**Pomegranate Fruit Disease Identification Using Machine Learning Models**

**Abstract:**

Pomegranate (Punica granatum) is a highly nutritious and economically valuable fruit crop, but its production is significantly threatened by diseases such as Anthracnose, Bacterial Blight, Cercospora, and Alternaria. Traditional disease detection methods rely on visual inspection, which is time-consuming and prone to human error. To address these challenges, this study explores the application of various machine learning (ML) techniques for automated pomegranate disease detection and classification. We evaluated the performance of multiple ML models, including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), 1-Nearest Neighbor (1NN), Random Forest (RF), Artificial Neural Network (ANN), and Feed Forward Neural Network (FFNN). The results demonstrated varying levels of accuracy. Among these, Random Forest outperformed others with the highest accuracy, while KNN and 1NN showed comparatively lower performance. These findings highlight the potential of ML techniques, particularly ensemble methods like Random Forest, in enhancing disease detection efficiency and accuracy. The study underscores the importance of adopting advanced ML models to support early disease identification, improve crop management, and mitigate economic losses in pomegranate cultivation.

\*Keywords\*: Pomegranate disease detection, Machine learning, Classification, Random Forest, Neural Network

**Introduction:**

Pomegranate is a globally significant fruit crop, renowned for its nutritional benefits and economic value. However, its cultivation faces major challenges due to diseases caused by fungal, bacterial, and viral pathogens, which adversely affect yield and fruit quality. Traditional diagnostic methods, which depend on manual inspection by farmers or experts, are often inefficient and subjective. To overcome these limitations, this study investigates the application of machine learning (ML) techniques for automated and accurate disease detection in pomegranates.

The study focuses on classifying five disease categories—Healthy, Anthracnose, Bacterial Blight, Cercospora, and Alternaria—using a diverse set of ML models. We evaluated the performance of Support Vector Machine (SVM), K-Nearest Neighbors (KNN), 1-Nearest Neighbor (1NN), Random Forest (RF), Artificial Neural Network (ANN), and Feed Forward Neural Network (FFNN).

These findings emphasize the potential of ML, particularly ensemble and neural network-based methods, in revolutionizing agricultural disease management. By enabling early and precise disease detection, these techniques can empower farmers to take timely corrective measures, thereby enhancing crop productivity and reducing economic losses. Future research could explore integrating larger datasets, advanced feature extraction methods, and hybrid models to further improve classification performance and adaptability to real-world farming conditions.

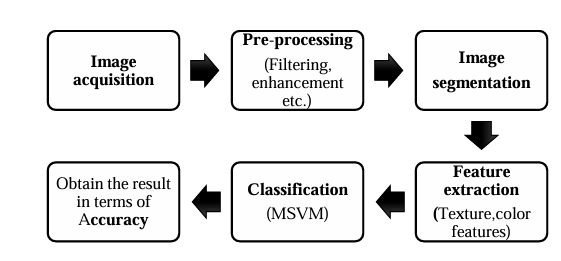
This study contributes to the growing body of research on smart agriculture, showcasing the transformative role of ML in addressing critical challenges in pomegranate cultivation.

