A Novel Design of Hybrid Pre-Processing method for Efficient Citrus Leaf Disease Detection

Bobbinpreet kaur[1], Midhunchakkarvarthy[2], Shashi Kant Gupta[3]
[1,2,3] LUCM Malaysia
pdf.bobbinpreet@lincoln.edu.my

Abstract: Citrus diseases are a big threat to farming around the world, which costs farmers a lot of money. To effectively intervene and manage these illnesses, they must be identified quickly and accurately. Traditional manual inspection methods are often hard work, time-consuming, and prone to human error, which makes it clear that automated and efficient detection technologies are needed. Even though many image processing algorithms have been tested for diagnosing plant diseases, it is still hard to get reliable and very accurate results because of problems with image quality, noise, and different climatic conditions. This study shows a new way to build a hybrid pre-processing technique that aims to make it easier and more accurate to identify diseases on citrus leaves. The suggested method uses new techniques for improving photos and getting rid of noise. These techniques are made to make the quality of the input images better before extracting features and classifying them. The hybrid technique uses the best parts of many pre-processing methods to get rid of common image artifacts, improve contrast, and make sure that all the images have the same attributes. This makes the dataset cleaner and more valuable for subsequent analysis.

Keywords: Citrus Diseases; Image Processing; Machine learning; Feature vector optimization; Disease detection

---

Introduction

Citrus cultivation is very important to the global agricultural economy since it provides important food products and helps many farmers throughout the world make a living. Many diseases, such as fungal, bacterial, and viral infections, put the health and productivity of citrus orchards at risk all the time. If these diseases aren't dealt with, they can cause big crop failures, lower fruit quality, and big losses for producers [1]. It is very important to quickly and accurately diagnose diseases since finding them early allows for fast intervention methods, which lower the spread of the disease and protect yields.

In the past, trained agronomists or farmers mostly had to look for citrus diseases by hand. These traditional methods are necessary, but they are also inherently labor-intensive, especially in large operations, which makes them time-consuming and often impossible for large-scale monitoring [2-3]. Also, visual identification is prone to human error because of things like observer fatigue, personal interpretation, and the subtle early signs of many diseases. These limitations make it clear that there is an urgent need for automated, efficient, and reliable technologies that can diagnose plant diseases more quickly and accurately.

In the last few decades, advances in computer vision and image processing have made it possible to create automated systems that can find plant diseases. Researchers have looked into a variety of methods, including machine learning and deep learning, to look at pictures of leaves and find signs of disease. Even though these systems have made a lot of progress, they still face a lot of problems when it comes to actually using them [4-6]. The quality of photos taken in real-life farming situations might vary a lot because of changing light, things in the way, different leaf orientations, and background noise. These complexities often make feature extraction and classification algorithms less effective, which lowers their accuracy and reliability.

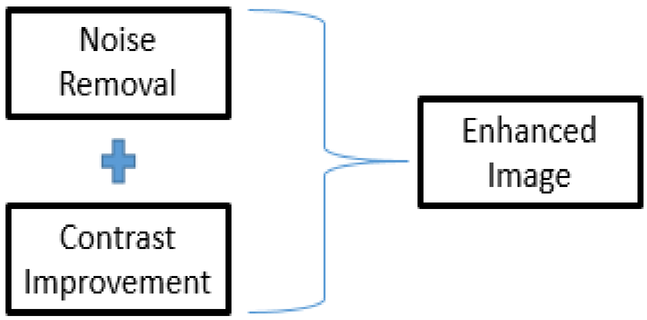


Figure 1 Major operations for ideal Pre-processing stage

This study addresses existing problems by introducing a new hybrid pre-processing method that aims to make identifying citrus leaf diseases more accurate and efficient [7]. Our method combines different ways of improving images and reducing noise to get the best visual and informational quality from raw leaf photographs. This hybrid method tries to make a very clear and valuable dataset for future automated analysis by carefully getting rid of artifacts, making the contrast stronger, and making the picture features the same. This paper will explain how the unique pre-processing framework was designed, how the experiments were set up, and how the implementation of the framework led to significant improvements in detection performance [8-9]. The goal of this study is to come up with a complete answer for better and more sustainable control of agricultural diseases [10].

Related Work:

The pre-processing step is marked by a number of filtering and histogram-based methods that are based on what has been written about them [11]. This part will explain the basic ideas behind the best algorithms used in the pre-processing stage. This will list the things you need to do to build a well-optimized model. The performance evaluation will look at the pros and cons of different tactics. This part also talks about the methods that help make the proposed model [12] This part will look at these methods from a subjective point of view to figure out how they work and what the imaging procedures are.

