International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I1P112

Volume 73 Issue 1, 146-154, January 2025


*Original Article*

# An Approach for Diagnosing and Differentiating Mango Fruit Diseases Using Hybrid CNN+SVM Classifier

Sandeep Kumar[1], Bhupesh Gupta[2], Neha Bhatia[3], K. Anuradha[4]

[1,2]*Department of CSE, MM (deemed to be) University, Mullana, Haryana, India.*
[3]*Department of IT, Panipat Institute of Engineering & Technology, Haryana, India.*
[4]*Department of AI & DS, KL University, Andhra Pradesh, India.*

[1]*Corresponding Author : sandeepgoel1986@gmail.com*

Received: 02 August 2024 Revised: 11 December 2024 Accepted: 17 December 2024 Published: 31 January 2025

***Abstract -** Fruits are an important source of nutrients needed to maintain excellent human health. The major cause of the notable decrease in crop yield is the considerable influence of fruit diseases, which result from poor maintenance techniques and the spread of fungal infections. Mango fruit is consumed widely around the world, although its quantity and quality might be affected by diseases. The laborious, time-consuming, heavily dependent on human labor, and inefficient aspects of the manual inspection process define it. The goal of this work is to create a hybrid model for mango fruit sickness detection that combines the strengths of a support vector machine and a resilient convolutional neural network. The salient features of both classifiers are combined in the study. In the proposed hybrid model, CNN serves as an automated feature extractor and SVM as a binary classifier. The suggested model's algorithm is trained and assessed using this dataset, which features images of four distinct mango fruit diseases and healthy mangoes. The findings show the efficacy of the suggested work with a 99% detection accuracy over the mango fruit disease dataset.*

***Keywords -** CNN, SVM, Mango Fruit Disease, Diagnosing, Differentiating.*

## 1. Introduction

Mangoes are an extremely lucrative fruit; hence, they are commonly farmed in tropical and sub-tropical climates worldwide. The pleasant aroma, tender flesh, and high nutritional content of mangoes have made mango enthusiasts the world over, and this has greatly benefited exporting nations as well as mango farmers. It is crucial to stress that the aesthetic attractiveness of a mango fruit is a key factor in determining its economic value. The most aesthetically beautiful mangoes are often exported, while the less aesthetically pleasing mangoes are kept for home use. Vitamin E, potassium, vitamin C, protein and calories, and niacin are the main types of these nutrients found in mango fruit. Mangoes and bananas are India's two most popular fruit crops, which are worth crores of rupees. On 16,363.48 hectares of land, 105,379.375 tons of mangoes are produced annually, according to the Central Statistical Agency [1]. Mango-related illnesses, however, can result in significant losses for a crop. Therefore, to detect the abrupt decline in yield and agricultural output levels, it would be imperative to diagnose mango illness as soon as feasible. Traditional methods of diagnosing early illnesses are labour-intensive, time-consuming, and often inaccurate. Implementing preventive measures on the farm would cause a hefty loss since the disease could easily spread and injure the mango plants. The capacity to swiftly and impartially evaluate the health state of mango trees is another benefit of automated detection and diagnosis systems. By refining the precision and efficacy of disease diagnosis and detection, this method enables farmers to stop the blowout of illness and guard their harvests promptly. Through automation it lowers the cost of disease control, gives farmers more precise information about the severity of the disease, and lessens the need for physical inspections. Mango production may be sustainably managed by automated technologies. They oversee the use of pesticides and other damaging substances [2]. An encouraging substitute is the automated diagnosis and detection of mango illnesses made possible by deep learning algorithms. Deep learning successfully combines picture categorization, object recognition, and natural language processing. Recently, it has been used to automate the diagnosis and detection of a number of ailments, most notably mango disease. The two biggest obstacles to machine learning adoption across many industries are data quality and choosing an existing tool appropriate for a certain task in the industry [3]. Since we also do not emphasise developing hi-tech network topologies or learning techniques here, our experience aligns with this general trend. Both healthy and sick mango fruits were among the 3333-five kinds of mango fruits used in this investigation. However, there are a lot of problems since the photographs are shot by people who work for global companies that process mangoes. Distractions include background noise that is out of focus, unequal fruit closeness and orientation, and variable lighting. The study being presented describes a hybrid deep learning model that reliably enables the involuntary identification and analysis of mango illnesses with a low number of false negative detections. Mango fruit image segmentation efficiently separates the region of interest for illness mango identification using threshold image segmentation. For effective feature extraction from segmented filters on the mango illness-affected photographs, hybridization between CNN and SVM is employed. As a high-level representative feature, CNN is mainly used to learn complex patterns [4]. Finally, characteristics are extracted using SVM to categorize infections in mango diseases. Training the SVM model with a well-annotated dataset of photos of infected mango disease yields significant classification results. This approach establishes a systematic and reliable framework for analyzing mango disease images by integrating CNN and SVM classification.

