IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125191

DETECTION OF OKRA DISEASE - A SURVEY

[1]Maanyatha Mahesh,[2]Suloni Praveen,[3]Swathi Meghana K R, [4]Supriya T C, [5]Shraddha C

Student, Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India[1-4]

Assistant Professor, Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India[5]

Abstract— The okra crop, commonly called lady's finger, is a commercial crop grown by farmers. However, when these crops become contaminated, they cause havoc in the lives of farmers as well as the economy. Plant disease problems are caused by challenges in agricultural activity and climate change. It produces financial troubles and losses for earnings, farmers, and the entire okra business. The difficulty persists if disease identification is performed without the use of appropriate procedures. Several studies have offered various methods for combating these illnesses. For detecting and classifying plant diseases, CNN, Deep learning, and other machine-learning algorithms are applied.

Keywords— Okra, CNN, Deep learning, Machine learning, detecting, classifying.

1. INTRODUCTION

One of the most important and fundamental pillars of any economy, and perhaps more so in India, is the agricultural sector. One of India's largest industries, agriculture employs a sizable portion of the country's people. Farmers that are on the cutting edge of agriculture typically rely on visual inspection and identification of various crop illnesses and problems. Crop experts employ the same procedure with a bit more expertise. However, both farmers and experts are only capable of precisely identifying a certain number of crops. Despite the use of methods that appear to be infallible, there is a chance that the outcomes will be inaccurate; they may have revealed improper assessment or be incorrect diagnosis. This takes time and wastes material and financial resources, which reduces the effectiveness of crop production, has an impact on farmers' investments, and could be seen as a blow to the farming industry and the country's economy. It thus reduces crop growth further, creating a shortage in the market and inflation in price. The use of image processing methods in automatic detection may prove to be a useful strategy that produces quicker and better results. Deep learning and the use of recognition models are simply two of the few contemporary methods that could deliver an ideal outcome. In this project, we'll employ deep convolutional networks, which are typically used to classify images of leaves, to try to tackle the challenge of detecting and identifying various crop/plant illnesses.

Digital image processing is the process of applying various operations to a picture in order to gain useful information from it or create a better image for additional analysis. It permits the use of more algorithms on the input data and can lessen processing noise and signal distortion. Image processing is a method for applying some action to a picture in order to enlarge it or extract important information. It is comparable to signal processing in that an image serves as the input, and the result could either be another image or the attributes of that image.

Image processing is a method for converting a physical image into a digital one so that you can edit, add to, or remove information from it. An image, such as a photo or image sequence, acts as the input in this type of signal distribution, while the output can either be another image or the attributes of the original image.

The following steps are significantly involved in image processing:

1. Importing the image using image acquisition software.
2. Reviewing and adjusting the image.
3. The output, which may be a revised image, or a report based on the analysis of the original image.

The application of computational models to the image processing of photographic files is referred to as digital image processing in computer science. Contrary to analog image processing, it enables the application of a larger variety of algorithms to the data entered and can also prevent problems like the build-up of noise and signal distortion during processing. It provides a platform for a selection of processes, such as the processing of analog and digital signals, picture signals, speech signals, and image signals. Digital image processing can be treated as a multidimensional system because images are specified over two dimensions. The three primary reasons that influence the development of digital image

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 8.102 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125191

processing are the evolution and development of computers, the development of mathematics, and the demand for a wide range of technology imaging services.

Image restoration (CT, MRI, PET), image reformatting (multi-plane, multi-view reconstruction), quick and easy image storage and retrieval, and images with high distribution are only a few of the benefits of digital image processing. There is furthermore viewed viewing (windowing, zooming). The time-intensive nature of digital image processing is one of its drawbacks. Depending on the system, it can be very expensive. It is only accessible to qualified users.

2. BACKGROUND AND RELATED WORK

Table 1: Represents the related work of the problem statement. It shows the different

Algorithms are used to analyze and detect many plant diseases. The accuracy percentage is also mentioned.

YEAR	AUTHOR	TITLE	ALGORITHM	ACCURACY	REMARKS
2021	Ankit Kumar Ghorai.S Mukyopadhyay	Nutrient deficiency and the detection of disease using image processing	CNN, ANN and SVM	89%	Special imaging has strong correlation with plant nutrition. PSF usually vary with growth stages [6].
2021	Abdul Rehman Zain Tariq	Detection of cucumber leaf Disease using the newly generated techniques of the Haralik features and the technique of the local tri-directional pattern	Tri directional pattern	91%	Feature selection techniques is proposed to maintain accuracy and reduce the computational time. Few features were ignored during the selection process [5].
2020	Norhalina Senan, Muhammad Aamir, Rosziati Ibrahim, N.S.A. Tanjuddin and W.H.N wan Muda	An efficient convolution neural network for paddy leaf disease and pest classification	CNN	96.60%	Proposed five layered CNN efficiently solves problem of paddy disease detection. Entropy and Information gain for learning high order information features are not used [3].
2020	Pranesh Kulkarni, Atharva Karwande, Tejas Kolhe, Soham Kamble, Akshay Joshi, Medha Wyawahare	Plant disease detection using Image Processing and Machine Learning	Image processing	93%	the proposed system is computationally efficient because of the use of statistical image processing and machine learning model. It is a computer vision-based system for plant disease detection with average 93% accuracy and 0.93 F1 score [8].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁘ Impact Factor 8.102 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125191

