SN Computer Science (2023) 4:668
https://doi.org/10.1007/s42979-023-01966-z

ORIGINAL RESEARCH

# Image Processing-Based Disease Prediction in Medical Plant Leaf

**P. Loganathan[1] · A. Muthukumaravel[2]**

Received: 19 February 2023 / Accepted: 23 May 2023


**Abstract**

The most crucial aspect of every nation is its agriculture. The production of and demand for medical plant leaf products are expanding quickly day after day. Farmers have a challenging duty because of the medical plant leaf disease prognosis in agriculture. Recapturing an image involves using a printout, scanned copy, projection, or Liquid Crystal Display (LCD) monitor to take it again after it has already been captured. In this research, a novel technique for selecting and extracting medical plant leaf disease feature sets using image processing is presented. MatLab is used for feature selection and scheme extraction. In businesses where the sorting process is heavily dependent on human labor, technology is crucial. To conduct the classification, Support Vector Machine (SVM) control topology is used. The suggested method is a pioneer in the use of image processing for classifying and evaluating medical plant leaves. One of the most sought-after techniques being used in the image processing business is support vector machine. Comparative to traditional methods, sorting and grading are more affordable and effective when SVM-based implementation is used.

**Keywords** Medical plant leaf · Quantization · Reconstruction · Edge detection · Edge map · ROC · Environment · Agriculture · Yield · MSE · PSNR · SSIM · Hybrid model · Technological · PSO · Image processing

## Introduction

With technological growth, most of the Agriculture Sector becomes an automated environment. This makes it easy implementation and easy to identify the medical plant leaf identification to users [1]. Digital cameras are used widely today. They produce real-like images. Printers and LCD monitors are used to retake images. Image recapture detection (IRD) is a digital forensic method used to differentiate between single-captured and recaptured images [1]. As various computer graphics and photo editing tools are available, determining whether the image is single-captured or recaptured is difficult. Human eyes are not able to distinguish such images. Images recaptured from printouts or LCD displays have different scenarios as they are taken from the reproduction surface [2].

IRD is also useful in improving the productivity of robot vision. Images can be modified to enhance quality known as composite images. Identifying such images is also a substantial function of IRD. Even in criminal investigation, IRD is an important task [2]. Whether a journalist presents a report, it is true or not, the images presented are real or not can be detected by IRD. IRD also plays an important role in intelligence analysis, such as driving unmanned vehicles, robot training, etc.

Various outstanding problems are raised in medical plant agriculture such as low utilization and high inputs of phosphate fertilizers causing a nutrient imbalance in the plants [3]. Farmers used old and traditional methods of medical plant cultivation but that requires an enormous amount of hard work and attention. Therefore, it provides difficulties in implementing traditional cultivation techniques and also time-consuming, and several processes are implemented which are not understandable by the farmers. Another problem is the determination of fertilizer elements is a major

This article is part of the topical collection “Industrial IoT and Cyber-Physical Systems” guest edited by Arun K Somani, Seeram Ramakrishnan, Anil Chaudhary, and Mehul Mahrishi.

✉ P. Loganathan
logup75@gmail.com

A. Muthukumaravel
dean.arts@bharathuniv.ac.in

[1] Department of Computer Applications, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India

[2] Faculty of Arts and Science, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India

Published online: 30 August 2023

part of Agriculture. Another problem is climate conditions; the agriculture process takes a lot of money invested, but sometimes due to improper climate conditions, the crops are getting damaged and a huge losses for farmers [4]. Temperature control and humidity control are the main problem of medical plant cultivation. Temperature plays an important role in agriculture; it includes photosynthesis, transpiration, absorption, respiration, and flowering, so these things are monitored by the temperature sensors; another factor is the humidity which includes various abnormal states like unbalance conditions and a humid atmosphere; these data are collected by the humidity sensors [4].

