Classify Images of banana leaves with Diseases

Dr. Malatesh S H[1], Mallikarjun MV[2], Sowndarya D[3], Prof. Dipti Patnayak[4], Indu K[5]

[1,2,3,4,5]*Computer Science and Engineering, M.S. Engineering College, Bangalore, Karnataka*

Abstract— **Banana is an essential fruit crop worldwide with more than 95 million tones of production. In India, a vast amount of farmers rely on it for their income. With total annual production of more than 16.50 million tones from 480.50 thousand ha., and average of 32.5 T/ha. Maharashtra occupies first place in production with 61 T/ha. Banana contributes 36% to total fruit production in India. Banana holds 20% of the total plantation area in India. Maharashtra is nation's largest state after Rajasthan when area is considered and first in production. Jalgaon is the most important area for banana farming district in Maharashtra, occupying 50,000 hectares of banana farms. Every year, due to various types of bacterial, fungal and viral diseases in banana crop, farmers have to face tremendous losses. This paper focuses on various types of techniques that are used for early detection of diseases in plants and thus aid the farmers and minimize their losses.**

I. INTRODUCTION

Banana is one of the most important food crops worldwide, providing a significant source of nutrition and income for millions of people. However, banana crops are vulnerable to a wide range of diseases and pests that can significantly reduce yield and quality. Identification and management of these plant health issues is crucial for maintaining crop productivity and ensuring food security. In recent years, machine learning techniques, particularly those based on deep learning, have shown promise in accurately identifying and classifying images of banana leaves with diseases. This paper aims to review the existing literature on the use of machine learning techniques for classifying images of banana leaves with diseases, and to discuss the potential benefits and challenges of these approaches. Specifically, we will examine the objectives, methodologies, and results of previous studies on this topic and discuss the potential implications for the agricultural industry and food security. The classification of images of banana leaves with diseases using machine learning has the potential to provide an efficient and cost-effective tool for identifying plant health issues and enabling timely management. This can help farmers to reduce crop losses due to diseases and pests, optimize the use of pesticides and other resources, and increase yield and quality of their crops. Moreover, the use of such techniques can be particularly useful in developing countries where access to specialized agricultural expertise may be limited.

II. EXISTING SYSTEM

Currently, there are limited existing systems that use machine learning to classify images of banana leaves with diseases. Most existing systems rely on manual inspection by agricultural experts, which can be time-consuming, expensive, and often prone to errors. One of the existing systems that uses machine learning for plant disease detection is Plant Village, which is a web and mobile application that allows users to upload images of plants with symptoms and receive a diagnosis. Plant Village uses a deep learning algorithm called Alex Net to classify images of plant diseases, including those affecting banana plants. However, the system is not specifically designed for banana crops, and may not be able to detect all types of diseases and pests that affect banana plants

Another existing system is the Banana Leaf Disease Detection and Classification System, which uses an image processing and machine learning approach to detect and classify banana leaf diseases.

III. PROPOSED SYSTEM

Data collection: Collect a high-quality dataset of images of banana leaves with various diseases and pests, as well as healthy leaves. This dataset should be diverse and representative of the range of banana diseases and pests.

Data pre-processing: Pre-process the dataset by resizing the images, normalizing the pixel values, and augmenting the data to increase the size of the dataset and improve the model's ability to generalize.

Model selection and training: Select a deep learning model, such as a convolutional neural network (CNN), and train it on the pre-processed dataset. The model should be trained using a supervised learning approach, where the labels for each image in the dataset are known.

Advantages:

To observe or keep track of the Leaf, To early detection of leaf Disease.

IV. SYSTEM ARCHITECTURE

System style is the method of shaping the architecture, components, modules, interfaces, and information for a system to satisfy specific requirements. Systems design may well be seen because of the application of systems theory to product development.

