Detecting Fruit Diseases using Deep Learning and Image Analysis

Tushar P Pareek
Department of Information Technology
SVKM's NMIMS, MPSTME
Shirpur, India
tusharpareek941@gmail.com

Deep Dobariya
Department of Information Technology
SVKM's NMIMS, MPSTME
Shirpur, India
deeppdobariya@gmail.com

Tanay Mahajan
Department of Information Technology
SVKM's NMIMS, MPSTME
Shirpur, India
tanaymahajan2002@gmail.com

Dhiraj Bhise
Asst. Professor
Department of Information Technology
SVKM's NMIMS, MPSTME
Shirpur, India
dhirajbhise11@gmail.com

***Abstract*—The implementation of several agriculture-related issues was made simple by advancements in computer vision technology. The detection of fruit diseases is one such issue. Using deep learning techniques, a lot of study has been done on many fruits, including apples, mangos, kiwis, passion fruit, and others. The most significant contributions made by this field in recent years are outlined in this review paper. In this review, we have performed a technical analysis of deep learning methods for predicting fruit illnesses. Along with the deep learning models utilized, the study also compares various picture acquisition, image pre-processing, and segmentation methods. The study discovered that the most accurate deep learning model can vary based on the system's computing capacity and the data used. In the article, directions for future research have also been covered.**

***Index Terms*—Fruits, CNN, Deep Learning, Machine Learning, Disease Detection**

I. Introduction

Fruit diseases represent a serious danger to crop health and agricultural productivity. Effective illness management depends on the early identification and accurate diagnosis of these conditions. In this study, we suggest an innovative method for fruit disease detection using Convolutional Neural Networks (CNN) and image processing methods. Our research focuses on using deep learning and computer vision to automatically identify fruit illnesses from digital photos. We start by compiling an extensive dataset of photos of numerous fruits suffering from various illnesses. To improve image quality and identify pertinent features, the dataset is pre-processed. The dataset is then used to develop and train a CNN architecture to discover discriminative patterns suggestive of diseases. We demonstrate the efficiency of our approach in precisely identifying and classifying fruit illnesses with high precision and recall through comprehensive trials and evaluations. Our findings demonstrate the potential of CNNs and image processing for fruit disease identification, showing notable improvements over current approaches. By offering [1] a reliable and effective method for automating the identification of fruit diseases, this research advances the field by facilitating early intervention and efficient crop management. The results of this study have implications for sustainable agriculture, lowering yield losses, and enhancing the general well-being and productivity of fruit crops. Future studies may examine the incorporation of this approach into intelligent farming technology for real-time disease monitoring and decision-making. Agriculture plays a critical part in harvesting any culture. In fact, agriculture is vital to human civilization. To achieve the objective of good output, field prevalence, and soil water management must be continually monitored. Product quality and quantity are greatly affected by fruit disease. Fruit disease is a major challenge in this area. This study examines diseases that limit fruit production. Image processing techniques are used to examine how fruit harvests are damaging. The proposed system's job is to identify flaws in the fruit images. The dataset can be gathered via smartphone cameras. For financial reasons [2], it is important to detect fruit illnesses as soon as possible. Deep Learning can identify and categorize diseases early on, which reduces disease transmission and boosts cure rates. Powdery mildew, rust, and black rot are prevalent signs of various diseases, which affect fruits such as apples and cherries, and others [3]. The manual investigation takes a long time, is labor dependent, and is prone to mistakes. Fruit color, shape, and texture such type of data can be extracted using artificial intelligence to help in the diagnosis of viruses. As a consequence, there are more errors committed when rating fruits for export. Researchers have introduced an image detection strategy to distinguish between healthy fruits and the one which is affected to improve accuracy and eliminate the faults in the human categorization process.

II. Objective

Technologies such as Convolutional Neural Networks and Image Processing have been the initial purpose of the study to

ISBN: 979-8-3503-0611-8/23/$31.00 ©2023 IEEE

create a reliable and effective system for identifying diseases in fruits. Fruit disease detection is essential for maintaining the health, production, and quality of fruits. For efficient disease management, preventing large-scale outbreaks so early disease detection is crucial. This research seeks to automate the disease detection process by utilizing CNN and Image Processing methods. The created system will examine digital fruit photographs and make use of deep learning algorithms to find visual signs and patterns linked to different diseases. The system will be taught to distinguish between fruit that is healthy and fruit that has diseases like powdery mildew, rust, black rot, and many more through training on a large dataset. The shortcomings of manual inspection methods, which are time-consuming, subjective, and heavily reliant on expert knowledge, give rise to the need for automated fruit disease diagnosis. By incorporating cutting-edge technologies, we can get beyond these restrictions and offer a more effective, reliable, and scalable solution. The findings of this study are expected to have a significant impact on farmers, agribusinesses, and the entire agricultural sector. Early and accurate disease identification allows for quicker responses, more focused care, and better disease management techniques. The overall goal of the project is to aid in the creation of an accurate, automated system for fruit disease diagnosis, and guarantee a supply of fruits of the highest caliber to meet the world's food demands.