To reduce the heavy loss and increase the output in agriculture, it requires new techniques and methods for farmers to detect the climate conditions before the cultivation process [5]. In agriculture, various parameters are increasing crucial, such as efficient resource management, timely monitoring, and persistent data. Spatial variables include soil properties, plant population, topography, fertilization, and irrigation and pest infestations. Various commercial products are available for crops, such as electronic sensors, Global Positioning Systems, harvester-mounted yield monitors, etc. The important reason for low crop production is the tiny size of the land holding and lack of handling modern technology in Agriculture. Precision Agriculture is one of the ways of adopting technology in the Agriculture sector. In today's agriculture, data are collected from sensors and examined from that results take decisions accordingly to increase medical plant crops, production [6].

This paper introduces an advanced scheme which enables the sorting of disease prediction based on image processing. The sorting makes use of a Support Vector Machine scheme to implement the process of disease prediction. A combination of an image processing schemes and microcontroller enables the proper and complete sorting.

## Background Study

We have surveyed 12 different IRD methods, such as wavelet decomposition, geometry-based, specularity distribution, forensic features, physics-based features, LBPV, image acquisition, multiple features, K-Singular Value Decomposition (K-SVD) learning, color moments and DCT coefficients, and edge profiles. We mentioned these methods because of their technology (extracted feature) differences. In wavelet decomposition [6], image statistics are used for the classification of retaken images. Geometry-based method [6] considers the image generation process as light, object, etc. In the specularity distribution method [7], spatial component of an image detects whether it is real or retaken. Various forensic features, such as Local Binary Pattern (LBP), Multiscale wavelet statistics, and color, are used [8]. Physics-based feature [7] we can say is a hybrid model of above-mentioned methods. LBPV [8] is an extension of LBP, as it considers rotation variation. Image acquiring process [8], i.e., use of a sensor, analog-to-digital camera converter, can also be used for detection. K-SVD [8] uses soft edges for the disclosure of retaken images. DCT is also used with color [8]. Edge profile [9] is an enhanced version of K-SVD where aliasing is also considered along with blurriness and the process is made completely automatic.

The output of the nation's agriculture greatly influences the Indian economy. Grape is a very popular fruit. All tropical, continental, and moderate climate zones are suitable for its easy cultivation. Various regions of India have various climatic conditions and kinds of soil. This renders fruit trees a significant crop with a high socio-economic value that is reproduced vegetatively. When the grapevine plant has illnesses, production and development will be poor. The illnesses are brought on by bacterial, viral, and viral diseases that are brought on by pests, rust, and microbes, among other things. Farmers diagnose these diseases based on their knowledge or with the assistance of experts through naked-eye inspection, which is a laborious and inaccurate task [10].

Over the past 5 years, nearly 300 research magazines were available on Google Scholar (GS) and more than 150 research magazines were offered in science direct journals with various DL concepts. Numerous publications on leaf disease prediction using ML and DL methods have been published recently. Because of the rising likelihood of developing leaf diseases, the current research utilizes cutting-edge IT techniques like DL algorithms, which can be used to forecast diseases from the grape leaf online dataset.

China is a developed agriculture-based society with a more than 2000-year tradition of growing grapes. China currently exports more grapes than any other country on the globe. Grape juice, raisins, wine, and other grape-related goods are also very valuable commercially. However, serious infections greatly reduce grape output and quality during the germination period, particularly in wet regions and times. To maintain the proper growth of grapes, fast and accurate diagnosis of grapevine leaf illness is therefore essential [11].

By substituting outdated practices with more effective ones, AI is revolutionizing farming and making the world a healthier place. Gadgets that can recognize and manage plants, weeds, pests, and illnesses through RS (Remote Sensing) are being developed and put into use thanks to AI and ML. The fungal disease reduces the yield and crop quality that are crucial to the Indian economy, including those that are used to produce food, fiber, and biofuel. DL technology helps farmers reach markets more quickly and with greater quality while also reducing waste [12].

The majority of the current methods for illness detection rely on visual identification. Unfortunately, not only is visual detection a tedious and time taken operation, but also the accuracy of identification falls short of the need. The misuse of insecticides that results from the incorrect diagnosis can ruin the grapes' growing conditions and degrade the fruit's quality [13].

