Machine Learning Techniques for Plant Disease Detection

Conference Paper · June 2021

DOI: 10.1109/ICOEI51242.2021.9453053

Divyanshu Varshney
National Institute of Technology, Kurukshetra

Burhanuddin Babukhanwala
National Institute of Technology, Kurukshetra

Javed Khan
National Institute of Technology, Kurukshetra

Deepika Saxena
Goethe-Universität Frankfurt am Main

Machine Learning Techniques for Plant Disease Detection

[1]Divyanshu varshney, [2]Burhanuddin Babukhanwala, [3]Javed Khan, [4] Deepika Saxena, [5]Ashutosh Kumar Singh

[1,2,3,4,5]Department of Computer Application , National Institute of Technology ,Kurukshetra, District – Kurukshetra (Haryana), INDIA

[1,2,3,4,5] opgmdivyanshu@gmail.com , burhanbabu98@gmail.com ,

jkhanjma@gmail.com, 13deepikasaxena@gmail.com, ashutosh@nitkkr.ac.in

Abstract. Undoubtedly, agriculture is an essential source of livelihood, which stands as a backbone of Indian economy. The plant production is severely affected due to various kinds of diseases, which if accurately and timely detected, could raise health standards and economic growth significantly. The traditional approach of disease detection and classification involves an immense amount of some time, an intense amount of labor and constant monitoring of the farm. By using disease detection methods, diseases caused by bacteria, viruses and fungi are often avoided. Within the upkeep of agricultural goods, crop protection plays a critical role. Techniques of Machine Learning are often used to identify the affected leaf images. The various machine learning algorithms used to identify plant disease are discussed in this study. It was done in various steps, such as image acquisition, image acquired, feature extraction, categorization of the illness and result display. This paper also needs to carry out an accurate study of various techniques for the identification of diseases of plants. The aim is to identify the plant diseases have used image analysis. It also, after detection of the illness, say the name of fertilizer to be used. The pests and insects accountable for the pandemic are also described.

Keywords: — Data mining, Image processing, Machine Learning, Plant disease detection

1. Introduction

In agricultural products, The main explanation for the decrease in quality is diseases. and output of agricultural goods. In required to treat but also manage them, crop diseases diagnosis for an initial point is therefore very important. Farmers make a great effort to choose the best plant seeds and also have the right climate for plant growth, but there are many diseases affecting plants that contribute to plant diseases. It is essential to discover the batches of plant diseases in agriculture at the early stage, which allows us to mitigate the damage, reduce production costs and increase profits. Many times, the human eye alone is not that successful in detecting the right illness. In the past, farmers used to pursue naked eye examination of experts with samples of infected plants or experts used to visit the field, and farmers took corrective steps to treat plant diseases based on their advice. It is very difficult to find trust worthy specialist in this methodology, and the solution does not work properly for the wide regions, the technique takes a long time. Often this approach is costly since it requires constant professional supervision. Agriculture is the root of every country's economy and it is therefore very important to accurately and promptly diagnose diseases of agricultural products. So to diagnose diseases, we need some automated, quick, reliable and less costly methods. Often, since it needs constant professional supervision, this approach is costly. Agriculture is the root of every country's economy and hence it is very important to accurately and promptly diagnose diseases of agricultural products. So to diagnose illnesses, we need some automated, fast, precise and less costly methods. Many applications related to agriculture have been developed for leaf recognition, detection of lead diseases, fruit diseases, etc. All of these systems involve photographic photographs that are taken by a digital camera. To fetch the relevant information for the analysis, image processing and analysis methods are then extended to captured images. In this article we present a survey of different strategies for crop diseases prevention and treatment plan using image analysis and machine intelligence to easily, randomly, automatically and accurately diagnose and identify plant leaf diseases.

The goal of this paper is to focus on the identification of plant diseases based on different approaches, such as image analysis. This can reduce efforts and use of pesticides.There are different characteristics of an image, such as grey, colour, shape, form, depth, motion, etc.[8]Significant improvements in crop yields too have occurred. There has also been some improvement in the field of agriculture with the development of technology. With the use of automatic identification of leaf diseases at early stage, the production of crops is increased. Various symptoms like spot on leaves and change in leaf color can be used to identify the disease at early stage. The image analysis techniques are effective, effective and accurate fields for detection of disease using plant leaves images in this area at the present time.

