Pomegranate Disease Detection Using CNN-LSTM Hybrid Model

Sahebgouda[1], Sumana Maradithaya[2] and Ambaji Jadhav[3]
[1]Department of Computer Science and Engineering, BLDEA's V.P.Dr.P.G.Halakatti College of Engineering and Technology, vijayapur
[2] Department of Computer Science and Engineering, M.S. Ramaiah Institute of Technology, Bengaluru,Karnataka
[3]Department of Electrical and Electronics Engineering, BLDEA's V.P.Dr.P.G.Halakatti College of Engineering and Technology, vijayapur

Keywords:

Pomegranate
Disease detection
Deep learning
CNN, LSTM
Machine learning

ABSTRACT

Pomegranate is a high-value fruit globally recognized for its nutritional benefits and applications in traditional medicine and cosmetics. India is a key player in the global pomegranate market, but the industry faces challenges such as diseases that affect crop productivity and economic losses for farmers. This study proposes a novel approach to pomegranate disease detection using a hybrid Convolutional Neural Network (CNN) and Long short- Term Memory (LSTM) model. The proposed model leverages CNNs for effective feature extraction and LSTMs for sequential data handling, achieving superior performance compared to traditional methods and other deep learning techniques. Experimental results demonstrate high accuracy, recall, precision, and F1 score. The Proposed model achieved an accuracy of 98.53% and loss of 0.0677. The study also explores the limitations of transfer learning approaches such as VGG16 and ResNet50, and larger models like AlexNet, which did not perform well in this context. The findings suggest that the hybrid CNN-LSTM model offers a scalable and adaptable solution for agricultural disease detection, with potential applications for various crops.

Corresponding Author:

Sahebgouda
Department of Computer Science and Engineering, BLDEA's V.P.Dr.P.G.Halakatti College of Engineering and Technology, Ashrama Road Vijayapur-586101,Karnataka,India

1. INTRODUCTION

Pomegranate (Punica granatum) is one of the most valuable fruits available. It is globally recognized as a "Super-food" owing to its nutritious characteristics [1]. Pomegranate fruit is rich in vitamins and dietary fiber [2]. Beyond its use as a food, pomegranate has all found many applications in various sectors, from traditional medicine to cosmetics. The fruit's extracts are utilized in the pharmaceutical industry for their potential in treating various medical conditions, such as cardiovascular diseases and cancer [3]. Moreover, the aesthetic and therapeutic properties of pomegranate make it a good ingredient in skincare and beauty products.

India is one of the leading countries in pomegranate production [4]. Contributing significantly to the global pomegranate market. India provides a suitable environment for pomegranate cultivation. However, the thriving pomegranate industry faces many formidable challenges like drought, climate change, etc. With diseases being a major factor contributing to yield losses. Pomegranate orchards are susceptible to various diseases caused by fungi, bacteria, and viruses. These diseases not only affect the fruit's quality but also pose a threat to the overall health of the trees, potentially leading to long-term economic losses for farmers. One of the major disease is the Bacterial blight, which affects all parts of the plant but is more destructive on fruits [5]. Other common diseases include Alternaria, Anthracnose, etc. affecting both plant and fruit

The Traditional methods of disease detection in agriculture rely on visual inspection and manual monitoring which are inherently time-consuming and often lack the precision required for early disease detection. Advent of Machine learning techniques, provides us an opportunity to implement the new approaches to disease detection in pomegranate cultivation. These new machine learning techniques enable us to build automation disease detection systems which are more efficient and accurate. Deep Learning techniques, which a subset of machine learning, is becoming increasingly popular in research as graphics processing technology advances and low-cost GPUs become more widely available.

This research proposes a hybrid model i.e. combining CNN with LSTM layer. This paper also compares other state-of-the-art models to the proposed model.

Fig. 1 shows the diseased pomegranate fruits with corresponding disease name below it.

Figure 1. Various diseases affecting Pomegranate

Correct Identification of diseases at early stages can help farmers to mitigate the losses. The proposed system in this project can be used by farmers for accurate detection of diseases.

2. LITERATURE SURVEY

Several studies have explored the use of machine learning and image processing techniques for pomegranate disease detection and classification. This section explores the findings from the literature survey, which summarizes information from more than 20 research publications.

These studies have used a variety of traditional machine learning techniques and newer deep learning techniques. This section is divided into three subsections each following a theme of machine learning like traditional techniques, deep learning techniques and emerging technologies.

2.1. Traditional Machine Learning Techinques

Earlier attempts to detect pomegranate disease relied heavily on traditional machine learning techniques such as KNN, k-means, and SVM. For example, Tejal Deshpande et al. proposed a system that uses k-means clustering for image segmentation and grades infected areas based on color and size. Color and size of the infected area are not reliable indicators of disease [6]. Their system can only detect healthy and bacterial blight. This limits their system's ability to perform multiclass classification tasks.

Several studies used K-means clustering for image segmentation, followed by SVM for classification. Manisha Bhangea and H.A. Hingoliwala proposed a system that uses K-means and SVM to classify pomegranate diseases with 82% accuracy[7]. Sulakshana A. Gaikwad et al. applied similar techniques for multi-fruit disease detection. Their system showed remarkable accuracy [8]. However, SVM is not suitable for larger datasets as it takes a longer time to train. Additionally, higher-dimensional data like images cannot be directly used with the SVM. This necessitates a need for feature extraction. Moreover, SVM fails when data is not linearly separable.

In their study, Pooja Kantale and Shubhada Thakare utilized the Ada-Boost Ensemble algorithm to classify pomegranate diseases, leveraging a dataset of 190 images. Their model achieved an impressive accuracy of 92.90% [9]. However, despite its high accuracy, the Ada-Boost algorithm has some drawbacks, such as its sensitivity to noise and high computational demands, especially when training on large datasets. These limitations can hinder its practical application in various agricultural scenarios.

Neural networks have also been applied in pomegranate disease classification. Mrunmayee Dhakate & Ingole A. B. developed a system that uses Artificial Neural Network as a classifier and showed an accuracy around 90% [10] with a dataset of 500 images. Miss. Kshamarani Purvimath & Dr.Pushpa B.Patil have focused on PNN(Probabilistic neural network) for disease classification and achieved an accuracy of 80.30% [11] with a dataset of 166 images, which is divided into 100 for training and 66 or testing. These systems are in their early stages and require image preprocessing, such as noise reduction to be effective.

