Machine Learning and Deep Learning Approaches Based Rose Plant Leaf Disease Recognition

Abstract. The most popular flower in the world, roses not only cheer people up but also support livelihoods. Diseases, however, can harm these priceless flowers' health and negatively affect both their quality and the growers' livelihoods. The increased occurrence of ailments in rose plants poses a severe danger to the ornamental flower industry and agricultural productivity. We describe a novel deep learning-based method for the automated diagnosis of leaf diseases in rose plants in this paper. A big dataset containing images of both healthy and damaged rose leaves was carefully picked to illustrate different disease types and stages. To analyze and identify the visual characteristics that correspond to various illnesses, we used a Convolutional Neural Network architecture, Support Vector Machine, and K-Nearest Neighbors architectures that were specifically intended for picture classification tasks. We address the interpretability and explainability of the model's predictions in addition to performance indicators, offering insights into the decision-making process. This work addresses a fundamental requirement for effective and long-lasting disease management in rose cultivation by bridging the gap between deep learning and plant pathology. CNNs are often the preferred choice due to their ability to automatically learn relevant features from raw pixel values.

Keywords: Rose plant, leaf disease detection, deep learning, convolutional neural networks, image classification, agricultural technology.

1 Introduction

The rose is renowned as the "Queen of Flowers" and is the most widely planted garden plant in the world. Unfortunately, many roses are susceptible to a range of diseases, reducing their usefulness in the environment. Black spot, powdery mildew, downy mildew, Botrytis blight, cankers (both common and brand), rust, crown gall, mosaic, and winter damage are all frequent rose diseases in Connecticut. Every year, the type of rose, the location, and the weather influence the intensity and distribution of these difficulties.[1]

A significant ability to produce roses commercially to meet the market's rising demand due to their widespread use. Globally, the cut rose industry contributes 46.54% (India), 23.68% (China), 2.02% (Italy), and so on [2]. Rose oil and water are used to make fragrances and cosmetics, which generate a lot of money in the commercial world. Every year, 111 hectares of rose flower cultivation produce 2423 tonnes of rose flowers in Bangladesh. [3]. Let us investigate the Rose plant (Rosa spp.). Roses are a popular ornamental plant with numerous species and varieties. They are, however, susceptible to a variety of ailments, the most frequent and well-known of which being Black Spot Disease.

Rose Black Spot Disease Diplocarpon rose is a pathogen. **Symptoms**

1. Black patches: On the upper side of leaves, dark black or purple patches with fringed or uneven borders form.
2. Yellowing: The area around the black spots turns yellow, and afflicted leaves may drop.
3. Reduced Vigour: Severe infections can diminish plant vigour, produce fewer blooms, and result in overall poor performance.

Causes

Warm, humid weather promotes the growth of black spot disease. The fungus survives on diseased leaves and spreads via splashing water or wind.

2 Literature Review

A study of the literature on machine learning-based methods for detecting diseases in rose plants indicates an increasing amount of research that uses different machine learning approaches to automatically identify and diagnose diseases in rose plants [10]. Comprehending the different ailments that impact rose plants, including powdery mildew, black spot, rust, and others, prepares the ground for creating machine learning models that identify diseases. Since each disease has its own set of symptoms, it is important to take these into account when developing a model. A number of academics have helped to compile databases that include pictures of both healthy and sick rose plants [11].

The technique of examining and identifying the various images that are available and producing the necessary output in the form of images or another comprehensive report is image processing, which is employed by the authors in the paper [4]. The image is first processed, after which it undergoes analysis, and lastly, it is thoroughly comprehended and assessed. Scholars have investigated methods including picture enhancement, normalization, and de noising to enhance the caliber and variety of input data, culminating in more resilient models. For efficient illness identification, it is essential to extract useful features from photos. Numerous feature extraction techniques, such as handcrafted features and deep features taken from convolutional neural networks that have already been trained, have been used in studies [12]. The model's capacity to distinguish between healthy and unhealthy plants may be impacted by the features selected.