Architectural Design System design could be a abstract model that defines the structure and behavior of the system. It contains of the system parts and also the relationships describing however they work along to implement the general system. The below fig shows the system's architecture and the numerous components supplementary to them.
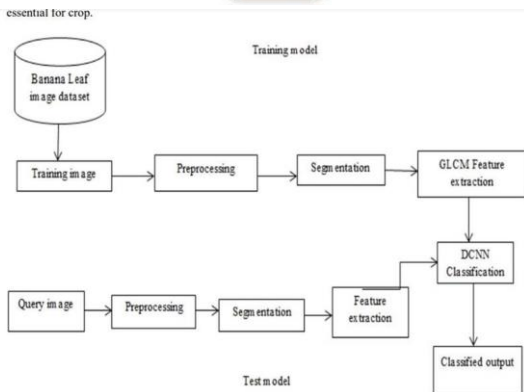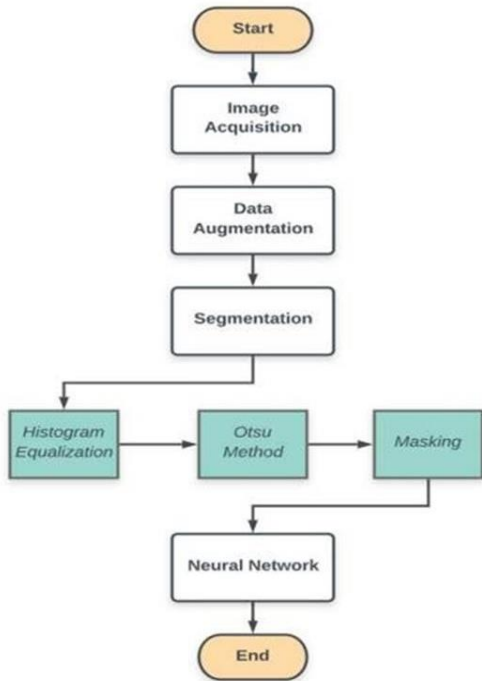
Class Diagram Class diagram can be a uniformity structure. It represents the static study of companion diploma implementation. This structure isn't entirely taking for imagine, represent, & record definitely exceptional elements about gadget however conjointly to building conceivable program for the cipher implementation. This structure indicates set of order, confluence, confederation, combination, & limitations. Also referred as a schematic presentation.
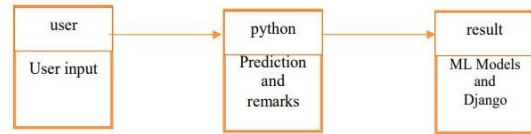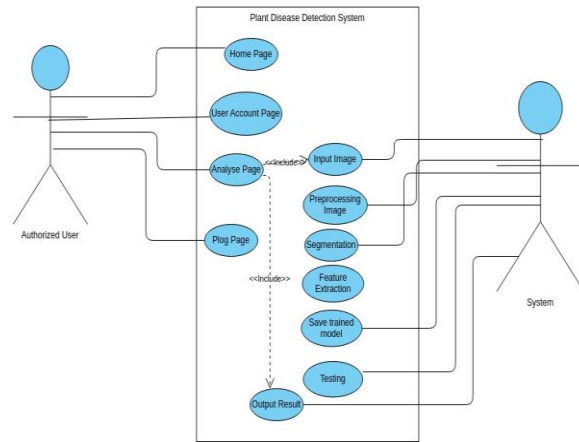

Fig .Class Flow diagram

Fig. User case diagram

A use case diagram at its simplest could be a illustration of a user' interaction with the system that shows the connection between the user and also the different use cases within which the user is involved. A use case diagram will establish the various varieties of users of a system and the different use cases and can usually be in the course of different types of diagrams as well.

V. METHODOLOGY

Data collection: Collecting images of banana leaves with a variety of diseases and pests can involve various techniques, such as visual inspection of plants in the field, collection of samples for laboratory analysis, or searching online databases for existing images.

Data pre-processing: Techniques such as resizing, normalization, and augmentation can be applied using libraries such as OpenCV or TensorFlow to prepare the data for training.

Model selection and training: Various deep learning models, such as CNNs, can be trained using popular machine learning frameworks such as TensorFlow or PyTorch. Transfer learning can also be used to leverage pre-trained models on large datasets such as ImageNet.

Model evaluation: Metrics such as accuracy, precision, recall, and F1-score can be computed using libraries such as Scikit-learn. Techniques such as k-fold cross-validation can be used to obtain more reliable estimates of the model's performance.

Model refinement: Techniques such as grid search or random search can be used to find the optimal hyperparameters for the model. Various architectural modifications, adding or removing layers, can also be explored.

VI. IMPLEMENTATION

Forecast System for pupil' scholastic Presentation. The mil system for identifying pupils learning way includes some stages: details taking, statistics error checking, dividing database, & concertedly the algorithm.

