# Fruit Disease Detection and Pesticide Recommendation

1st Samvedya Jedeshmukh
*dept. Computer Engineering (of Aff.)*
*PCCOE (of Aff.)*
Pune, India

2nd Sahil Kachole
*dept. Computer Engineering (of Aff.)*
*PCCOE (of Aff.)*
Pune, India

3rd Nishant Parakh
*dept. Computer Engineering (of Aff.)*
*PCCOE (of Aff.)*
Pune, India

4th Umesh Chaudhari
*dept. Computer Engineering (of Aff.)*
*PCCOE (of Aff.)*
Pune, India

5th Prof. Madhura Kalbhor
*dept. Computer Engineering (of Aff.)*
*PCCOE (of Aff.)*
Pune, India

***Abstract*—Fruit crops are a crucial component of agriculture worldwide because they offer important nutrients and financial advantages. However, a number of diseases can seriously harm fruit orchards, resulting in a decrease in output and quality. Manual examination and treatment is a time-consuming, labor-intensive, and sometimes unsuccessful procedure. Therefore, the creation of an automated system for fruit disease diagnosis and pesticide recommendation has the potential to completely alter how we cultivate fruits. This method reliably identifies and categorises illnesses impacting fruit crops using cutting-edge technology including machine learning, image processing, and data analysis. In addition to lowering the use of toxic pesticides and encouraging environmentally friendly and sustainable agricultural methods, it may provide the best pesticide treatment depending on the disease's symptoms. which can provide farmers access to real-time monitoring of crop health and environmental factors. Among the many advantages of a fruit disease detection system with pesticide recommendation are increased food safety, increased productivity, and less environmental impact. This technology has the potential to transform how we cultivate fruits and help to the conservation of biodiversity and the health of ecosystems. Therefore, to improve the system's precision and efficacy and make it more available to farmers throughout the world, research and development in this field are crucial.**

## I. INTRODUCTION

Fruit crops are an important part of the world's agriculture business since they offer valuable nutrients and financial advantages. However, a number of illnesses can negatively affect fruit harvests, resulting in a decrease in productivity and quality. Traditional hand examination and treatment techniques can take a long time, require a lot of work, and sometimes don't solve the issue. Consequently, a better method of identifying fruit illnesses and recommending suitable therapies is required. The creation of an automated system for detecting fruit diseases and pesticide recommendations has the potential to completely alter how we cultivate fruits. This method reliably identifies and categorises illnesses impacting fruit crops using cutting-edge technology including machine learning, image processing, and data analysis. The system may recommend the right pesticide treatment by analysing the disease's symptoms, therefore lowering the usage of toxic pesticides and fostering environmentally friendly and sustainable agricultural methods. A fruit disease detection system with pesticide recommendation has several advantages. By guaranteeing that the fruits customers eat are free of toxic substances and hazardous illnesses, it can improve food safety. By enabling farmers to swiftly identify and treat concerns before they become serious, this approach can also help farmers save time and resources. Additionally, by using less toxic pesticides, this method can help maintain ecosystem health and biodiversity.it give farmers access to real-time monitoring of crop health and environmental conditions, enhancing the effectiveness and precision of pesticide recommendation and disease detection. In this sense, a system for detecting fruit diseases and recommending pesticides has the potential to completely alter how we cultivate fruits and contribute to the growth of crops that are both healthy and of the highest calibre.

Farmers can gain from the deployment of a fruit disease monitoring system with pesticide recommendations in a number of ways. By swiftly recognising and resolving problems before they become serious, it can help them save time and resources. The device can assist farmers in using less toxic pesticides, which will have a positive influence on the environment and encourage eco-friendly and sustainable farming methods. Additionally, this method can increase food safety and assist farmers financially by assuring the production of healthy and high-quality crops.