The following research gaps were identified:

- Most studies rely on limited datasets that may not include various mango varieties or diseases, leading to models that are not generalizable across different conditions and regions.
- While many studies focus on specific diseases like anthracnose or powdery mildew, there is a lack of comprehensive models that can simultaneously detect and classify multiple diseases affecting mangoes.
- Variability in image quality and acquisition methods across studies poses a challenge. Establishing standard protocols for image collection and processing can improve model robustness.

### *1.1. Diseases In Mango Fruit*

Rot of Alternaria (black spot) In Australia, Egypt, India, Israel, and South Africa, post-harvest fruit rot is caused by Alternaria rot, commonly referred to as black spot. Black stains and post-harvest fruit rot during storage are caused by the fungus (germ). When the fruit is still developing, the germ enters and goes dormant. The fruit rots actively when the germs start growing again after it is picked and as it ripens more. Anthracnose is among the most dangerous mango illnesses in many regions where the fruit is cultivated. Where it rains during the mango blooming and fruit set stages is where anthracnose has its most destructive impact. In some situations, anthracnose can pose issues in the Northern Territory, even though it seems less of an issue because flowering and fruit sets take place during the dry season. A genus of fungi called Aspergillus contains a number of species that are known to infect plants, especially mangoes. Aspergillus niger and Aspergillus flavus are the most common culprits affecting mango fruits. The diseases caused by these fungi can lead to significant post-harvest losses due to their impact on fruit quality and shelf life.

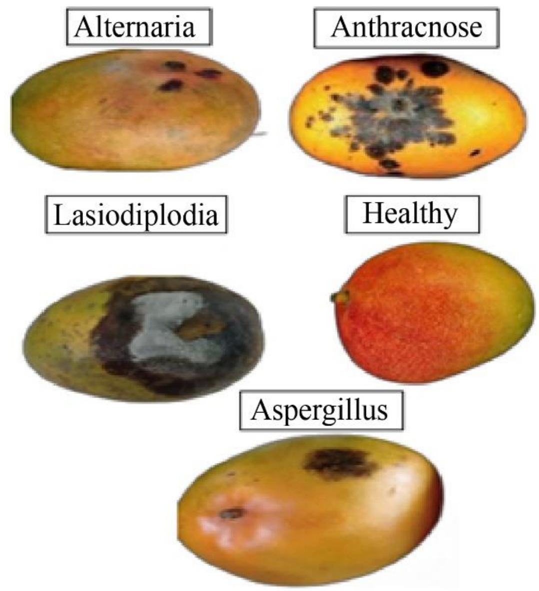


**Fig. 1 A sample image of healthy and diseased mango**

Lasiodiplodia is a fungal pathogen that causes significant mango tree diseases, particularly stem-end rot and dieback. This fungus is common in hot and subtropical areas and can severely damage mango orchards if not properly managed.

## 2. Related Work

Numerous approaches and strategies have been developed and put into use in recent years for the identification of diseases affecting the leaves of plants. These may be broadly divided into two categories: methods for locating or detecting diseases and methods for sorting or classifying diseases. Several techniques, such as feature fusion, segmentation, and picture classification, for mango, citrus, and tomato, were used. Similar to how they employ appropriate segmentation, feature extraction, and classification, these techniques are suitable for fruit, flower, and leaf diseases. Often referred to as the "king of fruits," mangos are a popular fruit in India and other countries. Mango slices, pulp, and fresh mangos are among the many varieties of mangos that are exported.

Nearly 27000 million tons of mango fruit were shipped globally in 2021–2022. Rabia Sallem et al. [5] This study uses a full-resolution convolutional network to identify mango leaf disease. By producing a high return on investment, the mango, the most popular fruit in the world, aims to lower disease rates. Owing to variations in symptoms, which they are currently detecting, and automated leaf segmentation of the illness. The capacity to recognize the sickness from any computer-aided system is the most crucial requirement for correctly segmenting it. Nithya et al. [6] primarily wanted to create a computer vision system that uses deep learning techniques to classify mango illness. The researchers employed a CNN to develop a model for identifying mango flaws. An online repository was used to gather 50 healthy and 50 diseased mangoes for the investigation. Data preparation techniques were used to increase the same dataset, remove noise, and enhance image quality. An adaptive Wiener Filter was utilized to eliminate noise from a histogram, and equalization was utilized to enhance contrast and quality. Ultimately, an 89.5 percent accuracy rate was attained by a CNN-based model for computer vision-based mango flaw identification. Omneya Attallah [7] The primary objective of this study is to classify tomato leaf diseases using CNN feature selection and transfer learning. This hyperparameter in high computability under the deep learning architecture has been used in many research that has employed the deep learning model in the classification complexity. Na-Eun Kim et al., [8] The categorization of tomato disease leaves using a convolutional neural network with lightweight attention is the basis of this study. Plant diseases provide a serious barrier to food safety and protection. Significant losses are necessary to shield crops from plant diseases promptly.