2019	Sagar Gaikwad, Sagar Shinde	Leaf disease detection using digital image processing with SVM Classifier	Image processing, SVM	91.65%	This displays an overview on various strategies for plant leaf sickness identification utilizing picture handling procedure. The k-closest neighbor technique is maybe the least difficult of all calculations for foreseeing the class of a test precedent [7].
2019	Aravind krishnaswamy rangarajan, Raja purushottaman and Anirudh Ramesh	Tomato crop disease classification using pre - trained deep learning algorithms	VGG16 And AlexNet	97.29% 97.49%	AlexNet provide good accuracy. VGG16 has least classification accuracy and processing time was more [4].
2015	D. Mondal. A.Chakravarty ,D.K kale and D.D Majunderc	Detection and classification technique of yellow vein masaic virus disease in okra leaf images using leaf vein extraction and naive Bayesian classifier	leaf vein extraction and naive Bayesian classifier	87%	Suitable for solving multi-class prediction models. Algorithm faces 'zero-frequency'problem [2].

3. LITERATURE REVIEW

Okra yellow vein mosaic, also known as the bhindi yellow vein mosaic virus, is a viral disease that affects okra plants and is brought on by a monopartite Begomovirus. It was initially discovered in 1924 in Sri Lanka and Bombay, India. It is the main factor limiting okra output. This disease is transmitted by whitefly. We collected data for 2000 images of okra leaves (diseased & healthy) and organized them into a train, validation, and test folders. The goal is to identify the yellow vein mosaic disease in okra leaves. Datasets are first obtained by having photos of the okra crops. Various studies outline the methods for categorizing and identifying plant diseases. Different algorithms, including those that use classification and regression approaches, are applied in machine learning and neural networks. Researchers [1] and others employ SVM Naive Bayes, logistic regression, and linear discriminant analysis to produce effective results and identify issues or infections in crops with fewer pictures collected. Unlike these Densenet is also used to serve the purpose. The training and testing sets are separated into two portions, making up the dataset. The dataset's feature vector is trained using the Densenet.

Rice disease images from the Kaggle repository(https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset) were used to collect the data for this research project. This dataset [2] includes one class of healthy photos in addition to three primary classes of paddy leaf images with the disease, including brown spot, hispa, and leaf blast. There are 3355 paddy photos in all. There were 523 photos for brown spot disease, 565 for hispa pests, and 779 for leaf blast disease in these images. Healthy paddy contributed to the remaining 1488 photo. After the segmentation of the main image into various sub-images containing regions of interest or features, these features are further extracted for the identification of disease. Feature extraction thus is a type of dimensionality reduction process that reduces the number of variable features and which in turn are actually necessary for disease detection. Features can be mainly classified into shape, color, texture, and edges. The feature shape includes area, perimeter, length, width, eccentricity, roundness, elongation, color axis length, solidity, concavity, and perimeter. The texture includes homogeneity, contrast, variance, correlation, energy, entropy, the moment of inertia and mean.

The PlantVillage dataset was used to obtain the images for 6 different tomato crop illnesses and healthy samples. In the segmented images of the dataset, all background pixels in the three channels (Red, Green, and Blue) are set to 0, except

DOI: 10.17148/IJARCCE.2023.125191

for the pixels related to the leaf, which are used in this study. For the selected diseases and healthy ones from the collection, a total of 13,262 segmented images were available. The 256 × 256 original input image was increased to 227 x 227 for the Alex Net model [3] and 224 x 224 for the VGG16 net. The design has a total of eight layers, the first five of which are convolutional layers and the final three of which are completely connected. To extract the most features possible, the first two convolutional layers are linked to overlapping

max-pooling layers. The fully connected layers are immediately coupled to the third, fourth, and fifth convolutional layers. The ReLu non-linear activation function is coupled to each and every output of the convolutional and fully connected layers. A softmax activation layer, which generates a distribution of 1000 class labels, is coupled to the final output layer.