Due to the constantly changing weather and climatic factors, illnesses are relatively common in crops. Crop diseases are typically difficult to prevent and have an impact on the development and output of crops. Effective illness detection and timely disease management measures are essential for ensuring massive yield and great quality. The widely cultivated grape plant in India is susceptible to various diseases that can harm the fruit, branch, and leaves. Leaf illnesses are the early signs of fungus, bacterium, and virus infection. Therefore, a system that can automatically identify the types of illnesses and perform the necessary activity is required [14].

## Proposed Methodology

Noise and unwanted features cause major hindrances to efficient recognition in any recognition system. In medical plant leaf recognition systems, the leaf is the prime barrier to accurate recognition. Hence, tackling these problems to increase the efficiency of the recognition system is of utmost importance. One of the main characteristics of an efficient recognition system is speed. Obtaining results in real time is a daunting challenge in recognition systems. The proposed method helps in this cause.

### Spectrum-Based Feature Extraction

Feature Extraction plays a vital role in image segmentation. Here, Feature Extraction is carried out with the combination of DFT and DCT. DFT helps in converting the image which is from the spatial domain to the frequency domain and necessary information is shifted toward the center position [9].

#### Feature Selection Using Proposed Position-Based Binary Particle Swarm Optimization (PBPSO)

Feature Selection plays next important role in selecting the important features used for classification. This is done by Proposed Position-based Binary Particle Swarm Optimization (PBPSO) which mainly targets on finding the right position in optimizing the best feature [15].

### Image Pre-processing

Image acquisition [16] is the first major step where the image is used for testing from the database. Pre-processing helps in removal of noise using various filtering techniques.

The steps are as follows:

- The RGB image is converted to black and white image.
- Then, Histogram equalization is applied on the gray image which helps in distributing the intensities over a wide range, thus increasing the contrast of the image.
- Using sharp masking method, the image is further enhanced to highlight its edges by sharpening.
- Then, the unsharped masked image will be added to the original gray image

$$\text{img} = \text{c1} * \text{img1} + \text{c2} * \text{img2}. \quad (1)$$

The algorithm is as follows:

Step 1: Initialize the cognitive, social, and inertial elements (c1, c2, and). Additionally, initialize the particle's Xit position within the search space.
Step 2: Assess each particle's performance using the fitness function while taking into account its current position.
Step 3: Compare each particle's performance to its best performance to get its personal best position (pbest), and then update the pbest.
Step 4: Update the global best gbest in a similar manner by evaluating each particle's performance in comparison to the global best particle.
Step 5: Adjust the velocity now.
Step 6: The position update is completed by combining the initial PSO and BPSO.
Step 7: Repeat step 2 as necessary to reach convergence or reach the end of the iterations.

### SVM-Based Disease Prediction Scheme

The basic of Image Processing is the extraction of features from the captured images and then analyzing them to understand the characteristics and then classifying them. The processes including optimization and data mining makes use of the feature selection process.

The proposed system is shown using a flowchart in Fig. 1.

The flowchart shows a brief description of the steps involved in the Capturing, Analysis, Processing, and Validation of the result and based on which the classification is performed. The main steps involved in the process are as follows:

- Image capturing
- Pre-processing
- Region segmentation
- K-Mean clustering
- Feature extraction