Otsu Threshold combined with k-means can be used for image segmentation before Image Classification. This approach was used by Rashmi Pawar & Ambaji Jadhav along with multi layer neural network, this system achieved an accuracy of 90% on 40 test images [12]. The study was limited to a small dataset. It suffers similar drawbacks as discussed above.

Authors Reddy et al. developed an ANN based approach for detecting diseases in grapes, apple and pomegranates. Their system used k-means clustering technique for image segmentation and color texture analysis to extract features using vectors like color, morphology, texture, and structure of holes on the fruit. The system achieved 93% accuracy [13]. The system uses color texture analysis for feature extraction and employs k-means clustering to segment an image. However, the authors did not provide details on the ANN architecture and hyperparameters.

SVM (Support Vector Machine) method is mostly used for binary classification but newer multiclass SVM has been developed and used in this field. Authors M. Venu Madhavan, et. al used K-means segmentation to extract region of interest(ROI) with multiclass SVM for classification. A simple histogram equalization is applied on extracted ROI to enhance the features. These features are used by Multiclass SVM to classify the leaf disease. Using this approach the authors achieved 98.07% accuracy [14].

2.2. Deep Learning Methods

With the advent of deep learning, field of image processing has significantly elevated the accuracy and efficiency of disease detection.

Authors Sharath D M et al. have proposed a system for detecting diseases in fruits like pomegranate, orange, grapes, and papaya using convolutional neural networks (CNN) [15]. Their approach involves image acquisition, preprocessing (resizing, denoising), segmentation using GrabCut algorithm, morphological processing and feature extraction. A CNN model is then trained on extracted features from the training dataset containing 12,891 images across various fruit diseases. The system achieved around 91% accuracy in detecting diseases like citrus canker, greening, etc in grapes, and bacterial blight, borers, cercospora in pomegranates.

Authors Shital Pawar et al. have proposed a leaf disease system for multiple plants using 15 layer CNN model [16]. Their system could detect diseases in 10 different plants with upto 93% accuracy. The authors did not report specific accuracy metrics for each plant species, making it difficult to assess the system's overall performance.

A newer more complex variant of CNN called Alexnet. AlexNet architecture has 8 layers, this model was employed by Authors Prashant B. Wakhare et al. and their model achieved an accuracy of 97.6% [17]. They have used a dataset of 1245 pomegranate leaf images. Their research was limited to detecting only two diseases which are bacterial blight and Alternaria.

An advanced region based CNN model called faster-RCNN was used in pomegranate disease, detection and classification in a research by authors Aziz Makandar & Syeda Javeriya [18]. Faster R-CNN is 250 times faster than RCNN. The model was designed to identify two common diseases—anthracnose and bacterial blight—using a custom dataset of annotated images. The authors acknowledged that low image resolution could lead to detection failures.

2.3. Emerging Technologies

Newer methods like transfer learning have emerged. In transfer learning, the pretrained models are utilized to fine tune on the pomegranate disease dataset. Commonly used pretrained models are VGG16, Inception-V3, Resnet, DenseNet. These models are the variants of CNN.

Authors Vaishali Nirgude & Sheetal Rathi in their work compared three cnn based architectures they are Resnet-18, Resnet-50 and Inception-V3 [19]. The authors collected 1,493 images of the pomegranate fruits and the leaves at different stages of disease development over six months. They augmented the dataset and evaluated the performance of these CNN architectures on the dataset. Authors used random cropping, random zooming, changing contrast, changing brightness, background removal, and noise addition in augmentation process. These models achieved 95.26%, 98.37% and 97% accuracy respectively [19]. The authors acknowledged the need to address the "Black Box" problem often associated with deep learning models. This could be handled by merging the Grad-CAM model. The authors did not provide specific implementation details to recreate the experiment.

Mohammed Saleh Al Ansari [35] used deep learning techniques such as AlexNet and VGG-16 to demonstrate the ability to automatically diagnose diseases in pomegranate plants using leaf images. Preprocessing techniques such as resizing and normalization were applied to the raw images. Both models were trained and tested on the preprocessed leaf image dataset. AlexNet achieved an accuracy of 89.57%, while VGG-16 achieved a higher accuracy of 95.23%. However, there are notable drawbacks to this study. The relatively small dataset of 559 images may not capture the full diversity of pomegranate leaf diseases potentially limiting the generalizability of the models. Additionally, the models were trained for only 12 epochs, which might not be sufficient for them to fully learn the features necessary for accurate disease diagnosis.

With advancement in image processing and deep learning methods new models were constructed which improved the performance of the disease detection systems. Performance of these deep learning techniques requires high quality data. Advanced image processing techniques can be applied to improve quality.

For example a research paper by Yan Qi et al., used MSRCR defogging algorithm and image normalization techniques to enhance the images. They used Canny SLIC algorithm for precise segmentation of diseased region. A focal loss function was used to improve the DenseNet169 architecture, resulting in an recognition accuracy of 98.98% for three types of grape, outperforming traditional DenseNet's 97.95% accuracy [21]. Additionally, the computational complexity and processing time associated with the proposed model are not discussed, which could be crucial for real-time applications.

Use of CNN in combination with other deep learning techniques is also becoming popular. These combinations of methods can improve the performance of models.

Authors M T Vasumathi et al. proposed an architecture which combined CNN-LSTM model to classify a set of about 6519 pomegranate fruit images [22]. Features like fruit color, number of spots, shape, size of diseased area, etc where extracted manually into a csv dataset. Their model achieved 98.17% on this dataset. The binary classification of normal or abnormal fruit limits its practical use.

Authors M T Vasumathi et al. in [23] extended their work to a multi-class classification using similar hybrid CNN-LSTM model and used dragonfly optimization algorithm for weight optimization. The same csv dataset was used in training of their model and achieved accuracy of 97.1%. Extracting the features manually is a time consuming task, this also limits its use in automated disease detection.