## II. Literature Survey

The author provided a model in reference [1] to find how much percent the fruit is impacted and recognize the fruit in the supplied image. To get better results in the classification and identification of fruit diseases Inception v3 model and Transfer Learning are used. The outcomes of this experiment are displayed in two different ways. 1. Classification of fruits 2. The percentage of diseases found. through the sampling of fruits including cherries, apples, and bananas. Fruit has been divided into four categories, A, B, C, and D, for the purpose of measuring the percentage of illness identification in this model. The grade of A means that the supplied image is usually good for eating, the grade of B shows that a small portion of the fruit is ruined, the grade of C indicates that half of the fruit is wrecked, and the grade of D indicates that the fruit is completely spoilt and unfit for consumption. The percentages for grades A through D range from 0 to 100%, 25 to 50% for grades B and C, 50 to 75 percent for grades C, and 100 percent for grades D.

The healthy and sick Citrus fruits and leaves may be distinguished using the suggested CNN-based leaf disease diagnostic model in article [2]. In this work, we attempted to diagnose illnesses from photos of citrus fruit and leaves using the CNN model. The following are the modules in our suggested model: Data collection, data preparation, and CNN model application are the first three steps. The proposed CNN model made use of two convolutional layers. While the second convolutional layer gathers high-level properties, the first convolutional layer removes low-level information from the image. This results in the disease classification of citrus fruit and leaves into Melanose, Black spot, canker, and scab classes. We investigated several machine and deep learning models on datasets of plant diseases and presented our results. The suggested CNN outperformed other classifiers in terms of accuracy, scoring 95.65% for citrus fruit/leaf disease classification experiments.

Convolutional Neural Networks (CNNs) and the Lab colour space model have been applied for the segmentation and classification of illnesses that damage the leaves of lemon plants, according to the authors of study [3]. The presented approach's major goal is to identify and categorise the condition. A healthy lemon leaf and three illnesses that affect lemon leaves were used to train and evaluate the Multiclass SVM algorithm and the CNN method. The model was able to recognise the presence of leaves and differentiate between healthy leaves and leaves with various illnesses that may be seen visually. Anthracnose, Citrus Canker, and Greasy Spot are the three illnesses. The findings of the experiment suggest that both methods can aid in the identification and categorization of illnesses on leaves. When the leaves were identified using the Multiclass SVM method, the accuracy rate was 83.6%, however the accuracy rate when using the CNN technique was 93.8%. As a result, the precision of the CNN algorithm's categorization of the leaves produced superior outcomes.

In the paper's suggested research effort, a machine learning-based intelligent system that can identify papaya disorders has been provided [4]. This study employed CNN with excellent accuracy (98.4%) and also used random forest, kmeans clustering, SVC, and SVC. The suggested methodology can assist the farmer in quickly identifying the issue and taking the necessary actions to decrease the infections. The current research will be expanded in the future to analyse a sizable dataset in an effort to identify the primary cause of papaya illnesses.

Avinash Kumar, Sobhangi Sarkar, and Chittaranjan Pradhan carried out a different study employing the technology of recommendation system in which they suggested the crops ideal for growing [5]. The approaches employed included the SVM classification algorithm, Decision Tree algorithm, Logistic Regression algorithm, and SVM classification algorithm. The rules derived from these models aid in the development of Recommendation Systems. An accuracy of 89.66% was attained in this model.

Zeel Doshi, Rashi Agrawal, Subhash Nadkarni, and Prof. Neepa Shah have also undertaken studies on agricultural diseases[6]. This paper introduces Agro Consultant, an intelligent system that will help Indian farmers decide which crop to grow based on the sowing season, the location of their farm, the characteristics of the soil, and other environmental factors like temperature and rainfall.

## III. Related Work

### A. Techniques Used

1. Image processing: Image processing is a field of computer science and engineering that deals with the analysis and manipulation of digital images. It involves developing algorithms and techniques that allow computers to extract useful information from digital images and to enhance their quality, clarity, and detail. The goal of image processing is to improve the visual appearance of images, extract relevant information, and transform images into more useful formats. Image processing is used to convert the RGB image into grayscale so that laser line is analysed so that obstacle is detected.