2. LITERATURE SURVEY

The previous research work done in this path did refer to several studies devoted to the subject in order to understand in detail about this survey. The survey of literature is done in sequential order:-

In 2008, Savvas Dimitriadis et al. [1] noted that it takes into account a factor of clarity agriculture (PA) that focuses on soil agricultural production. In this context, machine learning methods are used to obtain new insights immediately in the form of time expanding of decision for the effective utilization of biodiversity. To expand pre-existing understanding, the application model for machine learning proposed in the document is re-examined. As a step in this direction, it was demonstrated that an effective set of prediction models were developed which were used to forecast the state of the plant and to inhibit the undesirable effects of water stress in plants. Precision Agriculture (P.A.) is a framework of information based systems and ideals to improve resource efficiency and minimize environmental degradation by managing spatial and temporal variability. Data Mining is the method of extracting knowledge from previously unknown and theoretically valuable information.

In 2013, Asma Akhtar et al. [2] proposed a research paper presenting a technique for the quick and simple identification of plant illnesses. By gaining RGB images, the algorithm begin. Color transform is performed to RGB images during the next move. Using techniques to cluster K- means, images are then segmented. Features are derived from structure of data using the (GLCM) Gray level co- occurrence matrix. For classification, the As a tool, the Neural Network can be used. This technology's Complete Overall precision is 94 percent. The suggest method of integrating (DWT+DCT) characteristics for the identification of SVM Classifier gives the high reliability of 94.45%.

In 2015 Jagadesh D. Pujari et al.[3] Based on early prediction of leaf pathogen from clinical signs in the proposed method This valuable work in the actual life involves tasks such as image collection, post of images, selection of features, design of techniques for identifying fungal disease symptoms affected by various products in agricultural production, and finally production of architecture for the fungal infection effects for the Computer Vision System (CVS),With a first step, photos of fungus side effects impacted on difference ecology plants are gathered from pathology department. Picture treatment processes also used define and describe spores symptoms of the disease impacted on various forestry crops. The aim is to classify, classify and measure accurately first signs of diseases Plants diseases caused by bacteria, fungi, viruses, nematodes, etc. where its bacteria are the main life form disease. Medical conditions, emphasis was put on timely identification of fungal infection. In the actual world, farmers and farm professionals conceptually conduct a review of food production farms Such as cereals (with average classification is 85.33%), commercial crops (with average classification is 86.17%), fruits (with average classification is 94.085%), vegetables (with average classification is 91.54%) and the like caused by numerous diseases for identification.

In 2016, SharadaP.mohantyet al. [4] indicated this in the detection of image based plant disease. Here a public dataset is used by them. A fully belief network is trained to classify 14 different crops and 26 diseases using a public database of 54,306 pictures of healthy and infected leaf tissue gathered under realistic circumstances. The trained model on a carried testing dataset achieves an accuracy of 99.35 percent, displaying the viability of this method Overall, a clear path to smart phone diagnosis is helped plant diseases on a huge large scale by the method of teaching deep learning techniques on extremely high and Available to the public different images in this research. In many diverse domains, neural networks have previously been effectively implemented as examples of earlier part learning.

In 2016, SrdjanSladojevic [5] reported a latest approach to the study of a crop disease recognition system that is based on the classification of plant leaves using deep convolutionary networks. A software model will be developed to propose remedial measures for the management of pests or diseases in agricultural crops. Leaf-snap, which uses visual recognition from their leaves images in order to identify tree specifies, however as the system displayed in either paper classifies weed species rather than types plant. The development model has been responsible For the detection of leaves and differentiate respectively plant leaves and 13 different easily diagnosable ailments. Besides different class tests, on mean 96 percent, the testing findings on the proposed model accomplished accuracy between 91 percent and 98%.

In 2017 P.Ferentinos, [6], considered neural networks to detect and diagnose plant disease and used easy image datasets of healthy and pathologic trees via deep learning techniques. In this paper using an open archive of 87,848 pictures, model training is carried out and contains 25 new variants in a set of 58 specific categories of mixtures of host plants, as well as an early proponent of 87,848 images. Several model Frameworks have been instructed, with either the maximum performance trying to reach an accuracy rate of 99.53 percent in identifying the corresponding disease

.