In the study by authors Naseer, Aisha et al. [24] proposed a model to detect growth stages of pomegranate fruit. They utilized a dataset consisting of 5,857 images classified into five growth stages: Bud stage, Early-Fruit stage, Flower stage, Mid-growth stage and Ripe stage. The research also employed a novel feature engineering approach which is based on transfer learning. Authors introduced a transfer learning based CRnet based feature engineering to enhance the performance of the detection. CRnet is the hybrid of Convolutional Neural Network (CNN) and Random Forest (RF). CNN is used for extraction of spatial features from the images then these features were then inputted into the Random Forest (RF) model, creating a new probabilistic featured set that improved the detection accuracy of the pomegranate growth stages. The study utilized Synthetic Minority Over-sampling Technique (SMOTE) is used to address the class imbalance in the dataset. SMOTE generates synthetic samples for the underrepresented classes to balance the dataset, which enhanced its ability to generalize across different growth stages. While the proposed method achieved a high accuracy of 98%, it also faced some drawbacks. The computational complexity of the RF model was higher compared to other techniques, which could affect scalability and efficiency in large-scale applications.

Authors M. D. Nirmal et al. studied supervised and unsupervised machine learning approaches for detecting diseases in pomegranate leaves [25]. Their method utilized the K-means, an unsupervised technique, to isolate the diseased regions of the leaves. These segmented regions were then classified using supervised learning models, ResNet and MobileNet. The MobileNet model achieved a superior accuracy of 98.18%, while ResNet achieved 95.53% on the Mendeley database. Both ResNet and MobileNet are computationally intensive models, requiring substantial resources for training, especially on large datasets. The segmentation of each image adds an additional layer of complexity, further increasing the computational burden. This high resource demand limits the scalability of the approach, making it less feasible for deployment on large, diverse datasets or in resource-constrained environments typical of many agricultural settings.

The literature survey is summarized in Table I.

Table 1. Literature Survey

Ref	Part	Feature Extraction	Classifer	Accuracy	Critical Analysis
[6]	Both leaf and fruit	K- means clustering	Classification based on color and area	Not Mentioned	This system calculates the percentage of infection based on the size of the infected area. This is better than manual techniques but has lower accuracy.
[7]	Fruit	K- means clustering	SVM	82	This paper proposes a web based tool which uses Image processing techniques to determine if a pomegranate fruit is infected with bacterial blight or not. The svm used here is not suitable for multiclass classification.
[8]	Fruit	K- means clustering	SVM	Not Mentioned	check
[9]	Both leaf and fruit	GLCM, PSO	Ada-Boost Ensemble	92.9	This study used a very small dataset of 190 images of both fruit and leaf. There may be a chance of over fitting.
[10]	Both leaf and fruit	GLCM	Neural Network	90	This paper constructs a classifier from basic Artificial Neural Network(ANN) and uses the basic back-propagation algorithm to train ANN. 500 images of leaf and fruit were used as a dataset.
[11]	Fruit	Standard deviation, entropy, variance, smoothness, skewness, kurtosis, contrast	Probabilistic neural network(PNN)	80.3	PNN is a special class of ANN. This network is not widely used because they are slower than other methods.
[12]	leaf	Otsu thresholding, K means clustering	Multi layer neural network	90	The paper does not provide details of the model.

[13]	Multiple fruits including pomegranate	ANN	ANN	90	Accuracy of model can be improved by better preprocessing and restricting to classification of single fruit type
[14]	leaf	Otsu thresholding, K means clustering	Multi-Class SVM	98.07	Simple histogram equalization is applied on extracted ROI to enhance the features. Those features are used by Multiclass SVM to classify the leaf disease.
[15]	Multiple fruits including pomegranate	ANN	CNN	91	Authors used both leaf and fruit images. Dimensions of images were limited to 432x288 to reduce processing time. Morphological processing is also used to clean the image.
[16]	Leaf	K- means clustering	CNN, SVM	96.93, 96	This paper compared various segmentation techniques such as manual segmentation, k-means etc with CNN or SVM as classifiers. Authors found that manual segmentation with CNN as classifier gives highest accuracy.
[17]	Both leaf and fruit	Grabcut Segmentation and canny edge detection	CNN	93	Authors trained a CNN with 15 layers to detect diseases in 10 different plants with a dataset size of 50000 images. Their system also suggests pesticide for the detected disease.
[18]	Leaf	None	Alexnet	98.07	Authors of this paper developed a disease detection system for leaves using Alexnet and achieved high accuracy. Authors also compared this model with other CNN variants like resnet50, vgg16, etc and concluded that alexnet performed better.
[19]	Fruit	Statistical Analysis and DWT	Multiple ResNet50, ResNet18, InceptionV3	87.5, 97.92, 78.75	This study compared 3 CNN based models. These models were pretrained on the ImageNet dataset. Authors have mentioned the "Black Box" problem of Deep Learning models, which can be solved by integrating these models with the Grad-CAM model.
[20]	Leaves of multiple plants	None	Faster-RCNN	Not Mentioned	This Paper uses a variant of RCNN called Faster-RCNN. Authors collected and annotated the images. They also provided the detailed architecture of the model. Drawback is they have not provided proper metrics.
[21]	Fruit	None	Multi-Scale Improved DenseNet	98.98	Authors used the MSRCR based defogging algorithm and Canny SLIC algorithm for image segmentation. Authors improved the DenseNet algorithm by using Focal loss function, focuses more on processing samples, that are difficult or hard for classification in training process
[22]	Leaf	None	CNN-LSTM	98.19	Authors combined two deep learning techniques CNN and LSTM to construct a binary classifier. This model has an accuracy of 92.9%. Authors used a dataset of 6519 images. This system can be further developed to multi-class classification.
[23]	Fruit	None	CNN-LSTM	97.1	Authors proposed a system that uses a combination of various deep learning techniques- CNN, LSTM. Authors also used the dragon fly algorithm for optimization. This system was trained on a dataset of 6500 images for multiclass classification.

2.4. Research Gap

The literature that has been examined shows important progress in finding and classifying pomegranate diseases by using different machine learning and deep learning methods. However, even with these promising findings, there are still several gaps in research. Many studies have used relatively small datasets, which limits their ability to generalize; this leads to overfitting issues. Because of this, larger and more varied datasets are necessary to improve the strength and effectiveness of the models. These methods also need complex preprocessing techniques. Without proper preprocessing, models tend to give poor accuracy and often take longer to train. Significant preprocessing, which includes cropping, noise reduction and background removal, allows models to concentrate on the important parts of the fruit instead of the distracting noise found in images. Although these techniques are essential, they can also make the process more complicated.