2. Computer vision: Computer vision is an incredibly important field of study within computer science and engineering that involves the development of algorithms and techniques that allow computers to interpret and understand visual data from the world around them. This includes the analysis and processing of images, videos, and other types of visual data, with the goal of enabling computers to recognize objects, detect patterns, and make decisions based on what they "see". Computer vision has a wide range of applications, from industrial automation and robotics to medical imaging and surveillance. Its continued development has the potential to revolutionize many aspects of our daily lives, making it a field that commands great respect and admiration.

3. Hyperparameter tuning: Hyperparameter tuning is a critical process in machine learning that involves searching for the best combination of hyperparameters to optimize the performance of a model. Hyperparameters are variables that are set before the training of a machine learning model

and control its learning process. They determine the model's architecture, such as the number of hidden layers and nodes, learning rate, regularization strength, and batch size, among others. The process of hyperparameter tuning involves selecting the best hyperparameter values that lead to the best performance of the model, such as the highest accuracy or the lowest error rate. This process requires experimenting with different combinations of hyperparameters and evaluating the model's performance on a validation set to determine which combination provides the best results. Hyperparameter tuning is a crucial step in machine learning as it can significantly impact the performance of a model. By selecting the optimal hyperparameters, a model can achieve better accuracy, lower error rates, and faster convergence.

### *B. First Experiment*

Methodology: In this study, we developed a fruit disease detection system using the ResNet50 pre-trained model in the TensorFlow/Keras framework. The methodology can be divided into the following steps:

Data Preprocessing:

The input images were preprocessed using the ImageDataGenerator class from TensorFlow/Keras. The pixel values of the images were scaled down by a factor of 1/252. Dataset Preparation:

The training and validation datasets were prepared using the ImageDataGenerator class, which loaded the images from the respective directories. The dataset was resized to a uniform size of 180x180 pixels to ensure consistency. The dataset was batched with a batch size of 23 for efficient training. Exploratory Data Analysis:

A subset of images from the training dataset was visualized to gain insights into the dataset and understand the fruit disease classes. Model Architecture:

We used the ResNet50 pre-trained model as the backbone of our fruit disease detection system. The pre-trained model was initialized with weights trained on the ImageNet dataset. We froze the weights of the pre-trained layers to prevent them from being updated during training. The output of the pre-trained model was flattened, followed by a fully connected layer with 512 units and ReLU activation. The final output layer consisted of 4 units with softmax activation, representing the different fruit disease classes. Model Training:

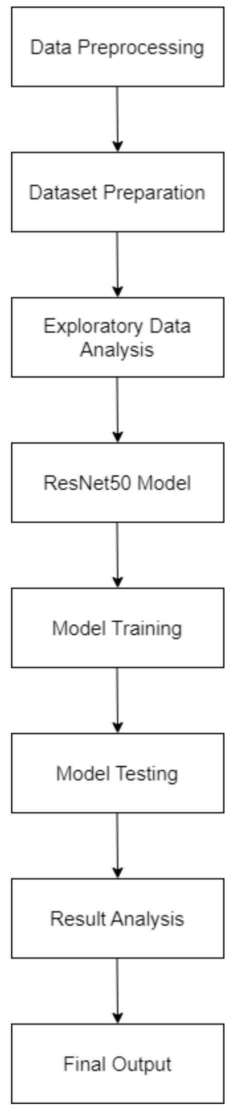
The model was compiled with the Adam optimizer and a learning rate of 0.001. The loss function used was sparse categorical cross-entropy, suitable for multi-class classification tasks. The model was trained for 11 epochs using the training dataset, with validation performed using the separate validation dataset. Evaluation and Performance Metrics:

The accuracy and loss curves were plotted to assess the model's performance during training. The accuracy and loss values were recorded for both the training and validation sets at each epoch. Model Testing and Performance Evaluation:

The trained model was evaluated on a separate test dataset obtained from the same fruit disease dataset. The test accuracy and loss were calculated to measure the generalization performance of the model. Results Analysis:

The obtained accuracy, loss, and other relevant metrics were analyzed to assess the effectiveness of the fruit disease detection system. The performance of the developed model was compared with other existing approaches and benchmarks, if applicable. The presented methodology provides a comprehensive framework for fruit disease detection using the ResNet50 pre-trained model. The next section will discuss the experimental results and provide an in-depth analysis of the model's performance.

Following is the proposed system's architecture :




The "Start Process" initiates the methodology for fruit disease detection. The "Data Preprocessing" step involves preprocessing the input images by scaling down their pixel values. The "Dataset Preparation" step prepares the training, validation, and test datasets using the ImageDataGenerator class. The "Exploratory Data Analysis" step visualizes a subset of images from the training dataset for insights and understanding. The "Model Architecture" step defines the architecture of the fruit disease detection model using the ResNet50 pre-trained model as the backbone. The ”Model Training” step involves compiling and training the model with the prepared datasets. The ”Evaluation and Performance Metrics” step evaluates the model’s performance by analyzing accuracy and loss curves during training. The ”Model Testing and Performance Evaluation” step tests the trained model on a separate test dataset and measures its accuracy and loss. The ”Results Analysis” step involves analyzing the obtained accuracy, loss, and other metrics to assess the effectiveness of the fruit disease detection system. The ”Final Output” represents the final outcome of the methodology, which includes the results, analysis, and conclusions of the fruit disease detection study. This flowchart provides a visual representation of the major steps involved in the methodology for fruit disease detection, outlining the sequence of actions from data preprocessing to the final output.

In our research, we explored transfer learning using three different pre-trained models, namely VGG-19, InceptionV3, and ResNet50. We experimented with different combinations of batch sizes and epochs to train the models and evaluated their performance based on the achieved accuracy. After careful analysis, we observed that the highest accuracy achieved was 80%. This result demonstrates the effectiveness of transfer learning in improving the accuracy of pre-trained models. Our findings suggest that transfer learning can be a valuable technique in various applications of computer vision.

### *C. Second Experiment*

The methodology employed in this research project involved the following steps to classify apple diseases using an ensemble learning approach, specifically the VotingClassifier:

Dataset Preparation:

The dataset consisted of a collection of apple disease images, categorized into separate train and test folders based on their respective disease types. All images were preprocessed by resizing them to a standardized dimension of 150x150 pixels to ensure consistency. Ensemble Classifier Construction:

Four base classifiers were selected for the ensemble classifier: DecisionTreeClassifier, RandomForestClassifier, KNeighborsClassifier, and SVC. Each base classifier contributed its own prediction to the ensemble classifier, and the final prediction was determined based on either a majority vote (hard voting) or a weighted average (soft voting) of the individual predictions. Training Phase:

The image arrays were flattened to create feature vectors, which were used as input for training the ensemble classifier. The flattened train images, along with their corresponding labels, were used to train the ensemble classifier. Hyperparameter Tuning:

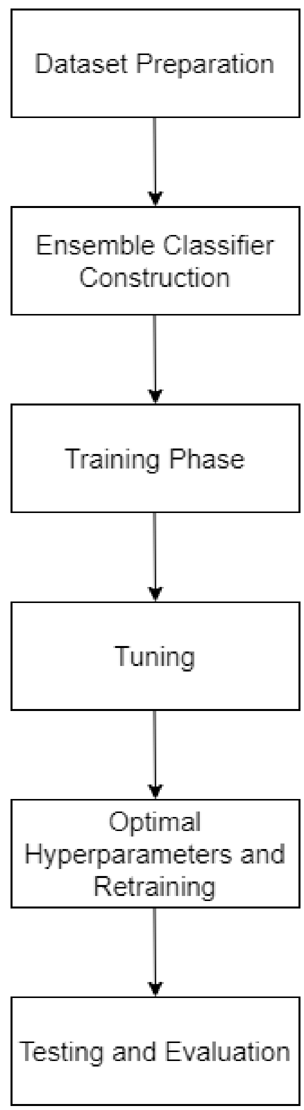
Hyperparameter tuning was performed to optimize the performance of the ensemble classifier. GridSearchCV, a cross-validated grid search, was employed to systematically explore different hyperparameter combinations. The hyperparameters considered for tuning were specific to each base classifier: For DecisionTreeClassifier: max_depth (options: None, 10, 20)
RandomForestClassifier: n_estimators (options: 50, 100, 200)
KNeighborsClassifier: n_neighbors (options: 3, 5, 7)
SVC: C (options: 0.1, 1, 10)
The performance of the ensemble classifier was evaluated for each hyperparameter configuration using 3-fold cross-validation. Optimal Hyperparameters and Retraining:

The best hyperparameters were identified based on the results of the hyperparameter tuning process. The ensemble classifier was retrained using the best hyperparameters to further optimize its performance. Testing and Evaluation:

The performance of the optimized ensemble classifier was assessed using the flattened test images. The accuracy metric was used to evaluate the classification effectiveness of the ensemble classifier.

Overall, the methodology encompassed the construction of the ensemble classifier, hyperparameter tuning, training and retraining, testing and evaluation, and comparative analysis to provide a comprehensive understanding of the ensemble learning approach applied to the classification of apple diseases.

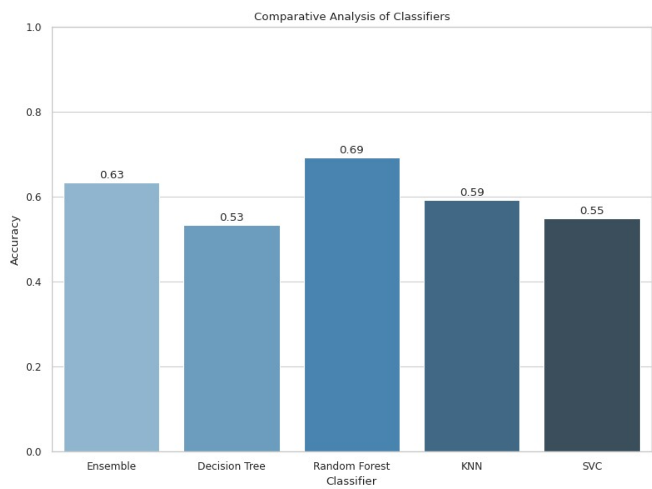
Following is the proposed system’s architecture :



### D. Comparative Study

The accuracy achieved by the ensemble classifier on the test set was compared with the performance of the individual base classifiers. The impact of different hyperparameters on the ensemble classifier's performance was analyzed and discussed. The VotingClassifier, as an ensemble approach, allowed for the combination of multiple classifiers to improve the accuracy and robustness of the classification model. By integrating the predictions of individual classifiers, the ensemble model captured a broader understanding of the apple disease classification problem.

Through hyperparameter tuning using GridSearchCV, the ensemble classifier was fine-tuned to identify the optimal combination of hyperparameters specific to each base classifier. This process ensured that the ensemble classifier was well-suited for the task of apple disease classification.




