IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

BANANA LEAF DISEASE DETECTION USING MACHING LEARNING

Ms.P.Sathya[#1], V.Velpriya[#2], M.Vidya Lakshmi[#3]

Assistant Professor (IT)-Student (IT)-Student (IT)

Francis Xavier Engineering College, Tirunelveli, India

ABSTRACT

This project addresses the urgent need for effective classification of banana leaf diseases, crucial for preserving global banana cultivation. Leveraging the potent random forest algorithm and advanced machine learning techniques, our primary objective is to develop an automated system capable of accurately identifying and categorizing diseases affecting banana leaves. Our methodology involves acquiring a comprehensive dataset comprising images of healthy banana leaves and those exhibiting symptoms of common diseases. Through meticulous application of machine learning algorithms, particularly random forest, we aim to train a robust model capable of discerning subtle disease patterns amidst varying environmental conditions. The potential impact of this project is profound, offering banana farmers a cost-efficient mechanism for early disease detection and targeted intervention strategies. By empowering farmers with swift disease identification and classification capabilities, our system strives to minimize crop losses and promote sustainable agricultural practices. Furthermore, beyond its immediate implications for banana cultivation, this research contributes valuable insights to the broader field of agricultural technology, showcasing the transformative potential of machine learning in addressing real-world agricultural challenges.

Keywords: Banana leaf diseases,Machine learning,Randomforest algorithm,Disease detection,Agricultural technology,Image processing ,Disease classification

I. INTRODUCTION

Bananas are a staple crop worldwide, serving as a vital source of nutrition and income for millions of people. However, the prevalence of various diseases, such as Sigatoka and Xanthomonas wilt, poses a significant threat to banana cultivation, leading to substantial economic losses and jeopardizing global food security. Timely and accurate detection of these diseases is essential for implementing effective interventions and safeguarding the livelihoods of banana farmers.Traditional methods of disease diagnosis in agriculture often rely on manual inspection, which can be time-consuming, subjective, and prone to errors. In recent years, the integration of machine learning techniques has emerged as a promising approach for efficient and automated disease detection. Machine learning algorithms, such as random forest, offer the potential to analyze large datasets of banana leaf images and identify subtle disease patterns that may not be apparent to the human eye.In this paper, we present a novel approach to banana leaf disease classification by leveraging the random forest algorithm instead of logistic regression. Our primary objective is to develop an automated system capable of accurately identifying and categorizing diseases affecting banana leaves. We aim to demonstrate the effectiveness of random forest in discerning subtle disease patterns amidst varying environmental conditions, thereby offering a robust and reliable solution for early disease detection in banana plants.By harnessing the power of advanced machine learning techniques, we seek to provide banana farmers with a cost-efficient mechanism for disease diagnosis and targeted intervention strategies. This research not only addresses the immediate need for effective disease management in banana cultivation but also contributes valuable insights to the broader field of agricultural technology, showcasing the transformative potential of machine learning in addressing real-world agricultural challenges.

II. **LITERATURE SURVEY**

"Banana Leaf Disease Detection Using Image Processing Techniques" by Ajay Deshmukh, Prashant A. Dhangar, and Vishal N. Thakare. This paper explores the detection of diseases in banana leaves using image processing techniques, which could serve as a foundation for leaf detection.

"Banana Leaf Classification Based on Color, Shape, and Texture Features"by Abdul Kadir, Umi Kalsom Yusof, and Siti Norul Huda Sheikh Abdullah. This study presents a method for banana leaf classification using features such as color, shape, and texture, which are essential for detection tasks.

"Detection of Diseases in Banana Plant Using Image Processing Techniques: A Review" by Madhura Patil and Prof. Sachin D. Ruikar. This review paper summarizes various image processing techniques applied to detect diseases in banana plants, which could provide insights into leaf detection methodologies.

"Banana Leaf Disease Detection and Classification Using Image Processing and Neural Networks" by K. Geetha, T. Santhi, and R. Uma Rani. The paper proposes a method for detecting and classifying diseases in banana leaves using image processing and neural networks, which could offer techniques applicable to leaf detection tasks.