Numerous studies have utilized single deep learning methods, which can be very effective; however, they are also time-consuming to implement and train. This is especially true for large models, like AlexNet and ResNet. Combining different deep learning techniques, such as CNN (Convolutional Neural Networks) and LSTM (Long Short-Term Memory), can potentially decrease both model size and training time while improving performance. Future research should focus on creating models that are resilient to variations (this is important). Achieving this could involve training models on datasets that cover a wide range of environmental conditions or using techniques like data augmentation to mimic these conditions during training. Although some models provide only binary classification, this limitation restricts their usefulness in situations where multi-class classification is more suitable.

3. PROPOSED METHODOLOGY

This section discusses the preprocessing of the dataset, developing the model and training the model. The below Fig. 5 shows the Architecture of Proposed Solution. For creating model and training model python programming language and many advanced libraries were leveraged.

4.1. Dataset Preprocessing

The Authors B, Pakruddin R and Dr. Hemavathy have created a standard dataset for training deep learning models [26] which is publicly available. This dataset contains 5099 images divided into 5 classes. The classes are Alternaria, Anthracnose, Bacterial Blight, Cercospora, Healthy. Below Table 2 shows the distribution of images in each class.

Table 2. Mendeley Dataset Details

Sl. No	Class Name	Image Quantity
1	Alternaria	886
2	Anthracnose	1166
3	Bacterial Blight	966
4	Cercospora	631
5	Healthy	1450

4.1.1. Resizing

Fortunately the dataset has all the images in square format. Images are of 3120x3120 pixels. The images were reduced to 512x512 pixels.

4.1.2. Background Removal

Images have fruit in the center and leaves in the background. Removal of background can help the model to focus on the fruit rather than background. It also helps to speed up the machine learning. Fig. 2 shows the output of the process. For Removing background rembg library was used [27].

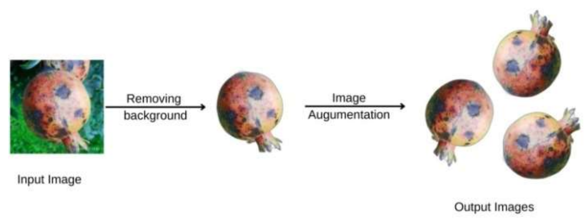

Figure 2. Background removal and Augmentation.

4.1.3. Augmentation

Computational techniques like scaling, rotating, shifting were used to generalize the model and reduce overfitting. The Table 3 shows the parameters used in augmentation.

Table 3. Augmentation Parameters

Sl. No	Parameter	Value
1	Rotation range	40
2	Width shift	0.2
3	Height shift	0.2
4	Zoom range	0.2
5	Horizontal flip	True
6	Fill mode	Nearest

4.2. Proposed Model

This paper proposes a CNN+LSTM model which is a hybrid model. In this suggested architecture, convolutional neural networks (CNNs) are used here to extract features from input image samples, while long short- term memory networks (LSTMs) utilize these features to model the temporal connections between the data points. This combined approach allows the model to effectively capture both local patterns, such as edges and corners, and global patterns, such as objects and scenes. Moreover, by incorporating LSTMs, the model is capable of understanding the temporal dynamics and dependencies that may exist between different parts of the image over time. Fig. 5 shows the system architecture of proposed model.

This integrated methodology enables the model to comprehensively understand the data and produce highly accurate predictions. The model can provide precise and reliable classification results by combining

CNNs' spatial feature extraction strengths with LSTMs' temporal modeling capabilities. Convolutional neural network (CNN) and a long short- term memory (LSTM) network are briefly described as follows.

4.2.1. Convolutional neural network

A simple neural network often lacks the capability to learn complex features required for sophisticated image recognition tasks such as disease classification. Deep learning architectures, such as Convolutional Neural Networks (CNNs), are more suitable for these tasks [28]. CNNs are a specific type of multilayer perceptron network that operate on the principle that complex features can be constructed by refining lower-level features at each successive layer. A CNN typically consists of multiple convolutional layers, multiple pooling layers, and fully connected (FC) layers. An example of a CNN architecture with its layers is illustrated in Fig 3.

The convolutional layer defines a set of kernels or filters that identify specific features from the input images. These features are represented as feature maps, also known as activation maps. Convolution operations apply these filters across the input using a parameter called "stride" [29].

Mathematically, the convolution operation is expressed as (1).

$$F(i,j) = (I * K)(i,j) = \sum\sum I(i+m, j+n)K(m,n) \tag{1}$$

where i here is the input matrix, K is a 2D kernel (filter) given by of size $m \times n$, and F is the resulting 2D feature map from the convolution operation denoted by $I * K$.

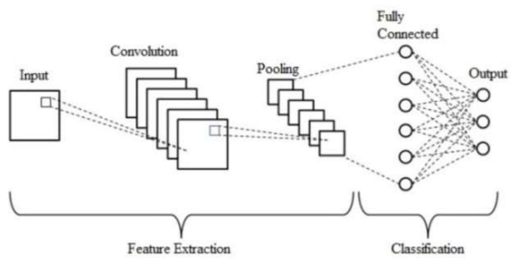

Figure 3. A typical architecture of the CNN

To introduce non-linearity into the feature maps, a Rectified Linear Unit (ReLU) layer is used. The ReLU activation function computes the output by applying a threshold at zero, which can be mathematically represented as (2).

$$f(x) = max(0, x) \tag{2}$$

The pooling layer, often following convolutional layers, performs downsampling of the input dimensions to reduce the number of parameters and computational load. Max pooling is one of the most common pooling technique used, where the maximum value within a defined region of the input is selected to form the pooled feature map [29]. Finally, the fully connected (FC) layer acts as a classifier, making decisions based on features extracted by the convolutional and pooling layers.

4.2.2. Long Short-Term Memory

Long Short-Term Memory (LSTM) networks are an advanced type of Recurrent Neural Network (RNN) designed to address the problem vanishing and exploding gradients that are associated with traditional RNNs [30]. LSTMs introduce memory blocks that replace conventional RNN units, allowing the network to retain long-term states and effectively connect previous information to current data. An LSTM network includes three primary gates: the input gate, forget gate, and output gate, which control the flow of information within the network. Fig. 4 shows the internal architecture of LSTM network, depicting how these gates function.

