



Rose Plant Leaf Disease Recognition Using Machine Learning Methodologies

**Egamamidi Rishika Reddy ^a, Sai Durga Satturi ^a,
Medavarapu Harshini ^a and Subhani Shaik ^{a*}**

^a *Department of Information Technology, Sreenidhi Institute of Science and Technology (Autonomous), Hyderabad, India.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: <https://doi.org/10.9734/ajrcos/2024/v17i11519>

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/126899>

Original Research Article

Received: 10/09/2024

Accepted: 12/11/2024

Published: 18/11/2024

ABSTRACT

The most popular flowers in the world are 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. In this paper, we describe a novel deep learning-based method for the automated diagnosis of leaf diseases in rose plants. 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 specifically intended for picture classification tasks. We address the interpretability and explainability of the model's

*Corresponding author: E-mail: subhanicse@gmail.com;

Cite as: Reddy, Egamamidi Rishika, Sai Durga Satturi, Medavarapu Harshini, and Subhani Shaik. 2024. "Rose Plant Leaf Disease Recognition Using Machine Learning Methodologies". Asian Journal of Research in Computer Science 17 (11):65-72. <https://doi.org/10.9734/ajrcos/2024/v17i11519>.

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 Douglas, S. M. (n.d.).

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 (Bhagat et al., 2019). 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 (Suresh et al., 2019). 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 is 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 vigor, 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.

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 (Koushik et al., 2023).

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 (Shaik and Ravibabu 2016). 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 (Yashwanth et al., 2023).

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 (Marcos et al., 2019). 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 (Chandra et al., 2023). The model's capacity to distinguish between healthy and unhealthy plants may be impacted by the features selected.

The authors (Hari et al., 2019) improve the SVM classifier to better detect plant diseases. Two datasets are included in the author-implemented SVM: a training dataset and a training 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. Bhutada et al. (2020) used CNN to diagnose a variety of grape plant problems. Their three distinct CNN architectures are compared in order to diagnose plant ailments. (Bhutada and Shaik 2020, Vijayalakshmi et al., 2023) 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 (Bhutada and Shaik 2020). They created a dataset of 159 coffee leaf images to test their CNN addition for rust recognition. In its implementation, this study (Vijayalakshmi et al., 2023) makes use of ten different plant disease classifications. The (Vijayalakshmi et al., 2023) accuracy rate of the job is 86%.

