IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

APPLICATION ON DETECTING AND TREATING THE DISEASE IN PLANT/CROPS

[1]J.Lethisia Nithiya, [2]G Ajay Simha Reddy [3]G Sai nithin, [4]G.N Yuva Simha Reddy,

[1]Assistant professor, School of Computing, Department of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Chennai, India- 600073.

[2, 3,4] Student , School of Computing, Department of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Chennai, India- 600073.

***Abstract*—Rapid improvements in deep learning (DL) techniques have made it possible to detect and recognize objects from images. DL approaches have recently entered various agricultural and farming applications after being successfully employed in various fields. Automatic identification of plant diseases can help farmers manage their crops more effectively, resulting in higher yields. Detecting plant disease in crops using images is an intrinsically difficult task. In addition to their detection, individual species identification is necessary for applying tailored control methods. A survey of research initiatives that use convolutional neural networks (CNN), a type of DL, to address various plant disease detection concerns was undertaken in the current publication. In this work, we have reviewed 100 of the most relevant CNN articles on detecting various plant leaf diseases over the last five years. In addition, we identified and summarized several problems and solutions corresponding to the CNN used in plant leaf disease detection. Moreover, Deep convolutional neural networks (DCNN) trained on image data were the most effective method for detecting early disease detection. We expressed the benefits and drawbacks of utilizing CNN in agriculture, and we discussed the direction of future developments in plant disease detection..**

***Index Terms*— machine learning; deep learning; plant leaf diseases; Agriculture**

I. Introduction

Plant diseases are one of the most critical elements impacting food production. They are responsible for a significant drop in the economic productivity of crops, as well as being an obstruction to this activity in some cases. According to the authors Gebbers, R.Adamchuk [1], disease management and control procedures must be carried out effectively to reduce output losses and ensure agricultural

sustainability, underlining the importance of continual crop monitoring paired with prompt and accurate disease

detection. In addition, as the world's population contin ues to rise, a significant increase in food production is required (FAO). This must be combined with the preservation of natural ecosystems through the use

of environmentally friendly farming methods. Food must keep a high nutritious value while still being secure worldwide as suggested by the author Miller, S.A [3]. This can be accomplished by using new scientific methodologies for leaf disease diagnosis and crop management, as well as applying these new technologies to large-scale ecosystem monitoring. The motivation for conducting this survey comes from the fact that CNN has recently been primarily used in agriculture, with CNN's growing popularity and success in solving many problems related to agriculture, and the fact that multiple research efforts using CNN to discuss various agricultural problems exist today. As a result of its success, CNN is perhaps the most popular and commonly used approach in agricultural research today.

Regarding image analysis, the current survey focuses on a particular subset of DL models and techniques since there are very few of this type of survey in the agricultural field, especially about CNN utilization. Thus it would beneficial to present and analyze rele vant work to help the authors conduct a more comprehensive review. A discussion about innovative and high-potential techniques for solving numerous difficulties in agriculture related to image and DL will be presented. In addition to reviewing recent research in this area, significant practical features of CNN based on images are presented to explain the technique's advantages and disadvantages further.

II. RELATED WORK

Many studies have been conducted to find an ideal solution to the problem of crop disease detection by creating techniques that can assist in identifying crops in an agricultural environment. This section will provide the most

recently reviewed studies on CNN's applicability in the broad field of agriculture; this section includes papers from peer reviewed articles that use CNN methods and plant datasets. Abade et al. [9] reviewed CNN algorithms for the detection of plant diseases. The authors studied 121 papers that were published between 2010 and 2019. PlantVillage was selected as the most widely used dataset, while TensorFlow was identified as the most frequently used framework in this review.

Dhaka et al. [10] outlined the basic methods of CNN models used to identify plant diseases using leaf images. They also compared CNN models, pre-processing approaches, and frameworks. The study also looks at the datasets and performance measures used to assess model efficiency. Moreover, Nagaraju et al. [11] also provided a review to find the best datasets, pre-processing approaches, and DL tech niques for various plants. They reviewed and analyzed 84 papers on DL's applicability in plant disease diagnosis. They observed that so many DL methods are limited in their ability to analyze original images and that effective model performance necessitates using a suitable pre-processing technique.