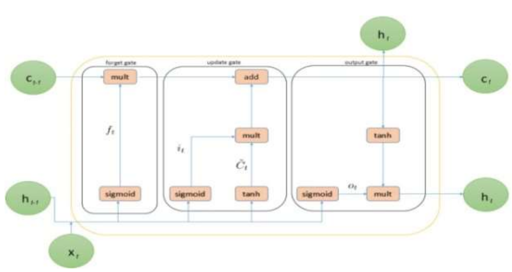

Figure 4. The internal structure of Long short-term memory

The input gate determines which part of the input to be added to the cell state. This process is defined by the following equations (3), (4), (5).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$
$$\tilde{C}_t = tanh(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$
$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{5}$$

Here, i_t is calculated by passing the previous hidden state h_{t-1} and the current input x_t through a sigmoid layer to decide which information to add. The new candidate cell state $\tilde{C}_t$ is obtained by passing h_{t-1} and x_t through a $tanh$ layer. The current cell state C_t is a combination of the old cell state C_{t-1}, modified by the forget gate ft and the new candidate cell state $\tilde{C}_t$.

The forget gate allows selective information passage using a sigmoid function, which can be expressed as (6).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{6}$$

The output gate determines the necessary states for continuation using the inputs h_{t-1} and x_t, as shown in the following equations (7) and (8).

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$
$$h_t = O_t tanh(C_t) \tag{8}$$

Where O_t is the output gate's activation, and h_t is the final hidden state output. W_i, W_f, and W_o represent weight matrices, while b_i, b_f, and b_o are the biases associated with the respective gates.

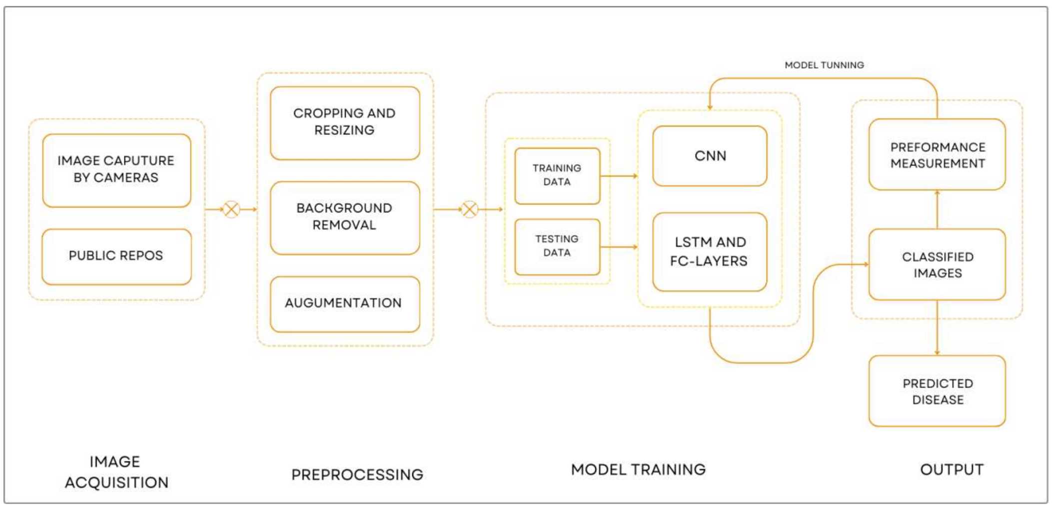

Figure 5. System Architecture of Proposed Model

4.2.3. Combined CNN-LSTM Model

The convolutional layers here are used to extract spatial features from images. Each layer applies a set of kernels (filters) to the input image, performing an element wise multiplications and summations. This operation is called as convolution which highlights specific features in the input image. The first convolutional layer (Conv2D) is having 32 kernels with a kernel size of 3x3 which uses the ReLU activation function. This layer processes the input images of shape (256, 256, 3). All subsequent convolutional layers all use ReLU activation function with each layer doubling the number of kernels.

Number of kernels used in each convolution layer is given in Table 4.

After the convolutional layers, the model incorporates a Long Short-Term Memory (LSTM) layer which captures temporal dependencies and sequence information. This information is useful for understanding the progression of disease symptoms over sequences of image features. Proposed model uses 128 LSTM units, the input to this layer is reshaped by a reshape layer. The reshape layer preceding the LSTM layer ensures that the input is in the correct format, transforming the 2D feature maps into a tensor suitable for LSTM processing. This transformation is critical for leveraging the LSTM's ability to process sequential data, as it enables the model to understand the progression of disease symptoms over a series of images.

The dense layers, also called as Fully Connected (FC) layers, perform the final classification based on the features extracted and processed by the preceding layers. The first dense layer in this model has 1024 units with ReLU activation function, which will introduce non-linearity. The subsequent dropout layer has a dropout value of 0.2 which helps in preventing overfitting by randomly setting a fraction of input units to 0 during training. The final dense layer has 5 units one for each class with softmax activation. The last layer does the classification of input into one of the five disease categories. Detailed Layer information of proposed model is given Table 4.

The model training utilizes 60% of the dataset, with 20% used for validation and the remaining 20% for testing. Training is performed with a batch size of 32 for 231 epochs. The model achieved a training accuracy of 98.92% with a loss of 0.0314 and a validation accuracy of 98.23% with a loss of 0.0519. The training and validation accuracy and loss for each epoch are depicted in Fig. 6.

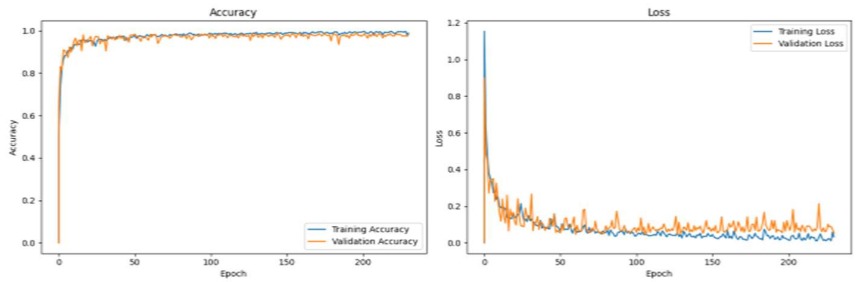

Figure 6. Accuracy and loss through each Epoch

Table 4. Proposed Model Summary.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 125, 125, 16)	4624
max_pooling2d_1 (MaxPooling 2D)	(None, 62, 62, 16)	0
conv2d_2 (Conv2D)	(None, 60, 60, 32)	4640
max_pooling2d_2 (MaxPooling 2D)	(None, 30, 30, 32)	0
conv2d_3 (Conv2D)	(None, 28, 28, 64)	18496
max_pooling2d_3 (MaxPooling 2D)	(None, 14, 14, 64)	0
conv2d_4 (Conv2D)	(None, 12, 12, 128)	73856
max_pooling2d_4 (MaxPooling 2D)	(None, 6, 6, 128)	0
reshape (Reshape)	(None, 36, 128)	0
lstm (LSTM)	(None, 36, 256)	394240
flatten (Flatten)	(None, 9216)	0
dense (Dense)	(None, 1024)	9438208
dropout (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 5)	5125

4.3. Other Models

Along with the proposed model 4 more models were trained for comparing results. Additional models VGG16, Alexnet, ResNet50. Same data with same ratio was used to train these other models.