The authors [5] improve the SVM classifier to better detect plant diseases. Two datasets are included in the author-implemented SVM: a training dataset and a train dataset. The original image is first taken, and after that, it is processed. Second, it separates the hue and saturation portions of the image as well as the black and background pixels. Thirdly, the diseased portion of the image is identified, and the healthy portion is separated from it. G. Suresh et al. [6] used CNN to diagnose a variety of grape plant problems. Their three distinct CNN architectures are compared in order to diagnose plant ailments. A. P. Marcos et al. [7, 8] worked on constructing a model that used leaf images and convolutional neural networks to detect plant ailments. For image processing, they employ transfer flow open-source libraries [7]. They created a dataset of 159 coffee leaf images to test their CNN addition for rust recognition. In its implementation, this study [8] makes use of ten different plant disease classifications. The [8] accuracy rate of the job is 86%.

3 System Architecture

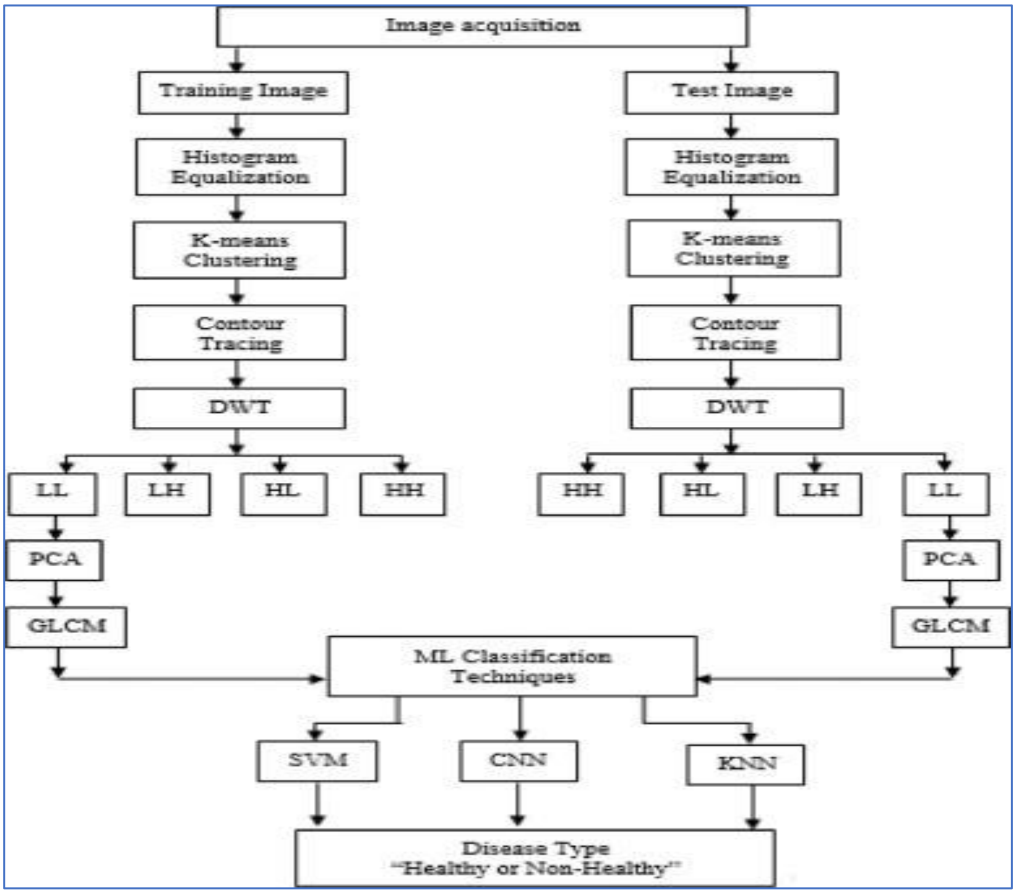

Fig. 1. Architecture [1]

The following architecture describes our proposed research theme of an image of the rose leaf as input. A general machine learning process towards predicting the output of the disease prediction. Training and testing of data are common for machine learning implementation. For classification purposes, we have to use machine learning and deep learning models for prediction. We predict it is healthy or unhealthy [13].

4 Methodology

4.1 Data Collection

Gather a comprehensive dataset containing images of healthy and diseased rose plants. The dataset should cover various diseases, stages, and environmental conditions. Create a collection of high-quality photos of rose plants.