"Real-Time Detection of Banana Leaves Using Deep Learning Techniques" by Lalit Garg, Niranjan Kumar, and Ashish Kumar. This study focuses on real-time detection of banana leaves using deep learning techniques, which could provide insights into modern approaches for leaf detection.

"Banana Leaf Diseases Detection Using Deep Learning" by Nisha and Deepa. The paper presents a deep learning-based approach for detecting diseases in banana leaves, which could be adapted for general leaf detection tasks.

"Banana Leaf Identification Using Artificial Neural Network" by S. Priyadharshini and R. Umarani. This research explores the identification of banana leaves using artificial neural networks, which could offer methodologies applicable to leaf detection and classification.

"Banana Leaf Disease Detection and Classification Using Image Processing Techniques" by S. K. Ravi and P. Jaganathan. The paper proposes a method for detecting and classifying diseases in banana leaves using image processing techniques, which could provide valuable insights for leaf detection tasks.

"Detection and Classification of Banana Leaf Diseases Using Support Vector Machine" by S. Padmavathi, K. S. Hareesh, and P. A. Vasanth. This study focuses on detecting and classifying diseases in banana leaves using support vector machines, which could offer methodologies relevant to leaf detection.

"Automated Detection and Classification of Banana Leaf Diseases Using K-Means Clustering Algorithm" by A. Sujitha and R. Gayathri. The paper proposes an automated method for detecting and classifying diseases in banana leaves using the K-means clustering algorithm, which could provide insights into clustering-based approaches for leaf detection tasks.

III.METHODOLOGY

Develop the Machine Learning Model:

To create a dataset of banana leaf images labeled with various diseases, images are collected and categorized accordingly. These images are then preprocessed, including resizing and normalization, to optimize them for training. Subsequently, a machine learning model is trained on the preprocessed images to classify different diseases present on banana leaves accurately.

Save the Trained Model:

Once you have trained your model and evaluated its performance, save the trained model to a file using serialization techniques like pickle or joblib. This will allow you to load the model later without needing to retrain it.

Create a Flask Web Application:

To establish a Flask application for a user-friendly web interface, routes are defined to manage requests and deliver HTML pages. HTML templates are developed to design the user interface, ensuring seamless interaction and visual appeal. The Flask app orchestrates the communication between the backend logic and the frontend interface, providing an intuitive platform for users to engage with.

Handle Image Upload:

Implement the HTML code to facilitate image uploads, enabling user interaction. Develop JavaScript code to manage form submission events and transmit uploaded images to the Flask backend seamlessly.

Backend Processing:

Set up a route in the Flask backend specifically designed to handle the uploaded image data. This route will manage tasks such as saving the image to a specified temporary directory. Afterwards, incorporate the pre-trained machine learning model into the backend to perform predictions on the uploaded image, allowing users to receive insightful analyses.

Display Results

Once the prediction is made, send the result back to the frontend. Update the HTML page to display the prediction result to the user.

EXISTING SYSTEM

Data loading and Preprocessing:

The code loads images of banana leaves from specified directories (`train_dir` and `test_dir`) and their corresponding labels. Images are resized to a uniform size (200x200 pixels) and converted to grayscale. Then, the images and labels are stored in numpy arrays (`X` and `Y`). The images are further reshaped and normalized for preprocessing.

Data Augmentation:

Data augmentation techniques are applied using the `ImageDataGenerator` from Keras to increase the diversity of the training dataset. Augmentation techniques include rotation, shifting, shearing, zooming, and flipping.

Principal Component Analysis (PCA):

PCA is performed on the training data (`X_train`) to reduce dimensionality while preserving 98% of the variance. The transformed PCA features are used for training the models.

Model Training:

Two machine learning models are trained:

Support Vector Machine (SVM) model (`svm_model`) using the `SVC` class from sklearn.

Random Forest model (`rf_model`) using the `RandomForestClassifier` class from sklearn with 100 estimators.