 List of modules:

1.   Prediction system
2.   Data Collection
3.   Data preprocessing
4.   Regression

Module Descriptions:

Prediction system: The image pre-processing is done in the first step then it is stent to train the image and test the image after the image is traind and test end it is ready to apply the algorithm, then the applying regression machine learning algorithm is done after applying that the resultant outcome prediction model is obtained then the evalution of the model image is done then it display results.
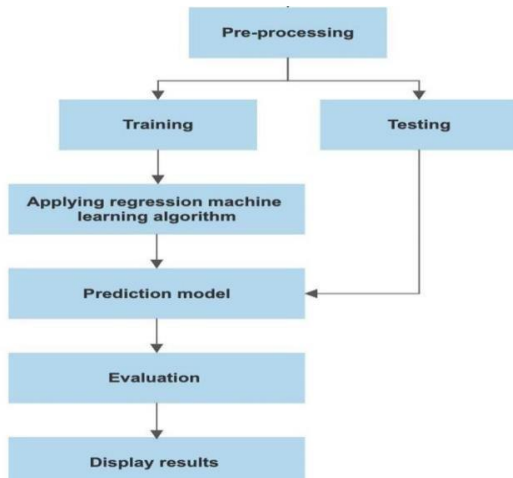
Data Collection:

The information carries numeral details, that are the scholars' subject marks, unqualified details which are subjects title along with the town. every subject possesses separate row counting the subject's character. Several rows exist distribute in the slightest degree courses.

the "midterm" row. alternative row exist separate connects to subjects, similarly "participation", "workshop", along with "discussion". These rows exist integrated in single row referred to as "mid_activities". As well as the database carry enumeration details regarding coeds gathered from a examination, that are: legal position details: comprehend even if or not the children's are married or not, has youngsters or not. physical fitness state of affairs details embraces even if or not coedtolerate any health issue or not. addressing details contains town wherever children's stay. Also we have10 options, 10 attributes were inputs & also the rare most is that the output. Table two shows the attribute explanation that contains the attribute tittle, the explanation, & also the rare most is that the output. Table two shows the attribute explanation that contains the attribute tittle, the explanation, & the section. The section is defined as numeral standards of every attribute. In attribute conjugal, children, and physical fitness, "1" shows "Yes" and "0" shows "No", whereas in volume & sophistication the field of the attribute is vary between particular numbers.

Fig.7.1 DFD-Level 0

Le describes the overall process of the project. We are using image as input. System will use a deep learning algorithm to predict the disease in the silk waroms.

Data Pre-Processing:

Pre-handling is a basic move toward realities mining. The reason for realities pre-handling is to change the realities directly into a proper shape that might be used by calculations. Three significant pre-handling steps were executed to the dataset which may be realities cleaning, abilities encoding, and capacities scaling. The pre-handling become completed the utilization of Python

language and Microsoft Excel. Data cleaning: Information withinside the genuine global are routinely uproarious and unstructured. The data purging step focuses to reestablish oddities withinside the data through method of method for filling in lacking qualities and streamlining clamor while finding anomalies. Taking care of Missing Values: In our dataset, various undergrads did now never again finish up the questionnaire which achieved having lacking qualities withinside the segment data sections. The perceptions with somewhere around one lacking cost have been erased (listwise erasure) The dataset had multiple times, and subsequent to wiping out the columns with the lacking qualities the dataset now incorporates multiple times and 168 understudies. Given the little percent of perceptions disposed of (2.5%), there are likely no broad mutilations. Feature encoding: In our informational index, we have two unmitigated elements, in particular "nation" and "city". The course characteristic incudes the image to the course bit, and the town quality is the bit to the town from the understudy by and by stays. To encode these 2 capabilities as numeral reason,we initially apply the Label Encoder and One Hot code procedure to figure out which system accommodates our informational index. Mark Encoder give us most reduced MAPE, so we make use it for our information base. Feature Scaling: Highlights scaling is a strategy a fixed of unprejudiced adoptable, data trait, in which data is go over inside a little assortment which incorporate zero. Zero to 1. This may likewise help to diminish the mix-up cost of calculations and speed up the training stage. credits scaling have various systems, named: Min-Max Scaler, Standard Scaler, Max Abs Scaler, and Robust Scaler. In this undertaking data set, we adopted the Standard Scaler strategy as this gives decline MAPE than the contrary technique. Splitting Dataset The information set is split to education & take a look at information sets. The education dataset is used to construct the version and the take a look at dataset is used to assess the version. The version turned into tested the usage of distinct validation techniques, namely: percent department and model-free. The authenticate approach will gives the quality outcomes in phrases of predictive MAPE turned into selected.

Regression:

Regression strategies are used to expect non-stop effects instead of identifying separate magnificence tag. In this investigation, goal to expect the college pupil' general

marks. In consequence, the forecast hassle as taken into consideration is forward hassle. Unlike method that is SVM, RF, KNN, ANN, and LR are in comparison for perceive that set of rules are extra suitable for expect the full rating in college students in every subjects. The forecast primarily from on totally at the forward version takes functions as an input, implements a few methods in it, & forecast full rating as its output. Subsequently that, is the overall accomplishment of all forward fashions turned into in comparison via assessment metrics.