III. Literature Review

Analysis of numerous research publications can reveal a common set of steps. Answering a few key questions while keeping the deep learning model, fruit, and processing aspects in mind was part of the strategy used to analyze various fruit diseases. In this article, the things which were discovered were [2]:

1) The fruit that has been used the most in the research.
2) Pre-processing models that were broadly utilized.
3) The segmentation method which was mostly utilized.
4) The deep learning model, which was used.
5) Data sets that were normally used by the researchers.

A. *Study of widely used datasets*

Most researchers prefer to use data from public records. People collect a wealth of information and make it available for the benefit of others; such datasets are called public repositories [2]. The most popular dataset contains 54,306 pictures of healthy and diseased plant leaves and fruits that were taken under regulated conditions and is part of the Plant Village Dataset. Since it is the largest dataset that is freely accessible, most studies have used it whole or in part. The apple, raspberry, cherry, grape, orange, peach, strawberry, and tomato are just a few of the fruits included in this dataset.

The New Plant Disease dataset, which includes a total of 15,200 photos categorised in 38 distinct categories and covers 14 different crops including tomato, apple, raspberry, strawberry, grape, orange, and various crops and vegetables, is another publicly accessible resource. Researchers also use this data in categories according to their requirements.

Researchers also gather data by taking pictures with digital cameras in controlled environments in labs, outside of open repositories. It has been found that data acquired in a controlled setting tends to perform badly in real-world applications but provides greater accuracy on the dataset that is accessible. The model performed well in the actual world despite the photographs being taken under non-uniform illumination intensities and cluttered field backgrounds, which led to a drop in accuracy.

Fruit 360 contains high-quality images containing fruits and vegetables. Included are several types of the fruits and vegetables listed below: Apples (varieties including Crimson Snow, Golden, Golden-Red, Granny Smith, Pink Lady, Red, and Red Delicious), Apricot, Avocado, Avocado that is ripe, Banana (varieties including yellow, red, and lady finger), Beetroot Red, Blueberry, Cactus fruit, Cantaloupe (two varieties), Carambula, Cauliflower, Cherry (varieties including Rainier), Cherry Wax (varieties including yellow, red, and Grape (varieties of blue, pink, and white), Grapefruit (varieties of pink and white), Guava, Hazelnut, Huckleberry, Kiwi, Kaki, Kohlrabi, Kumquats, and Kumquats are some of the fruits that are available. Lemon (regular, Meyer), Lime, Lychee, Mandarine, Mango (Green, Red), Mangostan, Maracuja, Melon Piel de Sapo, Mulberry, Nectarine (regular, Flat), Nut (Forest, Pecan), Onion (Red, White), Orange, Papaya, Passion fruit, Peach (different varieties), Pepino, Pear (various varieties, including Abate, Forelle, Kaiser,Monster,Red. Properties of data sets includes that there are a total of 90483 photos in this project. The total number of photos in the training set is 67692, each depicting a single edible item. The total number of photos in the test set is 22688, each depicting a single edible item. Size of the multi-fruits set: 103 photos (each image contains representations of more than one fruit or fruit class). The total number of classes is 131, all of which are fruits and vegetables. The dimensions of the image are 100x100 pixels.

Fig. 1. Images from datasets

B. *Diffrent Pre-processing Methods used on Datasets*

This technique involves improving and preparing the acquired images before they are sent into the detection system, image preprocessing is essential for the detection of fruit diseases. The quality and relevance of the images can be increased by the use of image preprocessing techniques, making them better suited for the precise detection of diseases [4]. Using image preprocessing to identify fruit diseases can be accomplished as follows:

- Noise Reduction: Removing noise from the photographs, which may be caused by camera sensors or ambient variables, improves their sharpness and clarity. To lessen noise and increase the signal-to-noise ratio, techniques like median filtering or Gaussian filtering can be used.
- Colour normalization: By uniformly distributing colours across photographs, you may make up for differences in camera settings and illumination. By ensuring constant colour representations through colour normalization techniques, the detection system is able to concentrate on disease-related features rather than variations in lighting or colours.
- Image enhancement: Increasing the contrast and brightness of the photos helps emphasize disease signs and expose minute details. It is possible to use methods like histogram equalization or adaptive contrast stretching to enhance image visibility and make illness signs easier to identify.
- Image Resizing and Cropping: Reducing computational overhead and focusing on the pertinent regions of interest for disease diagnosis can be accomplished by resizing photos to a standard resolution and cropping extraneous background areas.
- Segmentation: To separate the regions of interest from the backdrop, image segmentation techniques can be used. Segmentation facilitates the analysis of particular fruit regions and prevents unneeded interference from ancillary components.
- Image Registration: Aligning images from different angles or time points through image registration ensures consistent perspectives, enabling the detection system to track disease progression accurately.
- Data augmentation: Increasing the variety of training samples and enhancing the generalization and robustness of the CNN model can be accomplished by enhancing the dataset using picture preprocessing methods like rotation, flipping, or translation.
- Region-of-Interest (ROI) Extraction: The analysis can be improved and computer resources used less by locating and extracting regions from the fruit photos that specifically exhibit symptoms of the disease.

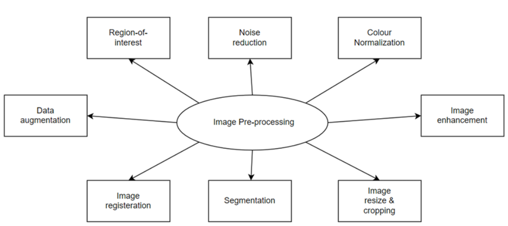

Fig. 2. Image Pre-processing techniques

The fruit disease detection system may obtain cleaner, more standardized, and informative images by utilizing various image preprocessing approaches, increasing the system's accuracy in identifying and classifying fruit diseases. The performance of the overall fruit disease detection project is ultimately improved by image preprocessing, which is a vital step in ensuring that the input data is adequately processed for future feature extraction and CNN model training.

C. Most used Image Segmentation Techniques by deep learning Models

The procedure of separating the elements which are there in an image is referred to as image segmentation. Figure 2 presents the popular segmentation approaches. To improve performance, the image must be segmented into relevant clusters. The major goal of this illustration is to distinguish between the fruit's foreground and background. The margins produced by segmentation aid in identifying the fruit's infection-prone areas. However, as deep learning models learn the picture segments on their own, this step is not required in works that use deep learning models. In models other than CNNs, clustering techniques, primarily k-means clustering, have been used. A different approach is OTSU thresholding [2]. The four significant methods have been covered in detail as follows:

- K-means clustering: [1]Given that the area the researcher is interested in could not be present in the entire image, as seen in the dataset, the image must be segmented. The k-means clustering approach is one of the researchers' go-to methods for detecting clusters of polluted areas because this data is not currently available, necessitating the use of an unsupervised algorithm for the analysis. The k means clustering is used as a segmentation strategy, resulting in clusters of the infected and healthy regions, making it simple for future detection. The programme employs a distance calculation method and further clusters regions with similar properties. Euclidean distance is the most often used formula for calculating distance.
- OTSU thresholding: Otsu's method, named after Nobuyuki Otsu [2], is used in computer vision to carry out programmed picture thresholding. By choosing a particular value, thresholding's main objective is to binarize an image into foreground and background. Here, the threshold value helps to reduce background noise in the image. By utilising this technique, the given image (typically in grayscale) is displayed as a histogram, and the threshold value is determined by the dispersion of the pixels on each side of the threshold. These thresholds divide the image into several halves.
- CNN for segmentation: While training and learning, this layer of the convolution neural network generates portions of the image. The section on deep learning models that follows in-depth discusses the convolution layer. The model attempts to distinguish between the healthy and unhealthy parts of the image after segmenting it. Utilizing various algorithms covered in the following section, the model is trained.
- Graph-based segmentation With this effective image segmentation method, pixels in the image are represented as nodes in a graph, and the connections between them are

shown as edges. This technique, which separates the image into regions based on pixel similarities to detect fruit diseases, makes it possible to distinguish between healthy and diseased parts of the fruit with high precision [2]. The graph partitioning technique divides the image into useful segments, capturing object boundaries and successfully managing image variability. It commonly uses graph-cut algorithms. The system distinguishes between healthy and diseased sections, allowing for accurate fruit disease identification. It does this by examining the properties of each segment, such as color histograms and texture descriptors. Graph-based segmentation is a valuable tool for precise and focused illness diagnosis despite needing careful parameter tweaking and maybe having increased computing complexity.