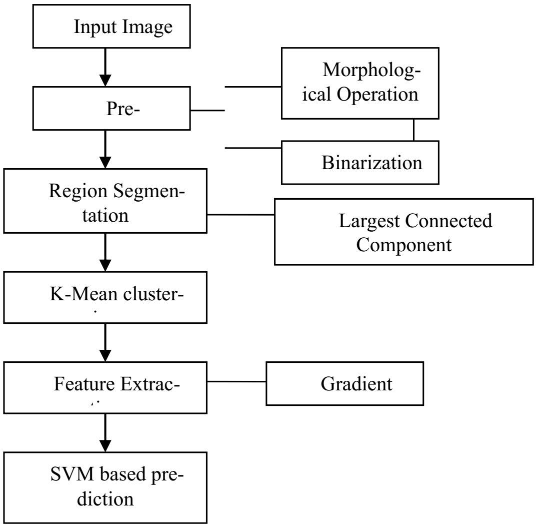


**Fig. 1** Proposed system architecture

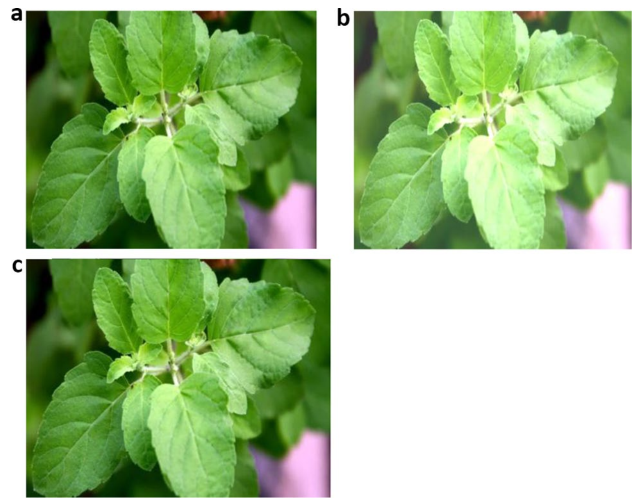


**Fig. 2** **a** Input image, **b** Morphological operations, **c** Pseudo-median result

- Model development and classification

### Importance of the Proposed Work

The pre-processing stage involves a combination of two processes which are Morphological operation and Binarization. The two fundamental morphological operations are dilation and erosion. It is also known as the Morphology opening result, while Erosion is also known as the Morphology Closing result. The two morphological opening and closing processes remove all the scattered dark and white pixels. In dilation, the image region is grown. In erosion, the image region is shrunk. In morphological operation, the system applies a structuring element to an input image creating an output image of the same size. The process of Dilation adds pixels to the boundary of the object in an image; on the other hand, erosion removes the pixels from the object's boundaries.

In the Pseudo-median filtering process, it replaces each entry with the median of the neighboring entries. The advantage of Pseudo-median operation is that it undertakes the removal of Gaussian noise which could not be effectively removed during the dilation and erosion processes. In Binarization, it converts the RGB image into a binary image based on threshold values. The threshold values are determined using the Otsu algorithm. The following Fig. 2b represents morphological operations and Fig. 2c shows Pseudo-median result.

### Region Segmentation

In Fig. 3, the region segmentation process is performed using the graph contour tracking method. Using the graph contour tracking approach, the system determines all connected components. This is then implemented using Matlab which removes objects that have an area greater than or less than defined threshold values; thereby enabling the extraction of objects which have an area in the defined region.

### K-Mean Clustering

The Fig. 4 shows K-Mean clustering. Clustering is a process of organizing objects into groups that is having identical properties. In K-Mean Clustering, the system partitions '*n*' observations into K-Cluster in which each observation belongs to the cluster with the nearest mean. This is implemented using an image Histogram, where the mean value is determined. The process involves breaking the image into K regions by reducing the number of colors to K and mapping each pixel to the closest color.

### Feature Extraction

The Fig. 5 shows the process of feature extraction is vital for disease prediction. Here, based on the gradient value at the Apex, Equator and Stalk region the R value, G value, and B value at the regions are recorded. The obtained value is then provided to the SVM algorithm for supervisory training.

### Model Development and Classification

The SVM then operates on the received data. The classifier has been defined to classify the result into two regions, namely, disease predicted and not predicted. The SVM works on the received data, and then, based on the trained classifier, the data are compared and then sorted to the

