**Paper No: SC-08** **Smart Computing**

# Novel deep learning approaches for crop leaf disease classification: A review

E. M. T. Y. K. Ekanayake*
*Postgraduate Institute of Science*
*University of Peradeniya, Sri Lanka*
e.thissa@gmail.com

R. D. Nawarathna
*Department of Statistics and Computer Science*
*University of Peradeniya, Sri Lanka*
ruwand@pdn.ac.lk

***Abstract*** **- To encourage sustainable progress, it is suggested that in a world connected by virtual platforms, modern society should merge big data, artificial intelligence, machine learning, information and communication technology (ICT), as well as the "Internet of Things" (IoT). When real-life problems are considered, the above technology processes are essential in solving the issues. Food is an essential need of human beings. Food supply has become crucial, and it is very important to increase the adequate cultivation of plants for large populations due to huge population growth. At the same time, farmers are struggling with a variety of food plant diseases that significantly affect the harvesting and production in agricultural fields. Nevertheless, the agricultural productivity of rural areas is directly involved with the increase in the economic growth of developing countries such as Sri Lanka, India, Myanmar and Indonesia. Early identification of crop disease, using a well-established modern technique, is vital. It necessitates a number of processes observing large-scale agricultural fields as a disease can infect different parts of the plant such as leaf, roots, stem and fruit. Most diseases appear in plant leaves and have the potential to spread them all over the field within a very short time. This paper reviews several state-of-the-art methods that can be used for plant leaf disease recognition with a special reference to deep learning based methods.**

***Keywords - attention mechanism, Deep Learning, disease identification, image processing, Machine Learning***

## I. INTRODUCTION

Most Asian economies are based on agriculture. When people enhance food plant productivity, this often results in a degradation of agricultural fields due to being ignorant of the natural environmental impact on the plantation process. Because of crop plant pathogens such as fungus, organism, virus, bacterial infections, phytoplasmas, plant disease cannot be neglected. Therefore, identification of the crop plant disease is the main objective in the agricultural field. When a disease arises because of the above pathogens in any type of plant systems, it may infect all parts of the plant, including its leaves, roots, stems, crowns, tubes, flowers, fruits and seeds. Consequently, the identification and classification of the disease at an early stage is crucial. Direct observation of the field by crop experts is a common approach in the detection and identification of crop diseases, but this solution is an obsolete method. In addition, identifying the disease by monitoring the fields by experts will be extremely expensive in the large-scale farming industry. To take a better solution, we can analyse images of the crop plant leaf disease using image processing technology. This may include extracts of the feature of the diseased area in terms of colour, texture, shape and other appearances from a measurable point of view in the plant area.

According to the level of expertise required, the cost of supervision will be high and time-consuming. A solution which uses an image processing technique, will assure more benefits in monitoring huge scale agricultural fields. Furthermore, this automatic identification of the crop disease, by analysing the symptoms of the related plant parts, makes the process both simple and economical. It will require computer vision to deliver an image-based programmed procedure control, the examination process, and the automation of robotic supervision.

Identifying crop diseases in a visual image is a difficult task and the accuracy of the identification can also become less valuable. This method can be used only in selected places. Using an automatic leaf disease identification technique will reduce the time and it will be more accurate, with less effort. When we consider the food plants, infected diseases are generally revealed by brown or yellow spots, early and late burn-patches, fungus, bacterial or virus diseases. The image processing technique is the way to measure the area affected by the disease, or determine the colour differences between a good location and the affected area. A few methods based on colour identification feature and K-means algorithm and threshold values are used for the segmentation process and identifying the disease.

The classification of a digital image process refers to the feature extraction information task from raster images. The resultant raster from the image classification process enables us to make a scale map. Supervised learning and unsupervised learning are the primary classification methods. Currently, there are a variety of ways to perform digital image classification interacting with thresholding methods. Most methods depend on colour identification, boundary detection, and the segmentation of digital images. Machine learning-based methods for crop disease identification and classification have become an important part of modern developments. Nowadays, most researchers tend to use new machine learning-based methods instead of traditional methods.

The main objective of this review is to suggest a better deep learning method for the identification and classification of plant leaf diseases at an early stage. In addition, it aims to compare and contrast the plant disease classification technologies with the latest deep learning methods, to verify the importance of the dataset of each method used, in order to assess the relevance of future enhancements for real world scenarios.

