Quest Journals
Journal of Software Engineering and Simulation
Volume 9 ~ Issue 12 (2023) pp: 43-51
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

A Comparative Study of Edge Detection Techniques to Identify Maize Leaf Diseases using Machine Learning

Khushi M Asudaria[1], Mohammed Faraaz Abdul Khadeer[2], Syed Shujauddin Rafai [3], Kondakindi Sphoorthy Reddy[4]
[1,2,3,4] Keshav Memorial Institute of Technology (KMIT), Hyderabad, India

Abstract. *Plant diseases cause significant loss of crop production and immense monetary loss to the farmer. This in turn affects the supply chain and increases prices for the end consumer. To reduce this loss, this paper proposes to use edge detection techniques and machine learning algorithms to classify the disease. Here the effectiveness of five techniques that utilize edge detection were tested, Sobel, Canny, Laplacian, Prewitt, and Roberts. All detectors were run through the same set of ML Classifiers, Gaussian Classifier, SVM, Decision Tree, Naive Bayes, Nearest Centroid, Complement Naive Bayes, Multilayer Perceptron Neural Network and AdaBoost Classifiers. This paper presents a comparative study of the edge detection techniques after running them through various ML models to obtain the best accuracy of classifying the disease correctly. This research describes effective techniques to identify plant disease. The procedure to identify disease consists of two main steps. 1. Applying Edge Detection techniques on the dataset. 2. Obtaining accuracy of different classifiers after applying various ML Models.*

Keywords: *Plant Disease Detection; Plant Disease Classification; Image preprocessing; Edge Detection; scikit learn; Machine-Learning.*

R*eceived 12 Dec., 2023;* ***R****evised 24 Dec., 2023;* ***A****ccepted 26 Dec., 2023* ***P****ublished with open access at* ***www.questjournals.org***

I. Introduction

The top four most widely grown crops worldwide include maize. It is a non-seasonal crop in India and is cultivated throughout the year. According to the authors of [16], maize production for India was 30,250 thousand tons in 2020. India's corn production increased at an average yearly rate of 4.67% from 5,101 thousand tons in 1971 to 30,250 thousand tons in 2020. Apart from human consumption and use as cattle fodder, it is also used in the production of corn starch and corn oil. A CIPHET study of 2012 - 2013 suggests a loss of 4.65% of the overall production of the corn crop annually. Hence, Maize is a vital crop for global food security, but its production is threatened by various diseases. Accurate and timely detection of maize leaf diseases is essential for effective disease management and prevention of crop loss. Corn is mainly affected by 3 types of diseases. They are Leaf blight, Common Rust, and Gray Leaf Spot, refer to Fig 1

(a) Healthy leaf **(b)** Leaf affected with Leaf Blight

(c) Common Rust **(d)** Gray Leaf Spot

Fig.1 Common Corn Crop diseases

Bipolaris maydis is a fungal disease-causing pathogen that leads to leaf blight. It thrives in warm and damp climates. Common Rust (*Puccinia sorghi*) is a species of Rust Fungus that infects corn; early symptoms include specks on leaves. These later turn into blister-like pustules on the plant tissue. These pustules change from brown to black. Gray Leaf Spot is a fungal disease that is caused by Cercospora. The symptoms of gray leaf spots are rectangular, brown lesions that are on the leaf. These disease-inducing spores reside in the topsoil and the conditions that promote it are humid, hot, and wet climates.

Manual identification and classification of diseases are slow and cumbersome; hence it is time-consuming and requires manpower. Hence, Kaur et al [4] states that Attempts to identify infection detected on leaves through automation have been under research for a long period of time.

Automated maize leaf disease detection systems have been developed to identify symptoms of diseases like northern leaf blight, gray leaf spot, and southern rust which can significantly reduce yield and quality. These systems use machine learning and computer vision techniques to analyze images of maize leaves and provide a diagnosis of the disease, enabling farmers to take prompt and appropriate action to mitigate its impact. In this context, maize leaf disease detection is a critical tool for ensuring food security, protecting the livelihoods of farmers, and promoting sustainable agriculture practices.