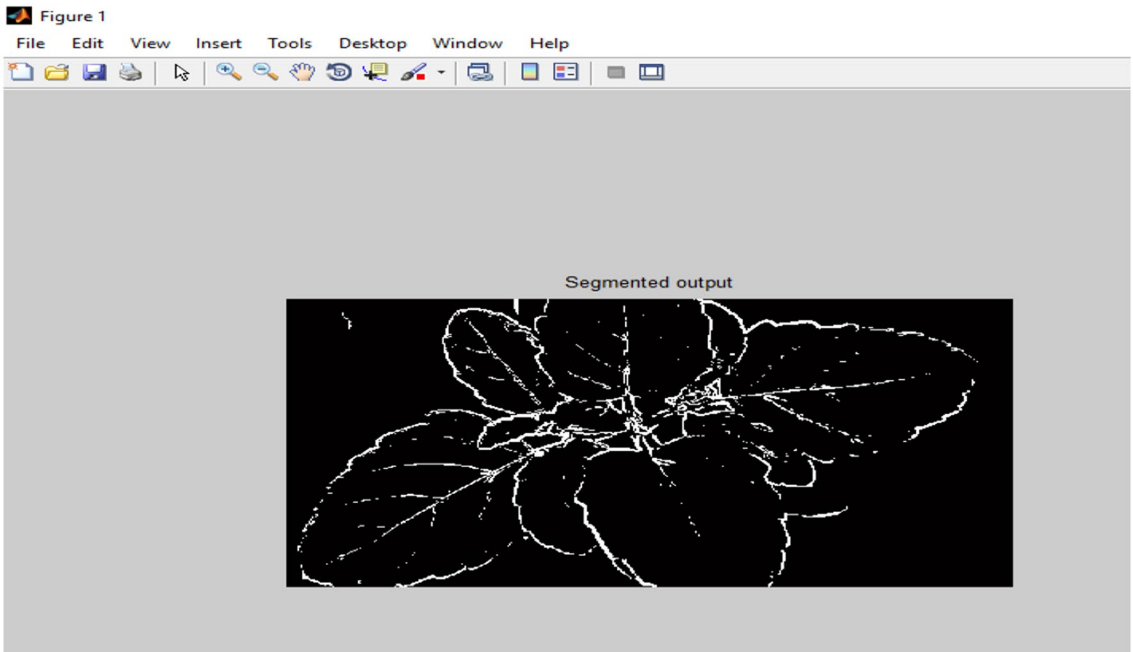


**Fig. 3** Region segmentation

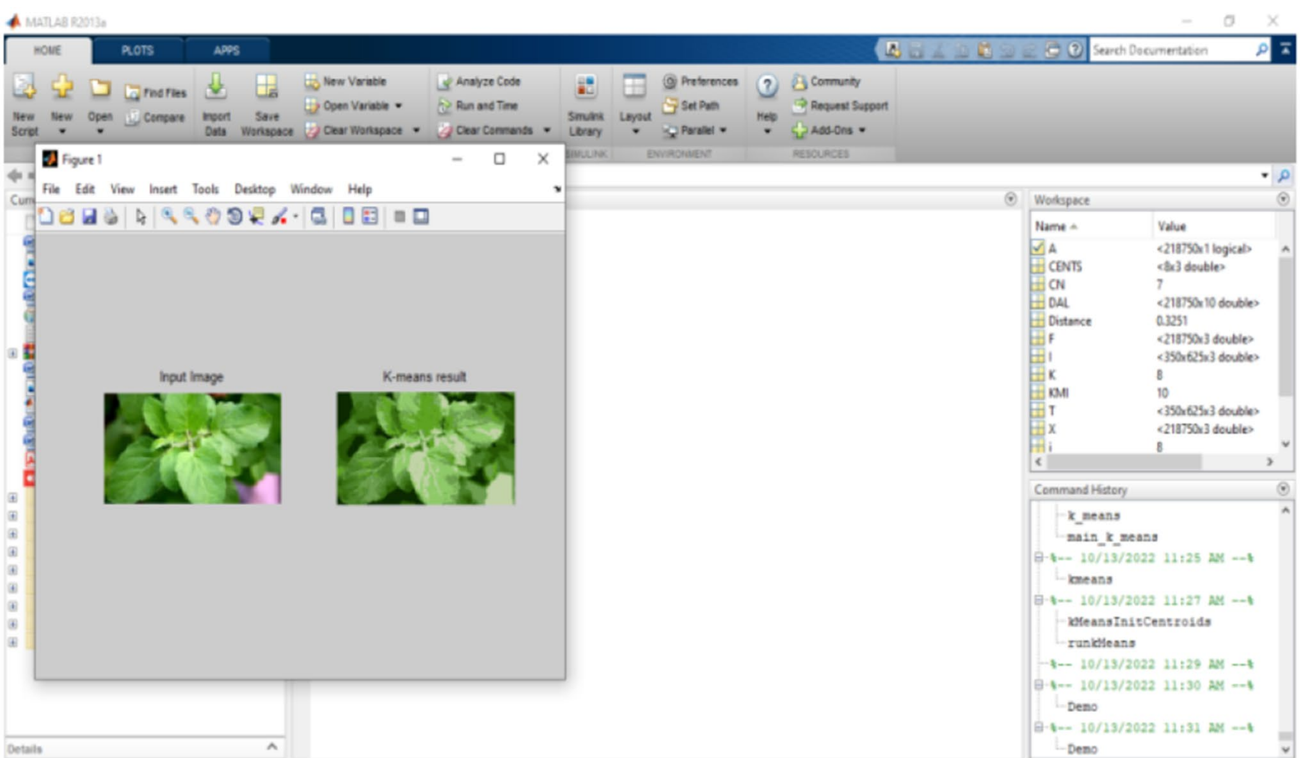


**Fig. 4** K-means clustering

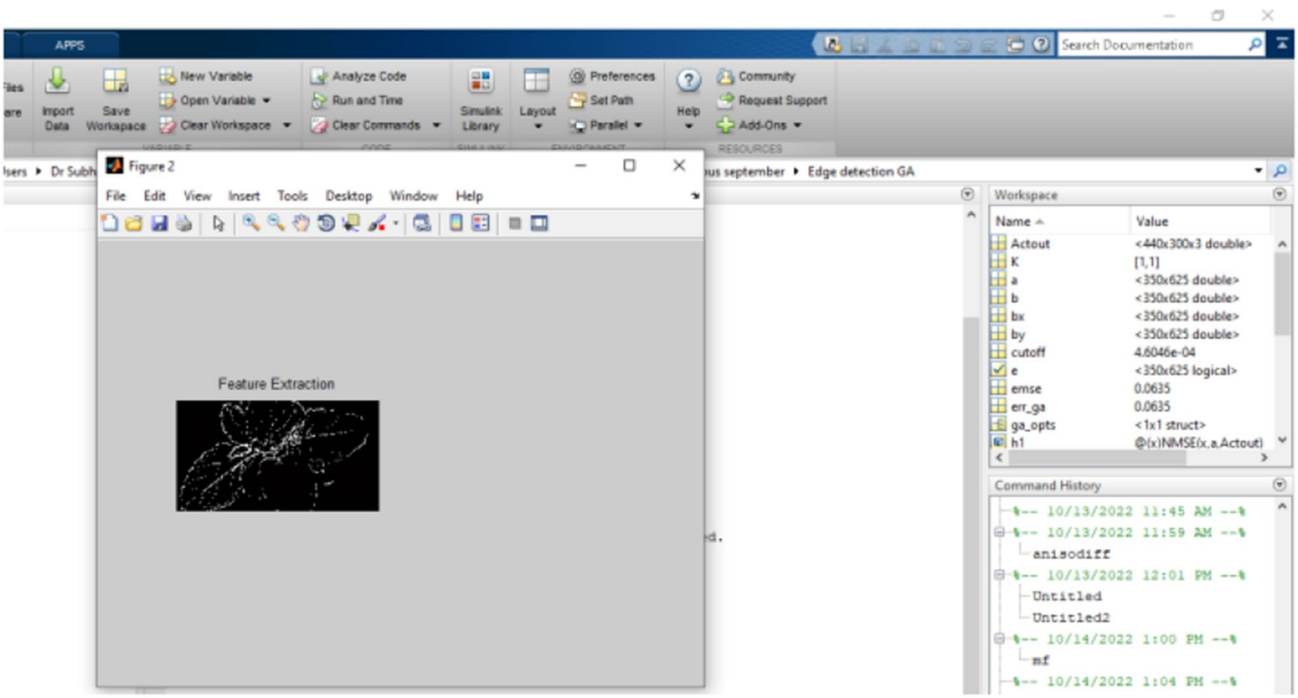


**Fig. 5** Feature extraction

defined classification. Once the class has been identified, a result is displayed which depicts if the medical plant leaf under analysis is disease predicted or not predicted.