**Literature Review:**

[1] The study highlighted the use of preprocessing methods like Gaussian filtering and histogram equalization to enhance image quality, followed by segmentation techniques such as K-means clustering and Otsu’s thresholding to isolate diseased regions. outperforming SVM (91.2%) and RF (89.5%) in distinguishing diseases like bacterial blight and fungal spots. [2] A hybrid approach combining image processing and machine learning (ML) to classify pomegranate leaf diseases such as bacterial blight, fungal spots, and Alternaria rot. SVM achieving the highest accuracy (94.7%). [3] The study utilized image preprocessing techniques (Gaussian blur, adaptive thresholding) to enhance disease visibility, followed by feature extraction using Histogram of Oriented Gradients (HOG) and deep learning-based feature selection. This results were effective such as SVM Accuracy: 93.5%, Decision Tree Accuracy: 88.6%. [4] An image processing-based approach for detecting bacterial blight in pomegranates, achieving 82% accuracy using SVM classification with features like color, morphology, and Color Coherence Vector (CCV). used morphology features for disease detection in grapes and apples, attaining 90% accuracy. [5] Combining image processing with sensor data for improved accuracy. The hybrid approach achieved higher accuracy (93.6%) compared to traditional image-only methods. [6] The study employed image preprocessing (noise removal, contrast enhancement) followed by feature extraction using texture (GLCM) and color histograms. A multilayer perceptron (MLP) neural network was trained for classification, achieving 91.2% accuracy. Outperformed SVM (87.5%) and k-NN (84.3%) in the same study. [8] Deep learning models like YOLO-v7 reported performance scores such as mAP@0.5 of 94.3% and recall of 88.8% for growth stage detection. traditional CNNs struggled in real-field classification tasks, with validation accuracies dropping to 56–65%. [9] Pre-trained CNN architectures (ResNet-50 and EfficientNet-B4) combined with handcrafted texture features (LBP, GLCM). deep learning's feature extraction strengths while enhancing robustness with traditional texture analysis. Hybrid approach reduced overfitting in small datasets. Optimized for edge devices with <2ms inference time per image. [10] A robust machine learning framework for pomegranate leaf disease classification, focusing on bacterial blight, fungal spots, and Alternaria leaf spot. the feature extraction follows Hybrid approach using GLCM (texture) and HSV color space features. Among these all the SVM (RBF Kernel) got 94.2% accuracy and another methadology is Gradient Boosting 91.5%. [12] Image processing-based approach for detecting pomegranate diseases (bacterial blight, fungal spots, and fruit rot) using traditional computer vision techniques. The feature extraction analyzes Color moments (HSV space) and texture features (GLCM). the accuracy of the model SVM (RBF Kernel 89.3%). [13] The detection of pomegranate diseases has gained significant attention due to its impact on yield and quality. Studies have demonstrated the efficiency of Convolutional Neural Networks (CNNs), such as ResNet50 and InceptionV3, achieving accuracy rates above 90% in real-field conditions. [16] The advancements in using deep learning for pomegranate disease detection. combined image processing with SVM, reaching 89.2% accuracy but noted scalability limitations. Faster-RCNN for localized disease detection, achieving 92.7% precision in real-time conditions. CNN-LSTM hybrid model, improving temporal pattern recognition with 93.8% accuracy.

**Proposed Methodology:**

This Proposed methodology outlines that is used to detect the pomegranate diseases using the machine learning models. The entire process includes the dataset used, in detailed explanation of the models and the techniques applied. The models are Support Vector Machine (SVM), K-Nearest Neighbors (KNN), 1-Nearest Neighbors (1-NN), Random Forest, Artificial Neural Network (ANN), Feed Forword Neural Network (FFNN). The methodology is designed to accurately find the five pomegranate disease classes like Healthy, Anthracnose, Bacterial Blight, Cercospora and Alternaria.



1. **Dataset Used:**

The dataset used in this study that comprises the high-resolution RGB images of pomegranate fruits that was categorized into six classes based on the disease type: Healthy images sample with no visible symptoms, Anthracnose was characterized by the dark fungal lesions on stem, leaves and fruits, Bacterial Blight was identified by the water-soaked lesions and black spots caused by the bacterial infection, Cercospora was marked by small, circular spots that resulting from the fungal infestation and Alternaria was a fungal disease produces the irregular brown to black spots on the leaves and fruits. To ensure the robust evaluation, the dataset was splitted into training and testing subset using an 80%-20% ratio. All the images were preprocessed by resizing them into a uniform dimension of 128 × 128 pixels to maintain the consistency across the dataset.