This paper is organized as follows. Section II presents a review of related literature. Section III summarizes the dataset, the proposed solutions' approach, and potential improvements. Section IV compares and contrasts the average accuracy of the different methods in brief. Finally concluding remarks are given in Section V.

## II. Review of techniques for crop leaf disease identification

An advanced attention mechanism is suggested in the paper [1] that successfully operates the informative areas of an input image. Also, the method explains the usage of transfer learning to construct some fine-grained image classification model based on a developed attention mechanism. Close-grained detailed image classification is an exacting task due to the difficulty in recognizing distinguishing features. When the input image is a fully represented object, finding a suitable method is not an easy task. In this particular classification, the model considers visual disturbance such as overlapping and external light. To use this model for crop leaf disease identification, it should concentrate on the detailed regents of the input images.

The researchers have experimented with transfer learning with the convolutional neural network in the experiment [2]. The model modified a network layout to increase the learning ability of the plant disease characteristics. The MobileNet with the squeeze and extraction (SE) section was used in this experiment. To increase the qualities of both, the pre-trained MobileNet and SE section were embedded in the developed network called SE-MobileNet. The SE-MobileNet was the model used for the identification of paddy leaf diseases. The speciality of the model was the double usage of the transfer learning technique, which helped in gaining the optimal solution. There were two phases in this experiment. The first phase was training the SE-MobileNet for the extracted layers, and the end of the convolutional layers were stopped with the pre-trained shared weights on the ImageNet. The second phase was training the SE-MobileNet model using the target input dataset.

A classification and identification technique model constructed in [3] can be used in classifying crop leaf diseases. In this experiment, before the feature extraction process, pre-processes were completed. In the pre-processing section, all the RGB images were converted into grey level images to the next step, which was the feature extraction of the input image. The elementary morphological functions were applied as the second step on the input image. Then, the input image was converted into a binary level image. In the next stage, if the pixel value of the binary image was zero, the pixel was converted into a responsible RGB image value. Finally, using the Naïve-Bayesian classifier [4], the disease was identified.

Another novel approach [5] presents for the detection and classification of rice leaf viruses. It used K-means clustering, multiclass support vector machine (SVM) [6] and particle swarm optimization (PSO) [7]. Grey Level Co-occurrence Matrix (GLCM) was used for the feature extraction process. The virus classification was done using a Support Vector Machine (SVM) classifier, and the recognition of the virus accuracy was enhanced by optimizing the data with PSO. The paper [8] The performance of 13 CNN models for rice disease detection in transfer learning and deep features plus the SVM method is evaluated in this work. When compared to other models, the statistical analysis findings, deep characteristics of resnet50 [9][10] and SVM classification model are superior. A comparison of all classification models based on CNN and conventional techniques was conducted.

An interesting model using RGB image acquisition is presented in the alternative experiment [11] to detect any type of plant disease affected by different agricultural crops. Converting the input RGB image format into Hue-Saturation-Intensity (HSI) format [12] and masking and removing the green pixels in the input image makes it accessible to the segmentation process, using Otsu's method [13]. Then, the texture features were calculated using the colour co-occurrence method and finally the disease was classified with the Genetic Algorithm [14].

The crop disease identification and classification process using a convolution neural network is presented in the paper [15]. This includes three convolution layers and three pooling layers followed by two fully connected layers. The results of the experiment clearly show the efficiency of the constructed model approach over the pre-trained models such as VGG16 [16], MobileNet and InceptionV3 [17].

The experiment [18] focused upon the leaf disease segmentation and classification of a few plants. Firstly, the disease area from the input images was segmented with an introduced superpixel cluster-based hybrid neural network. Texture, colour and shape were the main features whereby input images were classified under different classes. The experiment [19] tried to resolve the rough image dataset problem. The method initially limited the leaf area by applying the colour features of the input image. The classification process of the input leaf image depended on the structures of discriminatory characteristics. The property of the input image features showed a variety of patterns in the leaf area. Then, the researchers applied the feature discriminable characteristics with the Fisher vector in terms of different orders of the diversity of Gaussian distribution. In the paper [20], the EfficientNet [21] deep learning method experimented with in-crop leaf disease identification. The model performance was compared with several newly developed deep learning models. To train for the purpose, the researchers used the PlantVillage dataset in this experiment. The EfficientNet method and other deep learning models were trained with the transfer learning technique. In the transfer learning technique, each layer in the models was set up as trainable.