This paper demonstrates a correlation among edge identification techniques as a visual representation by comparing the various techniques, the combination of the best edge detection technique and the best machine learning classifier is obtained, which can be used to build an effective model for detecting corn crop diseases.

II. Literature Review

The main consensus of colleagues on this research study was to find the best classifier to detect a plant disease. Identifying the best classifier is a relatively simple task and has been done on several occasions previously.

Ramesh et al. [8] suggested a method for detecting plant diseases that involves preprocessing RGB (Red Green Blue) images by converting them to grayscale and then to HSV (Hue, Saturation, Value) for histogram calculation. A Random Forest classifier is used to classify the leaf as diseased or healthy with around 70% accuracy.

Deshpande et al. [6] proposed. Here, Haar wavelet based GLCM (Gray Level Co-occurrence Matrix) features and first-order features with SVM (Support Vector Machine) and KNN (K Nearest Neighbors) classifiers achieve an accuracy of up to 88%.

Mohammad Syarief et al. [7] utilized seven CNN architectures to classify images of maize leaves, including AlexNet, Visual Geometry Group 16, VGG19, Residual Neural Network 50, ResNet110, GoogleNet, and Inception-V3, along with three classification modes, SVM, KNN, and Decision Tree. They found that the leading classification outcomes were achieved using the AlexNet architecture with SVM, indicating that these methods were most effective for feature extraction and classification of maize leaf disease images. This comparative study proposes the inclusion of edge detection techniques.

III. Proposed Method

3.1 Image Acquisition

The images are acquired from Bangladeshi Crop Dataset. The initial testing was done on a dataset of 15000 select images from the Bangladeshi Crops dataset.

3.2 Image Preprocessing

During this step, Image preprocessing is done to reduce variables in the image. It is resized into a fixed size of 200x300, then the thresholds are set to 100 and 256.

Low Variance Thresholding feature selection method is used to eliminate the unnecessary features having low variance. After resizing, the aperture size is set as 7 and is run through a gaussian blur filter, finally it is converted to grayscale.

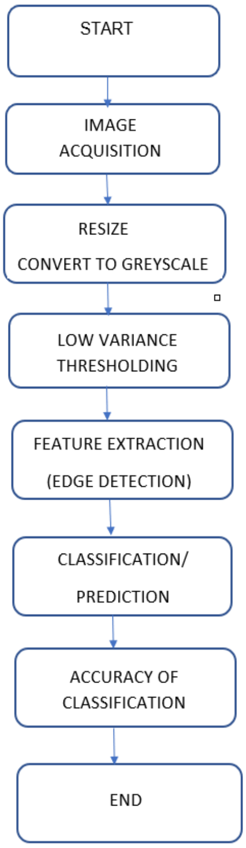


Fig 3.1 Flowchart of the proposed approach

3.3 Feature Extraction

An essential step in image preprocessing is Edge Detection. Muthukrishnan et al [1] describes the edge detection step as a fundamental tool for image segmentation. It is a devised step that aims to enhance and highlight relevant features and also diminishes the irrelevant details from our data, which in our case is the corn image.

The usefulness of each edge detection technique depends on its ability to detect meaningful edges. Edge detection makes it easier for Machine Learning algorithms to classify the diseases on the corn leaf.

Contrary to popular use, we are detecting the shape of the object. In this case, i.e., leaf, we are making the necessary edges of the features on the leaf which gives us an output that helps classify the disease.

The technique of detecting edges can be understood as filters that lower the quantity of data that needs to be processed while keeping important data that is required for the machine learning algorithm to recognize plant diseases. Most edge detectors have unique gradient approximations in the form of Gx and Gy, which are used to identify and highlight edges or boundaries in an image respectively. In the paper, the following 5 Edge Detection Techniques are used.