## Results and Discussion Analysis

The proposed method is simulated in MATLAB R2014a using different grayscale test images for various quality parameters, such as Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM), as follows [17].

### MSE Analysis

The MSE is computed using the following formula:

$$MSE = \frac{1}{M\,X\,N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(f_{ij}-\hat{f}_{ij}\right)^{2}. \quad (2)$$

From the above equation, **N**, **M** denote the width and the length of the given image, $\mathbf{f_{ij}}$ represents the given image pixel value, and $\hat{\mathbf{f}}_{\mathbf{ij}}$ indicates the denoised picture pixel data.

### PSNR Analysis

The value of PSNR is measured by Eq. (3)

$$PSNR = 10{*}\log\frac{L^{2}}{MSE}. \quad (3)$$

Here, $L$ indicates the image pixel's highest greyscale value.

### SSIM Analysis

The SSIM term retrieves three major attributes luminance, contrast, and structure from the given input image.

Table 1 represents the output quality measures between existing and proposed method.

Figures 6 and 7 represent the output analysis of two different medicine leaves. Figure 6 shows the results obtained from the first leaf which produces MSE of about 22.7635, PSNR of about 34.8008 and SSIM of about 0.9292, whereas the existing system produces MSE of about 49.9554, PSNR of about 31.1450, and SSIM of about 0.8808. From the result obtained, it is proved that our proposed system works better than existing system.

Figure 7 shows the results obtained from the first leaf which produces MSE of about 16.6443, PSNR of about 35.9182, and SSIM of about 0.9123, whereas the existing

**Table 1** Output quality measures between existing and proposed method

| S. no. | Input images | Quality measures | Existing method [18, 19] | Proposed method |
|---|---|---|---|---|
| 1 |  | MSE | 49.9554 | 22.7635 |
| | | PSNR | 31.1450 | 34.8008 |
| | | SSIM | 0.8808 | 0.9292 |
| 2 |  | MSE | 33.9021 | 16.6443 |
| | | PSNR | 32.8285 | 35.9182 |
| | | SSIM | 0.8488 | 0.9123 |

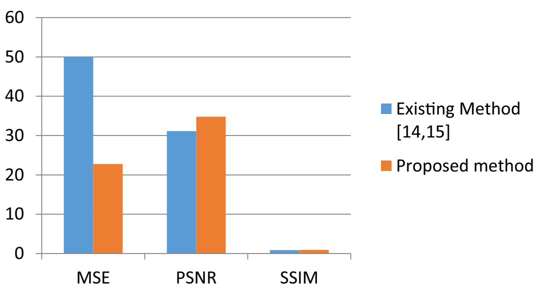


**Fig. 6** Output quality measures between existing and proposed method of first leaf

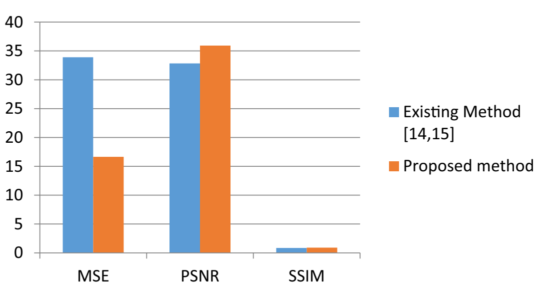


**Fig. 7** Output quality measures between existing and proposed method of second leaf

system produces MSE of about 33.9021, PSNR of about 32.8285, and SSIM of about 0.8488. From the result obtained, it is proved that our proposed system works better than the existing system.

The SVM-based disease prediction control scheme implemented in Matlab is providing the results as shown in the above table. The classifier has been defined to separate the medical plant leaf into two groups which are disease predicted and not predicted. The figures show the result that is displayed after the detailed analysis.