Kamilaris et al. [12] found that DL approaches were used to solve various agricultural challenges. According to the study, DL methods performed better than standard image processing techniques. They focused on weed monitoring devices in agricultural fields that were both remote sensed and ground-based. Weed monitoring is critical for weed control, according to them. They predicted that data acquired by various sensors would be saved in a public cloud and used in appropriate contexts at the optimal time. Balasundram, S [13] introduced a review for plant disease classification using a CNN. The data for plant leaf disease identification, highlighting existing problems and potential prospects. They also presented NN approaches for SDI development in a short time. They discovered that, as long as SDIs remain relevant for proper crop protection, they must be tested on various hyperspectral sensors at the plant leaf scale. The disease detection using CNN, focusing on potato leaf disease. They reviewed several papers and concluded that convolutional neural networks work better at detecting the disease. They also identified that CNN contributed significantly to the maximum possible accuracy for disease identification.

III. METHODOLOGY

In this work, we discussed the most recent research papers on applying DL in the agricultural field. Moreover, this work was accomplished through two essential stages: the first is the collection of 100 previous research works that discuss DL in its relationship to the agricultural field, and the second is a thorough examination and analysis of the collected work.

In the first stage, we looked for papers and articles published within the last five years using scientific databases such as Science Direct and Elsevier and web-based scientific indexing services. In addition, we conducted our searches for relevant papers using several keywords, the most prominent of which were agriculture, CNN and DL. Papers mentioning CNN but not applying it to the agricultural domain were thus removed. In the second stage, the papers chosen in the first stage were analyzed one by one, taking into account the following research questions:

- The approach used.
- The problem presented.
- The datasets used.
- The performance achieved.
- Limitations of the study, if any.

Have the authors compared their CNN-based approach with other technologies, and what is the difference in performance?

Examining how CNN performs is an essential aspect of this study. As a result, we reviewed and analyzed several relevant studies. We also compared CNN to other current technologies and summarized the most important advantages and disadvantages that affect CNN's performance. It should be noted that the current paper focuses on comparing techniques used for the same data and on the same scale. We also investigated and discussed the most significant problems and limitations identified by previous research.

A. CONVENTIONAL NEURAL NETWORK

The ANNs consist of three different layers: input, one or more hidden, and output layers. Neurons placed in hidden layers have an associated weight and a bias value. These values are multiplied by the input values and sent to an activation function. If the output value is greater than the specified threshold, that node carries the output value to the next layer of the network. Otherwise, no data is transmitted.

The process of spreading data in the network from one layer to the successive layer is called a feed-forward network. The ultimate objective is to minimize the cost function for any

input when tuning the model weights and bias. The process is depicted in Figure 1. CNNs, a form of multi-layer neural networks, are designed to extract dependencies in a grid-structured input such as images and text. The convolution operation applied in many intermediate layers is the most crucial property of CNNs. Similarly, a convolution operation is a dot-product of a set of grid-structured weights and another set of similarly structured inputs.

CNNs are now widely used for image classification, image segmentation, face recognition and object recognition. They have been successfully applied by many organizations in various domains such as health, web, mail services, etc. CNN can receive any data input, including images, video, sound, speech, and natural language. However, CNN is simply a stack of several layers (see Figure 2), pooling and fully connected layers, beginning with a convolution layer and progressing through the following layers: pooling, Relu correction, and ending with a fully-connected layer . As a result, each image received as input will be filtered, reduced, and corrected several times, to finally form a vector. The strength of the CNN is found in the convolution layer. The CNN will learn the most valuable filters for the task (such as detection). Another benefit is that several convolution layers can be considered: the output of one convolution becomes the input of the next one, and the pooling layer is another component of a CNN. It performs down sampling, which significantly reduces computational weight, memory usage, and the number of parameters. On the other hand, in fully Connected Layers, as the name implies, each layer has a complete connection with the layer that comes before it. We can use a “sigmoid” or “softmax” function with the last fully connected layer for class predictions.

As a result, the convolutional layers extract features from the input images, which are then reduced in dimensionality by the pooling layers. Typically, the fully connected layers use the high-level features learned to classify input images into predefined classes at the final layer. Moreover, the classification layer can extract features for classification anddetection tasks .