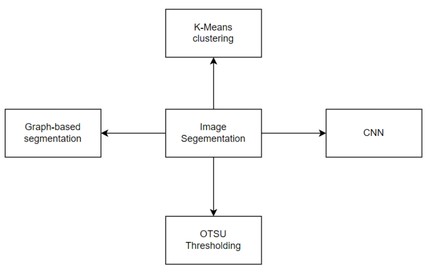

Fig. 3. Image Segmentation techniques

D. Comparative Analysis of various deep learning models used for fruit disease detection

Mohanty et al., 2016 [5], used noise removal, grayscaling, and resizing as pre-processing and segmentation techniques and used Google Net and Alex Net networks on Potato and Bell Pepper plants considering Plant Village Dataset achieved 85.3% Accuracy. And, the critical analysis showed that in an open setting, the trained model produced accuracy of only 31%. As a result, it is necessary to add a dataset with more leaves that have been photographed from more angles and under both controlled and uncontrolled climatic circumstances.

Thangaraj et al., 2020 [6] used Resizing, Flipping, and rotating as pre-processing and segmentation techniques and using Mobile Net networks on Avocado plant considering data collected using Cameras achieved 98% Accuracy. And, the critical analysis showed that the data is processed using a Convolution layer and a ReLu function to extract the features, and a SoftMax activation function is used to estimate the likelihood that each leaf would fall into one of the 8 classes that were studied.

Saha et al., 2020 [7] used Average Filter, Color Threshold as pre-processing and segmentation techniques and using CNN on Orange plant considering Data collected using Camera achieved 100% Accuracy. And, the critical analysis showed that the proposed model has to be expanded by adding datasets because there were only 20 to 30 photos in each category.

Katumba et al., 2020 [8] used Resizing as pre-processing and segmentation technique and using CNN on Passion Fruit plant considering Data collection in collaboration with Uganda National Research Institute achieved 0.351 precision. And, the critical analysis showed that to get good results, the CNN underwent 10,000 iterations. The experiment was carried out on a system with 56 GB of RAM, which sped up the procedure. The model's disadvantage is that it necessitates extensive calculation.

Sullaca et al., 2019 used Noise Reduction as pre-processing and segmentation technique and using CNN on Blueberry plant considering Data collected using cameras achieved 84% Accuracy. And, the critical analysis showed that the performance of these machine learning techniques when compared to deep learning models is better. Different algorithms were examined on the provided dataset.

Xie et al., 2020 [9] used Noise Reduction as pre-processing and segmentation technique and using CNN on Grape plant considering Grape leaf disease dataset achieved 99.53% Accuracy. And, the critical analysis showed that the researcher's suggested model is a ResNet variant with additional anchors in the layers.

Jiang et al., 2019 [10] used Convolution layer as pre-processing and segmentation technique and using Inception Model on Apple plant considering Apple Leaf Disease Dataset showed that Inception was improved by replacing the single 5X5 layer to two 3X3 layers. The inception model was also modified to a network which did not use conventional nXn layer.

Zhong et al., 2020 [11] used Dense121 network on Apple plant and considering Apple Leaf Disease Dataset achieved 99.31% Accuracy. And, the critical analysis showed that the model is trained on a small number of photos and must thus be expanded on a sizable amount of data, which is a limitation of the presented strategy for preventing overfitting and underfitting.

Polder et al., 2019 [12] used CNN on Potato and the experiment was set up in central polder field in Netherlands. Achieved 0.78% Precision. And the critical analysis showed that because the data utilised was unbalanced, the precision value was lower than that of the other models.

Ferentios et al., 2018 [13] used Resizing as pre-processing and segmentation technique and AlexNet, GoogleNet network on Multiple and considering public dataset of 87848 images 99.53% Accuracy. And, the critical analysis showed that it was possible to identify a wide range of plants and their ailments. But a more comprehensive real-time system needs to be created.

Ji et al., 2020 [14] used GoogleNet, VGGNet networks on Grape plant, considering Plant Village Dataset achieved 99.17% Accuracy. And, the critical analysis showed that to attain greater accuracy, a hybrid ResNet50 and Inception model has been developed.

Ma et al., 2018 [15] used Converting colorspace to HSV, resizing as pre-processing and segmentation technique and CNN network on Cucumber and considering Plant Village Dataset achieved 93.4% Accuracy. And, the critical analysis showed that on the machine learning models and the deep learning model, a comparison analysis has been done. The model performs better than machine learning models and prevents overfitting.