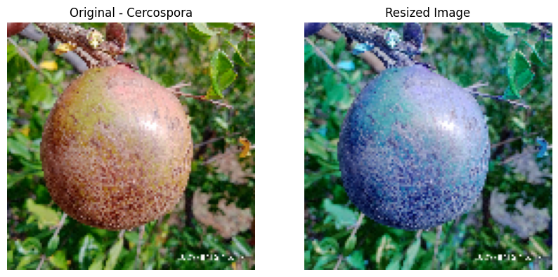
Fig1. Sample Images of Datasets

1. **Feature Extraction Techniques:**

Feature Extraction is the crucial step in enhancing the performance of the machine learning models, as it can transforms the raw image data into the meaning full patterns that can be efficiently classified. In this project several feature extraction methods were employed to capture the various aspects of the pomegranate fruits and leaves affected by the different diseases.

**Color features** : color histograms was used to quantify the distribution of colors across an image, that enabling the differentiation between the healthy regions and the areas affected by the fungal or the bacterial infections.

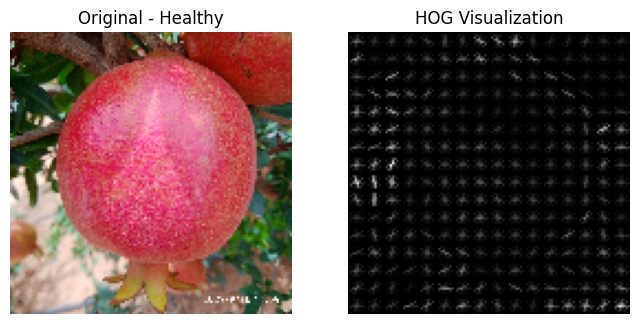
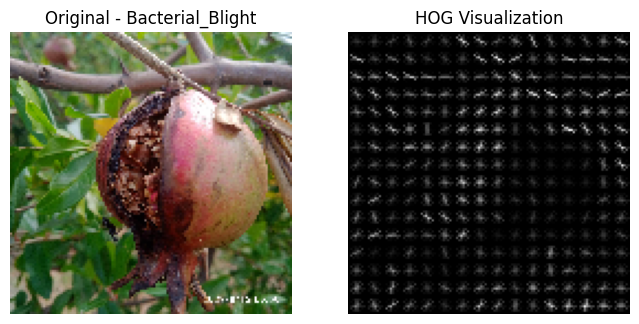
**Resizing:** All the images was resized to a uniform dimension of 128 × 128 pixels. Resize is a essential for feeding images into a machine learning algorithms like CNN, KNN, and SVM as the inconsistent image sizes can lead to errors during the training and evaluation.



1. (B)

Fig2. Resize Images of (A) Cercospora (B) Anthracnose

**Histogram of Oriented Gradients (HOG):** This technique is applied to capture the structural and shape characteristics of the images. HOG works by computing the gradients of the image intensity and creating histograms of the gradient directions, making it as highly effective for representing object contours and detecting the lesions on the surface of fruits and leaves.

****

1. (B)

Fig3. HOG Feature Extraction of (A) Healthy and (B) Bacterial Blight

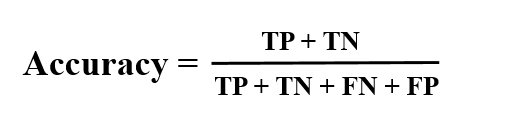
1. **Types of Models Used:**

The models used are K-Nearest Neighbors (K-NN), Support Vector Machine (SVM), 1-Nearest Neighbor (1-NN), Random Forest Classifier, Artificial Neural Network(ANN), Feed Forward Neural Network(FFNN). Each of these models follows an a distinct approach to learning and decision-making. These models allowed us to explore the different machine learning paradigms ranging from the traditional algorithms to the advanced deep learning techniques and to compare their performances on the given pomegranate disease dataset.