Accurate and automatically increasing disease detection capabilities are made possible by advances in phytopathology technology and computer vision. Prabhjot Kaur et al. [9] This article uses a reduction feature hybrid CNN to identify leaf diseases. India's economic growth and wealth are largely due to agriculture. When producing agricultural products, safety and loss from food destruction are taken into account. Identification of plant diseases over the long term is crucial for agriculture sustainability. Diverse illnesses and constraints of time owing to varied disease plant monitoring manually. Umit Atila et al. [10] This paper was created using deep learning with an efficient net model on plant leaf disease sorting. Plant leaves were sick through a visual check for diagnosis of the disease utilizing modern techniques and symptoms, which are visible for every kind of plant disease. Arya & Singh [11].

The investigation compares and contrasts convolutional neural networks applied using Alex Net concerning the diagnosis and identification of diseases in mangoes and potatoes, using almost 4004 images. Photos of mangoes were obtained using a local dataset, while potato images were sourced from the online repository Plant Village. It is based on experiments for comparing the efficiency and effectiveness of CNN and Alex Net architectures in identifying and classifying diseases of mangoes and potatoes. The study concludes that the value of 99% diagnostic accuracy for diseases of mango and potato by Alex Net is far better than CNN. Wongsila et al. [12] The author used a data set of a thousand images of healthy and diseased mangoes to train the CNN. The deep learning techniques help to separate those mangoes that are free from an anthracnose infection from those that are infected. Besides, on a held-out test set for the CNN, a whopping result of 97.62% excellent accuracy was recorded, meaning much improvement from previous approaches. The author developed a classification classifier that collects 1000 images of both healthy and ill mangoes. CNN attained an impressive precision rate of 97.62% on the test set, demonstrating remarkable advancement compared to previous approaches. Soumia Bensaadi et al. [13] This article classifies tomato plant ailments using an inexpensive convolutional neural network. Agriculture cannot serve as a source of energy and raw materials for industry without a robust food supply. The early detection tool used to solve this issue is a major component of the advancement of machine learning algorithms and computing technology in the field of agriculture.

Rabbia Mahum et al., [14] The foundation of this study is an efficient deep-learning model used to detect potato leaf disease. Agriculture significantly contributes to crop productivity and is essential in controlling potato illnesses. Humans have to work to reduce the source of plant disease by doing duties. Jinzhu Lu et al., [15] This study is based on an effective deep-learning model for detecting potato leaf disease. Agriculture is crucial to reducing potato diseases and makes a substantial contribution to crop output. Humans must perform their obligations in order to lessen the source of plant disease. The necessary crop is classified in light of the aforementioned challenges, and accuracy is assessed using an automated time detection method. Based on improved deep learning algorithm techniques, the components of the potato leaf visualization are proposed.

Singh et al. [16]. The author created a classification model using an MCNN trained on a dataset of 1,070 photos of anthracnose-affected mango leaves. The MCNN outperformed earlier approaches, achieving an accuracy of 96.89% on the test set. Maha Altalak et al., [17] This paper is based on deep learning technologies under the smart agriculture application. These food security goals achieve the country's economy, which has become inevitable under smart traditional agriculture. They applied the RNN (Recurrent Neutral Network) and CNN (Convolutional Neutral Network) under deep learning technology in various fields.

In agriculture, based on the recent article on deep learning method under the study analysis by the internet of things it has been monitored the agriculture parameter as we investigated it. Plant leaf disease may be identified and categorized using the agriculture system bits of intelligence we suggested based on a hybrid model of SVM and CNN. Ashok and Vinod [18] employed an android application in conjunction with a deep learning model to provide mango fruit disease detection. This process uses a CNN model that has been developed with a library collection of many photos of both affected and healthy mango fruits. Identification of mango fruit diseases is performed through the Android application with the help of this CNN model on mango fruit photos. This method was implemented using 1000 samples, whereby 500 were affected with various mango fruit diseases while the other 500 were healthy. The technique provided an accuracy level of 95% in detecting mango fruit diseases.