Lin et al., 2019 [16] used Background to black, to HSV, OTSU, masking as pre-processing and segmentation technique and CNN network on Cucumber and considering Cucumber Fruit Leaf Phenotype Automated Analysis Platform used to create samples achieved 96.08% Accuracy. And, the critical analysis showed that the number of samples were few.

Khan et al., 2021 [17] used Resizing as pre-processing and segmentation technique and ResNet34, CNN network on Apple and considering data collected using Camera in Kashmir Valley samples achieved 97.18%, 99.1% precision. And, the critical analysis showed that the model employed a gradient-based strategy in fewer epochs rather than the usual ReLu and SoftMax activation functions.

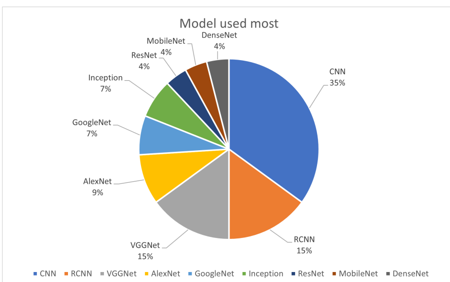

Fig. 4. Most used models in the field of fruit disease detection

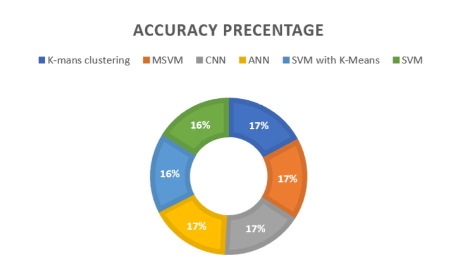

Fig. 5. %Accuracy achieved by various models

E. General Architecture

The steps taken to detect the disease are briefly described in this section [2]. Each step is discussed in each paragraph below as the section progresses through them in the sequence shown in Figure 1. There were five steps to the models that were used to detect fruit illnesses. Data preparation, which comes after data collection and is the initial stage, comprises picture resizing, flipping, rotating, and colour conversion to create an image that can be sent to the model for further processing. The following stage is image segmentation, which is optional when employing deep learning models because they can segment the image on their own during training. The last phase involves the classification model that generates a training model. This is then put to the test using datasets or real-time data.

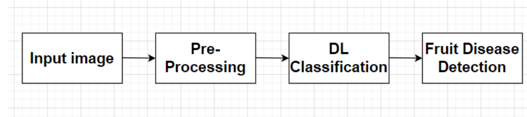

Fig. 6. Basic model architecture used in fruit disease detection

F. Models Performance Evaluation Metrics

Several parameters can be used to evaluate deep learning models. One of the parameters is model speed; a model's effectiveness depends on how quickly it operates. The training speed is influenced by both the model type and hardware requirements of the system being employed. Given that each research has a different experimental setting, this metric cannot be utilized to compare various proposed models [2].

The error rate is the second factor. The key to a successful model is low error rates. To reduce error rates, deep learning models employ backpropagation techniques. The most effective models are those that produce the fewest errors. Calculating the errors in the model between the actual value and what was expected using the mean square error along with root mean square error rate is an coherent method.

The most frequently utilized parameter for evaluating the deep learning models shown in Figure 4 is the third parameter, the confusion matrix. Actual true, Actual false, Predicted true, and Predicted False are the four parameters of the matrix. Using the confusion matrix, the values of numerous additional metrics, including precision, recall, accuracy, and F1 score, may be calculated.

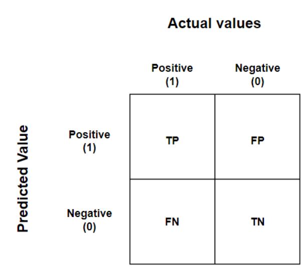

Fig. 7. Confusion Matrix

IV. Suggestions Based on Models Compared

We have come across several current technologies and techniques that are being used to detect disease in fruits that use CNN and Image Processing and after working on them here are a few suggestions that can be adopted to improve the process so some of them are:

1) Transfer Learning: It is a technique that can use pre-trained CNN models on huge image datasets. These models can be used as a starting point, and their performance and accuracy can be enhanced by fine-tuning them using a smaller dataset of fruit disease photos. By utilizing the learned features from comparable activities or domains, transfer learning can give an edge [6].