In 2018, Muhammad Sharif et al. [7] stated that the manufacturing of citrus fruits and their availability are badly affected by citrus diseases in the agricultural system. For the Citrus crop diseases categorization, a hybrid technique (composed of (a) lesion higher selectivity on citrus leaves and seeds; (b) classification of citrus diseases) is proposed The proposed approach exceeds old solutions and produces 97 percent share of the citrus disease image file set of data classification, 89 percent on the joint set of data and 90.4 percent on our nearby set of data.

a) Identification of injury spots on tropical fruit leaves and buds;

b) Identification of tropical fruit infections

Tropical fruit lesion areas are obtained by an improved measured classification step, which is executed over an improved source images. Then color scheme, softness, and geometries are merged in a code book. The chosen functions are loaded to the Multi class support vector machine (MSVM) for the classification purpose of tropical fruit disease. The approach developed is verified mostly on Tropical fruit Disease File set of data, the paired set of data (Infested with Plastic Surgery Town and Citrus Images Database), as well as our own collected image database. For the identification and tracking of tropical fruit infections, these datasets, such as interiority, bullet hole, fungus, zit, forestation and Melanose , were used.

In 2018, S. Sannakkiet et al. [8] suggested a classification that following are the main characteristics on the Subgroup system, using as characteristics the defected region and colouration. The main advantage is that it converts to L*a*b to remove layers of image color space and 97.30 is learned from nomenclature. The main disadvantage is in using for strictly limited crops only. With a contours used to limit the strength within the disease spot, the BPNN classifier resolves the different issues of race. Detection and Marking of Leaf Disease using Machine Learning Technology & Fuzzy Logic. They have used artificial neural network (ANN) as an algorithm, which mainly helps to confirm the severity of the diseased leaf. Pre-processing, extraction of attributes, training of classifiers and classification. Picture pre-processing brings all the scale of the image to a reduced uniform size. The object's appearance and the outline of the image are in this function descriptor, its intensity gradients are represented by This does not impact any conversions. This is done only because it is only possible to measure Over a single channel, Hu moments form the descriptor and Haralick features. The aim of this algorithm is to identify plant anomalies that occur in their barns or natural environments. The algorithm has been compared with other machine learning techniques for consistency 160 images of papaya leaves were used to train the model using the Decision tree classifier. With about 70% precision, the method could be rated. When a series of images are trained and other image data are used along with template matching such as SIFT, accuracy can be improved (Scale Invariant Feature Transform), SURF and DENSE together with BOVW (Speed Up Robust Features) (Bag Of Visual Word).

In 2019, Tejal Chandiwade et al. [9], have proposed a method that assists in plant disease detection and offers solutions that As a method of disease protection, it may be used. The Archive accessed It is adequately isolated from of the Internet and the various databases are isolated. Plant species are described and renamed to form an appropriate database and then to acquire a test database comprising of different plant diseases used to validate the consistency and confidence level of the project. The identification of plant disease can be achieved by examining the location on the infected plant's leaves. The method which was adopted to detect plant diseases is image processing using Convolutional Neural Network (CNN). A concept drone model is also intended to use for live surveillance of wide agricultural fields to which a large camera is mounted and to take images of plants operating as a software input, on the basis of which the software will inform us whether or not plant is healthy. The various approaches which are used in detecting the plant diseases are Support Vector Machine (SVM) and Artificial Neural Network (ANN). The main objectives of this papers were to detect crop disease and to provide remedy for the disease that is detected. This could promote disease control and increase productivity. The proposed system is python-based and gives an accuracy of about 78%. Using Google's GPU for processing, the accuracy and speed can be improved. The device may be mounted on drones in order to be able to conduct aerial surveillance of crop fields.