**Table 1. Literature review summary**

| Author Name | Techniques | Description | Limitations | Performance Accuracy |
|---|---|---|---|---|
| Rabia Saleem, et, al., [5] | CCA-based fusion and SVM | Leaf features extracted post-segmentation. | Identified disease for powdery mildew and sooty mould. Need to focus on remaining diseases. | 95.5% |
| Nithya et al[6] | CNN | A CNN-based computer vision system for classification in mango disease. | Relied on a small dataset, limiting generalizability. | 89.5% |
| Omneya Attallah [7] | CNN | The CNN condensed and representation of high level under the final fully connected layer to retrieve deep feature | Not identifying the severity of leaf disease. | 99.8% |
| Na-Eun Kim et, al., [8] | CNN, with lightweight attention | Classification of metrics accuracy standard using evaluated with model of performances based on CNN | Loss of information because of merging the leaf features | Average accuracy |
| Kaur et al.,[9] | Hybrid CNN | Features are extracted using the FC layer and variance technique | Training sets are complicated | 98.7% |
| Atila et.al.,[10] | CNN | CNN model is used for high resolution and augmented datasets were used. | Have to improve the dataset by increasing the number of classes. | 99.39% |
| Arya & Singh[11] | CNN | Demonstrated AlexNet's superior performance over CNN in diagnosing mango and potato diseases. | The reliance on specific datasets may limit broader applicability. | 99% |
| Wongsila et al.,[12] | CNN | A CNN effectively classified anthracnose-infected and healthy mangoes, demonstrating significant improvement over previous methods. | Limitations could be the model's potential overfitting to the specific dataset, limiting its generalizability to diverse or unseen mango varieties and conditions. | 97.62% |
| Bensaadi S et al., [13] | CNN and ANN | These methods avoid overfitting and use data augmentation by using large data sets to achieve accuracy. | Need to improve the ability to recognize new diseases | 97.04% |
| Mahum et al.,[14] | SVM | This method helps to identify 5 classes of potato leaf disease. | More training time is required, and fewer datasets are used. | 97.2% |
| Lu et.al., [15] | Deep learning (CNN) | Reviewed insufficient datasets | Insufficient datasets, No ideal robustness | 91.83% |
| Singh et al. 16 | MCNN model | An MCNN model classified anthracnose-affected mango leaves, but its small dataset may limit broader applicability. | The limitation could be focusing on a single disease, limiting the model's applicability to other mango leaf conditions or diseases. | 96.89% |
| Altalak et.al., [17] | CNN & RNN | Provides efficient solutions using hybrid deep learning systems | Need to evaluate the performance-based accuracy | 94.27% |
| Ashok and Vinod [18] | CNN | An Android application was developed using a CNN model to detect mango fruit diseases from a dataset of 1000 mango images. | The method is limited by its platform dependence on Android and potential sensitivity to variations in image quality and environmental conditions. | 95% |

## 3. Methodology

The method to detect and classify mango fruit diseases begins with input, preprocessing, extracting features, and classifying the diseases. The Research work, a hybrid model using CNN and SVM, is anticipated. The Mango Fruit dataset from Kaggle is used here. This dataset contains images of mangoes having diseases and healthy mangoes forming the five classes. Then, preprocessing images from the collected data is done by setting image dimensions to 82*82 pixels and setting the batch size to 32.

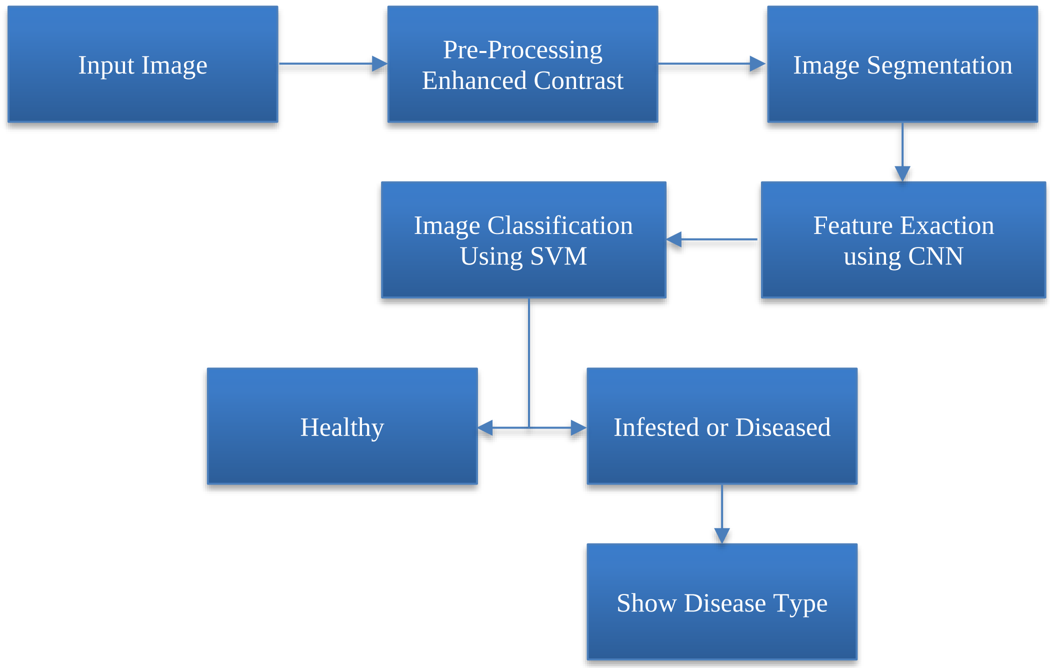


**Fig. 2 Flow diagram of anticipated methodology**

Then, data augmentation is done on the training set. All images are rescaled so that they have pixels between 0 and 1. The test data is not shuffled in order to maintain consistency in evaluation. Following feature extraction using the CNN model, SVM is employed for classification. Training, validation, and testing are the three categories into which the data set is separated in an experiment using a hybrid CNN+SVM model as part of the study. 90% of the data was used for training, 10% for validation, and 10% for testing. This allocation method's overall accuracy was quite good. The flow diagram for the suggested method is displayed in the accompanying graphic.