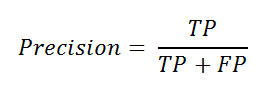
1. **K-Nearest Neighbors(KNN)** : KNN is a simple and effective classification technique. KNN works by comparing a new, unseen data point with the existing data points in the training set. The algorithm identifies the ‘k’ closest instances and classifies the new data based on majority class of these neighbors. Due to its simplicity, KNN is often used as the baseline model.
2. **Support Vector Machine(SVM)**: It is a powerful supervised learning model that seeks to find the optimal boundary, or hyperplane, in which that separates the data points of the different classes. SVMs are very well known for their ability to handle both the linear and non-linear classification tasks using the kernel functions. Their effectiveness, however, can be limited when dealing with the large datasets or the datasets with significant overlap between classes.
3. **1-Nearest Neighbor(1-NN)**: It is a special case of the KNN algorithm where only the single nearest data point is considered at when making an a prediction. 1-NN is the simplest form of the instance-based learning and then often provides surprisingly good results for small and well separated datasets. This can lead to the unstable predictions if the dataset is not clean or well distributed.
4. **Random Forest Classifier**: It is a ensemble learning method that constructs an multitude of an decision trees during training and outputs the class that represents the majority vote of the trees. This technique improves the classification accuracy by reducing the risk of overfitting. Random Forest also can handle the missing data and maintain accuracy when the large proportion of the data is missing.
5. **Artificial Neural Networks(ANNs)**: These are inspired by the structure and function of the human brain. ANNs consist of the interconnected layers of the processing units called neurons, which they work together to learn complex patterns in the data. They typically include an input layer, one or more hidden layers, and output layer. ANNs are highly versatile and have been successfully applied in each fields such as a image recognition, natural language processing(NLP), and biomedical diagnosis.
6. **Feedforward Neural Networks(FFNNs):** It represent the most basic type of artificial neural networks. In FFNNs information moves in only the one direction from input layer through the hidden layers and finally to the output layer without forming an any cycles or the loops. The FFNNs are the widely used due to their simplicity and the strong performance on a wide range of structural data problems.

**Performance Metrices:**

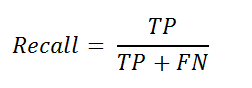
**Accuracy:** This is proportion of the correctly predicted instances (both true positives and true negatives) to the total number of instances.



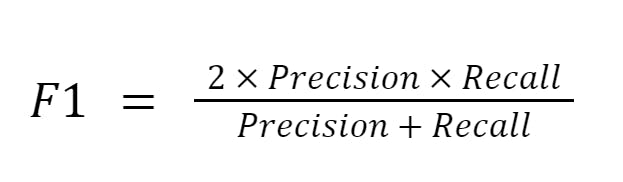
**Precision:** It indicates the proportion of true positive predictions to the total predicted positives. The high precision means the model is making the fewer false positive errors.



**Recall:** It measures the model’s ability to the correctly identify actual positives. The High recall indicates the fewer false negatives.



**F1-Score:** F1-Score is the harmonic mean of the precision and recall, by providing a balanced measure of the model’s performance, especially when there is uneven class distribution.



**Experimental Results:**

1. **Experimental results for KNN**

The K-Nearest Neighbors (KNN) model showed the very lowest performance on this experiment set for pomegranate fruit disease detection. Were it performed for “Alternaria” (F1-Score: 0.33), “Anthracnose” (f1-score: 0.71), “Bacterial Blight” (F1-Score: 0.64), “Cercospora” (F1-Score :0.52) and “Healthy” (F1-Score: 0.84). The overall accuracy was 0.68(68%) with macro avg(0.61) and weightavg(0.65) for this model.