## III. Materials, Methods and Proposed enhancement

The key components in the domain's reference study are the gathering of relevant material and the reviewing of the information with a competent analysis. In the first stage, the Google Scholar Web Science Indexing Facilities performed the keyword-based exploration for journal articles and conference papers. Two main search criteria were used to search the relevant articles. Those keywords were "plant disease classification " and " deep learning methods for plant disease identification" respectively. Initially, 10 articles were recognized. The selected articles were examined individually in the second stage. Key questions posed in analysis were: What was the dataset used? What were the disease categories in the dataset included? What methods were used and what was the level of average accuracy of the methodology they selected?

Table I shows a brief overview of the selected research papers on automatic crop disease identification, and their use of materials and methods. It summarizes the dataset, the methodology of the proposed solutions and future enhancement in the corresponding studies.

TABLE I. TABULAR LIST OF REFERENCE NUMBER OF REVIEWED PAPERS, THEIR METHODOLOGY AND FUTURE ENHANCEMENTS

| **Article Ref.** | **Dataset** | **Methodology** | **Future Enhancement** |
|---|---|---|---|
| [1] | PlantVillage public dataset | Transfer learning method and the NASNetLarge fine-grained model based on attention mechanism. | Train and test the model with more extensive image datasets from various geographical regions, field conditions, image capture modes, and multiple sources. |
| [2] | PlantVillage dataset and Fujian Institute of Subtropical Botany dataset | Twice Transfer learning and a modified deep CNN approach used the “MobileNet” with “Squeeze and Excitation” (SE) block. | Researchers want to use it on mobile devices to track and diagnose a variety of plant diseases. The model applies to other similar fields such as online defect assessment, molecular cell recognition, and identification of location from disparate pictures. |
| [3] | Not specified | K-means clustering [22], Basic Morphological functions, “Naïve Bayesian” classifier, “Colour Co-Occurrence” method. | None |
| [5] | Not specified | K-means clustering Multiclass SVM and “Particle Swarm Optimization” (PSO) technique. | Developing combinations of more algorithms with fusion classification methods to improve the recognition rate of the classification process. |
| [8] | 5932 field images | 11 CNN models in transfer learning approach and deep feature plus support vector machine (SVM) | Testing for more varieties of rice diseases and a more fine-tuned “Convolution Neural Network” model with the expectation of better performance. |
| [11] | Not specified | RGB to HSI conversion and thresholding. Segment the components using “Otsu’s method”. “Colour Co-Occurrence” method and “Genetic Algorithm” as a classifier. | None |
| [15] | PlantVillage public dataset | CNN based model | Due to the testing accuracy is lower; modify the model using a larger number of pictures and a different crop and procedure to improve the same model on the same dataset. |
| [18] | Shri Mata Vaishno Devi University Dataset | Seven different machine learning algorithms (LR, LDA, KNN, CART, RF, NB, SVM) with Simple linear iterative clustering (SLIC) [23], “Adaptive Linear Neuron” (ADALINE) [24], “Scale-Invariant Feature Transform” (SIFT) [25] | Improving the learning rate to increase the segmentation performance and adopting a deep neural network for classification using some nature-inspired algorithms. |
| [19] | Selected categories of PlantVillage public dataset | MLP [26] and SVM classifier | Classifying different plant diseases and improving the classification accuracy. |
| [20] | PlantVillage public dataset | EfficientNet deep learning model | Improved models enabling plant pathologists and farmers to identify plant diseases rapidly in mobile contexts. |

## IV. Results

The tabular list is presented below in Table II, including the accuracy value and classification technology that have been covered to achieve that level of accuracy. In addition, figure 1 represents accuracy values of the paper reference number in this review paper.