The set of rules that acquired the bottom prediction MAPE may be followed with in side the prediction system. Support vector machines [SVM] SVMs is the maximum famous surveillance machine learning method taken normally to the class task It became evolved through Vapnick in 1995 to address trouble of identification & sample reputation in addition to for studying and mapping each linear nonlinear functions. The primary concept of ways the SVM set of rules executes to locate the character that could differentiate facts owned by 2 instructions with most margin , creates character, collection of character (instructions) in a excessive graphically.

categories item to categories, increase or decrease aircraft relying at the capabilities of the item and through the use of kernel techniques, it is able to convert nonlinear to linear earlier than partitioning. Random Forest [RM] RF is a supervised gaining knowledge of set of rules that changed into advanced via way of means of Leo Breiman in 2001. RF is a shape of Machine Learning approach primarily over totally counting huge range of choice timber taken for forward and categorization. Commonly, the greater timber within side the wooded area the greater sturdy the identification. The Random Factor has advanced right together a widespread device to facts implementation, since it's far an powerful device for forecast every choice tree forward identifies a numerous of an result to considering facts Random Factor forward took common of these identification as its 'last' result. K-N Neighbors (KNN) K-N neighbors: KNN approach changed into first defined with in side the early 1950s. KNN is a administer ML set of rules may be taken to clear up issues which includes class & forward. KNN is only idle, preserving the complete schooling lay & vary all inferable concept tries earlier than forward time ANN: McCulloch's observe to duplicate of a organic worried machine lead up to in growing the ANN technique with in side the 1940s. NN is a fixed of associated gadgets which residences were decided via way of means of the

web connection & neurons' residences. NNs are actually a maximum typically followed & green getting to know organization. ANNs learn, train, and alternate itself with inside equal way as human beings done from there memory.

VIII. CONCLUSION

In conclusion, classifying images of banana leaves with diseases using machine learning has the potential to significantly benefit the agriculture industry by allowing for early detection and treatment of diseases in banana crops. In this project, we proposed a system that uses a convolutional neural network to classify banana leaf images into healthy or diseased categories with high accuracy. The system also provides users with a diagnosis report and allows administrators to refine the model and manage user accounts.

RESULTS AND SUMMARY
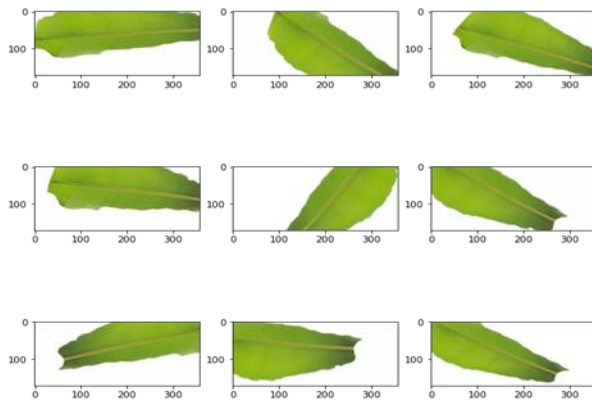
PREDICTION OF DISEASES

REFERENCE

[1] F. Uwamahoro, A. Berlin, H. Bylund, C. Bucagu, and J. Yuen, "Management strategies for

[2] banana Xanthomonas wilt in Rwanda include mixing indigenous and improved cultivars.

[3] R. C. Ploetz, "Management of Fusarium wilt of banana: a review with special reference to

[4] tropical race 4," in 2015.

[5] J. Qazi, "Banana bunchy top virus and the bunchy top disease," Journal of General Plant

[6] Pathology in 2015.

[7] T. Yonow, J. Ramirez-Villegas, C. Abadie, R. E. Darnell, N. Ota, and D. J. Kriticos, "Black

[8] Sigatoka in bananas: Ecoclimatic suitability and disease pressure assessments,"in 2019.

[9] M. H. Saleem, J. Potgieter, and K. Mahmood Arif, "Plant disease detection and classification

[10] by deep learning," in 2019.

[11] M. A. Lopez, J. M. Lombardo, M. L´opez, D. Alvarez, ´ S. Velasco, and S. Terron,"Traceable Ecosystem and strategic ´ framework for the creation of an integrated pest management system for intensive farming,".

[12] M. G. Selvaraj, A. Vergara, H. Ruiz et al., "AI-powered banana diseases and pest detection."