3. Summary Table

. Table 1- In this table we will discuss about some existing research paper

Year	**Literature Reference**	**Feature Selection Approach**	**Classification Approach**	**Datasets**	**Result (Accuracy %)**
2008	Savvas Dimitriadis	Navie Bayes(Classification)	Supervised	ETR photosynthetic Activity and E-Multi plant	88.18% correctly classified on Naïve Bayes Simple
2013	Asma Akhtar et al.	K mean KNN RNN	Unsupervised	Rose leaf samples(manual dataset)	85.41% in DCT*+DWT* 75.40% in DWT+Texture 75% in DCT+Texture
2015	Jagadeh D. Pujari et al.	PNN SVM Neuro KNN Black wise feature	Supervised/ Unsupervised	Image Data Set	86.17%commercial crop 85.33%Cereal crop 94.085% fruits 91.54% vegetables
2016	Sharada P.Mohanty	Neural Network	Supervised	Image Net	99.35% on an expected to hold testing dataset
2016	Srdjan Sladojevic	SVM, Neural Network, Decision Tree	Supervised	Plant Village Olive tree images Corn images Tomato	92-99% in Plant Village , 99% in Olive tree images, 97% in Corn

				Images Apple Images Rice Images Potato Images	images, 83% in Tomato Images, 98% in Apple images, 95% in Rice Images, 96% in Potato Images
2017	P.Femtions	Neural Network	Supervised	Alex Net, Google Net, Over-Feat, VGG	99.06% in Alex Net, 97.27% in Google Net, 98.96% in Over-Feat, 99.48% in VGG
2018	Sharif et al.	SVM(Classification)	Supervised	Citrus Infections Picture Gallery 2017, Dewdney and Trimmer 2009	97% on Picture Gallery Dataset 89% on combined Dataset
2018	S.Sannakkiet et al.	ANN SVM	Supervised	Image DataSet	70% Accuracy on Image Dataset
2019	Tejal Chandiwaeet et al	ANN SVM	Supervised	Mannual Dataset	78% Accuracy on Manual Dataset

4. Machine Learning Techniques

K-nearest neighbor (KNN) - The method for information processing is k-nearest neighbor. In machine learning methods, K nearest neighbor is a simple algorithm where the categorization is accomplished by recognizing the closest neighbor to query examples and using those neighbors to determine the question class.

Naïve Bayes - One of it's simplest and most powerful supervised learning is Naive Bayes, which allows to create fast machine learning algorithm that can make quick observations. It is a probability distribution, meaning it predictions on the basis of the likelihood of an item.

Support Vector Machine (SVM) - Support Vector Machine (SVM) is a set of related supervisory techniques used to classify and regress. For two-group classification problems, it employs different classifiers. They will category fresh text after having an SVM model collection of named training samples with each group. The support vector uses a mathematical equation, also referred to as an activation functions, which is a mathematical feature that compares the new information with the best picture from the learning patterns to estimate the mark of the intangible (apple or pear).

Artificial Neural Network - It is a networking device intended to enhance the normal human brain study and processing of data. It is the basis of Artificial Intelligence (AI) and addresses challenges that'd be complex and hard by human or mathematical criteria.

5. Challenges

Similarity of signs of different diseases.
It includes diseases, nutritional deficiencies, pests, phyototoxicity, excessive cold or heat, and varied mechanical damage. In order to resolve this challenge the working model needs to be re-train frequently with respect to the diseases.

In order to operate properly, the strategies suggested so far are generally constrained in their reach and rely on suitable capture situations..
This is attributed to the nature of diverse contexts that can not readily be isolated from the area of interest (usually leaf and stem).For Example: in a crop a disease named aster yellows is detected, now if it has another disease named bacterial wilt. Then our trained model won’t be able to identify it. The solution of this challenge would be Transfer Learning. In which the existing trained model can be used in such cases.

Growing demands for complete, healthy and diverse foodstuffs to sustain and boost living conditions for the growing population.
Nowadays, people are becoming more professional in terms of their diet i.e., what they eat. So, the food(crops) should be free from all the diseases.

6. Conclusion

In this review, a relative investigation is developed on 4 kinds of machine learning classifiers for plant identifying diseases. When compared to other classifiers, many authors use the SVM classifier for the classification of diseases. The outcome shows that more diseases are detected with high accuracy by the CNN classifier. In the future, The Naïve Bayes classification should only be used to classify diseases in crops and to help people automatically diagnose all types of diseases in crops using machine learning approaches from other classifiers, such as the decision tree.