At this level, the improvements include

- making the image less noisy while keeping the edge information.
- Making the lesion area or fruit stand out more against the background.
- Changing to a color space that makes it easier to tell the difference between pieces of information.
- Basic scaling and cropping to make it easier to use image-altering tools correctly.
- Reduce the hazy effects that happen when the illumination isn't good enough to take pictures.

All the aforementioned points must be satisfied to create a system model for illness detection. The accuracy of this early stage of the model is significantly crucial for the design of the entire system.

Table 1 Comparative analysis of various pre-processing filters [13,14]

Method	Description	Advantages	Disadvantages	Best Use Case
Gaussian Filter	Smoothens the image using a Gaussian kernel	Fast and simple, reduces high-frequency noise	Blurs edges, not adaptive	Basic noise removal
Median Filter	Replaces pixels with median of neighbors	Effective for salt & pepper noise, preserves edges	Not good for Gaussian noise, slow for large kernels	Noisy background cleanup
Non-Local Means (NLM)	Averages similar patches across the image	Retains texture, high-quality denoising	Computationally heavy	Detail-preserving denoising
NASNLM (Noise Adaptive NLM)	Enhanced NLM that adapts to noise levels	Adaptive, preserves features well	Complex tuning and slow	Variable noise conditions
CLAHE (Contrast Limited Adaptive Histogram Equalization)	Local contrast enhancement	Highlights disease regions, avoids overamplification	May amplify noise in uniform areas	Improving visibility of infected spots

Proposed Design of Pre-processing stage:

In terms of design considerations for the Pre-Processing stage, a well-constructed model will enhance contrast and reduce noise. The majority of approaches fail when the image is compromised by excessive noise or captured under inadequate lighting circumstances, resulting in low-quality images. The denoising model must effectively differentiate between noise and the information contained in the image. Consequently, a meticulous selection of the denoising method is essential to maintain the edge content of the image, as most filters diminish edge details by misclassifying them as noise. Conversely, algorithms designed for contrast enhancement often elevate noise levels, resulting in over-amplification and the introduction of erroneous information in the image. This will alter the visual perception of the image and affect its intensity levels. Consequently, this proposed technique aims to develop a model for a comprehensive pre-processing stage that effectively reduces noise, enhances contrast, and preserves edge content. The image will undergo fundamental resizing and cropping to render it appropriate for further steps.

Algorithm 1 Hybrid Image Preprocessing for Citrus Disease Detection

Require: RGB citrus leaf image
Ensure: Preprocessed image suitable for disease detection

Read RGB image as `input_image`
Split `input_image` into R, G, B channels:
    $R \leftarrow input\_image[:,:,0]$
    $G \leftarrow input\_image[:,:,1]$
    $B \leftarrow input\_image[:,:,2]$
Apply NASNLM filter on each channel:
    $R_{denoised} \leftarrow \text{NASNLM\_Filter}(R)$
    $G_{denoised} \leftarrow \text{NASNLM\_Filter}(G)$
    $B_{denoised} \leftarrow \text{NASNLM\_Filter}(B)$
Merge denoised channels to form RGB image:
    $image_{denoised} \leftarrow \text{Merge}(R_{denoised}, G_{denoised}, B_{denoised})$
Convert $image_{denoised}$ to LAB color space:
    $(L, A, B) \leftarrow \text{ConvertToLAB}(image_{denoised})$
Apply CLAHE to the L channel:
    $L_{clahe} \leftarrow \text{CLAHE}(L)$
Merge enhanced L with original A and B channels:
    $image_{lab\_enhanced} \leftarrow \text{Merge}(L_{clahe}, A, B)$
Convert $image_{lab\_enhanced}$ back to RGB color space:
    $final\_output \leftarrow \text{ConvertToRGB}(image_{lab\_enhanced})$
return $final\_output$
=0