B. Dataset:

Disease datasets typically encompass images and related data concerning the symptoms exhibited by various plant diseases. These datasets are frequently utilized to train machine learning models by enabling the identification and classification of different diseases based on patterns and features extracted from the images.This dataset is the collection of images depicting healthy and diseased crops, including tomatoes, potatoes, and grapes etc as of majorcultivation as taken into the consideration. The PlantVillage dataset consists of 54303 healthy and unhealthy leaf images divided into 38 categories by species and disease with both the test and train images as shown in the figure 3.1.

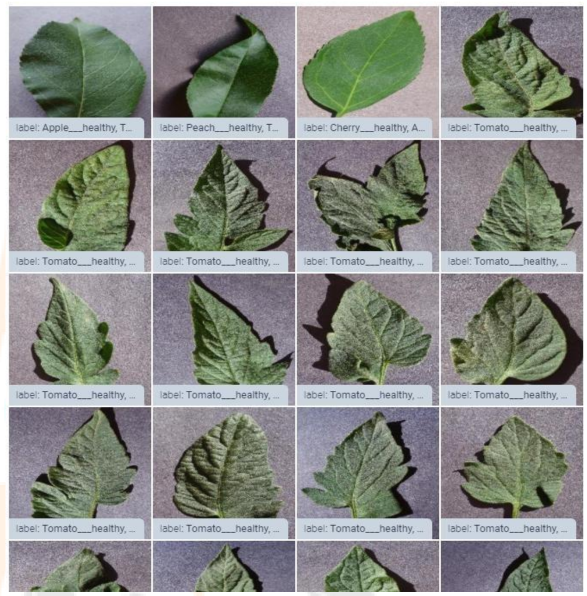

Fig 3.1 Dataset of plants with healthy and unhealthy

IV. PROBLEMS AND SOLUTIONS

This section may be divided into subheadings. It should provide a concise and precise description of the experimental results, their interpretation, and the experimental conclusions that can be drawn.

A.LIMITED PLANT DISEASE DATASET:

In the last five years, CNN has become increasingly used to detect plant leaf disease. However, CNN faces numerous hurdles. Therefore, in this section, we explained and summarized the problems and solutions encountered in developing CNN-based plant disease detection.

Datasets are critical for CNN models. As a result, the main challenge to using CNN for plant disease detection and classification is the requirement for large datasets. Thus, insufficient dataset size significantly impacts the practical implementation, which means that the results will be inaccurate no matter how efficient the model is. However, there are few publicly released datasets for agricultural researchers to work with, and in many cases, researchers must create their datasets of images from scratch. Furthermore, external environmental factors such as climate can readily affect data collection, making it time- consuming, and several days of work may be required. Moreover, to solve the problem of an insufficient and limited dataset, we provide the following solutions:

1. Data augmentation techniques: Data augmentation techniques increase the diversity of the data during training by artificially generating additional samples from the real dataset. Furthermore, image augmentation is a technique that creates new data from existing data to help train a deep neural network model. The most recent augmentation techniques With the expansion of the dataset, the accuracy improved as well. In another study, Barbedo used resizing and image segmentation methods to increase the size of the dataset from 1567 images to 46,409 images. The accuracy improved by 10.83% over the no expanded dataset.
2. Transfer learning: Transfer learning is a machine learning technique in which we reuse a previously trained model as the base for a new model on a new task. As a result of the new datasets, it will just retrain a few layers of pertained networks which helps to reduce the amount of data required.
3. Data sharing: Data sharing is another way of increasing datasets. Several studies are now being conducted worldwide on accurate disease detection. The dataset will become more accurate if the different datasets are shared. This situation will encourage more significant and satisfying study results. PlantVillage and Kaggle are the most commonly employed public datasets in the literature on DL methods for plant disease classification and detection.

C. IMAGE BACKGROUND:

From the section above, we found that almost all of the datasets used in training a CNN model for different studies used large datasets of images. However, one of the problems researchers face is the effect of image background on detection. Most of the time, this effect is unclear by the overlapping of many factors. The most remarkable is the interaction of plants with each other and the organization process. When images are collected in real-time conditions with a crowded background, some of the background features are similar to the area of interest; thus, a leaf segmentation technique is required in these conditions. Otherwise, the model will learn background features throughout training, leading to inaccurate classification results. On the other hand, some researchers are interested in organizing the image collection. In this case, the background is usually preserved because it creates relatively homogeneous backgrounds. It does not affect detection and may even improve detection accuracy.