2) Ensemble learning: To increase overall accuracy, ensemble approaches combine predictions from various models. The system can gain from many viewpoints and perhaps achieve improved accuracy by training numerous CNN models with various architectures or configurations, such as various depths or regularisation techniques, and integrating their outputs.
3) Multimodal Data Fusion: By combining several data modalities, such as hyperspectral imaging or near-infrared spectroscopy, extra knowledge on the health and sickness of fruits can be obtained. The detection system can capture a wider spectrum of illness signs by combining spectral and image data, providing more precise and trustworthy results.
4) Data Augmentation Techniques: Information expansion includes producing extra preparation tests by applying changes or annoyances to the current dataset. Strategies like revolution, scaling, trimming, and clamor expansion can assist with expanding the variety and size of the preparation information, which can work on the speculation and heartiness of the CNN model.
5) Domain-Specific Preprocessing: Preprocessing methods custom fitted to the qualities of natural product pictures and illnesses can improve the presentation of the framework. For instance, procedures like variety standardization, brightening rectification, or district-of-interest extraction can assist with normalizing the information and spotlighting illness-explicit areas, diminishing commotion, and further developing recognition precision.
6) Active Learning: Active learning approaches make the best use of human expertise by iteratively choosing informative samples for labelling. The system may intelligently choose difficult or ambiguous samples for manual annotation by including active learning in the training process. This enhances the model's performance with little assistance from experts.

V. Proposed Methodology

Deep learning models have become effective tools in recent years for attaining highly precise and effective fruit disease diagnosis. Convolutional Neural Networks (CNNs), which take use of developments in computer vision, are now the mainstay of such systems. With the help of visual symptoms, our deep learning models can efficiently distinguish between healthy and unhealthy fruits by automatically learning and extracting discriminative characteristics from fruit photos. Because of deep learning's adaptability, several CNN architectures, such ResNet, VGG, or Inception, can be combined and each has its own advantages for learning complex patterns and representations. Recurrent neural networks (RNNs) can also be used to sequentially analyze temporal data in the course of fruit diseases [18].

CNN based models have emerged as the most consistent and accurate for the fruit disease detection because of the various factors discussed below:

1) Feature Extraction: CNNs are created with the explicit purpose of learning and automatically extracting significant features from images [19]. These models, which identify various patterns, textures, and forms in the input images, are made up of many layers of convolutional filters. CNNs can recognize subtle visual signs and disease-related patterns for fruit diseases that may be difficult to detect using conventional image processing techniques.
2) Hierarchical Representation: CNNs are excellent at learning hierarchical representations of feature data. While higher layers capture more intricate and abstract aspects, lower layers just catch fundamental features like edges and colors. By using a hierarchical approach, CNNs are better able to distinguish between complex patterns and changes linked to various diseases.
3) Adaptability to Varying Data: The illumination, orientation, and disease stages of fruit disease photos may differ. It is generally known that CNNs can adapt to such fluctuations and generalize effectively across various data distributions. They are adaptable to a variety of fruit varieties and disease symptoms, which makes them useful for practical applications.
4) End-to-End Learning: CNN-based models are able to perform end-to-end learning without substantially depending on hand-engineered feature engineering since they learn directly from the unprocessed picture input. This makes the process of developing the model simpler and guarantees that the learned features are optimized specifically for the goal of identifying fruit diseases.
5) Transfer Learning: CNN models that have already been trained, such as those that have been trained on huge picture datasets like ImageNet, can be tailored to detect fruit diseases. Researchers can use transfer learning to adapt the concepts and representations they have learned from general picture classification tasks to more specialized tasks like illness detection. Even with scant labeled data, our method expedites training and enhances model performance.
6) Scalability: CNN-based models can scale to effectively handle big datasets. CNNs can be retrained or expanded to cover more disease classes as more labeled data becomes available or as new disease kinds appear, making them flexible and future-proof.
7) Real-Time and On-Device Applications: CNN models can be installed on edge devices and optimized for real-time inference, making them suited for on-device, in-field fruit disease detection systems [20]. Farmers and agricultural professionals may undertake quick and accurate disease assessments because to this capability.

According to Wang. H et al. [2] we can enhance the mask R-CNN technique to detect disease in fruits like apples, oranges, peaches, and pears as the research objects, and by this, we can get the insights like:

1) Improve Mask R-CNN for multi-scale feature fusion:

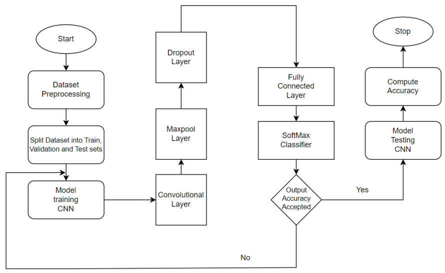

Fig. 8. CNN-based architecture for fruit disease detection

To the initial feature pyramid structure, add a bottom-up horizontal connection path [21]. The effect of multi-scale feature fusion is improved, and large-scale targets in the lesion detection process are effectively solved. It is not the best value to have poor positioning and network detection accuracy.