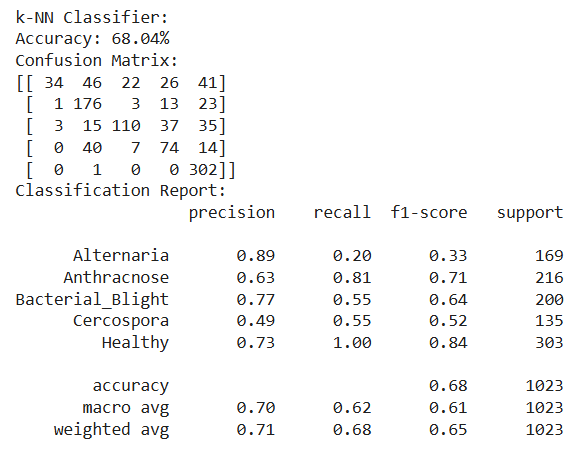
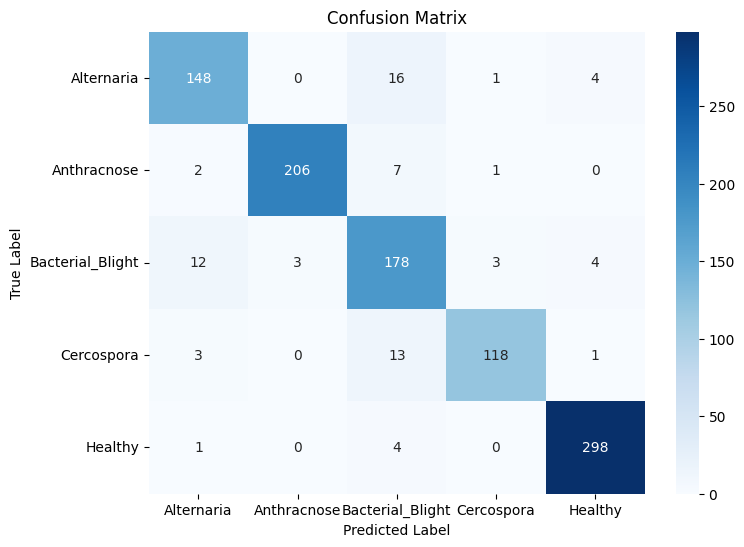
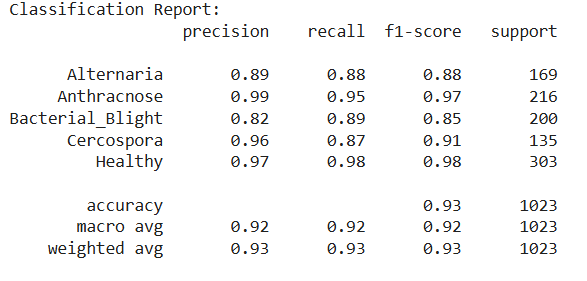


Fig4. Confusion Matrix and Classification Report

1. **Experimental Results for SVM**

The Support Vector Machine(SVM) model showed the very good performance on this experiment set for pomegranate fruit disease detection. Were it performed for “Alternaria” (F1-Score: 0.88), “Anthracnose” (f1-score: 0.97), “Bacterial Blight” (F1-Score: 0.85), “Cercospora” (F1-Score :0.91) and “Healthy” (F1-Score: 0.98). The overall accuracy was 0.9267(92.67%) with macro avg(0.92) and weighted avg(0.93) for this model.



1. (B)

Fig5. (A) Classification Report and (B) Confusion Matrix

1. **Experimental Results for 1-NN**

The 1-Nearest Neighbor(1-NN) model showed the very lowest performance on this experiment set for pomegranate fruit disease detection. Were it performed for “Alternaria” (F1-Score: 0.38), “Anthracnose” (f1-score: 0.71), “Bacterial Blight” (F1-Score: 0.68), “Cercospora” (F1-Score :0.49) and “Healthy” (F1-Score: 0.86). The overall accuracy was 0.6891(68.91%) with macro avg(0.63) and weighted avg(0.67) for this model.