References

1. M.Sharif, M. A. Khan, Z. Iqbal, M. F.Azam, M. I. U. Lali and M. Y. Javed, "Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection," Computers and Electronics in Agriculture, vol. 150, pp. 220-234, 2018.

2. S. Dimitriadis,C. Goumopoulos,"Applying Machine Learning to Extract New knowledge in Precision Agriculture Applications," Panhellenic Conference on Informatics, pp. 100-104, 2008.

3. A. Akhtar, A. Khanum, S. A. Khan and A. Shaukat, "Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques," International Conference on Frontiers of Information Technology, pp. 60-65, 2013.

4. K.P. Ferentinos,"Deep learning models for plant disease detection and diagnosis"Computers and Electronics in Agriculture, vol. 145, pp. 311-318, 2018.

5. S. P. Mohanty,D. P. Hughes and M. Salathe, "Using Deep Learning for Image- Based Plant Disease Detection," Frontiers in plant science, pp. 1-10, 2016.

6. S. Sladojevic, M. Arsenovic, A. Anderla, Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,"Computational intelligence and neuroscience, pp. 1-12, 2016.

7.] K. R. Gavhale and U. Gawande,"An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques," IOSR Journal of Computer Engineering, vol. 16, pp. 10-16, 2014.

8. J. Pujari, R. Yakkundimath and A. Byadgi, "Image Processing Based Detection of Fungal Diseases in Plants," International Conference on Information and Communication Technologies (ICICT), vol. 46, pp. 1802-1808, 2015.

9. S. Ramesh, R. Hebbar, P. Bhat and P. V. Vinod, "Plant Disease Detection Using Machine Learning," International Conference on Design Innovations for 3Cs Compute Communicate Control, pp. 41-45, 2018.

10. A. M. A. Karol, D. Gulhane and T. Chandiwade, "Plant Disease Detection using CNN & Remedy," International Journal of Advanced Research of Electrical, Electronics and Instrumentation Engineering, vol. 8, no. 3, pp. 622-626, 2019.

11. D. Saxena, and A.K. Singh, "A proactive autoscaling and energy-efficient VM allocation framework using online multi-resource neural network for cloud data center", Neurocomputing 426 (2021): 248-264.

12. Kumar, Jitendra, Deepika Saxena, Ashutosh Kumar Singh, and Anand Mohan.,"Biphase adaptive learning-based neural network model for cloud datacenter workload forecasting." Soft Computing (2020): 1-18.

13. Saxena, Deepika, and Ashutosh Kumar Singh. "Auto-adaptive learning-based workload forecasting in dynamic cloud environment." International Journal of Computers and Applications (2020): 1-11.

14. M. A. Hussein and A. H. Abbas, "Plant Leaf Disease Detection Using Support Vector Machine," Al-Mustansiriyah Journal of Science, vol. 30, no. 1, pp. 105-110, 2019.

15. N. K. Durga and G. Anuradha, "Plant Disease Identification Using SVM and ANN Algorithm," International Journal of Recent Technology and Engineering (IJRTE), vol. 7, pp. 471-473, 2019.

16. S. Pavithra, A. Priyadharshini, V. Praveena and T. Monika, "PADDY LEAF DISEASE DETECTION USING SVM CLASSIFIER," International Journal of communication and computer Technologies, vol. 03, no. 01, pp. 16-20, 2015.

17. E. Hossain, M. F. Hossain and M. A. Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease using KNN Classifier," International Conference on ELectrical, Computer and Communication Engineering (ECCE), pp. 1-7, 2019.

18. S. Sharma, V. Kaur and N. Dhillon, "PLANT DISEASE CLASSIFICATION WITH KNN-SVM CLASSIFICATION," International Journal of Advance Engineering and Research Development, vol. 5, no. 05, pp. 752-758, 2018.

19. D. P. Mohindru, G. Kaur and D. P. , "Simulative Investigation of Plant Diseases using KNN Algorithm," International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, vol. 7, no. 8, pp. 32-36, 2019.

20. M. Nababan, Y. Laia, D. Sitanggang, O. Sihombing, E. Indra, S. Siregar, W. Purba and R. Mancur, "The diagnose of oil palm disease using Naive Bayes Method based on Expert System Technology," Journal of Physics, pp. 1-5, 2015.