2) The algorithm's mAP value in detecting the target fruit surface lesions is above 95%, and the average detection speed is 2.6 frames per second when using GPU, according to the test results.
3) Despite the excellent detection accuracy of the proposed detection model in the paper, the detection speed is slightly slower. The next step will be to investigate how to increase detection speed while maintaining model detection accuracy.

VI. Proposed algorithm for implementation of the system

1) A bulk of input dataset is given for training.
2) A set of Test Datasets is given to test the accuracy of the training model.
3) Perform feature extraction using color attributes.
4) Utilise texture and colour properties to perform feature-level fusion..
5) Apply a classifier using a convolution neural network on the segmented image.
6) If the fruit has any ailment, proceed to step 7 instead. If not, proceed to step 8.
7) Use K-means clustering to separate images.
8) Print "Given fruit is Healthy".
9) Print the result with classified fruit disease.
10) Fetch the Proposed Remedy for the problem from the database.

VII. Pros and Cons For Proposed Model

In this section, we will look at some of the advantages that our model has over other models and also some of the expected difficulties or disadvantages that might occur during the entire course of the development of the model.

A. Advantages of proposed model

- Exactness is high, and it is Relevant for both low and high-pixel pictures.
- Upgrading the worth of natural product infection location likewise requires a couple of moments to give a careful outcome.
- The name of the illness is additionally found by featuring the impacted spots.

B. Problems that might occur in proposed model

- Limited Dataset: The need for a big and varied training dataset is one of the key drawbacks of applying Convolutional Neural Networks (CNNs) for early-stage fruit disease detection.
- High False Positive Rate: Similar to other machine learning models, CNNs are prone to false positives, which means they could mistakenly label a healthy fruit as being infected. Early-stage fruit illnesses frequently display modest signs that may be difficult for human observers to notice
- Variability in Disease Expression: The diversity in disease expression presents another difficulty in early-stage fruit disease detection with CNNs. The appearance and expression of symptoms can be influenced by a number of variables, including environmental conditions, fruit types, and disease progression stages.

VIII. Results And Discussion

Our Convolutional Neural Network (CNN) model for fruit disease identification performed well across a wide range of disease classes with excellent accuracy, demonstrating its capacity to generalize well to new data. However, difficulties such data paucity and probable over-fitting were noted, highlighting the necessity of further study to resolve these restrictions. Using CNNs to automate the identification of fruit diseases has great potential for increasing agricultural output and food security, but further research is needed to maximize the model's real-world applicability and make it useful for use in agricultural contexts.

1) Data Collection and Preprocessing:
   assemble a database of photos of both healthy and ill fruits. Resize, normalize, and enhance the data before preprocessing the photos to increase model resilience.
2) Data Splitting:
   The dataset should be partitioned into three distinct subsets, namely the training set, the validation set, and the test set. Typically, a distribution ratio of 80-10-10 or 70-15-15 is used.
3) Model Selection and Architecture:
   As the core of your model, use a CNN architecture like VGG, ResNet, or Inception. Add completely linked layers to the architecture to make it more suited to your particular goal.
4) Model Training:
   Utilizing an appropriate optimizer (such as Adam or SGD) and loss function (such as categorical cross-entropy), train the CNN model using the training dataset. Visualize loss and accuracy on the validation set to track

training progress. To avoid overfitting, use strategies like early halting.

5) Model Evaluation:
Metrics like accuracy, precision, recall, and F1-score may be obtained by evaluating the trained model's performance on the test dataset. Create a confusion matrix to see how well the model performs across various classes.

6) Visualization of Results:
To demonstrate the effectiveness of the model, provide visuals like confusion matrices, precision-recall curves, and ROC curves. Include examples of both instances where the model accurately detected illnesses and instances where it did not.

Fig. 9. Results Obtained

IX. Conclusion

Overall, it is clear from the research that despite having to deal with a number of obstacles, including a smaller and more accurate dataset, a high incidence of false positives, and heterogeneity in illness expression, we tried to find answers and thought about some of the methods to get around them. Data augmentation, which involves enhancing the training dataset by performing different transformations including rotation, scaling, flipping, and adding noise, was the method we examined. By doing this, you may broaden the dataset's variety, relieve the data shortage for uncommon diseases, and cut back on overfitting. The model might also be routinely updated by periodically being retrained with fresh data and adding samples of newly discovered illnesses. This made it possible for the model to accurately predict developing illnesses and stay current. Another alternative has been to use various data processing techniques. used image processing methods to improve image quality, clean up the background, and equalize lighting conditions. This enhanced the model's capacity for generalization and reduced variability brought on by external influences. The most effective technique we discovered and used was CNN-based image processing, which included steps like Image Acquisition, Image Pre-processing, and Image Segmentation. Using this method, we input the image, process it, then extract more information from it, classify it using DL, and finally identify the fruit disease.