Roberts Edge Detection. In 1965, a man named Lawrence Roberts came up with a method called "Roberts Edge Detection." The idea is to measure the modifications in the image to find where the edges are. It's like tracing a map with a pencil and feeling the bumps and ridges. The result is a set of numbers that tells you how steep the changes are between pixels. This can help us find important details in an image, like the edges of objects. Fig. 3b shows the maize leaf after the application of Roberts Edge Detection.

+1	0
0	-1

Gx

0	+1
-1	0

Gy

Fig. 3.2 Roberts Matrix

Sobel Edge Detection. Sobel is a method for detecting edges in pictures, which works by employing an approach called the Sobel approximation. This approach is geared towards identifying areas within the image that change rapidly and then flagging them as edges. It relies on a pair of distinctive filters, which are essentially tiny grids that are swept across the image to pinpoint these edges. The result of this process is a set of coordinates that shows precisely where the edges existing inside the image are located. Fig. 3c shows the maize leaf after the application of Sobel Edge Detection.

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

Fig. 3.3 Sobel's Matrix

Prewitt Edge Detection. The Prewitt operator is a way to find boundaries in pictures. The authors of [17] said that it works by looking at the dissimilarities among the luminosity of pixels next to each other in the picture. It doesn't work as well as another method called Sobel because it can make the edges look messy. Fig. 3d shows the maize leaf after application of Prewitt Edge Detection.

-1	0	+1
-1	0	+1
-1	0	+1

G_x

+1	+1	+1
0	0	0
-1	-1	-1

G_y

Fig. 3.4 Prewitt Matrix

Canny Edge Detection. The Canny Edge Detector finds edges in an image. It is better than other similar ones because it doesn't change the edges too much when it finds them. The program works by doing a few different steps. First, it makes the picture look smoother by using a special function called a Gaussian function. Then, it calculates how strong the edges are and in what direction they go. After that, it gets rid of any edges that are too weak or too close to each other. Finally, it sets a limit for how strong an edge has to be shown in the final picture. Fig. 3e shows the maize leaf after application of Canny Edge Detection.

Laplacian Edge Detection. The Laplacian operator is a tool that helps you find edges. It's like a special calculator that looks at each part of the picture and calculates how much it changes from the parts around it. The Laplacian operator is different from other edge detectors because it looks at how much the picture changes twice instead of just once like the other detectors. This helps it find more complex edges that might be harder to see with the other detectors. It can also detect edges in two different directions, some edges that point inwards and some edges that point outwards.

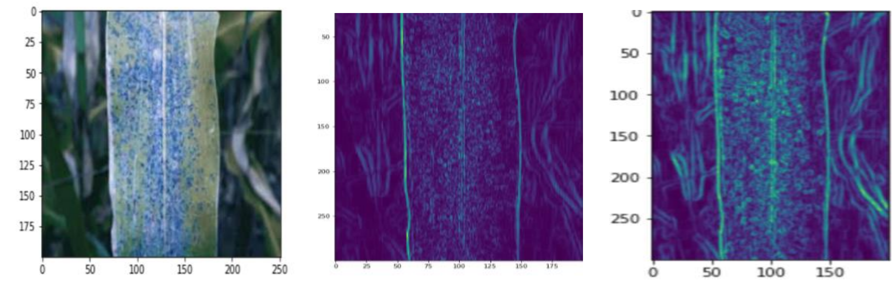
- Inward Edges (Negative)
- Outward Edges (Positive)

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \qquad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

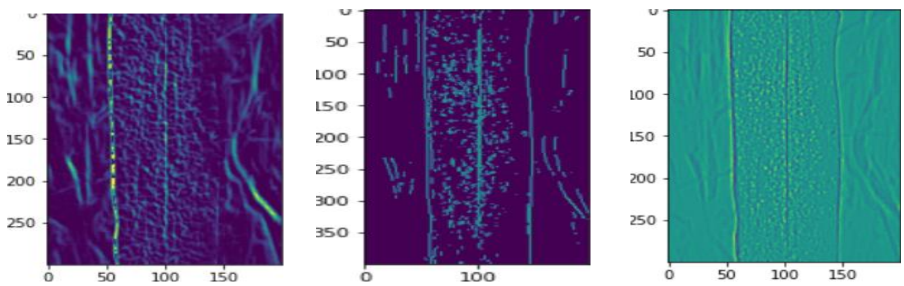
Gx Gy

Fig. 3.5 Laplacian Matrix

The original image and the images given by various edge detection techniques are given in the figure below. Fig. 3.6 shows the maize leaf after application of Laplacian Edge Detection.