V. PROPOSED SYSTEM

The proposed system was aimed to develop with the benefits of the farmers and agricultural sector. This system not only identifies the disease of the plant but also recommend the fertilizers and pesticides by considering the risk factors against the disease in an efficient way :

1. Machine Learning Model : In this system we used the CNN image classification technique that uses the image for recognition of patterns, color variation and spots on the plant leaf or the crop. By recognition of the features it tells the disease that occurred.

2. Diseases and treatment : By adding the mostly used crops and plant in the dataset with the major diseases that appears on crops and plant leafs. Along with the disease and also providing the treatment to the following disease that isdetected in the database.

WORKING MODEL:

A) Implemetation:

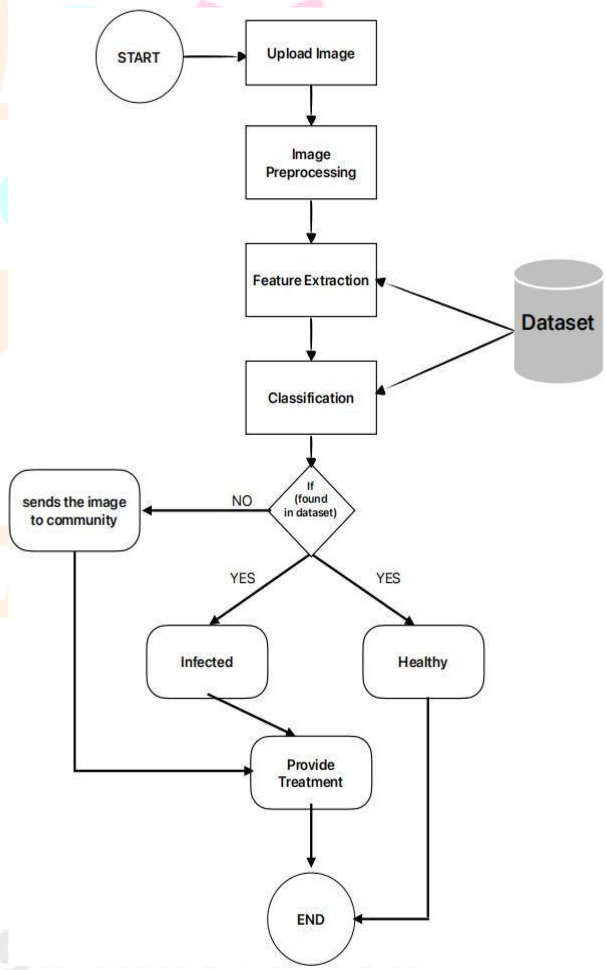

fig:5.1 Flow Diagram

The above figure:5.1 depicts the system's architecture, which is used to follow the model's flow and design. The home page of the mobile application where the crop or infected plant leaf image is uploaded to detect the disease.After that by using the CNN image classification which is done by color variation, pattern and spot recognition with other infected and healthy part of the leaf or the crop. By detecting the disease the following treatment is provided, if the unknown disease is known then the following image with problem is posted on the community of application developers. Agriculture experts and well-known farmers can give solution to their problem.

B) Unknown Disease Identified:

In case of the unknown disease is been identified through the YOLOv7 image classification then the image with problem description can be posted in the application community section followed with the cloud Mongodb and google firebase storage services. All the farmers and the users can surf through the application community to know the problems of other users who were uploading their different kinds of diseases can know about that too. Agriculture experts and well-known farmers can give solution to their problem.

C) Additional knowledge of plant diseases:

The additional feature of this proposed system in this application that it provides the knowledge of basic popular seen diseases with the more appropriate information with the correct use of pesticides and insecticides on a particular flowers, fruits or vegetables. By providing these knowledge to the farmers seen as most advantage to know more about diseases.

VI. OUTPUT

The machine learning model using the CNN image classification technique that describes the detection of the plant disease with name and confidence level as shown in fig:6 below.