References

[1] P. K. Devi *et al.*, "Image segmentation k-means clustering algorithm for fruit disease detection image processing," in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2020, pp. 861–865.

[2] S. Goel and K. Pandey, "A survey on deep learning techniques in fruit disease detection," *International Journal of Distributed Systems and Technologies (IJDST)*, vol. 13, no. 8, pp. 1–19, 2022.

[3] D. Bhise, S. Kumar, and H. Mohapatra, "Review on deep learning-based plant disease detection," in *2022 6th International Conference on Electronics, Communication and Aerospace Technology*. IEEE, 2022, pp. 1106–1111.

[4] J. Ahmad, B. Jan, H. Farman, W. Ahmad, and A. Ullah, "Disease detection in plum using convolutional neural network under true field conditions," *Sensors*, vol. 20, no. 19, p. 5569, 2020.

[5] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in plant science*, vol. 7, p. 1419, 2016.

[6] R. Thangaraj, D. Dinesh, S. Hariharan, S. Rajendar, D. Gokul, and T. Hariskarthi, "Automatic recognition of avocado fruit diseases using modified deep convolutional neural network," *International Journal of Grid and Distributed Computing*, vol. 13, no. 1, pp. 1550–1559, 2020.

[7] R. Saha and S. Neware, "Orange fruit disease classification using deep learning approach," *International Journal*, vol. 9, no. 2, 2020.

[8] A. Katumba, M. Bomera, C. Mwikirize, G. Namulondo, M. G. Ajero, I. Ramathani, O. Nakayima, G. Nakabonge, D. Okello, and J. Serugunda, "A deep learning-based detector for brown spot disease in passion fruit plant leaves," *arXiv preprint arXiv:2007.14103*, 2020.

[9] X. Xie, Y. Ma, B. Liu, J. He, S. Li, and H. Wang, "A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks," *Frontiers in plant science*, vol. 11, p. 751, 2020.

[10] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks," *IEEE Access*, vol. 7, pp. 59 069–59 080, 2019.

[11] Y. Zhong and M. Zhao, "Research on deep learning in apple leaf disease recognition," *Computers and Electronics in Agriculture*, vol. 168, p. 105146, 2020.

[12] G. Polder, P. M. Blok, H. A. De Villiers, J. M. Van der Wolf, and J. Kamp, "Potato virus y detection in seed potatoes using deep learning on hyperspectral images," *Frontiers in plant science*, vol. 10, p. 209, 2019.

[13] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and electronics in agriculture*, vol. 145, pp. 311–318, 2018.

[14] M. Ji, L. Zhang, and Q. Wu, "Automatic grape leaf diseases identification via unitedmodel based on multiple convolutional neural networks," *Information Processing in Agriculture*, vol. 7, no. 3, pp. 418–426, 2020.

[15] J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, and Z. Sun, "A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network," *Computers and electronics in agriculture*, vol. 154, pp. 18–24, 2018.

[16] K. Lin, L. Gong, Y. Huang, C. Liu, and J. Pan, "Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network," *Frontiers in plant science*, vol. 10, p. 155, 2019.

[17] A. I. Khan, S. Quadri, and S. Banday, "Deep learning for apple diseases: classification and identification," *International Journal of Computational Intelligence Studies*, vol. 10, no. 1, pp. 1–12, 2021.

[18] S. Shajahan, "Fruit disease detection using cnn."

[19] K. Gopalakrishnan and P. Vanathi, "Fabric defect detection using deep learning techniques," in *International Conference on Ubiquitous Computing and Intelligent Information Systems*. Springer, 2021, pp. 101–113.

[20] P. Sasikumar and S. Sathiamoorthy, "Visual place recognition using region of interest extraction with deep learning based approach," in *International Conference on Ubiquitous Computing and Intelligent Information Systems*. Springer, 2021, pp. 359–372.

[21] A. M. Rakhi, A. P. Dhorajiya, and P. Saranya, "Enhanced mask-rcnn for ship detection and segmentation," in *International Conference on Ubiquitous Computing and Intelligent Information Systems*. Springer, 2021, pp. 199–210.