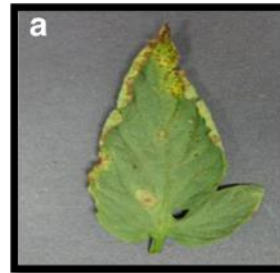

Fig. 1. Example image from PlantVillage dataset.
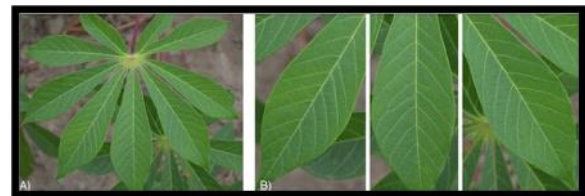

Fig. 2. Image taken in field conditions.

Fig. 3. Original leaf image (left) and cropped leaflet images (right).
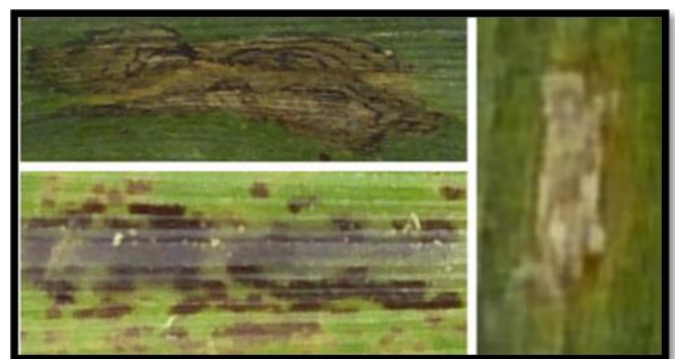

Fig. 4. Background elements mimics symptoms.

Fig. 5. Symptom regions in image divided into different images.

D. Variation on How Symptoms Appear

As stated by Kamilaris [9] most studies had images of upper surface of the plant leaf. But diseases can start

appearing in other plant parts such as the stem or fruit. According to Barbedo [23] some disorders can show varying symptoms based on the stage of the disease. Some disorders produce visually similar symptoms. In these cases, an image taken in the visual spectrum might not be enough for accurate detection. A plant can also show symptoms from multiple disorders. According to Bharali *et al.* [27] the plant is weakened by a disorder, it is more susceptible for other injections as it's immune system is weakened.

Most of these issues which are difficult to tackle. It is not practical to create datasets that cover all possible ways a symptom can show or all possible combinations of different symptoms although it should be attempted much as possible. One solution is to gradually increase the dataset size overtime with user captured images. Furthermore, if the model cannot make a prediction with high accuracy, the user could be instructed to perform additional tests to rule out the possible diseases. For example, according to Champoiseau [28], bacterial wilt in the tomato can be distinguished from other diseases that cause similar wilting by cutting the part of the infected stem and dipping it a transparent container with water. The infected stem would discharge a white ooze.

Other Capture Conditions

Images in the field can be taken in different light conditions depending on factors such as weather, cloud cover, time of day geographical region. Bharali *et al.* [26] have tried to control the illumination conditions by either conducting the experiments in laboratory conditions and Peressotti [29] by using artificial light sources. However, for deep learning techniques such controlling is not required. Future studies in the field should have images in the training dataset that are taken in different times of the day under different light conditions to account for this variation. A more specific problem related to illumination is specular light. According to Oppenheim [27] this is a high intensity reflection which occurs when light hits the surface at certain angles. Information is lost in areas of the image affected by this. The most practical way to avoid this is to position the camera in an angle that avoids it.

Angle at which the image is taken is also a factor to be considered. Peressotti *et al.* [29] states that ideally the leaf should be perpendicular to the camera. However, Oberti *et al.* noted that 40-60 degrees angle was the most appropriate when detecting powdery mildew in grapevine leaves. The author states that the reason could be due to the fact that at initial stages the filamentous structures of the fungus grow vertically in the plant tissue therefore it is detected better when the image is taken at a slanting angle. So, the type of disorder should be considered when deciding some factors such as the angle of capture.

IV. Discussion

Reviewing the existing literature, it is seen that deep learning models created with CNN architectures have obtained high classification accuracies for diagnosing plant diseases. But these models tend to not generalize for images captured under different conditions. The problem seems to lie on the high variation of multiple conditions in the images taken. This paper has tried to explore these different conditions and on how to mitigate the effect of them. Future research must focus on analysing these issues and finding out optimum solutions. More public datasets should be created in light of the limitations of the PlantVillage dataset for future studies.

Conflict of Interest

The authors declare no conflict of interest.

Author Contribution

Author Praveen performed the research involving existing implementations for deep learning for plant disease detection in the literature. Author Achala provided insight about deep learning models and techniques present in the literature and evaluated their performance. All authors have approved the final version.