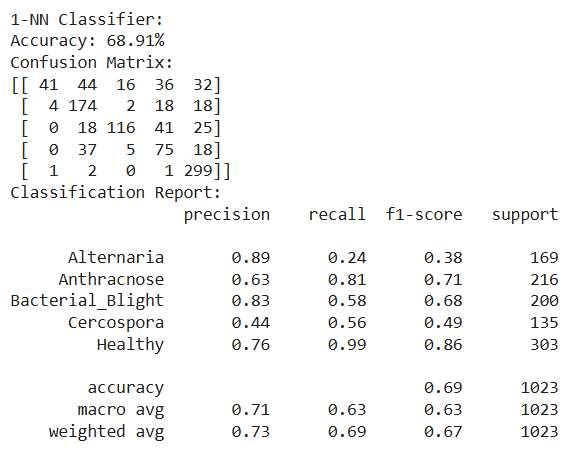


Fig6. Confusion Matrix and Classification Report

1. **Experimental Results for Random Forest Classifier:**

The Random Forest Classifier model showed the very highest performance on this experiment set for pomegranate fruit disease detection. Were it performed for “Alternaria” (F1-Score: 0.94), “Anthracnose” (f1-score: 1.00), “Bacterial Blight” (F1-Score: 0.89), “Cercospora” (F1-Score :0.90) and “Healthy” (F1-Score: 0.98). The overall accuracy was 0.9462(94.62%) with macro avg(0.94) and weighted avg(0.95) for this model. The 94.62% accuracy is the highest in this experiment.

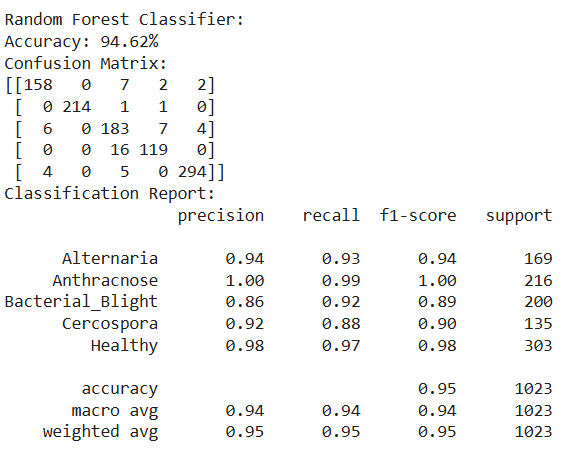


Fig7. Confusion Matrix and Classification Report

1. **Experimental Results for ANNs**

The Artificial Neural Network model showed the moderate performance on this experiment set for pomegranate fruit disease detection. Were it performed for “Alternaria” (F1-Score: 0.79), “Anthracnose” (f1-score: 0.83), “Bacterial Blight” (F1-Score: 0.79), “Cercospora” (F1-Score :0.45) and “Healthy” (F1-Score: 0.97). The overall accuracy was 0.81(81.04%) with macro avg(0.76) and weighted avg(0.80) for this model.

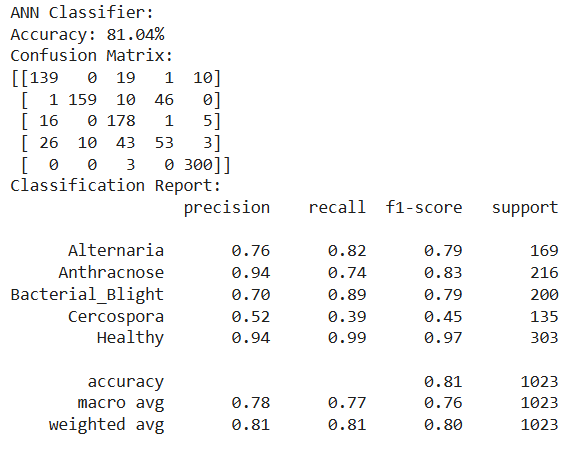


Fig8. Confusion Matrix and Classification Report

1. **Experimental Results for FFNNs**

The Feed Forward Neural Network model showed the moderate performance on this experiment set for pomegranate fruit disease detection. Were it performed for “Alternaria” (F1-Score: 0.72), “Anthracnose” (f1-score: 0.86), “Bacterial Blight” (F1-Score: 0.78), “Cercospora” (F1-Score :0.80) and “Healthy” (F1-Score: 0.93). The overall accuracy was 0.83(83.77%) with macro avg(0.82) and weighted avg(0.84) for this model.