COMPARISON TABLE I

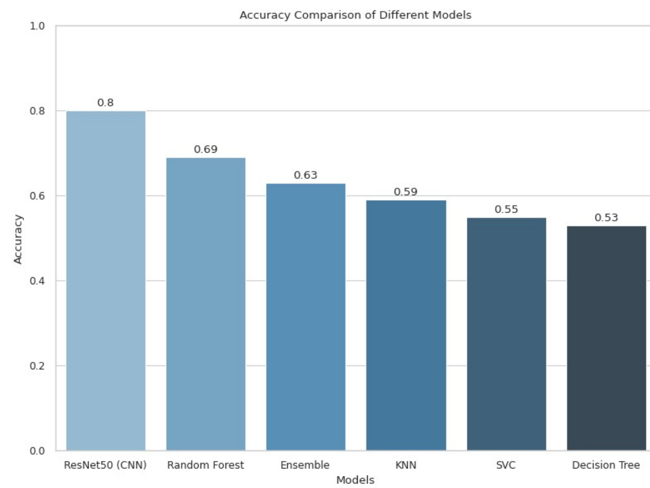
| SR. NO | ALGORITHM | ACCURACY |
|---|---|---|
| 1 | Ensemble | 63% |
| 2 | Decision Tree | 53% |
| 3 | Random Forest Classifier | 69% |
| 4 | KNN | 59% |
| 5 | SVC | 55% |

After comparing the accuracy of different classifiers, we gained valuable insights into their individual performances. Figure X presents the accuracy scores of the ensemble classifier, Decision Tree, Random Forest, KNN, and SVC models, reflecting their respective performances on the test set.

Based on the analysis and the comparison of the accuracy scores, it was determined that the Random Forest (RF) classifier outperformed the ensemble classifier and other individual classifiers. Therefore, the next steps involve further fine-tuning of the RF model to maximize its performance.

To achieve this, we utilized the GridSearchCV function from scikit-learn to perform a grid search over a predefined parameter grid. The grid search explored different combinations of hyperparameters for the RF model, including the number of estimators, maximum depth, minimum samples split, and minimum samples leaf. The grid search was conducted using 5-fold cross-validation to ensure robustness in the evaluation.

After performing hyperparameter tuning on the Random Forest (RF) model, the best set of hyperparameters was determined to be 'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 300. However, the accuracy achieved with these hyperparameters was only 69%.




COMPARISON TABLE II

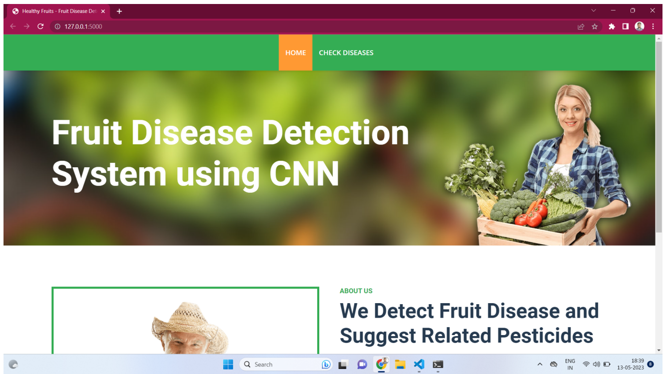
| SR. NO | ALGORITHM | ACCURACY |
|---|---|---|
| 1 | ResNet50 (CNN) | 80% |
| 2 | Random Forest | 69% |
| 3 | Ensemble | 63% |
| 4 | KNN | 59% |
| 5 | SVC | 55% |
| 6 | Decision Tree | 53% |

In comparison, the previous model that utilized the ResNet50 architecture yielded an accuracy of 80%. Therefore, based on the accuracy comparison, it is evident that the ResNet50 model outperforms the fine-tuned RF model.

As a result, the decision has been made to prefer the ResNet50 model over the RF model for the classification task at hand. The higher accuracy achieved by the ResNet50 model indicates its superior performance in capturing complex features and patterns within the image data.

