Advanced Disease Detection Techniques in Plants Using Leaf Disease Detection and Soil's Nutrient Deficiency

Divya Kumari[1*], H. J. Bharath[2], Jaisal Srivastava[3], J. Y. Manvith[4], Anusha Preetham[5]

[1,2,3,4]*Student, Department of Information Science & Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, India*
[5]*Assistant Professor, Department of Information Science & Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, India*

Abstract: **Plant disease prediction and detection has long been a major issue in the agricultural field. People with limited financial resources cannot afford to have their crops and fields inspected on a regular basis. As a result, a technique that can meet the needs of poor farmers will be a game changer. The proposed model in this paper has developed a system that can determine whether or not a plant is healthy. Furthermore, if the plan is unhealthy, the reason of the disease is determined by collecting two inputs from the plot where the diseased plant is present: plant leaves and soil samples. The model is taught to determine whether the disease is bacterial, fungal, or viral.**

Keywords: **Machine Learning, Fertilizer recommendation, Random Forest, Convolutional Neural Networks, SVM, HED.**

1. Introduction

Agriculture is a field which contributes to a major part of India`s economic growth and serves as the biggest employment source for all categories of people. Thus, the profit share of agricultural field is essential for everyone. The quality of crops grown or product generated collectively defines the profit rate and therefore farmers are concerned about their crop`s health status, soil nutrient content, water supply for irrigation, amount of sunlight available for them and so on. Among these the health of crops is of utmost importance as if after all the necessities are provided, if the plant dies or get infected by any pest or bacteria then the investment as well time will be wasted. Precision agriculture (PA) is a contemporary-day generation and technique that offers state-of-the-art techniques and gear for optimizing and modernizing agricultural discipline. Through use These state-of-the-art strategies, Economic improvement in agriculture Can be reached. PA may be used for the subsequent functions Many makes use of which include crop pests Identification, weed detection, Crop manufacturing and detection Plant sicknesses, etc. [2]. Plant disorder Detection could be very essential and farmers Always seeking out generation Who can do that sufficient Work with much less funding and time. Plant disorder detection may be finished through the usage of photograph processing strategies followed through system getting to know strategies to make the procedure easy and accurate. Digital photograph processing

gear and preprocessing techniques are hired through the researchers you acquire the excellent and preferred ouput that's frequently invisible to the human eye. The observations made through the human eye are typically procured for coming to the belief that whether or not the sicknesses are intense withinside the region of manufacturing or not. The full-size improvement has executed through the photograph processing withinside the discipline of agriculture [3].

One of the methods used is SVM i.e., linear SVM is a method used for classifying the data into two or more classes as it is multi class separator these two classes are one with most inefficient data and one with efficient data, thus reducing the issues in accuracy. The main objective behind this paper was to provide with a method to detect the disease present in the plant and to classify the plants into healthy and diseased plant. It`s goal to detect whether the disease is due to plant`s infection or due to lack of nutrients in the soil. The broad objective is to further classify what kind of attack is found on the plant like fungal, bacterial or pest attack. Depending on the kind of disease pesticide or fertilizer is recommended to the farmer to use in farm. The entire process involves IOT, image processing, machine learning, deep learning and many related techniques. The paper also states real-time, Cloud-based soil nutrient observation flame. Soil nutrients Humidity, pH, Temperature is measured by our suggestion Can be set up and made available Anywhere in the world with the help of Valid validation of cloud channels qualification. Significant contribution. The proposed job is done by the farmers Get Earth's real-time application Parameter analysis at the front door. Therefore, the proposed cloud-based floor Nutrient analysis setup Requests from farmers to add Soil nutrients and crop improvement Grow without expecting results from the test lab [7].

