EasyChair Preprint

№ 9643

Plant Pathology Disclosure

Pratiksha Bankar, Rucha Garud, Suraj Suryavanshi,
Vaibhav Matne and Vinod Wadne

March 6, 2023

Plant Pathology Disclosure

Pratiksha Bankar, Rucha Garud, Suraj Suryavanshi, Vaibhav Matne, Vinod Wadne, Computer Engineering,
(JSPM's Imperial College of Engineering and Research),
pbbankar2001@gmail.com, garudrucha@gmail.com, suravanshisuraj37@gmail.com,
matnevaibhav15@gmail.com,vinods1111@gmail.com,

***Abstract*—India is well known for its agricultural industry. Agriculture is the main sector in India. Hence, protecting the agriculture industry is one of most important aspect. In order to boost productivity, we must overcome the various diseases caused by bacteria, viruses, and other organisms. Implementing and understanding traditional methods of disease detection, such as monitoring, observing, or treating plants, is quite challenging. We must therefore design a system that yields precise results while being simple to understand. Using various machine learning and deep learning algorithms, we decided to develop an application or software to help in disease identification or digital image processing. As a result, this application is based on two techniques: deep learning for disease detection and recognition, and image processing techniques such as machine learning (ML). We will employ a variety of machine learning techniques, including image processing, image segmentation, classification, and feature extraction. Several diseases deplete the chlorophyll in leaves, resulting in brown or black spots on the leaf's surface. The basic idea behind this technique is to divide an image into different parts, extract different features from it, and then compare the current dataset with the dataset stored in a database to obtain the result. This can be accomplished through the use of feature extraction and classification. The features are extracted using the Grey Level Co-occurrence Matrix (GLCM), and they are classified using a classification algorithm such as SVM (Support Vector Machine).**

Android applications that use machine learning and deep learning techniques will undoubtedly increase agricultural productivity.

I. INTRODUCTION

INDIA has 159.7 million hectares of cultivatable land area, which is the second largest in the world. There are many obstacles in Indian agriculture like instability, plant disease, and much more. [1] To inspect tree leaves humans are bound to physically traverse large areas of the agricultural farms, hence, this may be more time-consuming and can be costly. Crop diseases cause serious financial losses to farmers. Many computer vision and artificial intelligence based approaches have been developed over the year for the automatic detection and identification of disease and pest attacks. One major flaw of these approaches is that they detect disease based on apparent characteristics of the disease. At this stage, the disease is usually untreatable or has already caused serious damage to the crops. The losses due to disease at this stage are untreatable. [2]There is a need for solutions that can predict the occurrence of the disease attack before it causes losses. Detection of plant leaf infections is an important task in agriculture, to overcome plant disease issues various researchers have used image processing along with ML/DL algorithms. Research has shown that for the recognition and categorization of plant diseases, image-processing techniques can be considered beneficial. So, in this study hereby we propose some recent advances in DL models that can be used to identify plant diseases. Moreover, the study identifies the various research gaps in the existing literature to make the vision of symptoms experienced in the plant leaf disease classification process in a better sense.

II. LITERATURE REVIEW

Johan and Arif., [1] The field of study, known as deep learning (DL), seems to offer a lot of potential for improved accuracy. In order to identify and categorize the signs of plant diseases, numerous developed or modified DL architectures are used in conjunction with a number of visualization techniques. Moreover, a number of performance indicators are employed to assess these structures and methodologies. This article offers a rough justification of the DL models used to depict various plant diseases. Also, several research holes are found that might be filled to increase transparency for identifying plant diseases even before their symptoms are plainly visible. S.Sladojevic., [3] The most popular methods for identifying plant diseases at the moment are artificial neural networks (ANNs) [4] and Support Vector Machines (SVMs). To improve feature extraction, they are used with various picture preparation techniques. ANN is an information-processing paradigm used in machine learning and cognitive science that was motivated by the way biological nerve systems, like the brain, process information. The brain is made up of a lot of neurons that are closely connected and cooperate to solve difficulties. [5]