4.3.1. Alexnet

Introduced by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012, AlexNet marked a significant milestone in the realm of deep learning. Its architecture revolutionized the computer vision field by demonstrating the power of Convolutional Neural Networks (CNNs) on large-scale image classification tasks. With five convolutional layers followed by three dense layers, it achieved remarkable accuracy, winning the ILSVRC in 2012 [31].

This model achieved 99.61% training accuracy 93.03% validation accuracy. The training and validation accuracy and loss for each epoch are shown in Fig. 7.

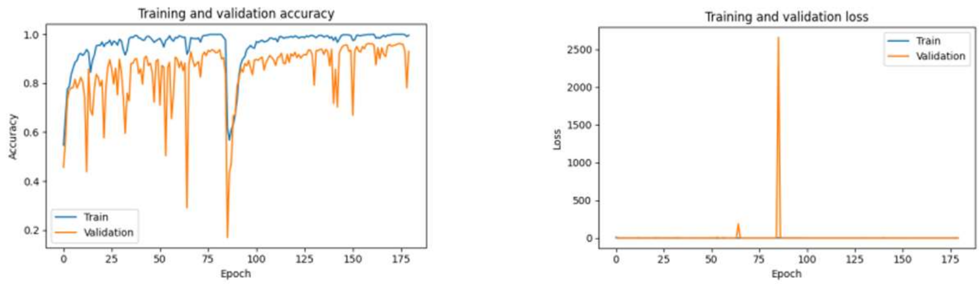

Figure 7. Accuracy and loss through each Epoch

4.3.2. VGG16

The VGG16 model, introduced by the Visual Geometry Group(VGG), University of Oxford, represents another significant advancement in field deep learning, particularly for tasks related to image classification [32]. With its 16-layer design, comprising 13 convolutional layers and then 3 Fully Connected (FC) layers, VGG16 demonstrated impressive capabilities in feature extraction and classification.

In many instances, researchers adopted transfer learning techniques, where pretrained VGG16 models were fine-tuned for specific tasks by freezing the weights of the convolutional layers and adding new fully connected layers for task-specific classification [33]. In this model, all layers were kept the same except for one additional fully connected layer. This change was made to adapt the pretrained model to a whole new task while still benefiting from its previously learned features.

This model achieved 99.80 training accuracy and 95.88 validation accuracy. The training and validation accuracy for each epoch are shown in Fig. 8.

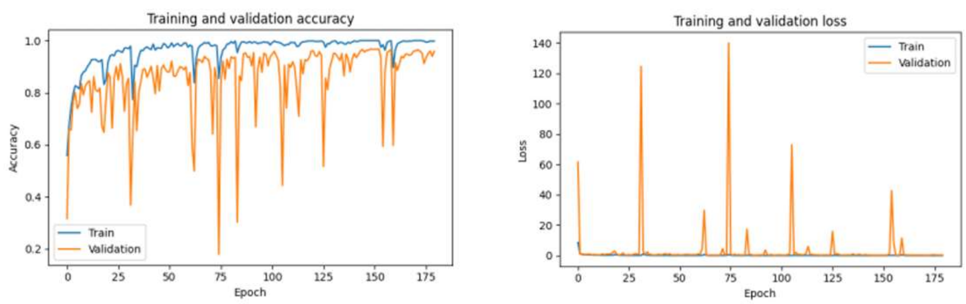

Figure 8. Accuracy and loss through each Epoch

4.3.3. ResNet50

ResNet50, an advancement of the ResNet architecture pioneered by Microsoft Research, exemplifies the ongoing refinement of deep learning frameworks [34]. Distinguished by its 50-layer structure, ResNet50 presents a heightened level of complexity, enabling nuanced feature extraction across diverse datasets and domains. In the conducted research, a customized version of ResNet50 was utilized, integrating four supplementary layers before the training phase. This augmentation aimed to further enhance the model's capacity to extract and leverage hierarchical features from input data, thereby potentially improving its performance on specific tasks.

This model achieved 99.08 training accuracy and 74.88 validation accuracy. The training and validation accuracy for each epoch are shown in fig. 9.

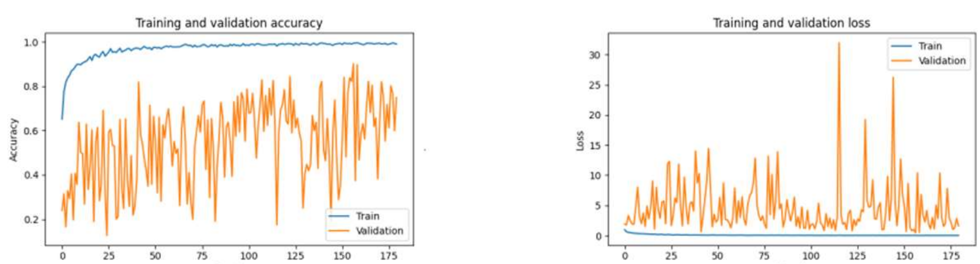

Figure 9. Accuracy and loss through each Epoch

4. RESULTS

The proposed CNN+LSTM hybrid model was compared to various models as discussed above using some of the important performance metrics. Metrics used are Classification Accuracy, Precision, Recall, F1-Score are described in Table 4. Additionally Confusion Matrix is also mentioned in this section.

Table 4.Preformance Metrics.