References

[1] R. N. Strange and P. R. Scott, "Plant disease: A threat to global food security," *Annu. Rev. Phytopathol.*, vol. 43, pp. 83–116, 2005.

[2] J. Swift, "Why are rural people vulnerable to famine?" *IDS Bull.*, vol. 20, no. 2, pp. 8–15, 1989.

[3] Federation of American Scientists, *The Economic Impact of Crop and Livestock Diseases*, 2014.

[4] M. Vurro, B. Bonciani, and G. Vannacci, "Emerging infectious diseases of crop plants in developing countries: Impact on agriculture and socio-economic consequences," *Food Secur.*, vol. 2, no. 2, pp. 113–132, 2010.

[5] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, 2016.

[6] M. N. K. Boulos, S. Wheeler, C. Tavares, and R. Jones, "How smartphones are changing the face of mobile and participatory healthcare: An overview, with example from eCAALYX," *Biomed. Eng. OnLine*, vol. 10, p. 24, Apr. 2011.

[7] C. Boissin, J. Fleming, L. Wallis, M. Hasselberg, and L. Laflamme, "Can we trust the use of smartphone cameras in clinical practice? laypeople assessment of their image quality," *Telemed. J. E Health*, vol. 21, no. 11, pp. 887–892, Nov. 2015.

[8] The Economist, "The rise of the cheap smartphone," *The Economist*, Apr. 05, 2014.

[9] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, Apr. 2018.

[10] J. Amara, B. Bouaziz, and A. Algergawy, "A deep learning-based approach for banana leaf diseases classification," *BTW*, 2017.

[11] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Front. Plant Sci.*, vol. 8, 2017.

[12] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.

[15] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, 2016.

[16] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, K. Boukhalfa, and A. Moussaoui, "Deep learning for plant diseases: detection and saliency map visualisation," in *Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent*, J. Zhou and F. Chen, Eds. Cham: Springer International Publishing, 2018, pp. 93–117.

[17] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A Robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, p. 2022, Sep. 2017.

[18] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, pp. 311–318, 2018.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *ArXiv14091556 Cs*, Apr. 2015.

[20] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, vol. 161, pp. 272–279, Jun. 2019.

[21] A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "Analysis of transfer learning for deep neural network based plant classification models," *Comput. Electron. Agric.*, vol. 158, pp. 20–29, Mar. 2019.

[22] H. Habaragamuwa, Y. Ogawa, T. Suzuki, T. Shiigi, M. Ono, and N. Kondo, "Detecting greenhouse strawberries (mature and immature), using deep convolutional neural network," *Eng. Agric. Environ. Food*, vol. 11, no. 3, pp. 127–138, Jul. 2018.

[23] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosyst. Eng.*, vol. 172, pp. 84–91, Aug. 2018.

[24] D. Oppenheim and G. Shani, "Potato Disease classification using convolution neural networks," *Adv. Anim. Biosci.*, vol. 8, no. 2, pp. 244–249, Jul. 2017.

[25] B. Liu, Y. Zhang, D. He, and Y. Li, "Identification of apple leaf diseases based on deep convolutional neural networks," *Symmetry*, vol. 10, no. 1, p. 11, Jan. 2018.

[26] P. Bharali, C. Bhuyan, and A. Boruah, "Plant disease detection by leaf image classification using convolutional neural network," in *Information, Communication and Computing Technology*, Singapore, 2019, pp. 194–205.

[27] J. G. A. Barbedo, "A review on the main challenges in automatic plant disease identification based on visible range images," *Biosyst. Eng.*, vol. 144, pp. 52–60, Apr. 2016.

[28] P. Champoiseau and T. Momol. (2009). R. solanacearum/ Bacterial wilt—Bacterial wilt of tomato. [Online]. Available: https://plantpath.ifas.ufl.edu/rsol/Trainingmodules/BWTomato_Module.html

[29] A. Clément, T. Verfaille, C. Lormel, and B. Jaloux, "A new colour vision system to quantify automatically foliar discolouration caused by insect pests feeding on leaf cells," *Biosyst. Eng.*, vol. 133, pp. 128–140, May 2015.

Praveen S. Muthukumarana is an undergraduate student following a degree in software engineering at the Informatics Institute of Technology, Colombo, Sri Lanka. He has worked as a trainee web developer for a year. He got interested in the field of plant disease detection while being involved in a project developing a system detecting plant diseases using images.

Achala C. Aponso holds a master degree in the field of artificial intelligence from the University of Moratuwa, Colombo, Sri Lanka. He is currently a lecturer at the Informatics Institute of Technology, Colombo, Sri Lanka. He worked as a software engineer at prior to his tenure as a lecturer. His research interests are in deep learning and computer vision.