In the experiment, a publicly available dataset of photographs of leaves from 14 plants was employed. The outcome work's outcomes were quite positive, with success rates typically exceeding 90%. Online sources (https://plantvillage.org/) were used to get the dataset for this study. The tomato plant leaves in the dataset total 1882 pictures and is classified into 5 classes. The categories are Leaf Mold Images, Early Blight Images, Late Blight Images, Bacterial Spots Images, and Healthy. Further features are extracted and provided to the classifier to predict different types of disease on different types of leaves. This method when implemented on okra crops classifies the different diseases infected on okra leaves. The recovered vector is saved in training sets and datasets, and Densenet is once more used for training. Based on the parameters provided to the model [4] , the generated feature vector will be used for the testing image of leaves. Then feature vectors are created, which under-trained and under-tested classifiers were supposed to save for predicting accurate results.

Images of plant disease samples are either taken under controlled conditions in a laboratory with a white background or in the field with a complex background, according to Meunkaewjinda et al. (2008) (Jadhav and Patil, 2016) [5]. For image acquisition, a variety of tools, including Android mobile phones, CCD color cameras, CCD multispectral cameras, and hyperspectral imaging systems are used [6]. Images of plant disease samples are acquired either in the field with complex background Meunkaewjinda et al., (2008) or under controlled conditions in the laboratory with white background (Jadhav and Patil, 2016). Various devices, namely Android mobiles, CCD color cameras, CCD multispectral cameras, and hyperspectral imaging systems are used in image acquisition. The quality of the image depends upon the camera and the orientation of the optical axis to the leaf plane.

Agriculture-related species Leaf is used as the camera's input. The camera lens provides a thorough description of the leaf, identifying its healthy and diseased areas. Using picture pre-processing, this leaf outlet is segmented and transformed into pixels. Five plant-related diseases are used to assess their program: early sear, cottony form, colorless shape, late burn, and little whiteness. The images captured in most of the cases may consist of noise due to weather and environmental changes, blurring, unwanted objects, poor contrast, and brightness. For instance, photographs sent by farmers from the field contain such noises. The image acquired is thus subjected to pre-processing like color space conversion, cropping, filtering, and enhancement. The noise-free image that has been improved, sharpened, and cropped is then divided into several sub-images. These elements of the input image are divided into pieces that are healthy and unhealthy. In order to segment a certain feature of interest, such as a spot, border, or colour distinction, it is necessary to consider that feature. Only the segments of our interest—the sick area and deformity or deficiency symptoms—will be used by the pipeline. The characteristics of these segments of interest include histogram features (begin, peak, and end), normal part, leaf spot, number of spots, spot and halo level of grey, relative spot area, spot and halo region around the spot, body of plant part, temperature, and moisture.

Typically, K means clustering [6] is employed to create image subgroups based on the Euclidean distance matrix for attributes.

The pixels are assigned to their nearest centroid (k) for grouping. The cluster centers are computed simultaneously by taking the mean of all the pixels in the cluster. These two steps are repeated until the whole clustering is completed.

[They used a public dataset for plant leaf disease detection called PlantVillage, which was curated by Sharada P. Mohanty et al. [7] for this purpose. The dataset comprises 87000 RGB photos of healthy and diseased plant leaves, divided into 38 classes. They have only chosen 25 classes to test the method on. Apples, corn, grapes, potatoes, tomatoes, and other plants are used. The convolutional layer's main goal is to extract features, and it does this by learning a useful feature representation from the input image. The neurons are positioned in the convolutional layer in accordance with feature maps. Every neuron in the feature map has its own receptive field, which is connected to the surrounding neurons in the layer above by a collection of randomly chosen weights known as a filter bank.

It convolves the input neurons to create the subsequent feature vector based on the learned weights, and then the outputs are forwarded using a non-linear activation function. A single feature vector typically has the restriction that all of the neurons have the same weights; however, distinct feature mappings within the same convolutional layer have varying weight matrices, allowing some meaningful features to be retrieved in stages.

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁘ Impact Factor 8.102 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125191

Three major illnesses that affect paddy were considered in this paper: Rice Blast, Brown Spot of Rice, and Sheath Blight of Rice. The organism Pyricularia oryzae, which causes Blast Disease, causes small, water-soaked bluish-green spots on the paddy leaves that have a grey center and a dark brown rim. The unclassified feature data extracted from the image are ranked to a particular class (disease) by various machine learning processes or classifiers. KNN or K- Nearest Neighbours algorithm [8] is the easiest classification technique under supervised machine learning used for prediction purposes. A vector line (in the case of 2D), plane (in the case of 3D), or hyper-plane (in the case of n dimensions; is prepared by using transformation of a group of data, i.e. kernel) is used to determine the class of a new unknown object or data in the support vector machine, another well-known supervised machine learning algorithm. Besides these, Alex Net, PNN, Fuzzy-C-mean clustering, Self-Organizing Map (SOM), Neural network SOM-NN, Convolutional neural network (CNN), etc. are different machine learning processes or classifiers used in the prediction of diseases of major crops.