According to M. Appalanaidu et al., [6]segmentation of the picture is used for background subtraction. KNN, ANN, and SVM methodology are used to manage the classification strategy. [7]Nearest distance between training and tested individuals is used to classify samples in KNN [6]. Models for extraction thresholding technique and morphological operation have been created by S. M. Nagashett et al. [7]. Then a multiclass SVM is used as a classifier. $L * A * B$ colour areas, which are a group of marks produced by analyzing the color and luminance components of various sections of the image, are supported for segmentation.The GLCM is used to extract features. [8] Considered samples of plant leaves taken using a camera, including rose/beans (bacterial disorder), lemon (sunburn disorder), banana (early scorch), and beans (fungal). The background maltreatment thresholding rule uses the inexperienced regions. Lastly, the metameric

image is induced using the genetic rule. In order to obtain useful alternatives from the metameric images, the color co-occurrence is tailored. For classification purposes, the SVM classifier uses the Minimum Distance Criterion. M. Saleem et al. [9] investigated a number of image processing methods to recognise plant illness. Many stages, such as image processing, image segmentation, feature extraction, and classifier, must be taken once the image is taken as an input in order to identify and categorize leaf diseases. The disease areas are identified by segmentation utilizing k-means clustering, and statistical Grey-Level Co-Occurrence Matrix (GLCM) is employed for feature and classification. [10] Similar to KNN, this study also aids in choosing the right algorithm by highlighting the differences between them.RBF (Radial Basis Function), BPN (Back-propagation Network), PNN (Probabilistic Neural Networks), and SVM (Support Vector Machine). The merits and disadvantages of these algorithms are identified in this paper.

[8]Machine learning (ML) technologies provide machines the ability to interact and behave like people. Also, it has the ability to make decisions for people. ML is useful for categorizing plant diseases. If the diseases are discovered before they spread to wide areas, the problems that the world has now in relation to the diseases impacting humans and plants can be lessened. [11]Today, ML is widely used throughout the world.The professionals can analyze plant diseases and identify their sources early on thanks to the various ML and DL techniques deployed. Using the GLCM technique, which reaches a 100% accuracy level when it inputs the classifier, the enhancement phase aided the texture feature extraction phase. The following characteristics are crucial for using SVN classification levels: contrast of intensity between adjacent pixels on the picture, correlation of adjacent pixels on the image, entropy (the clutter of intensity in a region), energy of uniformity of pixels, and homogeneity of comparable pixels. [12].

III. METHODOLOGY

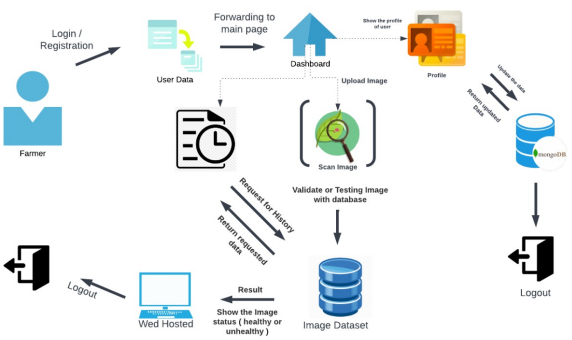

Fig. 1. Show the workflow of Plant Pathology Disclosure

The system's operation is shown in the above figure. When the application is opened for the first time, the user will see the front page, which is nothing more than the landing page, which shows the small details about the project and also contains the contact section. Over their will are options like "login" or "register." If the user is not already registered with the system, he or she must first register. After clicking on "Register," a registration form will get opened that contains some details like First name, Last name, Email, Password, Mobile, and Aadhar Card No. After submitting this form, an email will be sent to the user by the system to their registered email address. That email also contains a login form link, or the user can login by simply clicking on "login." After entering the correct details, the user will be directly redirected to the main page. If a user forgets their password, they can simply click on "forget password," and a dialogue box will open after submitting an email about their reset password, which will be shared with the user via email. After clicking on that link, a reset password page will open, which contains a new password and a confirm password box that will help the user change the password, but the password must be in standard format.