2. Literature Survey

This paper presents various studies Methods that can be used for Recognize, classify and recommend. The solution needed for plants Illness based on whether it is the cause Behind this is a nutrient or nutrient deficiency in the soil For the spread of all the pests and fungi above plant. Wear systematically Disease

---

detection process. Two broad categories are considered, that is, -leaf input and bottom input. First, process the sheet Input methods like KNearest Adjacent classifier, convolution Neural network, artificial neural network Network and support vector machine It is being used. In paper form, primarily based totally on CNN LVQ set of rules with the aid of using Melike Sardogan, four hundred schooling and one hundred take a look at datasets Taken from tomato leaves was suggested. The model was recognizable Diseased plants by classification Various bacteria, blight, septoria Leaf spots and curls of yellow leaves. Or Paper makes use of convolution Neural community version and studying Vector quantization (LVQ) algorithm. The accuracy completed turned into approximately 86% [1]. In some other article, plant leaves Mechanical sickness detection Learn from Amrita S. Tulshan. Or Disease detection K-nearest neighbor technique (KNN) Classification algorithm.it additionally Image preprocessing, picture segmentation, and Feature extraction. Or The accuracy received turned into approximately 98.5% [3]. On paper, tea leaf sickness Detection with the aid of using multi-lens picture Segmentation with the aid of using Somnath Mukhopadhyay, leaves taken into consideration. The technique used with tea leaves NSGAII-primarily based totally picture clustering, Multiclass SVM, PCA and capabilities Reduction [4]. It may generate the accuracy was about 83% Can recognize 5 different tea leaves Spider mite, spider mite, Thrips, Heropertis, Nikko Burning. Almost all soil analysis literature You can observe this soil nutrient Analysis is performed by Electrochemical sensor, optics-visual Visual acuity and electrical conductivity detection. Also, the most popular method in Soil survey literature is color sensor technology Controller- based method. Therefore, adapt the framework Based on color detection Controller-based method [7].

Below are the gaps in most of the work Could not be addressed:

1. Many existing technologies could not extract enough features from the photo. Necessity of rights combination of functions extraction technology requirement.
2. Images with overlapping sheets will be identified.
3. Images containing multiple objects handle.
4. Complex background, blurred photos and images caught low. It wasn't lighting consideration.
5. Dataset with huge number of images were not considered.
6. New techniques under machine learning need to be explored for better accuracy.

3. Proposed Solution

To overcome the routine method of going to the soil research canter and scientists and getting treatment, we provide users with a unique and efficient way to find their problems and give them the best possible solution to the illness of their crops with the highest accuracy. Offers the various solutions proposed are algorithms, interactive website, and backends using Python.

A. Algorithm Used

1) Decision Tree

Decision bushes are nonparametric supervised gaining knowledge of algorithms used for each category and regression tasks. It has a hierarchical tree shape such as root nodes, branches, inner nodes, and leaf nodes. Decision tree gaining knowledge of plays a grasping seek and makes use of a divide and rule method to pick out the quality divide factor withinside the tree. The splitting manner then repeats recursively from pinnacle to backside till all or maximum of the statistics are labeled into a selected magnificence label. Whether all records factors are labeled as a fixed of the equal type relies upon in large part at the complexity of the choice tree. Small bushes can effortlessly attain natural leaf nodes (that is, single-magnificence records factors). However, because the tree grows, it will become more and more tough to hold this purity, and generally too little records falls into a selected subtree. When this happens, it is known as records fragmentation and regularly ends in overfitting. As a result, choice bushes opt for smaller bushes, consistent with Occam's razor saving principles. That is, "the unit must now no longer be multiplied greater than necessary." In different words, choice bushes want to feature complexity simplest whilst needed. This is due to the fact the only rationalization is regularly the quality. Clipping is usually carried out to lessen complexity and save you overfitting. This is the manner of disposing of a department that has been cut up into much less crucial features. The version suitability can then be evaluated thru the manner of cross-validation. Another manner the choice tree can hold accuracy is to shape an ensemble thru a random wooded area algorithm. This classifier predicts greater correct results, especially if the character bushes aren't correlated with every different.

2) Naive Bayes

It is a probabilistic machine learning algorithm used in many classification functions based on Bayes' theorem. Gaussian Naive Bayes is an extension of Naive Bayes. We use other functions to estimate the data distribution, but it is easiest to implement a Gaussian or normal distribution because we need to calculate the mean and standard deviation of the training data. It is based on a probabilistic model, which makes it easy to code algorithms and make fast real-time predictions. Therefore, this algorithm can be configured to respond immediately to a user's request and is a common choice for solving real problems. But before we dig deeper into Naive Bayes and Gauss Naive Bayes, we need to know what conditional probabilities mean.