TABLE II. LIST OF REVIEWED PAPERS WITH ACCURACY VALUES AND USED METHODS

| **Article Reference Number** | **Classification Technology** | **Average Accuracy (%)** |
|---|---|---|
| 1 | NASNetLarge neural network model with Attention mechanism | 93.05% |
| 2 | Twice Transfer learning and the SE-MobileNet model | 99.33% |
| 3 | K-means clustering, basic morphological functions, Naïve Bayesian classifier, Colour Co-Occurrence method. | 87% |
| 5 | Multiclass SVM and Particle Swarm Optimization Technique | 97.91% |
| 8 | CNN based support vector machine (SVM) | 97.62% |
| 15 | Convolution Neural Network | 91.2%. |
| 18 | Computer Vision based approach | 98.57% |
| 19 | MLP and SVM classifier | 94.35% |
| 20 | EfficientNet deep learning model | 99.91% |

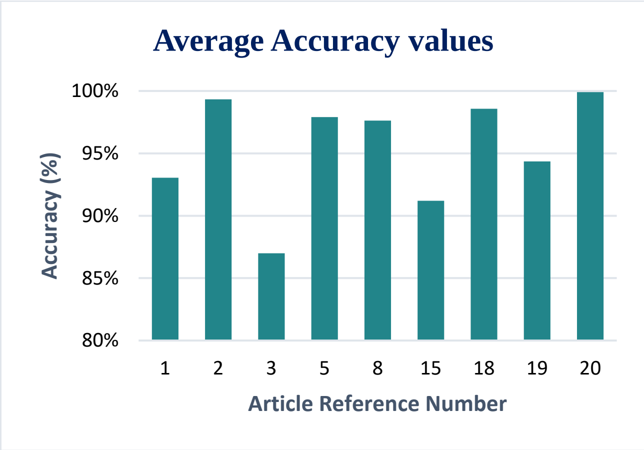


Fig. 1. Graph representation of accuracy values of reviewed papers

## V. CONCLUSION

This paper provides a survey of different disease classification methods that can be used for crop leaf disease identification. An algorithm machine learning technique for automatic detection and classification of crop leaf diseases is described later. Most researchers used the PlantVillage public dataset for the algorithms and testing methods. Therefore, diseases related to these plants were taken for identification. With shallow computational efforts, the optimal result was gained, which also demonstrates the algorithm's efficiency in the identifying and classifying plant leaf diseases. Identifying the crop leaf diseases in the early-stage or initial stage is the main advantage of those methods. To maximise the recognition rate in the classification process Artificial Neural Network, Computer Vision-based approach, a deep learning model can also be used.

## REFERENCES

[1] G. Yang, Y. He, Y. Yang, and B. Xu, "Fine-Grained Image Classification for Crop Disease Based on Attention Mechanism," Front. Plant Sci., vol. 11, no. December, pp. 1–15, 2020, doi: 10.3389/fpls.2020.600854.

[2] J. Chen, D. Zhang, M. Suzauddola, Y. A. Nanehkaran, and Y. Sun, "Identification of plant disease images via a squeeze-and-excitation MobileNet model and twice transfer learning," IET Image Process., no. May, pp. 1–13, 2020, doi: 10.1049/ipr2.12090.

[3] D. Mondal, A. Chakraborty, D. K. Kole, and D. D. Majumder, "Detection and classification technique of Yellow Vein Mosaic Virus disease in okra leaf images using leaf vein extraction and Naive Bayesian classifier," Int. Conf. Soft Comput. Tech. Implementations, ICSCTI 2015, pp. 166–171, 2016, doi: 10.1109/ICSCTI.2015.7489626.

[4] N. Boyko and K. Boksho, "Application of the naive Bayesian classifier in work on sentimental analysis of medical data," CEUR Workshop Proc., vol. 2753, pp. 230–239, 2020.

[5] Prabira Kumar Sethy, "Detection & Identification of Rice Leaf Diseases using Multiclass SVM and Particle Swarm Optimization Technique," no. 6, pp. 108–120, 2019.

[6] D. A. Pisner and D. M. Schnyer, Support vector machine. Elsevier Inc., 2019.

[7] F. Wang, H. Zhang, and A. Zhou, "A particle swarm optimization algorithm for mixed-variable optimization problems," Swarm Evol. Comput., vol. 60, p. 100808, 2021, doi: 10.1016/j.swevo.2020.100808.

[8] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep feature based rice leaf disease identification using support vector machine," Comput. Electron. Agric., vol. 175, no. May, p. 105527, 2020, doi: 10.1016/j.compag.2020.105527.

[9] I. Z. Mukti and D. Biswas, "Transfer Learning Based Plant Diseases Detection Using ResNet50," 2019 4th Int. Conf. Electr. Inf. Commun. Technol. EICT 2019, no. December, pp. 1–6, 2019, doi: 10.1109/EICT48899.2019.9068805.