The dashboard, which includes home, history and upload images, which is located on the main page after a user login successfully. After clicking on the upload image option, the user will see options like "capture" or "upload." After submitting the image, a result will be generated and shown on the same page, which includes stage of the disease, impact of disease on the plant and what care the farmer needs to take. Our system helps farmers better understand the amount of pesticide they need to use for plant health. It also informs the farmer whether or not the disease can be treated with organic fertilizer. If possible, which fertilizer is beneficial should also be mentioned.

Figure 2 illustrates the system's architecture or flow from login to logout.

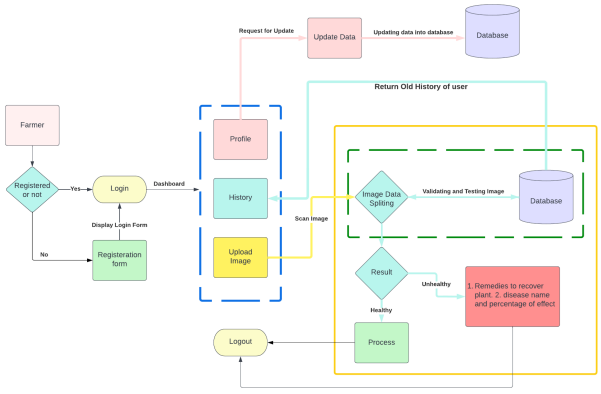

Fig. 2. Illustrations, System Architecchture.

IV. DATASET

The study's data [13] sets include descriptions of the leaves both before and after the diseases had an impact on them. Tables and images of the leaves that were taken in the fields are used to display the data. For the benefit of the readers' understanding, the data is analysed and categorised. A collection of information could be referred to as an information set (or dataset). [14]A data set is compared to one or more database tables in the case of unthinkable information, where each table's column corresponds to a certain variable and

each push is compared to a specific record of the data set by address. For each component of the information set, the information set keeps track of values for each of the factors, for instance, the protest's weight and height. A collection of archives or documents can also be included in information sets.// In machine learning, a dataset is a group of data points that may be analysed and predicted by a computer as a single entity.

V. Image Processing

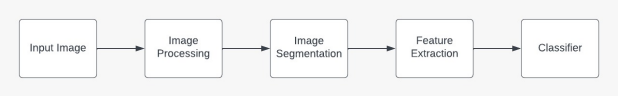

Fig. 3. Illustrations, The process after image upload or image received as input.

1. Image analysis: [15]Image analysis plays a crucial role in our system, as we wanted to extract details from images and use those details for comparison and generating results. It is one of the techniques used to perform operations on images in order to get useful information. An image can be the input for image processing, and the features or characteristics that go along with it can be the output. There are two techniques to process images. First, analogue processing digital images (DIP) [16]. For physical copies like prints and photos, analogue is employed. Digital photos can be edited using digital image processing (DIP). In our system, we are going to use A typical method in digital image processing (DIP) is image segmentation.

2. Image segmentation : Using the image segmentation approach, you can create many "image segments" from a single digital image. It enables the image's complexity to be reduced in the future, allowing for simpler image processing or analysis. An area of the image picked by the segmentation algorithm can be given to the detector instead of having to process the complete image. [16] It helps in processing only the damaged area of an image rather than the entire image. There are various types of image segmentation. We will, however, use region-based segmentation, which divides images into areas with similar properties. If an image is noisy, region-based segmentation techniques are recommended over edge-based segmentation techniques. [17] Binary thresholding is the process of transforming an image into a binary image, where each pixel is either black or white. All pixels with brightness levels below the threshold are black, and all pixels with brightness levels above the threshold are white. This is the purpose of the threshold value. As a result, the objects in the image are now separated into separate black and white regions.

3. Feature Extraction : All the data we collect through image segmentation is vast and takes various forms. [18] To understand this data, we need to process it, and then we use the next step in image processing, which is nothing but feature extraction. Feature extraction is the process of "dimensionality reduction," which involves breaking down a large initial set of raw data into smaller, easier-to-manage groups. [19]After dividing all the data into various segments, those data will have a large number of variables With the selection and combination of variables into features, this technique helps to extract the best features from those enormous datasets. These properties are simple to comprehend.