3) Support Vector Machine

SVMs are one of the most now no longer unusual place supervised getting to know algorithms used for every kind and regression problems. However, it's far mainly used for tool getting to know kind problems. The reason of the SVM set of regulations is to create most efficient lines or preference boundaries that could divide n-dimensional place into instructions simply so new records elements can be effects placed in the suitable elegance withinside the future. This tremendous preference boundary is known as the hyperplane. The SVM selects extrema/vectors to help create the hyperplane.

The set of regulations is known as a help vector tool because of the truth the ones intense times are known as help vectors.

4) Random Forest

Random forests, or random choice forests, are ensemble mastering techniques for type, regression, and different duties that paintings with the aid of using constructing numerous choice timber for the duration of education. For type duties, the Random Forest output is the magnificence decided on with the aid of using maximum timber. For regression duties, a median or common prediction for every tree is returned. Random Decision Forest modifies the choice tree addiction of overfitting education sets: 587–588 Random forests usually carry out higher than choice timber, however are much less correct than gradient-more advantageous timber. However, statistics houses can effect performance.

5) XgBoost

It is an abbreviation for Extreme Gradient Boosting proposed via way of means of researchers on the University of Washington. This is a library written in C ++ that optimizes gradient boosting schooling. Boosting is ensemble modeling. This is a method that tries to construct a robust classifier from a big wide variety of susceptible classifiers. This is completed via way of means of constructing the version the usage of a chain of susceptible fashions. First, create a version from the schooling data. Then a 2d version is created that attempts to restoration the mistakes that exist withinside the first version. This method maintains and fashions are introduced till the entire schooling dataset is efficaciously anticipated or the most wide variety of fashions is introduced.

Gradient Boosting: Gradient boosting is a famous boosting algorithm. With gradient boosting, every predictor corrects the mistake withinside the preceding predictor. In comparison to Adaboost, schooling example weights aren't adjusted. Instead, every predictor is skilled the usage of the residual mistakes of its predecessor as a label. There is a method referred to as the Gradient Boosting Tree, the primary learner of that's CART (Classification and Regression Tree).

XGBoost is an implementation of the gradient boosted selection tree. The XGBoost version dominates many Kaggle competitions. In this algorithm, the selection tree is built in a sequential format. Weights play an essential position in XGBoost. All unbiased variables are assigned weights and the weights are dispatched to the selection tree that predicts the outcome. The tree will increase the burden of the variables that have been incorrectly anticipated and sends those variables to the second one selection tree. Then gather those man or woman classifiers / predictors to generate a greater effective and greater correct version. You can address regression, classification, ranking, and custom prediction issues.

6) Convolutional Neural Network

A convolutional neural community (CNN) is a deep studying neural community designed to technique established arrays of records which includes plots. CNN may be very glad to seize the layout of the enter picture, which includes lines, gradients, circles, or even eyes and faces. This belonging makes convolutional neural networks extraordinarily sturdy to pc vision. CNNs may be run at once on sub processed pictures with none pre-processing. Convolutional neural networks are feedforward neural networks, not often up to 20. The power of convolutional neural networks comes from a specific kind of layer referred to as the convolutional layer. The CNN incorporates many layers of convolutions which are stacked on pinnacle of every other, each of which could apprehend greater complicated shapes. The third or 4th layer convolution can apprehend handwritten numbers, and the twenty fifth layer can distinguish human faces. The time table on this vicinity is plenty of duties which includes machines seeing the arena like humans, spotting it in a comparable way, picture and video recognition, picture inspection and classification, media updates, advice systems, and greater. Is with a view to use the knowledge. Natural language processing, etc.

7) Holistically Edge Detection

Proposed Holistic Nest Edge Detector (HED) Addresses two important issues: (1) Overall image training and Predictions inspired by fully convolutional neural networks, for frame-by-frame classification (system is Generates an edge map image directly with the image as input as output); (2) Nested multiscale feature learning inspired by deeply functioning, deeply monitored networks Layer monitoring to "guide" the initial classification results. we Find the preferred properties of this foundation the technique shows that HED is accurate and computationally efficient.