[10] M. O. Ramkumar, S. S. Catharin, V. Ramachandran, and A. Sakthikumar, "Cercospora identification in spinach leaves through resnet-50 based image processing," J. Phys. Conf. Ser., vol. 1717, no. 1, 2021, doi: 10.1088/1742-6596/1717/1/012046.

[11] M. S. Arya, K. Anjali, and D. Unni, "Detection of unhealthy plant leaves using image processing and genetic algorithm with Arduino," EPSCICON 2018 - 4th Int. Conf. Power, Signals, Control Comput., pp. 1–5, 2018, doi: 10.1109/EPSCICON.2018.8379584.

[12] W. Yi, Z. Jing, and G. Shuang, "Hue–saturation–intensity and texture feature-based cloud detection algorithm for unmanned aerial vehicle images," Int. J. Adv. Robot. Syst., vol. 17, no. 3, pp. 1–8, 2020, doi: 10.1177/1729881420903532.

[13] P. Yang, "An improved Otsu threshold segmentation algorithm Wei Song *, Xiaobing Zhao and Rui Zheng Letu Qingge," vol. 22, no. 1, pp. 146–153, 2020.

[14] A. Tarafdar, B. P. Kaur, P. K. Nema, O. A. Babar, and D. Kumar, "Using a combined neural network ─ genetic algorithm approach for predicting the complex rheological characteristics of microfluidized sugarcane juice," Lwt, vol. 123, p. 109058, 2020, doi: 10.1016/j.lwt.2020.109058.

[15] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, "ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network," Procedia Comput. Sci., vol. 167, no. 2019, pp. 293–301, 2020, doi: 10.1016/j.procs.2020.03.225.

[16] A. Krishnaswamy Rangarajan and R. Purushothaman, "Disease Classification in Eggplant Using Pre-trained VGG16 and MSVM," Sci. Rep., vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-59108-x.

[17] C. Wang et al., "Pulmonary image classification based on inception-v3 transfer learning model," IEEE Access, vol. 7, pp. 146533–146541, 2019, doi: 10.1109/ACCESS.2019.2946000.

[18] S. S. Chouhan, U. P. Singh, U. Sharma, and S. Jain, "Leaf disease segmentation and classification of Jatropha Curcas L. and Pongamia Pinnata L. biofuel plants using computer vision based approaches," Meas. J. Int. Meas. Confed., vol. 171, p. 108796, 2021, doi: 10.1016/j.measurement.2020.108796.

[19] Y. Kurmi, S. Gangwar, D. Agrawal, S. Kumar, and H. S. Srivastava, "Leaf image analysis-based crop diseases classification," Signal, Image Video Process., vol. 15, no. 3, pp. 589–597, 2021, doi: 10.1007/s11760-020-01780-7.

[20] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," Ecol. Inform., vol. 61, no. June 2020, p. 101182, 2021, doi: 10.1016/j.ecoinf.2020.101182.

[21] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 36th Int. Conf. Mach. Learn. ICML 2019, vol. 2019-June, pp. 10691–10700, 2019.

[22] P. Govender and V. Sivakumar, Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019), vol. 11, no. 1. Turkish National Committee for Air Pollution Research and Control, 2020.

[23] C. Wu et al., "Fuzzy SLIC: Fuzzy Simple Linear Iterative Clustering," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 6, pp. 2114–2124, 2021, doi: 10.1109/TCSVT.2020.3019109.

[24] M. Jannati, S. H. Hosseinian, B. Vahidi, and G. jie Li, "ADALINE (ADAptive Linear NEuron)-based coordinated control for wind power fluctuations smoothing with reduced BESS (battery energy storage system) capacity," Energy, vol. 101, pp. 1–8, 2016, doi: 10.1016/j.energy.2016.01.100.

[25] T. Lindeberg, "Scale Invariant Feature Transform," Scholarpedia, vol. 7, no. 5, p. 10491, 2012, doi: 10.4249/scholarpedia.10491.

[26] V. N. Ghate and S. V. Dudul, "Optimal MLP neural network classifier for fault detection of three phase induction motor," Expert Syst. Appl., vol. 37, no. 4, pp. 3468–3481, 2010, doi: 10.1016/j.eswa.2009.10.041.