4. Image classifier :

Image classification is the process of arranging and grouping pixels or vectors. Within an image, based on specific rules, classification helps identify the particular image. [14] Further classification can be supervised or unsupervised. Image classification is simply the process of labelling an image with a label from a predefined set of categories [1].

The comparisons made by researchers in this sector are based on scanning photographs of both healthy and unhealthy plants. Plant defects can be found using deep learning (DL). Digital images are represented by the function I(x, y) or I(x, y, z), where "I" denotes pixel intensity and (x, y) or (x, y, z) denotes the image coordinates (for binary/grayscale or RGB images, respectively). [20] In order to categorize illnesses according to how they affect plants, sick plants' leaves are examined using pixel-wise processes. Plant diseases can be recognized by looking at the patterns on these leaves. It's possible that the computer won't be able to directly process pictures taken in the fields. Images are converted into data that can be understood by machines and computers. Many diseases affect different areas of the leaf; some target the whole leaf, while others target the edges or stalk. Using an image-based detection method, it is possible to classify the diseases by analysing the photographs of the leaves.

[21]Background noise should be reduced before feature extraction in order to get accurate results. Hence, the RGB comes first and is likely the most crucial. The image is grayscaled before being smoothed with a Gaussian filter. Afterwards, Otsu's thresholding algorithm binarizes the image. To fill up the tiny gaps in the foreground, a morphological change is used to the binarized image. [22]After foreground detection, a segmented image is subjected to the bitwise AND operation to create an RGB image. The original colour images and binarized versions are then combined. leaf. After segmentation, the shape, texture, and colour aspects of the image are retrieved. Using contours, the leaf's area and perimeter are computed.

All the points along the boundaries of objects with the same form, colour, or intensity are connected by contour lines. [13] With an RGB image, the mean and standard deviation of each channel are also calculated. estimated. We translated the image to HSV colour space and calculated the ratio of pixels with hue (H) channel pixel intensities ranging from 30 to 70, as well as the total number of pixels in one channel, to determine the quantity of green colour in the image. By deducting the green colour component from 1, the non-green portion of the image is calculated. SVM is one of the machine learning algorithms that is used to categorise data into distinct groups using different categories.

In order to categorise two classes, SVM creates a decision boundary that is a hyperplane between them. [23] SVM's

primary goal is to separate a given dataset. SVM is implemented using a variety of libraries, including TensorFlow, Keras, Colab, etc.

[2]A technique called RF (Random Forest) combines the predictions of various decision trees. The combined outputs of all the trees are used as the output. Each tree is trained using a random subset of all the data provided.

VI. CONCLUSION

Data is collected from the "new plant disease detection" website, which has data of different diseases on different plants. We have a dataset which contains the images of leaves of plants. Images of both healthy and unhealthy plant leaves are available in this dataset. After Upload of image it is compared with the images of leaves present in the dataset. By image processing and machine learning algorithms these images are compared with the uploaded image. If the image got matched with any image present in this dataset then disease got detected. If the plant is healthy it also gets by this comparison. The accuracy is not that much in this case. Thus, neural networks used to increase accuracy. India's commercial growth depends on agriculture.This helps in increasing the production by detecting diseases on plants at an early stage. It helps farmers to automatically get the health information of his field.

REFERENCES

[1] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease detection and classification by deep learning," *Plants*, vol. 8, no. 11, 2019.

[2] S. Ibad, W. H. Suristiyanti, M. N. A. Farah, N. Rijati, and C. Supriyanto, "Application of grayscale co-occurrence matrix (glcm) method for classification of quality type of guava leaves as traditional medicine using neural network algorithm," in *2022 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2022, pp. 310–313. DOI: 10.1109/iSemantic55962.2022.9920397.

[3] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanović, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, 2016.

[4] K. Nagasubramanian, S. Jones, A. Singh, S. Sarkar, A. Singh, and B. Ganapathysubramanian, "Plant disease identification using explainable 3d deep learning on hyperspectral images," *Plant Methods*, vol. 15, Aug. 2019. DOI: 10.1186/s13007-019-0479-8.