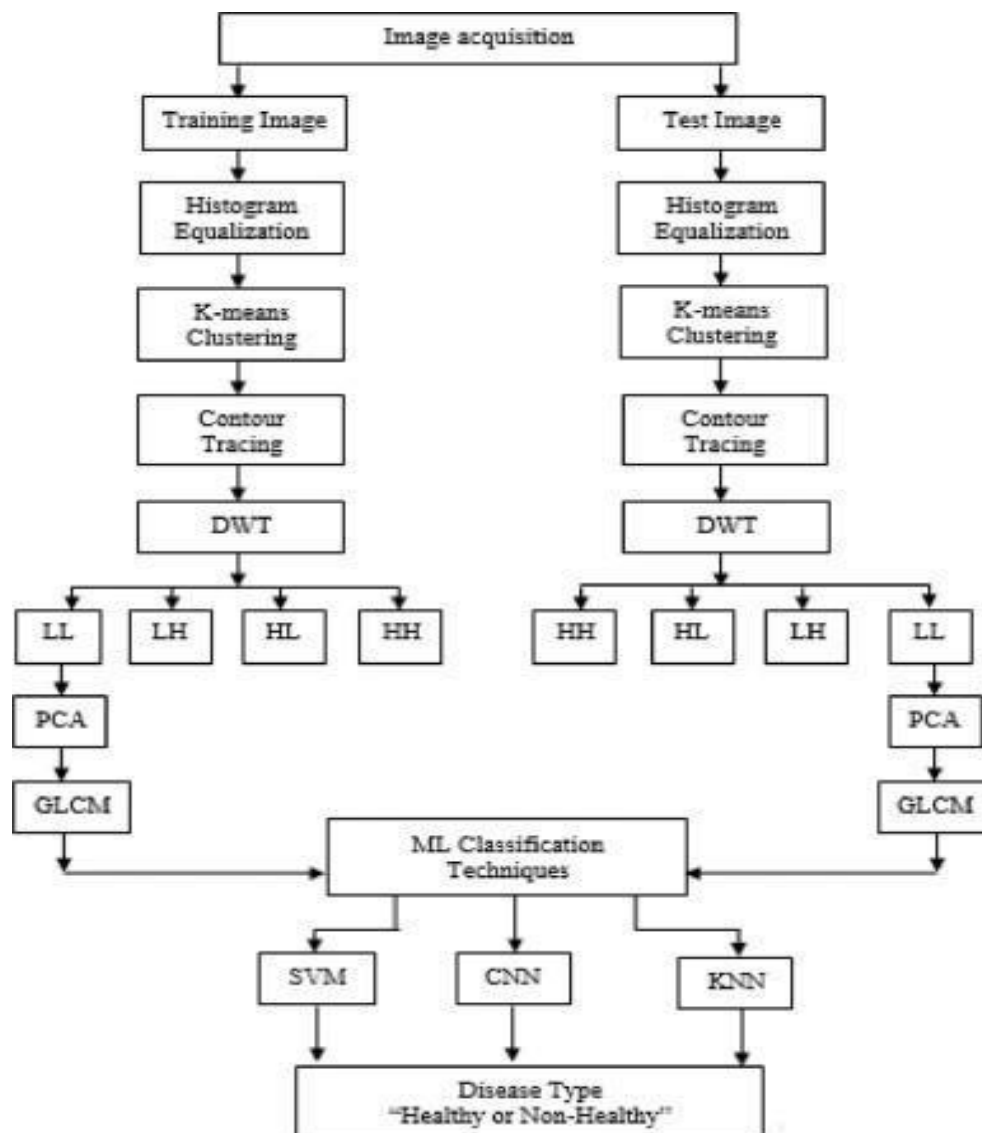


Fig. 1. Architecture Douglas, S. M. (n.d.)

3. METHODOLOGY

3.1 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 spots, and powdery mildew. Understand the common diseases that affect rose plants. Common diseases include black spots, powdery mildew, rust, and others.

3.2 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 (Shravya et al., 2019).

3.3 Feature Extraction

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 color, texture, and shape when extracting relevant features from segmented photos. The performance of the model can be greatly influenced by feature engineering (Santosh and Shaik2019).

2. Use pre-trained CNNs like VGG16, ResNet, or Inception. From big datasets like ImageNet, these networks have already learned hierarchical characteristics.
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.

3.4 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.

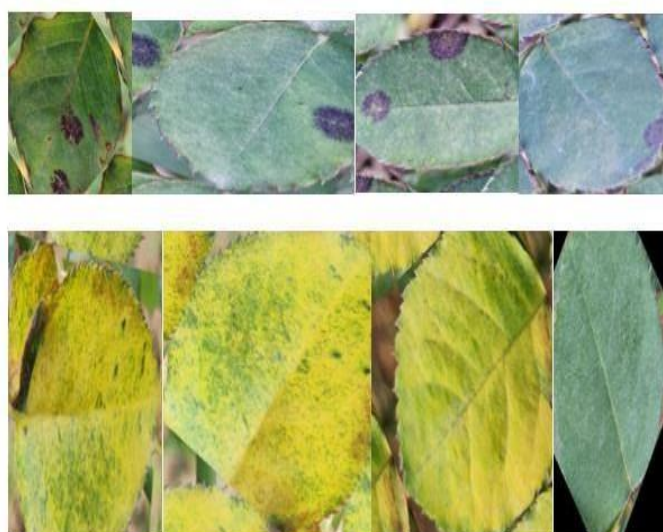


Fig. 2. Dataset

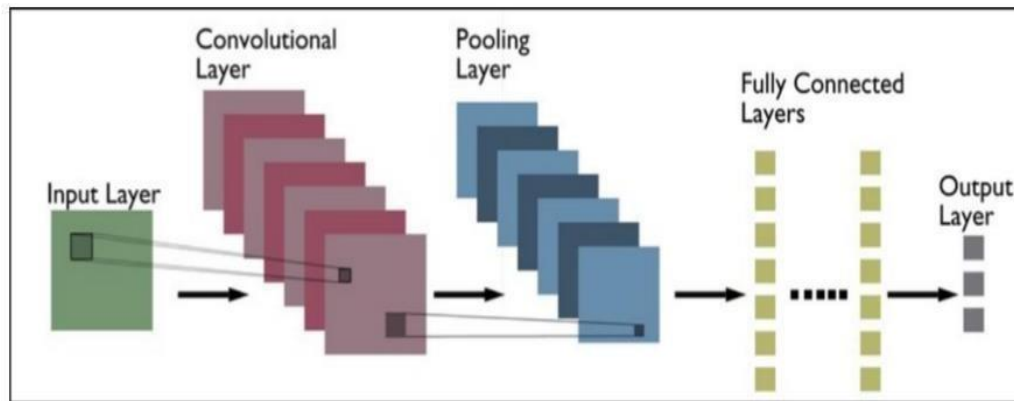


Fig. 3. CNN model Architecture

3.5 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.

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.

3.6 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.

3.7 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.

3.8 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.

3.9 Model Evaluation

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

3.10 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.