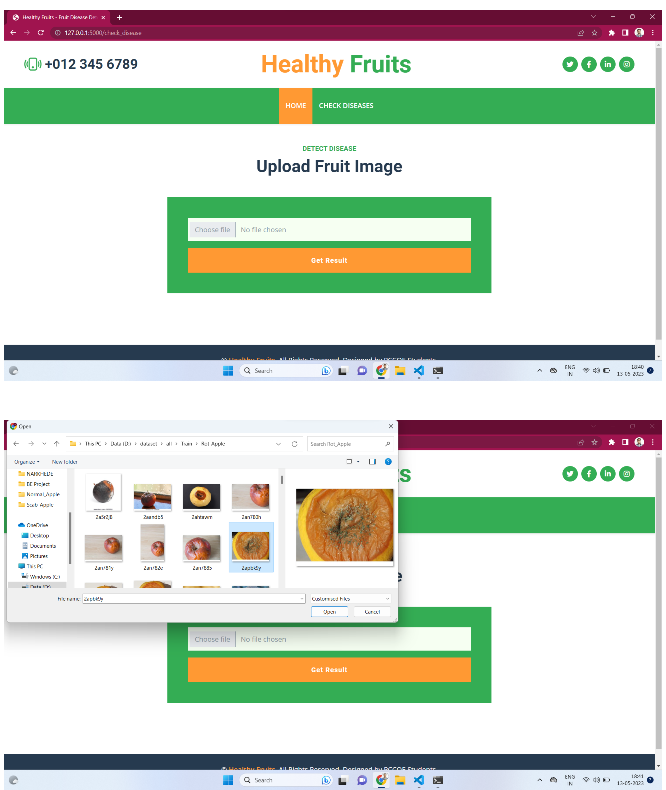
## IV. Results

1. Home Page :




Fig[1]

2. Interface for uploading fruit image:




Fig[2]

3. Detected Result:


Fig[3]

## V. Conclusion

In conclusion, a system for detecting fruit diseases and recommending pesticides can be very advantageous for the agricultural sector. This system can precisely identify different fruit illnesses and recommend the best pesticide treatment by utilising cutting-edge technology like machine learning, image processing, and data analysis. Farmers may save time and money by being able to rapidly recognise and deal with problems before they get out of hand. By advising the use of less hazardous pesticides, this approach can also encourage the adoption of sustainable agricultural methods that are kind to the environment. Overall, a system for detecting fruit diseases and recommending pesticides has the potential to completely change how we cultivate fruits and contribute to the growth of crops that are healthy and of good quality.

## VI. Future Scope

There are several areas of future work in fruit disease detection and pesticide recommendation systems, including:

Improving accuracy: There is a need to improve the accuracy of fruit disease detection systems and pesticide recommendation systems to reduce false positives and false negatives. This can be achieved by using more advanced machine learning algorithms and data sources, including satellite imagery, weather data, and soil data.

Integrating multiple data sources: There is a need to integrate multiple data sources, such as sensor data, remote sensing data, and weather data, to provide more comprehensive and accurate recommendations for pest management and disease control.

Developing customized solutions: There is a need to develop customized solutions for different crops, regions, and farming practices to provide tailored recommendations that take into account local conditions and farmer preferences.

Automating decision-making: There is a need to automate decision-making in fruit disease detection and pesticide recommendation systems to reduce the workload on farmers and improve the speed and accuracy of recommendations. This can be achieved by developing intelligent agents that can make decisions based on real-time data.

Addressing environmental concerns: There is a need to develop fruit disease detection and pesticide recommendation systems that address environmental concerns, such as reducing the use of pesticides, minimizing the impact on non-target species, and promoting sustainable farming practices. Overall, the future work in fruit disease detection and pesticide recommendation systems is focused on improving the accuracy, efficiency, and sustainability of pest management and disease control in the agriculture and food industry. Overall, the future work in fruit disease detection and pesticide recommendation systems is focused on improving the accuracy, efficiency, and sustainability of pest management and disease control in the agriculture and food industry.