B. Python Libraries

1) NumPy

It usually written in C, is a Python extension module. It is described as a Python package used to perform various numerical calculations and operations on multidimensional and one- dimensional array elements. Computations using numpy arrays are faster than using regular Python arrays. The NumPy package was created in 2005 with the help of Travis Oliphant by incorporating the functionality of its predecessor module, Numeric, into all other modules, Numarray.
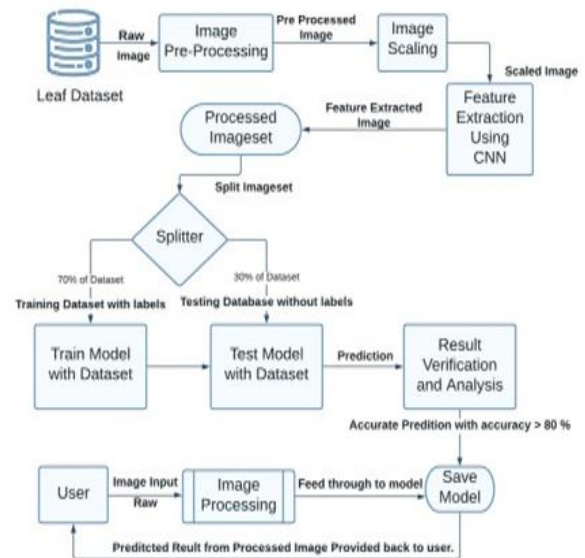

Fig. 1. Block diagram of the proposed model [11]

2) Online PlantX website for plant disease detection and solution recommendation

The website designed here is for detecting the diseases in the plant's leaf or due to soil nutrients deficiency. The website is user friendly and interactive. It makes use of machine learning algorithms to make recommendations for fertilizers and pesticides as per need.

The recommendations made here are highly accurate. The soil nutrient deficiency is calculated using N, P, K values of it. It also takes the PH value into consideration and analyses the soil's health status.

The leaf detection layout is designed with options of picture uploading or taking direct picture from the camera. It analyses the picture using HED techniques and classifies it using CNN algorithm. It detects a leaf as healthy, non-healthy, if diseased then what disease and then recommends solution accordingly.

4. Conclusion

In this article, the leaf turbulence detection and kind method is based on a convolutional neural community the use of the HED method to a dataset of 500 images. This is a version of the depth survey based primarily on the collected images and aspect detection of soil patterns. Performs a visible function extraction from each pattern consisting of leaf and soil patterns, which together produce the final output. However, his older version is no longer optimal for recognizing the disease and at the same time offers the technique by recommending fertilizers and pesticides. This release addresses and addresses the issues and limitations faced by many authors. Photo processing performance has been improved over previous experiments, and the complexity of the IoT has been slightly reduced.

Research on soil sampling papers was still underway, so this version should end again with the prescribed strategies and methods.

References

[1] Sardogan, M., Tuncer, A., & Ozen, Y. (2018, September). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 382-385).
[2] Sujatha, R., Chatterjee, J. M., Jhanjhi, N. Z., & Brohi, S. N. (2021). Performance of deep learning vs machine learning in plant leaf disease detection. Microprocessors and Microsystems, 80, 103615.
[3] Tulshan, A. S., & Raul, N. (2019, July). Plant leaf disease detection using machine learning. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6).
[4] Mukhopadhyay, S., Paul, M., Pal, R., & De, (2021). Tea leaf disease detection using multi-objective image segmentation. Multimedia Tools and Applications, 80(1), 753-771.
[5] Nandhini, S., & Ashokkumar, K. (2021). Improved crossover-based monarch butterfly optimization for tomato leaf disease classification using convolutional neural network. Multimedia Tools and Applications, 80(12), 18583-18610.
[6] Lu, J., Hu, J., Zhao, G., Mei, F., & Zhang, C. (2017). An in-field automatic wheat disease diagnosis system. Computers and electronics in agriculture, 142, 369-379.
[7] Varsha Kiran Patil, Aniket Jadhav, Someshwar Gavhane, Venkatesh Kapare, 2021
[8] K. Spandana and S. Pabboju, "Applications of IoT for Soil Quality," 2020.
[9] Swapnil Sunil Raut and Vidya Chitre, "Soil monitoring and Testing using IoT for fertility level and crop prediction," 2020
[10] P. Sukumar, T. Kavitha, A. Deepika, V. Jashnavi, "Real Time soil fertility analyzer using IoT."
[11] D. Kumari, H. J. Bharath, J. Srivastava, J. Y. Manvith, and A. Preetham, "Improved Plant Disease Detection Techniques using Convolutional Neural Networks: A Survey", in *International Journal of Research in Engineering, Science and Management*, vol. 5, no. 4, pp. 92–95, Apr. 2022.