(a)Original Leaf (b) Roberts ED (c) Sobels' ED



(d) Prewitt's ED (e) Canny ED (f) Laplacian ED

Fig. 3.6 Outputs for different edge detection techniques **a.** Original leaf. **b.** Leaf with Roberts Edge Detection. **c.** Leaf with Sobel Edge Detection. **d.** Leaf with Prewitt Edge Detection **e.** Leaf with Canny Edge Detection **f.** Leaf with Laplacian Edge Detection

3.4 Classification/Prediction

Following the extraction of the features, we applied some of the popular classification approaches listed below to categorize the leaf. The classifiers were chosen for their prowess on this dataset, their capacity to handle non-linear relationships, and their interpretability i.e., how easily they can be understood.

Support Vector Machine. SVM algorithm belongs to the category of machine learning models that require supervision during the training process. Though it can be used for both classification and regression, it is simple and easy to use for classification. This algorithm generates a 2-D space, referred to as a hyperplane, which exists in an n-dimensional area. A point is plotted on the graph that can fall on any side of the plane, and the side on which they fall determines the class to which the leaf disease belongs. Since it is a powerful classifier, it can be used to learn the boundary between the edge and non-edge pixels based on the gradient information.

Decision Tree. Decision Tree is comparable to SVM since it is a supervised ML algorithm as well. It is one of the easiest to infer machine learning algorithms where tree representations are based on the features available. The leaf nodes contain all the possible outcomes. Based on the if-else statements, the tree is traversed till the leaf node is reached. Terms like 'Entropy' and 'Information Gain' are used in determining the class of the given data point. It can be utilized to learn complex decision boundaries between edge and non-edge pixels, especially when the features have non-linear relationships with edge detection.

Naive Bayes. Naive Bayes is a supervised, eager learner, machine learning algorithm which uses probability to classify the given data point. The theorem of Bayes is a mathematical concept that serves as the foundation of an approach used to classify traits into different categories. This method operates on the idea that each pair of traits being classified is different from the others. This method is useful because it does not require a lot of training data to work and can quickly predict which category a trait belongs to. Naive Bayes is computationally efficient and requires minimal parameter tuning, making it suitable for real-time edge detection applications.

Gaussian Processes Classifier. The non-parametric machine learning algorithm known as Gaussian Processes Classifier has the capability to handle binary classification tasks. These are a generalization of the Gaussian probability distribution. Compared to SVM, they are capable of predicting highly calibrated probabilities. It also allows for principled uncertainty quantification, which can be valuable in scenarios where the quality of edge detection is critical.

Nearest Centroid. The Nearest Centroid Classifier is the simplest classifier in machine learning. Though it is underrated and underutilized, it is powerful and extremely efficient for a few specific machine-learning classification applications. It works on simple principles, the centroid for each class is computed during training. By computing the minimal distance between the selected data point and the centroid of each class, the nearest centroid classifier determines the class of the testing sample. Although Nearest Centroid may not handle complex decision boundaries well, it can perform reasonably when the classes are relatively well-separated, making it suitable for simpler edge detection tasks.

Complement Naive Bayes (CNB). As the multinomial Naive Bayes classifier does not work very well on uneven datasets, we use the CNB classifier. This classifier is well-suited for uneven data, as it reduces the time taken for an image to belong to a particular class by comparing it with various occurrences parallelly, rather than characterizing it for each class individually. In edge detection, the number of non-edge pixels often heavily outweighs the number of edge pixels, resulting in class imbalance. CNB can be used to tackle this issue.