There are 14 different types of crops in the public dataset called Plant Village, which can be downloaded from Kaggle. They decided to use a tomato leaf [9] for this project. This dataset on tomato crops has 10 different classes. One of them is in a healthy class, while the other nine are ill classes. In this effort, 300 photos were captured for each class. Because empirical studies demonstrate that utilizing 20–30% for testing and 80–20% for training allows a model to achieve the greatest outcome, the data split ratio was set at 70:30 for training and testing, respectively. Three thousand photos in total were used for training and testing. The leaf is subjected to image processing, including scaling, filtering, and segmentation. Only features like the redness of the leaf and swollen leaves are extracted throughout the feature extraction process. Based on a diagnostic of the leaf made by a comparison algorithm, the controller displays the leaf ailment. Ethernet and Wi-Fi modules were employed.

Shah et al. (2016) [10] suggested creating an image database using familiar photos from the live farm, which would be directly captured in digital form with numerical values for use in digital image processing. The Redmi Note 5 camera's high pixel intensity was employed by Ramesh and Vydeki (2020) to take pictures of both healthy and sick rice leaf portions. Resized to 256 * 256 Pixels were all the leaf samples collected from the rural region. Rice leaf illnesses, such as bacterial leaf blight, brown spot, and leaf smut, are represented in the secondary dataset made available by the UCI machine learning repository under the image acquisition step (https://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Diseases). Eight pixels surround the center pixel in the nearest neighbor pixel. Only a small number of nearby pixels offer more pertinent information since they are closer to the center. In this case, it is thought that the eight nearby pixels make up a pattern. The central pixel, together with the two closest neighbourhood pixels, are compared with each individual neighbourhood pixel. The pixels that are two in number and adjacent are either horizontal or vertical pixels because they are both the most likely to be considered neighbours.

CONCLUSION

The Survey has helped in understanding the diseases and their diversity of them. Much Research has proposed methods to detect plant disease without manually doing it. Detecting using ML and Deep learning algorithms has reduced time wastage on manually detecting the bacterial attack. Many works proposed have successfully achieved an accuracy level of 93%-98%. We are attempting to train such a model to produce better, more efficient, and more accurate outcomes than previously presented initiatives, thereby assisting farmers, crops, and nutrient levels.

REFERENCES

[1] www.Google.com

[2] D. Mondal, A. Chakraborty, D. K. Kole, and D. D. Majumder, "Detection and classification technique of Yellow Vein Mosaic Virus disease in okra leaf images using leaf vein extraction and Naive Bayesian classifier," 2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI), 2015, pp. 166-171.

[3] Norhalina Senan, Muhammad Aamir, Rosziati Ibrahim, N. S. A. M Taujuddin and W.H.N Wan Muda, "An Efficient Convolutional Neural Network for Paddy Leaf Disease and Pest Classification," International Journal of Advanced Computer Science and Applications(IJACSA), 2020, 11(7).

[4] Aravind Krishnaswamy Rangarajan, Raja Purushothaman, and Anirudh Ramesh, "Tomato crop disease classification using pre-trained deep learning algorithm," Procedia Computer Science, Volume 133, 2018, pp. 1040-1047.

[5] Detection of cucumber leaf diseases using the newly generated techniques of the Haralik features and the technique of the local tri-directional pattern by Abdul Rehman, and Zain Tariq (2021)

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁘ Impact Factor 8.102 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125191

[6] Nutrient Deficiencies and the detection of disease using image processing(2021) “Ankit Kumar Ghorai,S Mukhopadhyay”

[7] Leaf Disease Detection using Digital Image Processing with SVM Classifier(2019)"Sagar Gaikwad, Sagar Shinde"

[8] Plant Disease Detection Using Image Processing and Machine Learning "Pranesh Kulkarni, Atharva Karwande, Tejas Kolhe, Soham Kamble, Akshay Joshi, Medha Wyawahare"

[9] Paddy Leaf Disease Detection Using Image Processing and Machine Learning(2019) "Dr. Neha Mangla, Priyanka B Raj, Soumya G Hegde, Pooja R"

[10] Detection and Classification of Tomato Crop Disease Using Convolutional Neural Network (2022)by "Gnanavel Sakkarvarthi,Godfrey Winster Sathianesan ,Vetri Selvan Murugan ,Avulapalli Jayaram Reddy ,Prabhu Jayagopal 3 and Mahmoud Elsisi"

[11] Literature Review of Diagnosis Rice Leaves Diseases Using Image Processing(2021) "R. I. L. Jayasooriya, Samantha Mathara Arachchi"