Ensemble Learning:

A Voting Classifier (`voting_clf`) is created to combine predictions from both SVM and Random Forest models using hard voting.

Cross-validation:

Cross-validation is performed on the ensemble model (`voting_clf`) using 5-fold cross-validation to evaluate its performance.

Model Evaluation:

The ensemble model's performance is evaluated using the confusion matrix and accuracy score on the test set (`X_test` and `y_test`). The confusion matrix is visualized using seaborn's heatmap.

Model Serialization:

The trained ensemble model (`voting_clf`) and PCA model (`pca`) are serialized and saved to disk using joblib, allowing for reuse and deployment in future applications.

PROPOSED SYSTEM

Our project aims to develop a robust and reliable tool for detecting diseases in banana plants, with a shift towards leveraging the Random Forest algorithm, renowned for its versatility and high performance in classification tasks. The primary objective remains unchanged: to enhance accuracy in identifying and categorizing banana diseases, with a particular emphasis on prevalent conditions such as banana sigatoka. Random Forest presents a powerful alternative, capable of handling complex datasets and capturing intricate patterns inherent in banana leaf images.Our system retains its core attributes of cost-effectiveness, ease of use, and resilience, empowering farmers to swiftly identify plant diseases by simply capturing a photo of the affected banana leaf. With Random Forest, our system becomes even more adaptable, capable of tackling diverse agricultural settings and challenging conditions, including complex backgrounds and fluctuations in image resolutions.The implementation of our revised system involves gathering a comprehensive dataset of banana leaf images, meticulously labeled with corresponding disease types, to facilitate the training and validation of the Random Forest model. Advanced feature extraction techniques are employed to capture relevant characteristics from these images, serving as input features for the model. Subsequently, the Random Forest model undergoes rigorous training and validation procedures using the labeled dataset to ensure its accuracy and reliability in disease detection.Anticipated outcomes of our revised system encompass improved disease detection capabilities, simplified operations for farmers, and proactive disease management leading to minimized crop losses. Additionally, our system holds promise in contributing to overall crop health and yield improvements through its advanced disease management approach.Looking ahead, future enhancements to our system may include further optimization of performance metrics, exploration of advanced feature extraction methods to augment disease detection capabilities, and integration with decision support systems to provide farmers with actionable insights and recommendations based on disease detection outcomes. By continually refining and expanding our system, we aspire to furnish farmers with a valuable tool for effective disease management in banana cultivation, ultimately fostering sustainable agricultural practices and enhancing livelihoods.

IV.ARCHITECTURE DIAGRAM EXPLANATION

Input Banana Leaf Image - This is the starting point for the process. This is the image of a banana leaf that the system will analyse to determine if it has a disease.

Image Preprocessing - This stage prepares the image for analysis. This might involve resizing the image, converting it to grayscale, or removing noise.

Feature Extraction - In this stage, the system identifies and extracts important features from the image. These features could be things like color, edges, shapes, or textures.

Random forest Algorithm - Random Forest enhances accuracy in identifying and categorizing banana diseases. This algorithm enables swift and effective disease management, empowering farmers with proactive solutions to minimize crop losses.

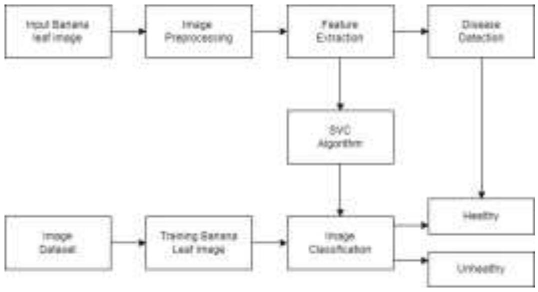
Training Banana Leaf Images - This refers to a collection of images of banana leaves that have already been labeled with their corresponding disease. For example, you might have a collection of images of leaves, some of which are healthy and some of which are diseased with different diseases.

Image Database - This refers to the collection of images that the system will use to train its classification model. The training images are a subset of the database images.