## REFERENCES


[1] Nikhitha M, Roopa Sri S, Uma Maheswari B, "Fruit Recognition and Grade of Disease Detection using Inception V3 Model", 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). June 2019

[2] Asad Khattak, Muhammad Usama Asghar, Ulfat Batool, Muhammad Zubair Asghar ,Hayat Ullah, Mabrook Al-Rakhami , (Member, Ieee), And Abdu Gumaei, "Automatic Detection of Citrus Fruit and leaves Diseases Using Deep Neural Network Model", Institute of Electrical and Electronics Engineers(IEEE). July 2021

[3] Balambigai Subramanian, Jayashree.S,Kiruthika.S, Miruthula.S, "Lemon Leaf Disease Detection and Classification using SVM and CNN", International Journal of Recent Technology and Engineering (IJRTE), Volume-8 Issue-4. November 2019

[4] Md. Ashiqul Islam, Md. Shahriar Islam, Md. Sagar Hossen, Minhaz Uddin Emon, Maria Sultana Keya, Ahsan Habib, "Machine Learning based Image Classification of Papaya Disease Recognition", Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA-2020).

[5] Avinash Kumar, Sobhangi Sarkar and Chittaranjan Pradhan "Recommendation System for Crop Identification and Pest Control Technique in Agriculture, : International Conference on Communication and Signal Processing, April 4-6, 2019, India.

[6] Zeel Doshi, Rashi Agrawal, Subhash Nadkarni, Prof. Neepa Shah "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms", : 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)

[7] Asmaa Ghazi Alharbi, Muhammad Arif, "Detection and Classification of Apple Diseases using Convolutional Neural Networks", 2020 2nd International Conference on Computer and Information Sciences (ICCIS). November 2020.

[8] Wahyuni Eka Sari, Yulia Ery Kurniawati, Paulus Insap Santosa. "Papaya Disease Detection Using Fuzzy Naïve Bayes Classifier", 3rd International Seminar on Research of Information Technology and Intelligent Systems(ISRITI-2020).

[9] Prof. Shripad S.Veling ,Mr.Rohit ,S. Kalelkar ,Miss. Likhita V. Ajgaonkar3, Miss. Nivedita V. Mestry, Mr. Nilesh N.Gawade, "Mango Disease Detection by using Image Processing", International Journal for Research in Applied Science & Engineering Technology (IJRASET). April 2019

[10] Vani Ashok & D. S. Vinod, "A Novel Fusion of Deep Learning and Android Application for RealTime Mango Fruits Disease Detection", Advances in Intelligent Systems and Computing book series (AISC,volume 1171). August 2020

[11] Yousef Abbaspour-Gilandeh, Abdollah Aghabara, Mahdi Davari and Joe Mari Maja," Feasibility of Using Computer Vision and Artificial Intelligence Techniques in Detection of Some Apple Pests and Diseases", Multidisciplinary Digital Publishing Institute (MDPI) - Publisher of Open Access Journals. January 2022

[12] Ginne M. James and S. Sujatha," Categorising Apple Fruit Diseases Employing Hybrid Neural Clustering Classifier", Categorising Apple Fruit Diseases Employing Hybrid Neural Clustering Classifier. December 2020

[13] Ruhika Sharma and Sukhpreet Kaur," Convolution Neural Network based Several Orange Leaves Disease Detection and Identification Methods", International Conference on Smart Systems and Inventive Technology.2019

[14] Md. Tarek Habib, Anup Majumder, Rabindra Nath Nandi, Farruk Ahmed and Mohammad Shorif Uddin," A Comparative Study of Classifiers in the Context of Papaya Disease Recognition", Proceedings of International Joint Conference on Computational Intelligence.2020.

[15] K. Kurumatani, "Time Series Prediction of Agricultural Products Price Based on Time Alignment of Recurrent Neural Networks," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 2018, pp. 81-88, doi: 10.1109/ICMLA.2018.00020.

[16] Keke Zhang, Qiufeng Wu, Yiping Chen, Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN, Computers and Electronics in Agriculture, Volume 183, 2021, 106064, ISSN 0168-1699