Leaf Disease Classification Using SVM And CNN Algorithms

[1] Jayaprakash Y M, Lecturer, Department of Computer Science and Engineering, Government Polytechnic, Nagamangala, Mandya.

ABSTRACT: To detect leaf diseases, we are utilizing image processing with a Convolution neural network (CNN). A convolutional neural network (CNN) is a form of artificial neural network that is specifically intended to process pixel input and is used in image recognition. The data collection consists of five plant leaf diseases. These images are preprocessed by converting the HIS color to RGB images. Using the K-mean cluster algorithm, the'H' elements of the affected portion of the leaf are segmented. Using the GLCM algorithm, feature extracted from cluster part of leaf image. The features are tested, and train then sorted based on maximum accuracy. With the best linear combination SVM and CNN get input of selected of mage features. There are70%images for training, and 30% images are being tested. To evaluate the data properly, the K-fold method is used. The CNN deep learning algorithm offers95.33 % accuracy, and the SVM machine learning algorithms offer 85.07 % accuracy. Therefore, CNN is said to be better than SVM. In future other leaf disease detection is also performing using various dataset with different diseases.

KEYWORDS: Diseases, Feature extraction, Convolution, Biological system modeling, Micro organisms, Training

I. INTRODUCTION

India is an agrarian nation with approximately 70% of its population relying on agriculture for their livelihoods. For farmers, the process of selecting appropriate crops and finding effective herbicides and pesticides is diverse and extensive. The presence of diseases in plants has a significant detrimental impact on agricultural products, leading to a substantial reduction in both their quality and productivity [1]. Therefore, strategies and technical knowledge and the field became an important matter to be mastered. The systematic and structured should be developing so that they will use by operators to increase the overall production. In most cases, disease symptoms are seen on the leaves, stem, and fruit [2]. It needs to be detected that disease but very time appointing an expert would may be more costly. Continuous monitoring with naked-eye observation is not possible for a farmer. So, we use digital image processing techniques[3].

Fig1:leaf (Taxonomy, Characteristics)

Particular medications can be custom fitted to battle particular pathogens if plant illnesses are accurately analyzed and recognized early and ecological increases[4].Initially, the input RGB image is converted into the HIS model. The k- mean segmentation is an efficient segment of the disease affected area of leaves of crops. The GLCM is a features extraction algorithm that measures the feature values from the ROI disease-affected image from the „H" component[5]. Exact disease finding is a difficult task for farmers which results in loss of income to the farmers and the state[6].In the classical method, expert people detected diseases in leaves by naked eyes which are very expensive for farmers [7].

India heavily relies on its agricultural sector, with over 60% of the population engaged in farming activities. As a result, the country's economy is largely dependent on agriculture. Unfortunately, India has been facing various challenges that have negatively impacted the agricultural sector. The change in weather patterns and the increasing incidence of plant diseases have led to a reduction in crop yields. This has had devastating consequences for farmers who are facing economic hardships. One of the biggest challenges in managing plant diseases is the difficulty in detecting them at an early stage. Typically, diseases are only noticeable when they have spread widely, making it challenging to implement effective control measures.To address this issue, technology can be used to provide solutions that aid in the early detection of plant diseases.

Plant diseases can be caused by various factors such as bacteria, fungi, or viruses, and can range in severity from mild leaf or fruit damage to crop destruction. There are several types of plant diseases such as black spot, powdery mildew, Fusarium Wilt, Gray Mold, Leaf Blight, White Mold, Scab, and Fire Blight. For this project, we collected leaf images from Kaggle and performed image preprocessing to remove noise and convert RGB images to gray scale. To classify these images, we used convolutional neural networks (CNNs), which have proven to be effective in image classification tasks. By leveraging technology to detect and diagnose plant diseases, we can help farmers take appropriate control measures to mitigate the spread of the disease and prevent significant crop losses. This, in turn, can contribute to improving the livelihoods of farmers and boosting the country's economy.

II. LITERATUREREVIEW

Leaf images affected by the disease are pre-processed by resizing the image or color conversion of an image or calculate the histogram of an image. To remove the noise from image filters isused and enhance the image quality. The color, shape, and texture features are calculated by different segmented methods, k-mean is the most popular method used for finding the region ROI of the image. In some work combination of texture, color, and shape features are used to classify the various diseases on a leaf. The texture features are mostly used in the disease detection of the leaf. These calculated values were given to the classifier and the leaf disease was classified by this method. Machine learning classifiers are used for leaf disease classification. After studying the previous work, found the research gap for our research work.