Metrics	Description	Formula	Graph	Remarks
Classification Accuracy	This is one of the critical metrics for measuring performance. It refers to the percentage of correct predictions made by the classification model.	$\frac{Numberofcorrectpredict}{TotalnumberofPredicti}$	Figure 10	The proposed model achieved the highest test accuracy of 98.53% with a test loss of 0.0672. Resnet50 performed the worst among all the models.
Precision	Precision is the ratio of correct predictions(true positives) to total predictions	$\frac{TruePositives}{(TruePositives + FalsePo}$	Figure 11	The proposed model achieved precision of 0.99 which is highest compared to other models.
Recall	Recall is also known as sensitivity. It measures the ratio of true positives correctly classified as Positive to the total number of positive predictions	$\frac{TruePositives}{(TruePositives + FalseNe}$	Figure 12	The proposed model has the achieved highest recall value.
F1 Score	F1 Score is a harmonic mean of precision and recall	$\frac{2 \times Precision \times Recall}{Precision + Recall}$	Figure 13	It shows that the proposed model has outperformed all other models.

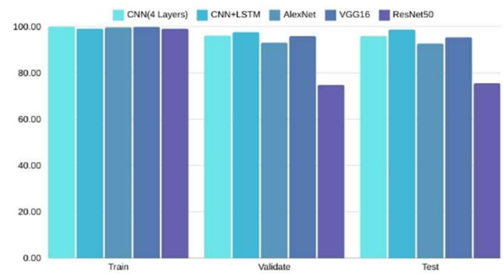

Figure 10. Test Accuracy

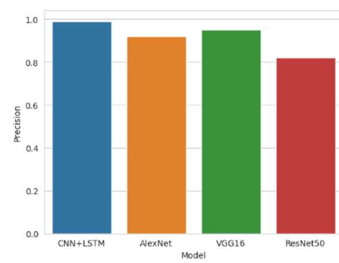

Figure 11. Precision

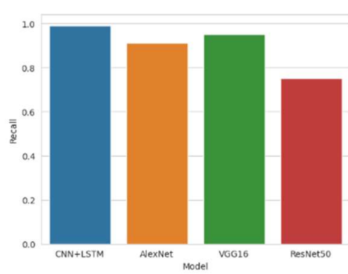

Figure 12. Recall

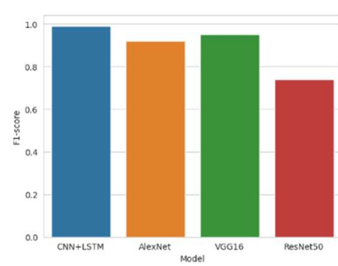

Figure 13. F1-Score

4.1. Confusion Matrix

Confusion Matrix provides a detailed breakdown of the actual versus predicted classification. Fig. 14 shows the confusion matrix of Proposed Model.

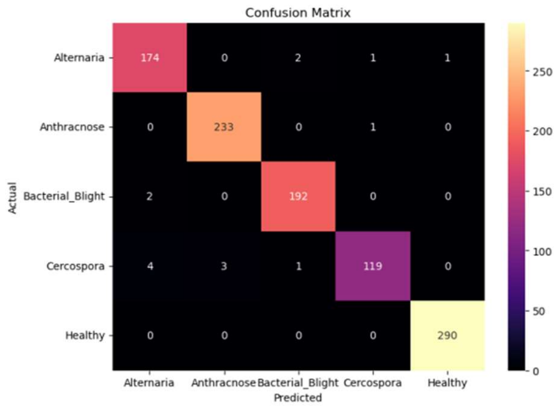

Figure 14. Confusion Matrix

From the above metrics, it can be observed that the proposed CNN+LSTM model outperformed the other CNN based model Alexnet and state-of-art transfer learning models ResNet and VGG16. Also it is to be noted that the proposed model uses fewer layers (CNN as well as fully connected) compared to other models.

5. CONCLUSION

This study proposed a novel approach for detecting pomegranate fruit disease detection by using a hybrid model. By combining the feature extraction of CNN with sequential handling of LSTM, the proposed model demonstrates higher performance compared to other models that were considered for this study.

Experiment results show that the proposed model CNN-LSTM archives higher accuracy and precision, significantly improving the disease detection and classification capabilities. This advancement can help the agriculture and food industry in reducing economic costs.

The study provides a comprehensive evaluation of various machine learning models, highlighting the efficacy of hybrid models in agricultural disease detection. Similar approaches could benefit other crops which offer scalable and adaptable solutions.

However there are limitations also exist. The dataset, while extensive, is region-specific and limited to few varieties of pomegranate fruit. Expanding the dataset to include other varieties and other locations can increase the robustness of the model. Additionally, integrating advanced image preprocessing techniques and exploring other hybrid combinations may yield better results.

The study also adds the ineffectiveness of transfer learning approaches like VGG16 and ResNet50 and the performance issues of larger models like AlexNet indicating that larger architectures may not always perform better in this context.

In conclusion, the CNN-LSTM hybrid model marks a significant advancement in agricultural disease detection, offering a promising tool for improving the sustainability and productivity of pomegranate cultivation and potentially revolutionizing agricultural practices.

REFERENCES

[1] Fahad, Shah & Faisal, Shah & Naushad, Mahnoor. (2020). Pomegrantes Economics and Medicinal Aspects in the World. SSRN Electronic Journal. 10.2139/ssrn.3597891.
[2] Mena, Pedro, et al. "Pomegranate fruit for health promotion: Myths and realities." Funct Plant Sci Biotechnol 5 (2011): 33-42.
[3] Ismail, Tariq, Piero Sestili, and Saeed Akhtar. "Pomegranate peel and fruit extracts: a review of potential anti-inflammatory and anti-infective effects." Journal of ethnopharmacology 143.2 (2012): 397-405.
[4] Ganeshkumar, D., Kalyan Bhattacharyya, Lakshmi Narsimhaiah, and Santosh S. Yumnam. "Performance of pomegranate export from India." Economic Affairs 61, no. 4 (2016): 575-580.
[5] Heber, D. (2011). Pomegranate ellagitannins. In Phytochemicals in Nutrition and Health (pp. 103-115). Springer.
[6] Deshpande, Tejal & Sengupta, Sharmila & Raghuvanshi, K. (2014). Grading & Identification of Disease in Pomegranate Leaf and Fruit. 10.13140/RG.2.2.12639.84643.
[7] Bhange, Manisha & Hyder Ali, Hingoliwala. (2015). Smart Farming: Pomegranate Disease Detection Using Image Processing. Procedia Computer Science. 58. 280-288. 10.1016/j.procs.2015.08.022.
[8] S. A. Gaikwad, K. S. Deore, M. K. Waykar, P. R. Dudhane, and Prof. G. Sorate, “Fruit Disease Detection and Classification,” International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 12, Dec. 2017, e-ISSN: 2395-0056
[9] Pooja Kantale , Shubhada Thakare, 2020, Pomegranate Disease Classification using Ada-Boost Ensemble Algorithm,

INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 09 (September 2020)
[10] M. Dhakate and Ingole A. B., "Diagnosis of pomegranate plant diseases using neural network," 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Patna, India, 2015, pp. 1-4, doi: 10.1109/NCVPRIPG.2015.7490056.
[11] Miss. K. Purvimath and Dr. P. B. Patil, "POMEGRANATE FRUIT DISEASE DETECTION AND CLASSIFICATION," International Journal of Engineering Applied Sciences and Technology, vol. 1, no. 7, Jun. 2016, ISSN No. 2455-2143, pp. 102-106
[12] Pawar, Rashmi & Jadhav, Ambaji. (2017). Pomogranite disease detection and classification. 2475-2479. 10.1109/ICPCSI.2017.8392162.
[13] G. Nagarjuna Reddy , L.Gowthami , P.Eshwar,P.Sarvothama Reddy ,K.Madhu Sudhan Reddy."ANN based Fruit Disease Detection using Texture Analysis", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.10, Issue 4, page no.e187-e193, April-2023, Available : http://www.jetir.org/papers/JETIR2304425.pdf
[14] M. Venu Madhavan, D. Ngoc Hoang Thanh, A. Khamparia, S. Pande, R. Malik et al., "Recognition and classification of pomegranate leaves diseases by image processing and machine learning techniques," Computers, Materials & Continua, vol. 66, no.3, pp. 2939–2955, 2021.
[15] D. M. Sharath, S. A. Kumar, M. G. Rohan, Akhilesh, K. V. Suresh and C. Prathap, "Disease Detection in Plants using Convolutional Neural Network," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 389-394, doi: 10.1109/ICSSIT48917.2020.9214159.
[16] Sánchez, M.G., Miramontes-Varo, V., Chocoteco, J.A., Vidal, V. (2020). "Identification and Classification of Botrytis Disease in Pomegranate with Machine Learning." In: Arai, K., Kapoor, S., Bhatia, R. (eds) Intelligent Computing. SAI 2020. Advances in Intelligent Systems and Computing, vol 1229. Springer, Cham. https://doi.org/10.1007/978-3-030-52246-9_43
[17] S. Pawar, S. Shedge, N. Panigrahi, A. P. Jyoti, P. Thorave and S. Sayyad, "Leaf Disease Detection of Multiple Plants Using Deep Learning," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 241-245, doi: 10.1109/COM-IT-CON54601.2022.9850899.
[18] Wakhare, Prashant B., Jayash A. Kandalkar, Rushikesh, S. Kawtikwar, Sonali A. Kalme and Rutvija V. Patil. "Development of Automated Leaf Disease Detection in Pomegranate Using Alexnet Algorithm." Current Agriculture Research Journal (2023): n. Pag.
[19] Nirgude, Vaishali & Rathi, Sheetal. (2021). "A ROBUST DEEP LEARNING APPROACH TO ENHANCE THE ACCURACY OF POMEGRANATE FRUIT DISEASE DETECTION UNDER REAL FIELD CONDITION." Journal of Experimental Biology and Agricultural Sciences. 9. 863-870. 10.18006/2021.9(6).863.870.
[20] Makandar, Aziz & Javeriya, Syeda. (2021). Faster-RCNN Based Deep Learning Model for Pomegranate Diseases Detection and Classification. Computational Intelligence and Machine Learning. 2. 10-19. 10.36647/CIML/02.02.A002.
[21] Y. Qi, C. Zhichao and L. Lan, "Fruit Disease Identification Based on Improved Densenet Fusion Defogging Algorithm," 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2021, pp. 313-317, doi: 10.1109/ICCWAMTIP53232.2021.9674088.
[22] Vasumathi MT, Kamarasan M. (2021) "An Effective Pomegranate Fruit Classification Based On CNN-LSTM Deep Learning Models." Indian Journal of Science and Technology. 14(16): 1310-1319. https://doi.org/10.17485/IJST/v14i16.432
[23] M. T. Vasumathi and M. Kamarasan, "An LSTM based CNN Model for Pomegranate Fruit Classification with Weight Optimization Using Dragonfly Technique," Indian Journal of Computer Science and Engineering, vol. 12, no. 2, pp. 371-384, 2021.
[24] Naseer, Aisha & Amjad, Madiha & Raza, Ali & Munir, Kashif & Abdelsamee, Nagwan & Alohali, Manal. (2024). ―A Novel Transfer Learning Approach for Detection of Pomegranates Growth Stages. ―IEEE Access. 10.1109/ACCESS.2024.3365356.
[25] M. D. Nirmal, Pramod P. Jadhav & Santosh Pawar (13 Jan 2023): ―Pomegranate Leaf Disease Detection Using Supervised and Unsupervised Algorithm Techniques‖, Cybernetics and Systems, DOI: 10.1080/01969722.2023.2166192
[26] B, Pakruddin; R, Dr. Hemavathy (2023), "Pomegranate Fruit Diseases Dataset for Deep Learning Models", Mendeley Data, V1, doi: 10.17632/b6s2rkpmvh.1
[27] "rembg", PyPI, May 24, 2024. https://pypi.org/project/rembg/
[28] Gu J., Wang Z., Kuen J., Ma L., Shahroudy A., Shuai B., Liu T., Wang X., Wang G., Cai J., Chen T. Recent advances in convolutional neural networks. Pattern Recogn. 2018 doi: 10.1016/j.patcog.2017.10.013.
[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778
[30] Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int J Uncertain Fuzziness Knowledge-Based Syst. 1998;6:107–116. doi: 10.1142/S0218488598000094.
[31] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems, vol. 25, pp. 1097-1105, 2012.
[32] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2014
[33] Q. Yan, B. Yang, W. Wang, B. Wang, Y. Zhang, & J. Zhang, "Apple leaf diseases recognition based on an improved convolutional neural network", Sensors, vol. 20, no. 12, p. 3535, 2020. https://doi.org/10.3390/s20123535
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
[35] Dr. Mohammed Saleh Al Ansari, University of Bahrain: January 2024 DOI: 10.17762/ijritcc.v11i9.9597 ―A Machine Learning Approach to Pomegranate Leaf Disease Identification‖ Article in International Journal on Recent and Innovation Trends in Computing and Communication.