### *3.1. Mango Fruit Images Collection*

There are 3333 photos in total collected from the Indian Agricultural Research Institute, Karnal (Haryana, India), which are publicly accessible internet data sets. Also, a total of 350 photos were collected from the web dataset that was publicly accessible. In an effort to increase the dataset, the image augmentation approach was applied to all those insufficient images for feature extraction phases. By forming an adapted version of an existing image, an image augmentation increases the dimensions of the training data set. To create "new" photos in this experiment, the source images were rotated, cropped, zoomed in, flipped, and moved. Therefore, 4171 augmented images are included in this experiment: 200 shots, 150 annotated photos of anthracnose illness, 175 images of Alternaria mango disease, 125 images of Aspergillus disease, and 170 images of Lasiodiplodia mango make up a dataset of healthy mangos. Training, validation, and testing were conducted using different datasets.

### *3.2. Data Pre-processing*

Eliminating blur or poor-quality photos would purge the data of extraneous information. Enhanced image-data quality by removing unnecessary or non-informative noise while keeping important information. In order to save processing time and computational expenses, we also used Open CV to resize dataset photos into a standardized format of 224 × 224 pixels during the data-processing step. They were converted into a NumPy array to make working with the photographs easier for Karas.

### *3.3. Data Segmentation*

Images must be broken into smaller pieces called segments whose feature comparisons worked on as intensity, color, and textures [19]. The paper basically focuses on basic threshold-based segmentation as a fundamental input to image segmentation, under which intensity weights separate pixels with respect to any threshold value determined beforehand. For this, such segmentation is appropriate for segmenting light images from objects of the background or other images.

It is simple and fast during threshold-based segmentation. But while converting any mango image into binary using the gray scale, part of it still contains the background part matching it. At the same time, the best segmentation means complete separation of the image (background and foreground) and its full lossless reconstruction. Binary inverse thresholding techniques are used in this research. A pixel's value is adjusted to zero if it surpasses the specified threshold and to its maximum value; otherwise, it is adjusted to its maximum value in accordance with the binary inverse threshold technique.

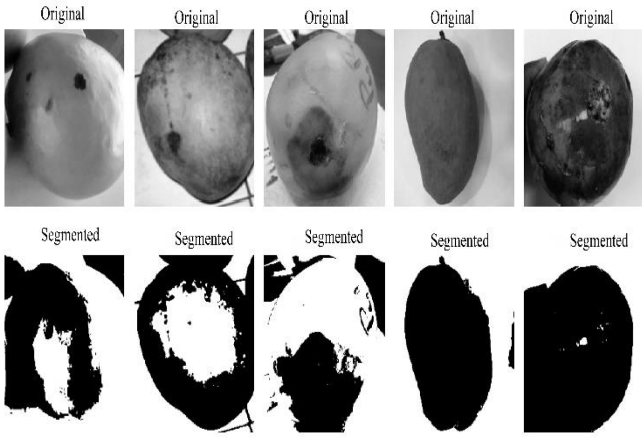


**Fig. 3 Segmented images**

### *3.4. Augmentation of Data*

Since deep learning models require a large dataset to train, this model is trained using a publicly accessible web dataset. The pre-learned model will then be re-trained using the collected dataset. Further, to enlarge the dataset and train the suggested model, data augmentation is carried out. Thus, for each collected dataset, flip and rotate at 45, 90, 135, 180, 270, and 360 degrees [20]. Here again is augmentation. This may also be referred to when training a deep learning model as transfer learning. The Kaggle data science bowl's online dataset, which provides photographs of mango fruit, will be used here. With respect to the quantity of the relevant dataset above, this is very insufficient for the intended model. Therefore, the previously mentioned data augmentation method will be applied. This is what regularly happens whenever we find a sparsity in a data set. Expand the normal size dataset and then the mango disease dataset by rotating and arcing 45, 90, 135, 180, 270, and 360 degrees out of the 6000 collections. After dividing that dataset into test, validation, and training datasets, we transfer the learned parameters using locally collected datasets. Data augmentation involves various methods, including cropping, adding noise, translating, rotating, and flipping. These methods include rotation and flipping, which are used to enhance a dataset. Others only alter the original raw dataset by adding or removing data. Rotation and cropping were hence the techniques that were selected.

### *3.5. Feature Extraction*

Convolution neural networks are used extensively to compute the input that is multidimensional in nature. Along with many characteristics by which weight computations flow during the training phases, such is the defining shape of the neural network. CNNs use these features for model training and validation. Through their three feature extraction layers, the convolutional, pooling, and completely connected ones, they learned to classify mango diseases and handle images of disorder processing. The flattened convolution and maximum pooling creations are fed into the densely connected layer with 256 neurons. The output is then linked to the dense layer, and the sequential model is finished using the convolution network layers from the Python tensor flow library's Keras API. This CNN model was developed with consideration for the following layers.