3.11 Deployment

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

3.12 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 regularly. To adjust to changes in disease patterns or environmental factors, the model should be updated as needed.

4. RESULTS AND DISCUSSION

We are using three different models to identify the diseases.

```

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]]
    
```

Image 1. CNN model

We got an accuracy of 60%.

```

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
    
```

Image 2. SVM model

We got an accuracy of 60%.

```

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
    
```

Image 3. KNN model

We got the accuracy as 50%.

Table 1. Comparative study of methods

S. No.	Name of the Algorithm	Accuracy Rate (%)
1	CNN	60
2	SVM	60
3	KNN	50

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.

Accuracy point of view CNN and SVM both methods perform same accuracy rate. However, research will give different results based on data size, training of data, and number of attributes. So many features impact predicted results.

5. 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.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of this manuscript.

ACKNOWLEDGEMENT

The authors of the first three members found the problem and then implemented a total project within one month. The fourth author has given guidance to students and the final implementation

was done by him. There is no financial support from any organization for this project.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

- Douglas, S. M. (n.d.). Department of Plant Pathology and Ecology, The Connecticut Agricultural Experiment Station. 123 Huntington Street, P.O. Box 1106, New Haven, CT 06504-1106.
- Bhagat, A. A., Badgujar, C. D., Bhosale, S. S., & Supe, V. S. (2019). An economic analysis for export of fresh cut rose flowers from India. *Journal of Pharmacognosy and Phytochemistry*, SP2, 291-298.
- Suresh, G., Gnanaprakash, V., & Santhiya, R. (2019). Performance analysis of different CNN architectures with different optimizers for plant disease classification. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp. 916-921). IEEE.
- Marcos, A. P., Rodova, N. L. S., & Backes, A. R. (2019). Coffee leaf rust detection using convolutional neural networks. In 2019 XV Workshop de Visão Computacional (WVC) (pp. 38-42). IEEE.
- Hari, S. S., Sivakumar, M., Renuga, P., Karthikeyan, S., & Suriya, S. (2019). Detection of plant disease by leaf image using convolutional neural networks. In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN) (pp. 1-5). IEEE.
- Bhutada, S., & Shaik, S. (2020). IPL match prediction using machine learning. *IJAST*, 29(5).
- Singh, M., Bhutada, S., Vijayalaxmi, K., & Shaik, S. (2022). Online social networks fake news on COVID-19 in the South India region. *NeuroQuantology*, 20(10).
- Vijayalakshmi, K., & Shaik, S. (2023). Predicting employee attrition methodologies of K-fold technique. *I. J. Mathematical Sciences and Computing*, 1, 23-36.
- Shaik, S., Kumar, P. S., Reddy, S. V., Reddy, K. S., & Bhutada, S. (2023). Machine learning-based employee attrition prediction. *Asian Journal of Research in Computer Science*, 15(3), 34-39.

- Shaik, S., & Ravibabu, U. (2016). Detection and classification of power quality disturbances using curvelet transform and support vector machines. In 5th IEEE International Conference on Information Communication and Embedded Systems (ICICES-2016), S. A. Engineering College, Chennai, India, 25th-26th February 2016.
- Yashwanth, K., Mahitha, K., Bhavesh, S., Iyengar, N. C. S. N., & Shaik, S. (2023). Detection of crime hotspots using deep learning algorithms. In 6th International Conference on Soft Computing and Signal Processing (Springer Conference), Malla Reddy College of Engineering and Technology (Autonomous), Hyderabad, June 23-24, 2023.
- Chandra, D. M., Teja, D., Kumar, S. K., Sreevidya, N., & Shaik, S. (2023). WhatsApp chat analysis and spam message detection using machine learning algorithms. In 2nd International Conference on Data Science and Artificial Intelligence (Springer Conference) (ICDSAI), 24-25 April 2023, Lendi College of Engineering (Autonomous), Vizianagaram, AP.
- Koushik, P., Ruchitha, P., Adithya, D., Shaik, S., & Bhutada, S. (2023). Deep learning approach for expression-based songs recommendation system. In 2nd International Conference on Data Science and Artificial Intelligence (Springer Conference) (ICDSAI), 24-25 April 2023, Lendi College of Engineering (Autonomous), Vizianagaram, AP.
- Shravya, C., Pravallika, & Shaik, S. (2019). Breast cancer prediction using machine learning techniques. International Journal of Innovative Technology and Exploring Engineering, 8(6).
- Santosh, P., & Shaik, S. (2019). Heart disease prediction with PCA and SRP. International Journal of Engineering and Advanced Technology, 8(4).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the publisher and/or the editor(s). This publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© Copyright (2024): Author(s). The licensee is the journal publisher. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
<https://www.sdiarticle5.com/review-history/126899>