Plant leaf samples are collected from open farms with a different session at the early stage, middle stage, and last stage of the disease. These samples are collected from July to February of the year. In this study, four diseases and one healthy leaf were observed in plant leaves. The affected diseases are harmful to crop production. There is a variety of disease spot which are same with different disease spot which makes confusion to recognize the disease[9].The wrong prediction of the disease may go in the wrong direction regarding spraying pesticide or chemical treatment. The loss of money and time in this is very useless [10]. Leaf plants affected by fungal, bacterial, and viral diseases are captured by the digital camera.

Several research papers have been conducted on using image processing and machine learning algorithms for plant disease detection and classification. In one paper by Pushkara Sharma, Pankaj Hans, and Subhash Chand Gupta, a dataset of over 2000 images was used and divided into 19 different classes. Gaussian Blur was used for noise removal, and images were converted from RGB to HSV for image preprocessing. K-means clustering was used forsegmentation. Four classifiers were tested, including logistic regression, KNN, SVM, and CNN. The CNN classifier showed the highest accuracy of 98%. [6] Another paper by Ms. Deepa, Ms Rasmi N, and Ms. Chinmai Shetty used machine learning techniques to identify plant leaf disease. Gray cooccurrence matrix (GLCM) was used for feature extraction, and K-means clustering was used for clustering. SVM was used as the classifier, and four classes were defined, including Alternaria Alternata, Anthracnose, Bacterial Blight, and healthy leaves.

[8] In a paper by Vaishnavi Monigari, G. Khyathi, and T. Prathima, a dataset of over 20,000 images of diseased and healthy plant leaves was used and classified into 15 classes to train the CNN. Open CV framework was used for image processing, and image augmentation was used to increase the number of images in the dataset. The developed model achieved 90% accuracy and could distinguish healthy leaves from eight diseases.

[2] Finally, a paper by Marwan Adanan Jasim and Jamal Mustafa AL-Tuwaijari focused on plant leaf disease detection and classification using image processing and learning techniques for tomato, pepper, and

potato leaves. The dataset used consisted of over 20,000 images, and CNN was used for classification, including 12 classes for diseased leaves and 3 classes for healthy leaves. The model achieved an accuracy of 98.29% for training and 98.029% for the testing dataset. [4]

III. METHODOLOGY

To detect plant leaf diseases, the first step is to collect a dataset of images with both healthy and diseased leaves that accurately represent real-world scenarios. The collected images are preprocessed by removing noise and irrelevant information through techniques such as normalization and data augmentation. The selected model for this task is typically a Convolutional Neural Network (CNN). The dataset is then split into training, validation, and testing sets,and the model is trained on the training set while monitoring its performance on the validation set. After training, the model is evaluated on the testing set to measure its accuracy, precision, recall, and F1-score. If the model is not performing well, optimization techniques such as transfer learning or data augmentation can be used to improve its performance. Once the model is optimized, it can be deployed in a real-world scenario by integrating it into an application. This process involves a cyclical approach of data collection, preprocessing, model selection, training, evaluation, optimization, and deployment until the desired level of accuracy is achieved. Data Collection: The first step is to collect a dataset of plant leaves with and without diseases.

The dataset should be representative of real-world scenarios where the model will be deployed. Data Preprocessing: Once the dataset is collected, it needs to be preprocessed to remove any noise or irrelevant information. This may involve techniques like data cleaning, normalization, and augmentation. For instance, the collected images are pre- processed to convert RGB images into grayscale images and then into an array form. Model Selection: The next step is to select an appropriate deep learning model. Convolutional Neural Networks (CNNs) are commonly used for image classification tasks such as plant leaf disease detection. The chosen CNN comprises several layers, including Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D.Model Training: The selected model needs to be trained using the preprocessed dataset. This involves splitting the dataset into training, validation, and testing sets.

Fig2:Leaf Disease Detection and Classification Using Convolutional Neural Network

The model is then trained on the training set while monitoring its performance on the validation set. Model Evaluation: After the model is trained, it needs to be evaluated on the testing set to measure its accuracy, precision, recall, and F1- score. This step determines the model's performance on unseen data. Model Optimization: If the model is not performing well, optimization techniques like transfer learning, fine tuning, or data augmentation can be used toimprove its performance. For instance, additional layers can be added to the CNN to improve its accuracy. Model Deployment: Once the model is optimized, it can be deployed in a real-world scenario. This involves integrating the model into an application, such as a mobile app or a web service. This methodology involves a cyclical process of data collection, preprocessing, model selection, training, evaluation, optimization, and deployment until the desired level of accuracy is achieved.