Provide pictures of both disease-free plants and ones with common ailments like rust, black spot, and powdery mildew. Understand the common diseases that affect rose plants. Common diseases include black spot, powdery mildew, rust, and others.

Fig. 2. Dataset

4.2 Data Pre-processing

Pre-process the images to enhance model performance. Apply techniques such as resizing, normalization, and data augmentation to address issues like varying lighting conditions. Clean and pre-process the images, including resizing, normalization, and augmentation. This enhances the model's ability to generalize to different conditions. Remove any irrelevant or corrupted images from your dataset. Check for duplicates and eliminate them to avoid bias in the training process [14].

4.3Feature Extraction

4.3.1 Utilize feature extraction methods to capture relevant information from images. Handcrafted features or features taken from pre-trained convolutional neural networks (CNNs) can be included. Take into account colour, texture, and shape when extracting relevant features from segmented photos. The performance of the model can be greatly influenced by feature engineering [15].

4.3.2 Use pre-trained CNNs like VGG16, ResNet, or Inception. From big datasets like ImageNet, these networks have already learned hierarchical characteristics.
4.3.3 Remove the final classification layer and use the output from the last convolutional layer or a global average pooling layer as features.

Use texture analysis techniques, such as Gabor filters or local binary patterns, to extract information about the texture patterns present in the leaves.

1.1 Data Splitting

To train and evaluate the machine learning model, divide the dataset into training, validation, and test sets. Validate the model using a separate dataset not seen during training. Test the model on new, unseen images to evaluate its generalization ability. A common split is around 70-80% of the total dataset. Typically, allocate about 15-20% of the dataset to the validation set. Allocate the remaining portion (e.g., 10-15%) of the dataset to the test set [16].

1.2 Model Selection

Choose a suitable machine learning model for classification. Consider models like Support Vector Machines, Random Forests, or deep learning models such as CNNs.The models that we are using are [17]:

1.3 CNNs (Convolutional Neural Networks): CNNs are a sort of deep neural network that is meant to interpret structured grid data, such as photographs. They are extremely adept at tasks such as image recognition. CNNs use convolutional layers to learn feature spatial hierarchies automatically.

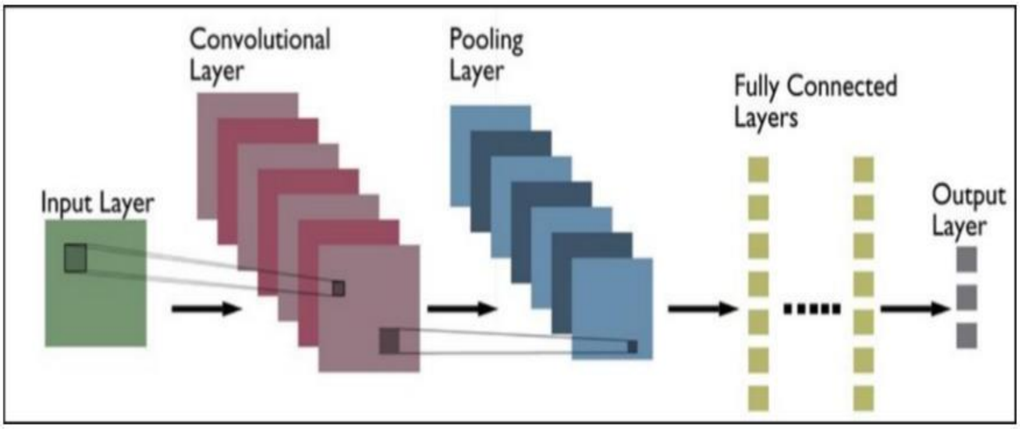

Fig. 3. CNN model Architecture

1.4 **SVM (Support Vector Machine):** The powerful machine learning approach Support Vector Machine (SVM) handles regression, outlier identification, and linear or nonlinear classification. When attempting to determine the biggest possible separation hyperplane between the various classes present in the target feature, SVM algorithms perform very well [18].