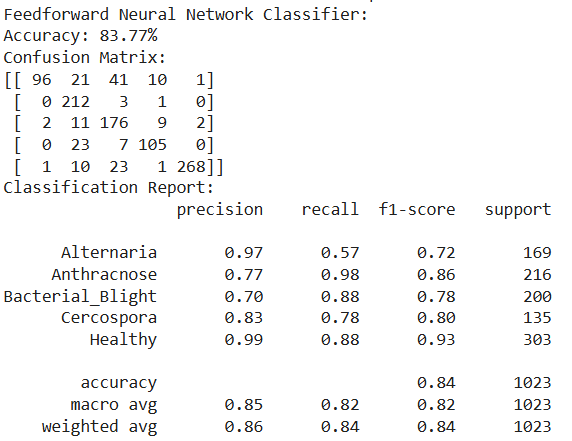


Fig9. Confusion Matrix and Classification Report

**Conclusion:**

In this experiment we focused on the detection of the pomegranate fruit disease through the development and evaluation of the Machine Learning and Deep Learning models By utilizing the dataset with the comprising five distinct classes They are : Healthy, Anthracnose, bacterial Blight, Cercospora, and Alternaria. With the approach of involving the data pre-processing, feature extraction and model training. And we successfully implemented the Six different ML and DL algorithms of KNN, SVM, 1-NN, Random Forest Classifier, ANNs and FFNN. And your experimental results say that the Random Forest Classifier performed the best compared to five other models. The Future work should be more diverse datasets so we can improve the more further classification accuracy.