Multi-layer Perceptron Classifier. The multi-layer perceptron is a type of feed-forward neural network classifier which consists of 3 layers.

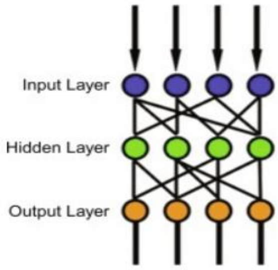


Fig. 3.7 Multi-Layer Perceptron Model

The multi-layer perceptron classifier can recognize and differentiate the testing data automatically, once it has been trained on the given dataset. This advantage of the classifier has been used in order to minimize the dissimilarities among the obtained and likely estimated accuracy values when trying to classify an image to a class. In other words, it recognizes patterns in unseen data.
The design of a neuron is shown below.

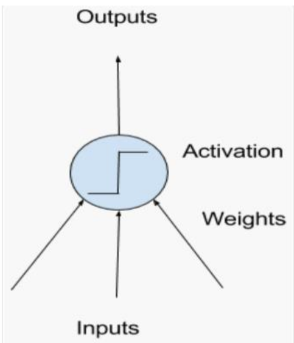


Fig 3.8 Model of a Simple Neuron

MLP can capture complex non-linear relationships in the data, which is crucial for edge detection tasks where edges often exhibit intricate patterns and structures.

AdaBoost. The AdaBoost or Adaptive Boosting algorithm is an ensemble boosting classifier. Kusomo et al. [3] mentions ensemble methods are used to avoid overfitting of data. As AdaBoost is a boosting classifier, it highlights relevant details that have failed to be identified by a certain classifier in an instance, therefore increasing overall accuracy. AdaBoost can be combined with simple classifiers to improve their performance, especially in scenarios where the edge detection problem is challenging, and the data is imbalanced. One-level decision trees are the most commonly used algorithms with AdaBoost. These trees are also called **Decision Stumps.**

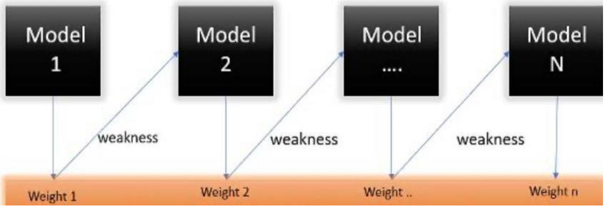


Fig. 3.9 AdaBoost Model

To obtain the performance metrics of each Machine Learning classifier, sci-kit-learn's metrics module was used.

IV. Results and Discussions

The extracted features in the preceding step are classified using the above classification techniques. Our dataset consists of three categories of maize leaf diseases, and one category of healthy leaf images. Obtained results of accuracy in the classification are presented in the table.

The tables compare the accuracy of 8 ML models after the application of 5 edge detection techniques as the data pre-processing step.

In the Table 1, out of the first four ML models, for Canny Edge Detection, the Decision Tree Classifier has given us a high accuracy of 90.75%. For the Sobel Edge Detection, we observe SVM gives us an accuracy of 94.00%. When we use Robert Edge Detection, again the Decision Tree Classifier gives us an accuracy of 91.17%. The Prewitt Edge Detection gives us the accuracy of 92.25% with the SVM classifier. Similarly, in Laplacian, the highest accuracy of 92.75% is obtained using the SVM classifier.