- **Pooling Layer: -** These images are sent to the max pooling layer by the pooling layer after the convolution layer. The max pooling layer determines the kernel size, stride length, window size, and kind of pooling operation. It features a 2x2 window size filter at the top pooling layer. Additionally, it facilitates downsampling, which is the process of making input pictures smaller. It reduces the number of image parameters and increases the cost of building the CNN model. The proposed method is based on regular pooling and max pooling subsampling approaches. When the pooling layer uses 2 x 2, each feature map will be processed using the 'MAX' function to treat the feature dimension.

- **Initiation: -** Numerous activation functions, including Sigmond, Soft-Max, Tanh, and ReLu, are used in convolutional neural networks. In the paper, we classify constitutional images using ReLU activation. A function known as ReLU was used to prevent and solve the gradient descent problem. ReLU-based neural network models perform better and are easier to train than models that use other activation functions, such as sigmoid or hyperbolic tangent activation functions.

- **Pool size selection: -** A 3x3 filter size characteristic has been identified to extract certain features from the image of the disease.

- **Flatten Layer: -** The Flatten layer, which comes after the max pooling layer, classifies the layer's input that is fully connected by altering it. The fully connected layer swiftly receives and processes the resultant feature map. Input images are fed into the completely linked layer after completing the convolution pooling and flattening stages. The flattened layer converts 2D data into 1D data. Finally, the images from the flattened dataset are categorized in the fully connected layer.

- **Optimizer & reduced overfitting: -** Adam optimizer is used here as it is user-friendly, efficient and effective. To decrease the overfitting of the training datasets, dropout possibilities are set to 0.2, 0.25 and 0.3 prior to the fully connected layers.

### *3.6. Classification*

SVM, one of the most used supervised machine learning methods, may be used for both degradation and cataloguing. Its main objective is to deal with binary classification issues; however, it may be modified to operate with an MCSS. SVM creates a hyperplane in a multidimensional space to differentiate between the various classes. SVM generated the optimal hyperplane repeatedly to reduce cataloguing errors. One of the most popular supervised machine learning techniques, SVM, may be applied to both cataloguing and degrading. Although it may be adjusted to work with an MCSS, its primary goal is to address binary classification problems. In a multidimensional space, SVM constructs a hyperplane to distinguish between the different classes. To lower cataloguing mistakes, SVM repeatedly produced the ideal hyperplane.

## 4. Results & Discussion

Because a lot of data is easily accessible for training and testing procedures that would allow it to apply holdout validation, the approach was used for the study instead of cross-validation. After training models were finished, models based on CNN and SVM were assessed based on how well they performed on the testing dataset. The modes' performance was evaluated using several widely used criteria, including [21] F1 score, sensitivity (recall), accuracy, and precision. Accuracy may be defined as either the model's overall prediction precision or the training (validation) data's classification accuracy. Using a confusion matrix, the effectiveness estimate was performed, and the results included true positives, true negatives, false positives, and false negatives.

### *4.1. Convolutional Neural Network Model Evaluation*

The CNN model extracts the features of mango diseases, and an improvement in the accuracy is also found, which is 85%. In order to enhance data quality, eliminate noise, and expand datasets, image processing, image segmentation, and augmentation approaches are used for the pictures before using the CNN technique to construct the model. Following the CNN data processing, the batch size=32 was used, and Adam optimizer and ReLU activation were further applied. The features extracted were measured using performance parameters. Table 2 shows the achievement of 85% accuracy when using CNN for feature extraction. After feature extraction, the SVM classifier is used to classify all sets of 5 diseased and healthy classes. The hybrid approach of using SVM classification after the CNN model yielded better results. After conducting experiments, 99% accuracy was obtained, which shows that the combination of CNN and SVM constructed was effective and efficient in detecting diseases in mango fruits.

**Table 2. Feature extraction using CNN model**

| Disease | Precision | Recall | F1-Score |
|---|---|---|---|
| Alternaria | 0.86 | 0.7 | 0.77 |
| Anthracnose | 0.89 | 0.71 | 0.79 |
| Aspergillus | 0.74 | 0.9 | 0.81 |
| Healthy | 0.91 | 0.95 | 0.93 |
| Lasiodiplodia | 0.9 | 0.93 | 0.91 |
| Accuracy | | | 0.85 |
| Macro Avg | 0.86 | 0.84 | 0.84 |
| Weight Avg | 0.86 | 0.85 | 0.84 |

**Table 3. Disease classification using SVM model**

| Classification Report | Precision | Recall | F1-Score |
|---|---|---|---|
| 0 | 0.99 | 0.99 | 0.99 |
| 1 | 0.99 | 0.98 | 0.99 |
| 2 | 0.98 | 0.99 | 0.99 |
| 3 | 0.98 | 0.99 | 0.98 |
| 4 | 0.99 | 0.99 | 0.99 |
| Accuracy | | 0.99 | |
| Macro Avg | 0.99 | 0.99 | 0.99 |
| Weighted Avg | 0.99 | 0.99 | 0.99 |

Table 3 shows the classification report and achievement of 99% accuracy on performance parameters when using SVM for classification.