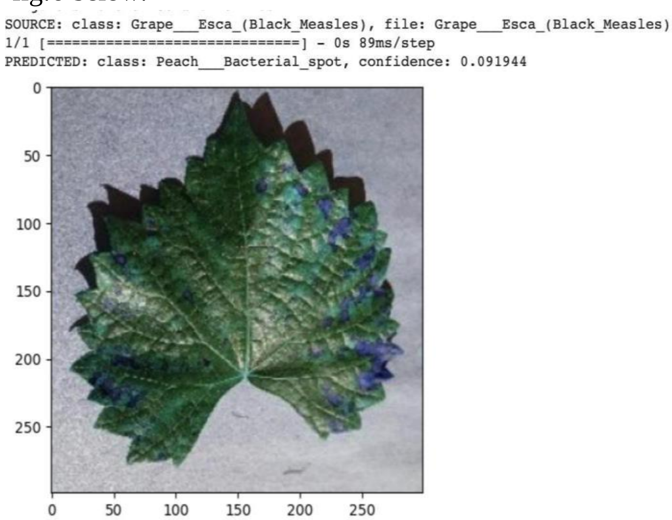

Fig.6 Detection of Plant Disease with confidence interval

This CNN based method for plant disease classification using the leaves of diseased plants. Building such a neural network with high efficiency is a complex task. Transfer learning can be employed to achieve greater efficiency . Inception v3 is one of the models available that inherently have the capability to classify images and further can be trained to identify different classes. Also by dataset classification using contour method, the training set can be chosen to ensure proper training of model for all features. This provides better feature extraction than randomly classifying the dataset.

VII. CONCLUSION

A CNN methods are widely used in the detection of plant diseases. It has solved the problems of traditional object detection and classification methods. In this study, we presented a detailed review of CNN-based research on plant leaf disease detection in crops over the last five years. We addressed highly related research articles to present a comparative analysis of various CNN models. Most studies used CNN approaches, and they note that pre-training models compared with training from scratch models on plant leaf datasets can quickly improve performance accuracy, especially if there is a sufficient dataset for each class to train the models. However, an essential future impact would be to develop highly efficient detection approaches employing large datasets with different plant leaf diseases. This would also address the class imbalance by requiring large generalized datasets.

REFERENCES

[1] Gebbers, R.; Adamchuk, V.I. Precision agriculture and food security. Science 2010, 327, 828–831.

[2] Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. Front. Plant Sci. 2016, 7, 1419.

[3] Miller, S.A.; Beed, F.D.; Harmon, C.L. Plant disease diagnostic capabilities and networks. Annu. Rev. Phytopathol. 2009, 47, 15–38.

[4] LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444.

[5] Najafabadi, M.M.; Villanustre, F.; Khoshgoftaar, T.M.; Seliya, N.; Wald, R.; Muharemagic, E. Deep learning applications and challenges in big data analytics. J. Big Data 2015, 2, 1–21.

[6] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

[7] Dhaka, V.S.; Meena, S.V.; Rani, G.; Sinwar, D.; Ijaz, M.F.; Woz´niak, M. A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. Sensors 2021, 21, 4749.
[8] Nagaraju, M.; Chawla, P. Systematic review of deep learning techniques in plant disease detection. Int. J. Syst. Assur. Eng. Manag. 2020, 11, 547–560.
[9] Abade, A.; Ferreira, P.A.; de Barros Vidal, F. Plant diseases recognition on images using convolutional neural networks: A systematic review. Comput. Electron. Agric. 2021, 185, 106125.
[10] Fernández-Quintanilla, C.; Peña, J.; Andújar, D.; Dorado, J.; Ribeiro, A.; López-Granados, F. Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops? Weed Res. 2018, 58, 259–272.
[11] Lu, J.; Tan, L.; Jiang, H. Review on convolutional neural network (CNN) applied to plant leaf disease classification. Agriculture 2021, 11, 707.
[12] Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. Comput. Electron. Agric. 2018, 147, 70–90.
[13] Golhani, K.; Balasundram, S.K.; Vadamalai, G.; Pradhan, B. A review of neural networks in plant disease detection using hyperspectral data. Inf. Process. Agric. 2018, 5, 354–371.