Table 1. Edge Detection vs. ML Algorithms

Edge Detection	SVM	DT	NB	GPC	NC	CNB	MLP	AdaBoost
Canny	89.75	90.75	85.00	84.00	64.25	74.75	84.25	59.25
Sobel	94.00	90.25	73.25	79.25	65.00	62.00	**94.75**	49.00
Roberts	89.67	91.17	78.00	90.00	67.84	68.84	90.17	55.67
Prewitt	92.25	91.50	80.00	83.50	70.00	72.00	25.25	66.25
Laplacian	92.75	89.00	80.25	81.00	79.75	77.78	85.75	73.50

From the remaining four ML models, for Canny Edge Detection, the MLP Classifier has given us an accuracy of 84.25%. In the case of Sobel Edge Detection, we can see that the MLP classifier gives us the best accuracy of 94.75%. Once more, the MLP classifier provides us with a high 90.17% accuracy when we apply Robert Edge Detection. Prewitt Edge Detection with the Complement Naive Bayes classifier has a 92.25% accuracy rate. In order to attain the highest accuracy of 85.75% in Laplacian Edge Detection, the MLP classifier is utilized once again.

A quick insight into our work, an attempt was made to enhance the image by updating their sharpness and contrast, in hope for an improved accuracy for more models was a failure as with extra details, the noise in the image increased causing significant decline in the accuracy. Observing this effect, we decided on an optimal value for maintaining a good accuracy in the model.

We reach a conclusion from the above analysis as follows, the Sobel Edge Detection Method along with a Multi-layer Perceptron classifier, has achieved a significantly higher accuracy of **94.75%**, when compared to other methods.

V. Conclusion and Future Scope

The aim of this paper is to study the output and understand the effect of different types of edge detection techniques applied on leaf images followed by data analysis of various ML models applied to each of the edge detection techniques separately. This insight helps us determine which ML model can be leveraged when applying a type of edge detection, to detect the corn crop disease effectively.

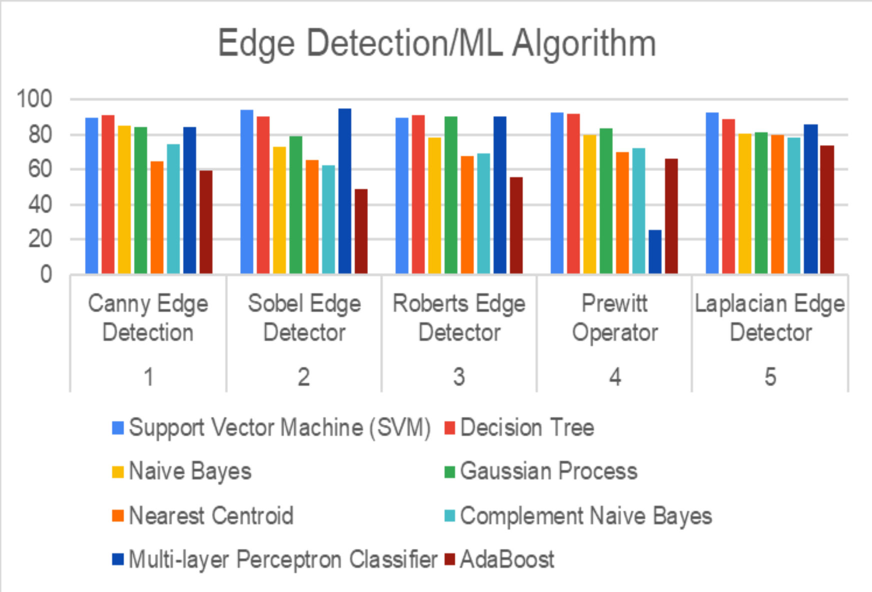


Fig 5 Edge Detection & ML Algorithm vs. Accuracy

Fig 5 represents the summary of our research work in which the different colored bars each represent different classifiers. Each group of such bars are applied to various edge detection techniques. The classifiers are numbered from 1 to 5 on the X-axis of the graph and the scale on the Y-axis in the range of 0 to 100 represents the accuracy with which the crop disease is best detected.

From the table we can conclude that the Sobel Edge Detector is the best edge detection technique for Multi-layer Perceptron Classifier giving us the best accuracy of **94.75%**.

Furthermore, this paper methodology can be extended to detect other crop leaves affected by similar diseases and integrate to an application which can alert the farmer to remove out the affected leaves from the crops. Thus, saving the crop from any further damage.