1.5 KNN (K-Nearest Neighbor): KNN is a simple classifier in machine learning that works by determining the nearest neighbors to query samples and then using these neighbors to determine the query's class. When the technique is applied, it defines comparable measures and, as a result, the sample collection category.

1.6 Model Training

Train the selected model on the training dataset. Fine-tune hyperparameters to optimize model performance. Train the selected model using the pre-processed and feature-extracted data. Utilize a labeled dataset for supervised learning. Fine-tune hyperparameters for optimal performance.

1.7 Model Evaluation

Evaluate the trained model on the validation set to assess its performance. Adjust the model, if necessary, based on validation results.

1.8 Performance Metrics

Use appropriate metrics for model evaluation, such as accuracy, precision, recall, F1 score, and area under the ROC curve. In disease detection, metrics like precision, recall, and F1 score can provide insights into how well your model is identifying diseases while minimizing false positives and false negatives. Adjust your evaluation strategy based on the importance of different types of errors in your specific use case.

1.9 Deployment

Deploy the trained model for practical use, whether in a web application, mobile app, or an embedded system for real-time disease detection.

1.10 Continuous Improvement

Keep track of the model's performance over time. To improve accuracy, update the model as needed using fresh data and retraining. Monitor the model's performance in the field on a regular basis. To adjust to changes in disease patterns or environmental factors, the model should be updated as needed.

5 Results And Analysis

We are using three different models to identify the diseases.

5.1 CNN model

```
Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         8
           1       0.60      1.00      0.75        12

    accuracy                           0.60        20
   macro avg       0.30      0.50      0.37        20
weighted avg       0.36      0.60      0.45        20

Confusion Matrix:
 [[ 0  8]
 [ 0 12]]
```

We got an accuracy of 60%.

5.2 SVM model

```
Accuracy: 0.6
Confusion Matrix:
[[3 3]
 [5 9]]
Classification Report:
              precision    recall  f1-score   support

           0       0.38      0.50      0.43         6
           1       0.75      0.64      0.69        14

    accuracy                           0.60        20
   macro avg       0.56      0.57      0.56        20
weighted avg       0.64      0.60      0.61        20
```

We got an accuracy as 60%.

5.3 KNN model

We got the accuracy as 50%.

```
Accuracy: 0.5
Confusion Matrix:
[[2 4]
 [6 8]]
Classification Report:
              precision    recall  f1-score   support

           0       0.25      0.33      0.29         6
           1       0.67      0.57      0.62        14

    accuracy                           0.50        20
   macro avg       0.46      0.45      0.45        20
weighted avg       0.54      0.50      0.52        20
```

CNNs are frequently the go-to option for image-based leaf disease identification since they can automatically extract pertinent characteristics from raw pixel values. SVMs could work well with smaller datasets or in situations where efficient feature engineering is possible. KNN is simple to use and may not work as well with high-dimensional image data, but it can be taken into consideration for simpler applications.

6 Conclusion

When comparing Support Vector Machines, Convolutional Neural Networks, and k-nearest Neighbors for leaf disease detection, several parameters such as data type, dataset size, feature representation, and processing resources must be considered. CNN and SVM give the same accuracy but CNN will perform better on larger datasets. To achieve the best results, it is recommended that you test out all three strategies and adjust the settings according to your needs and dataset. Performing extensive analysis and cross-validation is essential to choose the best algorithm for your leaf disease detection task. CNNs are frequently the go-to option for image-based leaf disease identification since they can automatically extract pertinent characteristics from raw pixel values. SVMs could work well with smaller datasets or in situations where efficient feature engineering is possible. KNN is simple to use and may not work as well with high-dimensional image data, but it can be taken into consideration for simpler applications.