### *4.2. Visualization & Evaluation*

By visualizing the trained model, issues like overfitting and underfitting can be identified. The following confusion matrix provides a detailed explanation of the model's performance in different classes. The assessments listed below are essential for comprehending the model's advantages and disadvantages in terms of categorizing various mango illnesses.

- Plotted a confusion matrix.
- Training loss and accuracy graph over epochs.
- Validation loss and accuracy graph over epochs.
- Comparisons with other existing pre-trained models are carried out.

**Table 4. SVM confusion matrix**

| | | | | | | |
|---|---|---|---|---|---|---|
| True Label | 0 | 654 | 1 | 2 | 3 | 1 |
| | 1 | 2 | 503 | 3 | 3 | 3 |
| | 2 | 1 | 0 | 719 | 6 | 0 |
| | 3 | 1 | 3 | 4 | 790 | 3 |
| | 4 | 1 | 0 | 5 | 3 | 622 |
| | | 0 | 1 | 2 | 3 | 4 |
| | | Predicted Label | | | | |

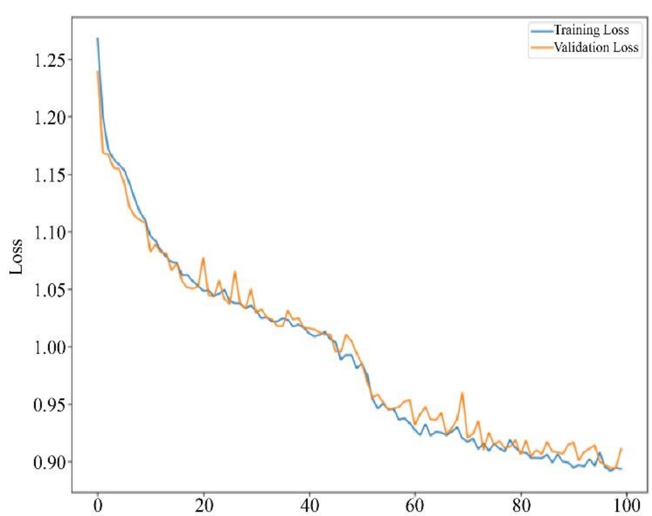


**Fig. 4 CNN performance graph for loss**

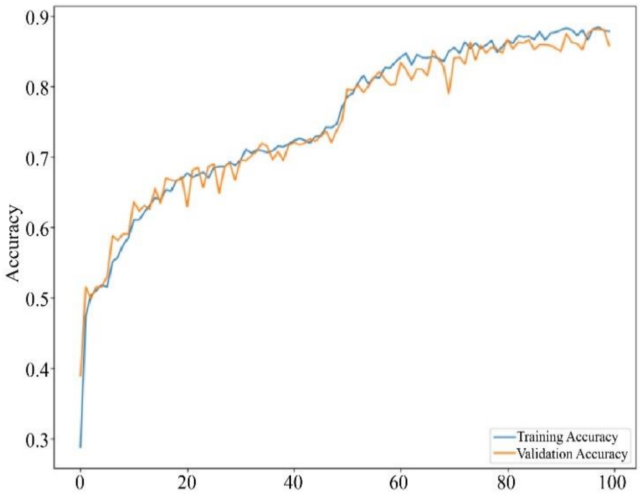


**Fig. 5 CNN Performance graph for accuracy**

**Table 5. Comparative analysis of performance measures in various pre-trained models and proposed CNN+SVM model**

| Performance Measurements | Model | | | | |
|---|---|---|---|---|---|
| | CNN-HOG | CNN | L-CNN | CNN-FOA | CNN+SVM |
| Recall (%) | 82 | 79 | 69 | 94.3 | 99 |
| Precision (%) | 92 | 93 | 78 | 95 | 99 |
| Accuracy (%) | 95 | 82 | 86 | 97 | 99 |
| F1-Score | 68 | 63 | 57 | 94.2 | 99 |
| ROC AUC Score (%) | 73 | 84 | 83 | 98 | 99 |
| MCC (%) | 81 | 79 | 79 | 90 | 99 |

After the suggested model was implemented, it was necessary to evaluate its success by comparing the outcomes with those of other models. The findings were compiled and developed in the table 5 after the suggested models were compared to the CNN+HOG, CNN, L-CNN, and CNN-FOA models [22]. The findings demonstrate that the suggested model works better in terms of accuracy, recall, and precision. As a result, it was proposed that this model is the most accurate in forecasting mango illnesses.

## 5. Conclusion & Future Work

The mangos are highly prized agricultural products that are traded widely around the world. The assessment of mango fruit quality has conventionally been carried out by manual techniques, entailing a substantial time and labour commitment. Numerous non-destructive techniques, such as inspecting the fruit inside out, have been developed to solve these problems.