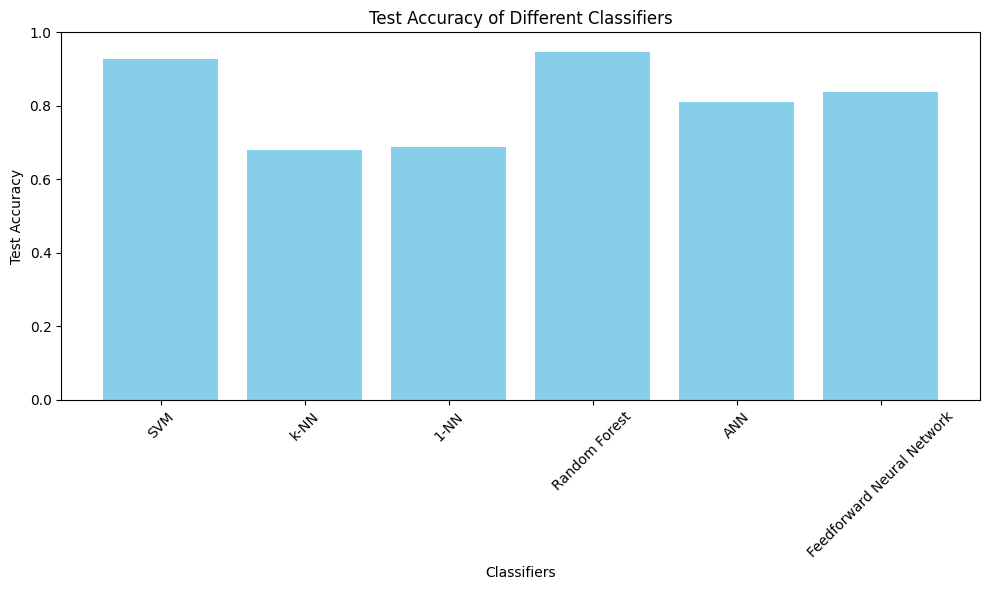


Fig10. Test Accuracy of Different Models

**References:**

1. P. Kantale and S. Thakare, "A Review on Pomegranate Disease Classification Using Machine Learning and Image Segmentation Techniques," *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2020, pp. 455-460, doi: 10.1109/ICICCS48265.2020.9121161.
2. Nirmal, M. D., Jadhav, P., & Pawar, S. (2022). Classification of Pomegranate Leaves Diseases by Image Processing and Machine Learning Techniques. *Cybernetics and Systems*, 1–15. https://doi.org/10.1080/01969722.2022.2145448
3. P. Wakhare, S. Neduncheliyan and P. B. Mane, "Machine Learning for Accurate and Efficient Pomegranate Fruit Disease Detection: A Novel Approach to Improve Crop Yield and Quality," *2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, Pune, India, 2023, pp. 1-5, doi: 10.1109/ICCUBEA58933.2023.10391966.
4. JOUR Smart Farming: Pomegranate Disease Detection Using Image Processing Bhange, Manisha Hingoliwala, H.A. Procedia Computer Science 58 280 288 2015 2015/01/01/ Second International Symposium on Computer Vision and the Internet (VisionNet’15) 1877-0509 <https://doi.org/10.1016/j.procs.2015.08.022>
5. S. Pawara, D. Nawale, K. Patil and R. Mahajan, "Early Detection of Pomegranate Disease Using Machine Learning and Internet of Things," *2018 3rd International Conference for Convergence in Technology (I2CT)*, Pune, India, 2018, pp. 1-4, doi: 10.1109/I2CT.2018.8529583.
6. 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.
7. JOUR A comprehensive standardized dataset of numerous pomegranate fruit diseases for deep learning B․, Pakruddin R․, Hemavathy Data in Brief 54 110284 2024 2024/06/01/ 2352-3409 <https://doi.org/10.1016/j.dib.2024.110284>
8. A. Naseer, M. Amjad, A. Raza, K. Munir, N. A. Samee and M. A. Alohali, "A Novel Transfer Learning Approach for Detection of Pomegranates Growth Stages," in *IEEE Access*, vol. 12, pp. 27073-27087, 2024, doi: 10.1109/ACCESS.2024.3365356
9. S. K. Banerjee, P. Chakraborty and S. Nath, "Hybrid Transfer Learning based Pomegranate Fruit Disease Identification," *2024 International Conference on Big Data Analytics in Bioinformatics (DABCon)*, Kolkata, India, 2024, pp. 01-06, doi: 10.1109/DABCon63472.2024.10919383.
10. M. D. Nirmal, P. Jadhav and S. Pawar, "Pomegranate Leaf Disease Classification using Feature Extraction and Machine Learning," *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2022, pp. 619-626, doi: 10.1109/ICOSEC54921.2022.9951907.
11. H. S. Rana, M. N., S. S. Pokhare, R. A. Marathe and J. Rajan, "Convolutional Neural Network Based Approach for Automatic Detection of Diseases from Pomegranate Plants," *2024 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, Mangalore, India, 2024, pp. 66-72, doi: 10.1109/DISCOVER62353.2024.10750668.
12. S. D.M., Akhilesh, R. M.G., S. A. Kumar and P. C., "Disease Detection in Pomegranate using Image Processing," *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, Tirunelveli, India, 2020, pp. 994-999, doi: 10.1109/ICOEI48184.2020.9142972
13. Conference Proceedings Disease Detection for Pomegranate: A Review Deshpande, Priya Kore, Sharada Conference on Plant Disease Detection, Pune, India 2022
14. Journal Article Pomegranate Fruit Disease Prediction Using Machine Learning Sri, P Ramya Prabanith, T Saraswathi, P Kirubakaran, S
15. Sichuan Zhitu Linghai Intelligent Technology Co., Ltd., Chengdu 610213, China  [**https://doi.org/10.3390/e27020137**](https://doi.org/10.3390/e27020137)
16. Gaurav Mishra, Jagruti Nagaonkar, Harsh Parmar and Gaurav Joshi Web Conf., 328 (2025) 01051 <https://doi.org/10.1051/epjconf/202532801051>
17. A dataset of pomegranate growth stages for machine learning-based monitoring and analysis Zhao, Jifei Almodfer, Rolla Wu, Xiaoying Wang, Xinfa Data in Brief 50 109468 2023 2023/10/01/ 2352-3409 <https://doi.org/10.1016/j.dib.2023.109468>