[5] M. Appalanaidu and G. Kumaravelan, "Plant leaf disease detection and classification using machine learning approaches: A review," in Apr. 2021, pp. 515–525. DOI: 10.1007/978-981-33-4543-0_55.

[6] M. Appalanaidu and G. Kumaravelan, "Plant leaf disease detection and classification using machine learning approaches: A review," in Apr. 2021, pp. 515–525. DOI: 10.1007/978-981-33-4543-0_55.

[7] S. M. Nagashetti, S. S. Biradar, S. D. Dambal, C. G. Raghavendra, and B. D. Parameshachari, "Detection of disease in bombyx mori silkworm by using image analysis approach," *2021 IEEE Mysore Sub Section International Conference (MysuruCon)*, pp. 440–444, 2021.

[8] "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 305–310, 2022, International Conference on Intelligent Engineering Approach(ICIEA-2022), ISSN: 2666-285X.

[9] "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," *Smart Agricultural Technology*, vol. 3, p. 100 083, 2023, ISSN: 2772-3755.

[10] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, and R. Pramodhini, "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 305–310, 2022, International Conference on Intelligent Engineering Approach(ICIEA-2022), ISSN: 2666-285X.

[11] D. Singh, A. K. Singh, S. Chaurasia, and L. Varshney, "Plant disorder and disease detection using image processing," in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022, pp. 2259–2262. DOI: 10.1109/ICACITE53722.2022.9823616.

[12] G. Kumar and P. Bhatia, "A detailed review of feature extraction in image processing systems," Feb. 2014. DOI: 10.1109/ACCT.2014.74.

[13] M. Yu, X. Ma, H. Guan, M. Liu, and T. Zhang, "A recognition method of soybean leaf diseases based on an improved deep learning model," *Frontiers in Plant Science*, vol. 13, 2022.

[14] A. Yadav, P. K. Yadav, S. Toomula, S. Jaiswal, and P. Patro, "Qualitative texture analysis on detection of plant disease," in *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2021, pp. 1758–1761. DOI: 10.1109/ICECA52323.2021.9676026.

[15] S. Ramesh Maniyath, V. P V, N. M, *et al.*, "Plant disease detection using machine learning," Apr. 2018, pp. 41–45. DOI: 10.1109/ICDI3C.2018.00017.

[16] R. M. Prakash, G. Saraswathy, G. Ramalakshmi, K. Mangaleswari, and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2017, pp. 1–4. DOI: 10.1109/ICIIECS.2017.8275915.

[17] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523–3542, 2022. DOI: 10.1109/TPAMI.2021.3059968.

[18] D. Das, M. Singh, S. S. Mohanty, and S. Chakravarty, "Leaf disease detection using support vector machine," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, 2020, pp. 1036–1040. DOI: 10.1109/ICCSP48568.2020.9182128.

[19] P. Kartikeyan and G. Shrivastava, "Hybrid feature approach for plant disease detection and classification using machine learning," Jun. 2022, pp. 665–669. DOI: 10.1109/AIC55036.2022.9848939.

[20] M. Vaishnnave, K. Devi, P. Srinivasan, and G. Arut-PerumJothi, "Detection and classification of groundnut leaf diseases using knn classifier," Mar. 2019, pp. 1–5. DOI: 10.1109/ICSCAN.2019.8878733.

[21] K. S. F. Azam, F. F. Riya, S. Al Mamun, and M. Kaiser, "A novel approach of detecting image forgery using glcm and knn," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 2021, pp. 125–129. DOI: 10.1109/ICICT4SD50815.2021.9397057.

[22] S. Omar, R. Jain, and V. Bali, "Leaf disease detection using convolutional neural network," in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, vol. 1, 2022, pp. 53–56. DOI: 10.1109/COM-IT-CON54601.2022.9850950.

[23] S. Pawar, S. Shedge, N. Panigrahi, A. P. Jyoti, P. Thorave, and S. Sayyad, "Leaf disease detection of multiple plants using deep learning," in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, vol. 1, 2022, pp. 241–245. DOI: 10.1109/COM-IT-CON54601.2022.9850899.