In this paper, we developed a deep learning-based CNN+SVM model to detect mango diseases into 5 classes that are Rot of Alternaria, Anthracnose, Aspergillus, Lasiodiplodia and Healthy and computed the accuracy. Further, the same is compared with other existing models, and significant positive results were found. Our model combines the power of CNN for feature extraction with traditional machine learning (SVM) for classification. With this approach, a potentially more robust and accurate disease detection model is developed. Future work could involve experimenting with different CNN architectures, exploring other SVM kernels, or incorporating additional data preprocessing techniques to further improve the model's performance.

## References


[1] Charles Samuel Mutengwa, Pearson Mnkeni, and Aleck Kondwakwenda, "Climate-Smart Agriculture and Food Security in Southern Africa: A Review of the Vulnerability of Smallholder Agriculture and Food Security to Climate Change," *Sustainability*, vol. 15, no. 4, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Subhasis Bhadra, and Allen R. Dyer, *Resilience and Well-Being among the Survivors of Natural Disasters and Conflicts*, Handbook of Health and Well-Being, Springer Nature, pp. 637-667, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Sulogna Chakma et al., "Adapting Land Degradation and Enhancing Ethnic Livelihood Security through Fruit Production: Evidence from Hilly Areas of Bangladesh," *Agro-Biodiversity and Agri-Ecosystem Management*, Singapore: Springer, pp. 217-238, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] V. Laxmi, and R. Roopalakshmi, "Artificially Ripened Mango Fruit Prediction System Using Convolutional Neural Network," *Intelligent Systems and Sustainable Computing: Intelligent Systems and Sustainable Computing*, Singapore, pp. 345-356, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Rabia Saleem et al., "Mango Leaf Disease Identification Using Fully Resolution Convolutional Network," *Computers, Materials and Continua*, vol. 69, no. 3, pp. 3581-3601, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] R. Nithya et al., "Computer Vision System for Mango Fruit Defect Detection Using Deep Convolutional Neural Network," *Foods*, vol. 11, no. 21, pp. 1-11, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Omneya Attallah, "Tomato Leaf Disease Classification via Compact Convolutional Neural Networks with Transfer Learning and Feature Selection," *Horticulturae*, vol. 9, no. 2, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Anil Bhujel et al., "A Lightweight Attention-Based Convolutional Neural Networks for Tomato Leaf Disease Classification," *Agriculture*, vol. 12, no. 2, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Prabhjot Kaur et al., "Recognition of Leaf Disease Using Hybrid Convolutional Neural Network by Applying Feature Reduction," *Sensors*, vol. 22, no. 2, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Umit Atila et al., "Plant Leaf Disease Classification Using Efficient Net Deep Learning Model," *Ecological Informatics*, vol. 61, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Sunayana Arya, and Rajeev Singh, "A Comparative Study of CNN and AlexNetfor Detection of Disease in Potato and Mango Leaf," *International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, Ghaziabad, India, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] S. Wongsila, P. Chantrasri, and P. Sureephong, "Machine Learning Algorithm Development for Detection of Mango Infected by Anthracnose Disease," *Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering*, Cha-am, Thailand, pp. 249-252, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Soumia Bensaadi, and Ahmed Louchene, "Low-Cost Convolutional Neural Network for Tomato Plant Diseases Classifiation," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 1, pp. 162-170, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Rabbia Mahum et al., "A Novel Framework for Potato Leaf Disease Detection Using an Efficient Deep Learning Model," *Human and Ecological Risk Assessment: An International Journal*, vol. 29, no. 2, pp. 303-326, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Jinzhu Lu, Lijuan Tan, and Huanyu Jiang, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," *Agriculture*, vol. 11, no. 8, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Uday Pratap Singh et al., "Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease," *IEEE Access*, vol. 7, pp. 43721-43729, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Maha Altalak, "Smart Agriculture Applications Using Deep Learning Technologies: A Survey," *Applied Sciences*, vol. 12, no. 12, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Vani Ashok, and D.S. Vinod, "A Comparative Study of Feature Extraction Methods in Defect Classification of Mangoes Using Neural Network," *2nd International Conference on Cognitive Computing and Information Processing*, pp. 1-6, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[19] Siddharth Singh Chouhan, Uday Pratap Singh, and Sanjeev Jain, "Web Facilitated Anthracnose Disease Segmentation from the Leaf of Mango Tree Using Radial Basis Function (RBF) Neural Network," *Wireless Personal Communcations*, 113, pp. 1279-1296, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Aayush Sharma, Harjeet Kaur, and Deepak Prashar, "Generative Adversarial Networks-Based Approach For Data Augmentation In Mango Leaf Disease Detection System," *International Conference on Communication Systems and Network Technologies (CSNT)*, India, pp. 816-821, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Byeongjun Min et al., "Data Augmentation Method for Plant Leaf Disease Recognition," *Applied Sciences*, vol. 13, no. 3, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Ramalingam Kalaivani, and Arunachalam Saravanan, "A CONV-EGBDNN Model for the Classification and Detection of Mango Diseases on Diseased Mango Images utilizing Transfer Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14349-14354, 2024. [CrossRef] [Google Scholar] [Publisher Link]