References

1. Dr. Sharon M. Douglas Department of Plant Pathology and Ecology the Connecticut Agricultural Experiment Station123 Huntington StreetP. O. Box 1106 New Haven, CT 06504-1106.
2. A. A. Bhagat, C. D. Badgujar, S. S. Bhosale and V. S. Supe, "An economic analysis for export of fresh cut rose flowers from India," Journal of Pharmacognosy and Phytochemistry SP2: 291-298, 2019.
3. M. A. Haque, M. A. M. Miah, S. Hossain and M. Alam, "Profitability of rose cultivation in some selected areas of Jessore district," Bangladesh J. Agril. Res, vol. 38(1), pp. 165-174, 2013.
4. Anand. H. Kulkarni and Ashwin Patil R. K, "Applying image processing technique to detect plant diseases", International Journal of Modern Engineering Research (IJMER), Vol.2, Issue.5, Sep-Oct. 2012 pp-3661-3664 ISSN: 2249-6645 2012.
5. Rajleen Kaur, Sandeep Singh Kang, "An Enhancement in Classifier Support Vector Machine to ImprovePlant Disease Detection", IEEE 3rd International Conference on MOOCs, Innovation and Technology inEducation (MITE), 2015, pp. 135-140.
6. G. Suresh, V. Gnanaprakash, R. Santhiya, "Performance Analysis of Different CNN Architecture with Different Optimisers for Plant Disease Classification," In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), pp. 916- 921, IEEE, 2019.
7. A. P. Marcos, N. L. S. Rodova, A. R. Backes, "Coffee leaf rust detection using convolutional neural network", In 2019 XV Workshop de Visão Computational (WVC), pp. 38-42, IEEE, 2019.
8. S. S. Hari, M. Sivakumar, P. Renuga, S. Karthikeyan, S. Suriya, "Detection of plant disease by leaf image using convolutional neural network," In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), pp. 1-5, IEEE, 2019.
9. Dr. Sunil Bhutada and Subhani Shaik, "IPL Match Prediction using Machine Learning", IJAST, Vol.29, Issue 5, April-2020.
10. Monisha Singh, Dr. Sunil Bhutada, Dr. K Vijayalaxmi, Dr. Subhani Shaik," Online social networks fake news on covid-19 in the South India Region", Neuo Quantology, Vol.20, Issue 10, August 2022.
11. Dr. Vijayalakshmi K and Dr. Subhani Shaik, "predicting employee attrition methodologies of K-fold technique" I. J. Mathematical Sciences and Computing, March2023, 1, 23-36.
12. Subhani Shaik, P. Santhosh Kumar S. Vikram Reddy K. Sai Srinivas Reddy and Sunil Bhutada," Machine Learning based Employee Attrition Predicting", Asian Journal of Research in Computer Science, Volume 15, Issue 3, Page 34-39, March, 2023.
13. Subhani Shaik and Dr. Uppu Ravibabu, "Detection and Classification of Power Quality Disturbances Using curvelet Transform and Support Vector Machines", in the 5th IEEE International Conference on Information Communication and Embedded System (ICICES- 2016) at S. A. Engineering college, Chennai, India on 25th -26th, February 2016.
14. Yashwanth, K. Mahitha, S Bhavesh, Dr. N Ch Sriman Narayana Iyengar, Dr. Subhani Shaik," Detection of Crime hotspot using Deep learning algorithms", 6th International Conference on Soft computing and Signal Processing (Springer Conference), Malla Reddy College of Engineering and Technology (Autonomous), Hyderabad, June 23-24 2023.
15. D. Mani Chandra, D. Teja, S. Kiran Kumar, N Sreevidya, Dr. Subhani Shaik," WhatsApp Chat analysis and spam message detection using Machine learning algorithms", 2nd International Conference on Data Science and Artificial Intelligence (Springer Conference) (ICDSAI) 24-25 April 2023, Lendi College of Engineering (Autonomous), Vizianagaram, AP.
16. P. Koushik, P. Ruchitha, D. Adithya, Dr. Subhani Shaik, Dr. Sunil Bhutada," Deep learning approach for expression-based

songs recommendation system", 2nd International Conference on Data Science and Artificial Intelligence (Springer Conference) (ICDSAI) 24-25 April 2023, Lendi College of Engineering (Autonomous), Vizianagaram, AP.

17. Ch. Shravya, Pravallika and Subhani Shaik," Brest cancer prediction using machine learning Techniques", International Journal of Innovative Technology and Exploring Engineering, Vol. 8, Issue 6, 2019.
18. P. Santosh and Subhani Shaik," Heart disease prediction with PCA and SRP", International Journal of Engineering and Advanced Technology," Volume-8, Issue-4, 2019.