The goal is to Convolutional Neural Networks(CNNs) are widely adoptedfor analyzingand classifying digital images, especially in the field of plant leaf disease detection. These algorithms are designed to effectively capture and process image features through multiple layers of filters and nonlinear operations. The CNNs are proficient in handling large datasets and can dynamically learn new features from them in a supervised manner. By learning the important features from the input images, CNNs can make accurate predictions about the presence of diseases in plant leaves. Keras is a high-level API used to build and train deep neural networks for various machine learning tasks, including image processing. It offers a user-friendly interface for constructing complex neural network models by providing pre-built layers and modules.

Keras is written in Python, which is a popular programming language in the field of machine learning. With Keras, developerscan easily build deep neural networks, including CNNs, without having to worry about the low-level details of the underlying hardware and software. OpenCV is a popular open-source library that provides various computer vision and deep learning algorithms for image processing, including feature extraction and classification. OpenCV is written in C++ and supports multiple programming languages, including Python. It has a wide range of functions and toolsforimageanalysis,segmentation,andobjectdetection.OpenCViswidelyusedinthefieldofcomputervisiond ue to its ease of use, efficiency, and cross-platform compatibility. By using OpenCV, developers can implement various image processing techniques and algorithms to enhance the accuracy of plant leaf disease detection models.

IV. RESULTS AND ACCURACY

The pre-processing stage in plant leaf disease detection involves converting the plant leaf images to grayscale and binary format. This process simplifies the images and reduces their complexity, making them more suitable for deep learning model processing. Grayscale conversion helps in reducing the color channels to a single channel and improves contrast, highlighting the edges of the leaves for easier identification of patterns and features. Converting grayscale images to binary format involves thresholding the pixel values to either black or white, further simplifying the images and highlighting the edges of the leaves more prominently. By isolating the regions of interest, i.e., the plant leaves and their features, and removing irrelevant information like the background or soil, binary conversion simplifies the image and improves its clarity. The result is a noise-free, simplified image that highlights the important features of the plant leaves, making it easier for the deep learning model to identify and classify different diseases accurately, leading to better overall performance of the system.

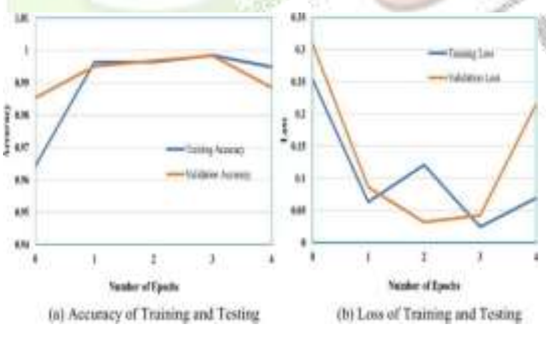

Fig3:leafdiseasedetectionandclassificationusingmodifiedtransferlearningmodels

The solanum nigrum leaf disease detection using deep learning approach project involves using convolutional neural networks (CNNs) and OpenCV for the detection of Solanum nigrumleaf diseases. The steps of the project involve image acquisition, pre-processing, feature extraction using OpenCV, CNN structure design, and image classification. The project achieved an accuracy of around 97% after training the CNN for 20 epochs. Additionally, the project involved predicting the appropriate pesticides and medicine based on the detected disease. The accuracy of the prediction was around 92%. Overall, the project demonstrated the effectiveness of deep learning algorithms and computer vision techniques in the field of agriculture for automated Solanum nigrum disease detection and treatment recommendation. Once the deep learning model has detected the Solanum nigrum disease accurately, it can be linked to a database of known Solanum nigrumdiseases and their respective treatments. This database can contain information about the disease symptoms, affected Solanum nigrumparts, and recommended treatments, including the appropriate medication. The model can then use this information to predict the appropriate medication for the detected Solanum nigrum disease. This can be done through rule-based systems, decision trees, or other machine learning algorithms. The predicted medication can then be displayed to the user, along with relevant information such as dosage and application instructions.