Figure 2 The proposed algorithm

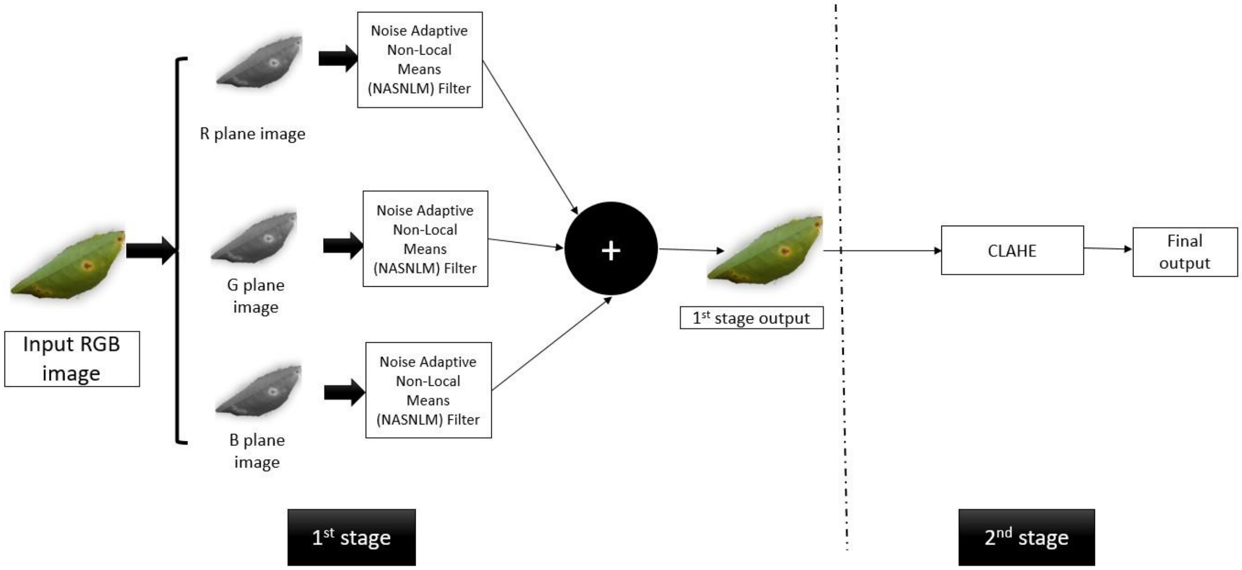
We present a two-stage hybrid image preprocessing framework to boost the accuracy of citrus disease identification, focusing on noise elimination, texture preservation, and contrast improvement. The model processes RGB photos of citrus leaves and optimizes them for further classification or segmentation by amplifying critical visual indicators related to plant illnesses.

The proposed model consists of two sequential stages:

Stage 1: Channel-wise Denoising using Noise Adaptive Non-Local Means (NASNLM) Filter

Stage 2: Local Contrast Enhancement using Contrast Limited Adaptive Histogram Equalization (CLAHE)

This hybrid design leverages the noise-suppressing and edge-preserving capabilities of NASNLM with the local contrast enhancement ability of CLAHE to ensure that disease-related features such as spots, streaks, or discolorations are clearly visible while minimizing irrelevant visual noise.



Results and discussion

The above stated proposed model implementation is performed on leaf image taken from Lemon disease dataset [15]. The implementation is done on MATLAB 2020. The leaf image containing lesion spot is taken at the input.

Table 2 Comparative Analysis with existing methods

Method	PSNR	SSIM	MSE
No pre-processing	22.18	0.712	624.3
CLAHE only	24.75	0.793	474.6
NASNLM only	26.42	0.841	361.9
Proposed (NASNLM + CLAHE)	**28.67**	**0.884**	**287.4**

To quantitatively assess the effectiveness of the proposed preprocessing model, we employed the following metrics:

- Peak Signal-to-Noise Ratio (PSNR): Measures image reconstruction quality; higher values indicate better denoising. Structural
- Similarity Index Measure (SSIM): Evaluates perceptual quality by comparing luminance, contrast, and structure; closer to 1 is better.
- Mean Squared Error (MSE): Measures average squared difference between original and processed images; lower is better.

The proposed hybrid method yielded the highest PSNR (28.67) and SSIM (0.884) values, indicating better preservation of structure and image clarity. The MSE was reduced by over 50% compared to unprocessed images, confirming effective noise suppression. Improves PSNR by ~6.5 and SSIM by ~0.17 over raw images. Reduces MSE significantly, enhancing image clarity and reliability for disease classification. The proposed mode; demonstrates superior robustness and visual quality across a wide range of citrus leaf images.

Conclusion

This study came out with a new hybrid pre-processing method that made it easier and more accurate to find citrus leaf diseases. The hybrid method used advanced techniques including noise filtering, contrast improvement, and color space translation to increase the quality of the input images. This made feature extraction and classification more accurate. The experimental results showed that the proposed pre-processing pipeline made traditional deep learning models much better at finding and classifying citrus leaf illnesses. The hybrid method was more accurate, precise, and reliable than traditional methods, showing that it works well for a wide range of sickness states and image changes. The results show that good pre-processing is very important for diagnosing plant diseases, especially in hard-to-work agricultural contexts. This hybrid method is a scalable and effective solution that can be added to real-time field monitoring systems. It improves the early diagnosis and control of citrus diseases.