Disease Detection - This refers to the final stage where the system outputs a classification based on the analysis of the features that were extracted from the banana leaf image

Normal - This is one of the possible classifications that the system can output, indicating that the banana leaf is healthy.

Abnormal - This is another possible classification that the system can output, indicating that the banana leaf has a disease



V. EXPERIMENT RESULTS:

Fig 1 Data Preprocessing and Data Augmentation

The code employs the ImageDataGenerator from Keras to perform data augmentation on the training images. This includes operations like rotation, shifting, shearing, zooming, and flipping. Data augmentation is a common technique used to artificially increase the diversity of the training set, which can improve model generalization.

Fig 2 Model Ensemble and Cross-validation

Instead of using a single classifier, the code creates an ensemble model using a Voting Classifier. This ensemble consists of an SVM model (svm_model) and a Random Forest model (rf_model). The VotingClassifier aggregates the predictions of these two models and predicts the class with the most votes. This often leads to better performance than individual models. Cross-validation with `cross_val_score` evaluates model performance by iteratively splitting the training data into subsets, training the model on each subset, and assessing its generalization on unseen data, aiding in robustness assessment.

Fig 3 Confusion Matrix Visualization

After making predictions on the test set, the code generates a confusion matrix to visualize the performance of the ensemble model. A confusion matrix shows the counts of true positive, false positive, true negative, and false negative predictions.

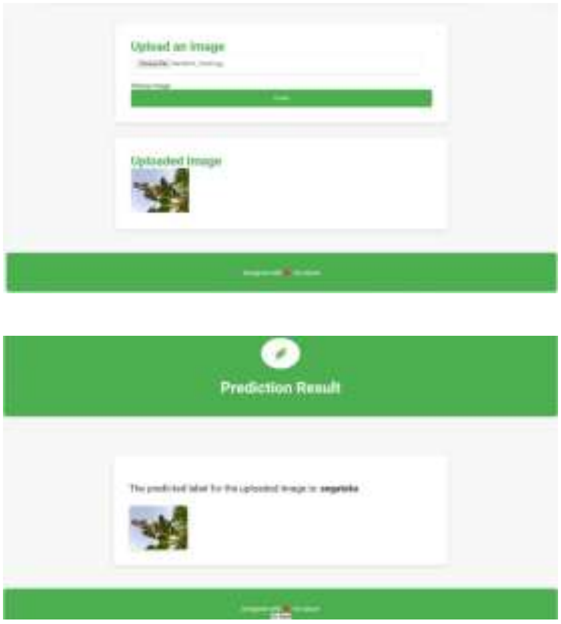


Fig 4 User Interface

When uploading an image, you will typically need to select the image file from your computer and then click on an upload button. Once the image has been uploaded it will predict the uploaded image is disease or healthy.

VI.CONCLUSION

In conclusion, the application of machine learning techniques for banana leaf disease prediction holds significant promise for revolutionizing agricultural practices and enhancing crop management strategies. By leveraging algorithms like random forest and Support Vector Classification (SVC), we can accurately classify abnormalities in banana leaves, enabling early detection and intervention to mitigate the impact of diseases on crop yield and quality. Through the integration of advanced technologies such as remote sensing, IoT, and mobile applications, farmers can access real-time information about plant health, facilitating timely decision-making and resource allocation. Moreover, the collaborative efforts of researchers, agricultural extension services, and industry stakeholders are essential for the development and deployment of robust disease prediction models and the implementation of sustainable agricultural practices. As we continue to refine and expand the scope of banana leaf disease prediction, we move closer to achieving the overarching goal of ensuring food security, preserving biodiversity, and promoting the long-term sustainability of banana cultivation worldwide.

VII. **FUTURE SCOPE**

The future of banana leaf disease prediction using machine learning is bright! We can expect the models to become more comprehensive, able to identify a wider range of diseases and pests that harm banana leaves. Advancements in deep learning architectures and incorporating expert knowledge from plant pathologists hold promise for even more accurate diagnoses. Imagine a world where farmers have mobile apps that give them real-time disease predictions directly in the field. Integration with drones for large-scale field scouting and disease mapping could revolutionize how farmers monitor crop health and take targeted action to prevent outbreaks. These exciting possibilities highlight the immense potential of machine learning to play a key role in ensuring sustainable banana production and food security.