V. CONCLUSION

The detection and diagnosis of Solanum nigrum leaf diseases is a major concern in agriculture. Farmers need to track crop fields and recognize signs of disease at early as possible. The images processing is an aid to the identification and classification of leaf diseases. For the identification of leaf disease, there are three image features, i.e., texture, color, and shape. Textures are the most important feature out of the three. Image features value entered in CNN. The k-fold methods are used for train the model greater accuracy evaluation. The image features provide input for the identification of a disease and are reliable using the GLCM algorithm. CNN provides 92.53 % accuracy. Furthermore, the project's ability to predict suitable medicines for the detected diseases is a valuable addition to the agriculture industry, as it helps farmers make informed decisions about the most effective treatment for their crops. With further advancements in technology and the integration of precision agriculture techniques, the future of Solanum nigrum leaf disease detection and agriculture can become more efficient, sustainable, and productive.

REFERENCES

[1] Lu, J., Tan, L., & Jiang, H. (2021). Review on Convolutional Neural Network (CNN) Applied to Solanum nigrum Leaf Disease Classification. Agriculture, 11(8), 707. doi:10.3390/agriculture11080707

[2] Monigari, Vaishnavi. (2021). Solanum nigrumLeaf Disease Prediction. International Journal for Research in Applied Science and Engineering Technology. 9. 1295-1305. 10.22214/ijraset.2021.36582.

[3] Sujatha, R., Chatterjee, J. M., Jhanjhi, N., & Brohi, S. N. (2021). Performance of deeplearning vsmachine learning in Solanum nigrumleaf disease detection. Microprocessors and Microsystems, 80, 103615. doi:10.1016/j.micpro.2020.103615

[4] Jasim, M. A., & AL-Tuwaijari, J. M. (2020). Solanum nigrumLeaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques. 2020 International Conference on Computer Science and Software Engineering (CSASE). doi:10.1109/csase48920.2020.9142097

[5] G Ramkumar, Amirthalakshmi T.M, R. Thandaiah Prabu, A. Sabarivani, "An Effectual Solanum nigrumLeaf Disease Detection using Deep Learning Network with IoT Strategies", Annals of RSCB, pp. 8876–8885, Apr. 2021.

[6] Sharma, P., Hans, P., & Gupta, S. C. (2020). Classification Of Solanum nigrumLeaf Diseases Using Machine Learning And Image Preprocessing Techniques. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence). doi:10.1109/confluence47617.2020.9057889

[7] Ajra, H., Nahar, M. K., Sarkar, L., & Islam, M. S. (2020). Disease Detection of Solanum nigrumLeaf using Image Processing and CNN with Preventive Measures. 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE). doi:10.1109/etcce51779.2020.9350890

[8] Deepa, R. N and C. Shetty, "A Machine Learning Technique for Identification of Solanum nigrumDiseases in Leaves,"20216thInternationalConferenceonInventiveComputationTechnologies(ICICT),Coimbatore,India, 2021, pp.481-484,doi:10.1109/ICICT50816.2021.9358797.

[9] G. Bhogade, R. Mawal, A. Patil, and P. S. R. Desai, "Leaf Disease Identification using Image Processing," vol. 7, no. 5, pp. 11489–11491, 2017.

[10] P.Rokita,"Applicationof imageprocessing techniquesin computergraphicsalgorithms," Comput.Networks,vol. 29, no. 14, pp. 1705–1714, 1997, doi: 10.1016/s0169- 7552(97)00085-8.

[11] J. D. Pujari, R. Yakkundimath, and A. S. Byadgi, "Image processing Based Detection of Fungal Diseases in Solanum nigrum s," Procedia Comput. Sci., vol. 46, no. Icict 2014, pp. 1802–1808, 2015, doi: 10.1016/j.procs.2015.02.137.

[12] J. Baghel and P. Jain, "K-Means Segmentation Method for Automatic Leaf Disease Detection," vol. 6, no. 3, pp. 83–86, 2016.

[13] J.BaghelandP.Jain,"DiseaseDetectioninSoyaBeanusingK-MeansClusteringSegmentationTechnique,"Int. J.Comput.Appl.,vol.145,no.9,pp.15–18,2016,doi:10.5120/ijca2016910758.

[14] U. Mokhtar, M. A. S. Ali, A. E. Hassenian, and H. Hefny, "Tomato leaves diseases detection approach

based on SupportVectorMachines,"201511thInt.Comput.Eng.Conf. TodayInf.Soc. What'sNext?, ICENCO 2015,pp.246– 250, 2016, doi: 10.1109/ICENCO.2015.7416356.