References

[1]. Muthukrishnan, Ranjan, and Miyilsamy Radha. "Edge detection techniques for image segmentation." International Journal of Computer Science & Information Technology 3.6 (2011): 259, doi: 10.5121/ijcsit.2011.3620.

[2]. Badage, Anuradha. "Crop disease detection using machine learning: Indian agriculture." Int. Res. J. Eng. Technol 5.9 (2018): 866-869.

[3]. Kusumo, Budiarianto Suryo, et al. "Machine learning-based for automatic detection of corn-plant diseases using image processing." 2018 International conference on computer, control, informatics and its applications (IC3INA). IEEE, 2018, doi: 10.1109/ic3ina.2018.8629507.

[4]. Kaur, Sukhvir, Shreelekha Pandey, and Shivani Goel. "Plants disease identification and classification through leaf images: A survey." Archives of Computational Methods in Engineering 26.2 (2019): 507-530, doi: 10.1007/s11831-018-9255-6.

[5]. Applalanaidu, Majji V., and G. Kumaravelan. "A review of machine learning approaches in plant leaf disease detection and classification." 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). IEEE, 2021, doi: 10.1109/icicv50876.2021.9388488.

[6]. Deshapande, Anupama S., et al. "Fungal disease detection in maize leaves using haar wavelet features." Information and Communication Technology for Intelligent Systems. Springer, Singapore, 2019. 275-286, doi: 10.1007/978-981-13-1742-2_27.

[7]. Syarief, Mohammad, and Wahyudi Setiawan. "Convolutional neural network for maize leaf disease image classification." Telkomnika (Telecommunication Computing Electronics and Control) 18.3 (2020): 1376-1381, doi: 10.12928/telkomnika.v18i3.14840.

[8]. Ramesh, Shima, et al. "Plant disease detection using machine learning." 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C). IEEE, 2018, doi: 10.1109/icdi3c.2018.00017.

[9]. Nagasubramanian, Gayathri, et al. "Ensemble classification and IoT-based pattern recognition for crop disease monitoring system." IEEE Internet of Things Journal 8.16 (2021): 12847-12854, doi: 10.1109/jiot.2021.3072908.
[10]. Shruthi, U., V. Nagaveni, and B. K. Raghavendra. "A review on machine learning classification techniques for plant disease detection." 2019 5th International conference on advanced computing & communication systems (ICACCS). IEEE, 2019, doi: 10.1109/icaccs.2019.8728415.
[11]. Punam Thakare (2011) "A Study of Image Segmentation and Edge Detection Techniques", International Journal on Computer Science and Engineering, Vol 3, No.2, 899-904, DOI:10.5121/ijcsit.
[12]. Wiesner-Hanks, Tyr, et al. "Image set for deep learning: field images of maize annotated with disease symptoms." BMC research notes 11.1 (2018): 1-3, doi: 10.1186/s13104-018-3548-6.
[13]. Tichkule, Shivani K., and Dhanashri H. Gawali. "Plant diseases detection using image processing techniques." 2016 Online International Conference on Green Engineering and Technologies (IC-GET). IEEE, 2016, doi: 10.1109/get.2016.7916653.
[14]. Pawar, Pooja, Varsha Turkar, and Pravin Patil. "Cucumber disease detection using artificial neural network." 2016 international conference on inventive computation technologies (ICICT). Vol. 3. IEEE, 2016, doi: 10.1109/inventive.2016.7830151.
[15]. "India Maize Production, 1960-2020 - Knoema.Com." Knoema, https://knoema.com/atlas/India/topics/Agriculture/Crops-Production-Quantity-tonnes/Maize-production. Accessed 19 Mar. 2023.
[16]. "Digital Image Processing Algorithms Using MATLAB - GeeksforGeeks." GeeksforGeeks, GeeksforGeeks, 22 Feb. 2021, https://www.geeksforgeeks.org/digital-image-processing-algorithms-using-matlab