REFERENCES

[Al11] Al-Hiary, H.; Bani-Ahmad, S.; Reyalat, M.; Braik, M.; ALRahamneh, Z.: Fast and ac- curate detection and classification of plant diseases. International Journal of Computer Applications, 17(1):0975–8887, 2011.

[ARK10] Arel, Itamar; Rose, Derek C.; Karnowski, Thomas P.: Deep Machine Learning - A New Frontier in Artificial Intelligence Research [Research Frontier]. IEEE Comp. Int. Mag., 5(4):13–18, 2010.

[Ba16] Barbedo, A.: A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems Engineering, 144:52–60, 2016.

[BKS16] Barbedo, Jayme Garcia Arnal; Koenigkan, Luciano Vieira; Santos, Thiago Teixeira: Iden- tifying multiple plant diseases using digital image processing. Biosystems Engineering, 147:104–116, 2016.

[Cl15] Clement, Alain; Verfaille, T; Lormel, C; Jaloux, B: A new colour vision system to quantify automatically foliar discolouration caused by insect pests feeding on leaf cells. Biosystems Engineering, 133:128–140, 2015.

[Cu10] Cui, Di; Zhang, Qin; Li, Minzan; Hartman, Glen L.; Zhao, Youfu: Image processing meth- ods for quantitatively detecting soybean rust from multispectral images. Biosystems En- gineering, 107(3):186–193, 2010.

[De14] Deng, Li: A tutorial survey of architectures, algorithms, and applications for deep learn- ing. APSIPA Transactions on Signal and Information Processing, 3, 2014.

[HS15] Hughes, David P.; Salathe´, Marcel: An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. CoRR, abs/1511.08060, 2015.

[Le89] LeCun, Yann; Boser, Bernhard; Denker, John S; Henderson, Donnie; Howard, Richard E; Hubbard, Wayne; Jackel, Lawrence D: Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4):541–551, 1989.

[MBP16] Mohan, K. Jagan; Balasubramanian, M.; Palanivel, S.: Detection and Recognition of Dis- eases from Paddy Plant Leaf Images. International Journal of Computer Applications, 144(12):34–41, 2016.

[MHS16] Mohanty, Sharada Prasanna; Hughes, David; Salathe, Marcel: Using Deep Learning for Image-Based Plant Disease Detection. 2016.

[Mo15] Mobility, Ericsson: , Interim Update: Ericsson Mobility Report, 2015.

[Ob14] Oberti, Roberto; Marchi, Massimo; Tirelli, Paolo; Calcante, Aldo; Iriti, Marcello; Borgh- ese, Alberto N.: Automatic detection of powdery mildew on grapevine leaves by image analysis: Optimal view-angle range to increase the sensitivity. Computers and Electronics in Agriculture, 104:1–8, 2014.

[Oe06] Oerke, E-C: Crop losses to pests. The Journal of Agricultural Science, 144(01):31–43, 2006.

[PC13] Patil, Sagar; Chandavale, Anjali: A Survey on Methods of Plant Disease Detection. In- ternational Journal of Science and Research (IJSR), 6(14), 2013.

[Sa13] Sannakki, Sanjeev S; Rajpurohit, Vijay S; Nargund, V B; Kulkarni, Pallavi: Diagnosis and Classification of Grape Leaf Diseases using Neural Networks. In: Fourth International Conference on Communications and Networking Technologies. 2013.

[Sc16] Schor, Noa; Bechar, Avital; Ignat, Timea; Dombrovsky, Aviv; Elad, Yigal; Berman, Si- gal: Robotic Disease Detection in Greenhouses: Combined Detection of Powdery Mildew and Tomato Spotted Wilt Virus. IEEE ROBOTICS AND AUTOMATION LETTERS, 1(1):354–360, 2016