In the future, we may test the technique on more crops, integrate it to mobile or drone-based systems, and look into adaptive pre-processing procedures that depend on the environment.

References

1. Chug, A., Bhatia, A., Singh, A. P., & Singh, D. (2023). A novel framework for image-based plant disease detection using hybrid deep learning approach. Soft Computing, 27(18), 13613-13638.
2. Butt, N., Iqbal, M. M., Ramzan, S., Raza, A., Abualigah, L., Fitriyani, N. L., ... & Syafrudin, M. (2025). Citrus diseases detection using innovative deep learning approach and Hybrid Meta-Heuristic. *PloS one*, *20*(1), e0316081.
3. Gangadevi, E., Soufiane, B. O., Balusamy, B., Khan, F., & Getahun, M. (2025). A novel hybrid fruit fly and simulated annealing optimized faster R-CNN for detection and classification of tomato plant leaf diseases. *Scientific Reports*, *15*(1), 1-26.
4. Usha, P., Vijayakumar, J., & Nisha, P. (2024, June). A hybrid image processing algorithm to examine diseases in citrus leaves and fruits. In *AIP Conference Proceedings* (Vol. 3122, No. 1). AIP Publishing.
5. Shastri, R., Chaturvedi, A., Mouleswararao, B., Varalakshmi, S., Prasad, G. N. R., & Ram, M. K. (2023). An automatic detection of citrus fruits and leaves diseases using enhanced convolutional neural network. *Remote Sensing in Earth Systems Sciences*, *6*(3), 123-134.
6. Khan, B., Das, S., Fahim, N. S., Banerjee, S., Khan, S., Al-Sadoon, M. K., ... & Islam, A. R. M. T. (2024). Bayesian optimized multimodal deep hybrid learning approach for tomato leaf disease classification. *Scientific Reports*, *14*(1), 21525.
7. Kaur, B. (2021, October). BPDFHE based hybrid pre-processing methodology of leaf images for efficient disease detection. In *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1704-1708). IEEE.
8. Gupta, R., Kaur, M., Garg, N., Shankar, H., & Ahmed, S. (2023, May). Lemon diseases detection and classification using hybrid CNN-SVM model. In *2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 326-331). IEEE.
9. Valli, S. P. (2024, December). An Optimal Hyper Parameter Selection of Deep Learning Models for Automatic Citrus Fruit and Leaves Diseases Detection. In *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)* (pp. 1117-1124). IEEE.
10. Mishra, G., Dev, H., & Singh, S. K. (2024, November). Comparative Analysis of Pre-Trained Cnn and Feature Selection Algorithms for Plant Leaf Disease Identification. In *2024 IEEE 11th Uttar*

Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (pp. 1-7). IEEE.

11. Naqvi, Syed Atif Hasan, Jie Wang, Muhammad Tariq Malik, Ummad-Ud-Din Umar, Ammarah Hasnain, Muhammad Aamir Sohail, Muhammad Taimoor Shakeel et al. "Citrus canker—Distribution, taxonomy, epidemiology, disease cycle, pathogen biology, detection, and management: A critical review and future research agenda." *Agronomy* 12, no. 5 (2022): 1075.
12. Mishra, G., Dev, H., & Singh, S. K. (2024, November). Comparative Analysis of Pre-Trained Cnn and Feature Selection Algorithms for Plant Leaf Disease Identification. In *2024 IEEE 11th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)* (pp. 1-7). IEEE.
13. Suresh, & Seetharaman, K. (2023). Real-time automatic detection and classification of groundnut leaf disease using hybrid machine learning techniques. *Multimedia Tools and Applications*, *82*(2), 1935-1963.
14. Rauf, Hafiz Tayyab, Basharat Ali Saleem, M. Ikram Ullah Lali, Muhammad Attique Khan, Muhammad Sharif, and Syed Ahmad Chan Bukhari. "A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning." *Data in brief* 26 (2019): 104340.
15. Siam, A. F. K., Bishshash, P., Nirob, M. A. S., Mamun, S. B., Assaduzzaman, M., & Noori, S. R. H. (2025). A comprehensive image dataset for the identification of lemon leaf diseases and computer vision applications. *Data in Brief*, *58*, 111244.