## Conclusion

One of the most well-known industries worldwide is the Indian medical plant sector. Due to the great quality of the medicinal plants supplied from India, there is a sizable international market for them. However, as processing typically requires human labor, correctly identifying medical plant leaves for export is a time-consuming and expensive operation. With a practical and quick disease prediction and sorting strategy using the most recent SVM technology, this research offers exporters an option. The most popular technology, image processing, is successfully used in this work. The study's findings demonstrate that disease sorting and prediction have been successfully applied, making them a quick and affordable alternative for the fruit business. In future, deep learning algorithms, such as CNN, RNN, and LSTM, can be used which will give more results than proposed SVM.

## Declarations

**Conflict of Interest** We have no conflict of interest to declare. On behalf of all co-authors, the corresponding author shall bear full responsibility for the submission.

## References

1. Francis et al., Disease detection and classification in agricultural plants using convolutional neural networks—a visual understanding. International Conference on Signal Processing and Integrated Networks, 2019;1063–1068.

2. Lee, et al. How deep learning extracts and learns leaf features for plant classification. Pattern Recogn. 2017;71:1–13.
3. Shen et al., Msr-net: Low-light image enhancement using deep convolutional network. arXiv preprint. arXiv:1711.02488. 2017
4. Grinblat G, et al. Deep learning for plant identification using vein morphological patterns. Computers Electron Agric. 2016;127:418–24.
5. Mohanty, et al. Using deep learning for image based plant disease detection. Front Plant Sci. 2016. https://doi.org/10.3389/fpls.2016.01419.
6. Kose N, Dugelay J-L. Classification of captured and recaptured images to detect photograph spoofing. Proceedings International Conference on Information, Electron, 2012;1027–1032
7. Bhardwaj S, Mittal A. A survey on various edge detector techniques. Elsevier Proc Technol. 2012;4:220–6.
8. Rhoma EM, Abobaker AM. Mean square error minimization using interpolative block truncation coding algorithms. Proceedings of IEEE International Conference on Education Technology and Computer (ICETC), Shanghai, 2010;5:42–45
9. Ng T-T, Gao X, Qiu B, Chang S-F. Single-view recaptured image detection based on physics-based features. IEEE International Conference on Multimedia Expo (ICME), 2010;1469–1474. https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset
10. Sonar PS, Wankhade NR. Grape Leaf Disease Identification Using Machine Learning Techniques. International Research Journal of Modernization in Engineering Technology and Science (Peer-Reviewed, Open Access, Fully Refereed International Journal), 2022;04(07)
11. Xie X, Ma Y, Liu B, He J, Li S, Wang H. A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks. Front Plant Sci. 2020;11:751.
12. Sharma G, GouravShrivastava. Crop disease prediction using deep learning techniques—a review. ICTACT J Data Sci Mach Learn. 2022;03(02):312–5.
13. Liu B, Ding Z, Tian L, He D, Li S, Wang H. Grape leaf disease identification using improved deep convolutional neural networks. Front Plant Sci. 2020. https://doi.org/10.3389/fpls.2020.01082.
14. Jaisakthi SM, Mirunalini P, Thenmozhi D. Grape leaf disease identification using machine learning techniques. https://typeset.io/papers/grape-leaf-disease-identification-using-machine-learning-19xyut1nip, (2019)
15. Cao H, Kot AC. Identification of Recaptured Photographs on LCD Screens. IEEE International Conference on Acoustic Speech Signal Processing, 2010;1790–1793
16. Yu H, Ng T-T, Sun Q. Recaptured photo detection using specularity distribution. IEEE Inernational Conference on Image Process. (ICIP), 2008;3140–3143
17. Ng T-T, Chang S-F, Hsu J, Xie L, Tsui M-P. Physics-motivated features for distinguishing photographic images and computer graphics. Proceedings 13th Annual ACM International Conference on Multimedia, 2005;239–248
18. Farid H, Lyu S. Higher-order wavelet statistics and their application to digital forensics. IEEE workshop on Statistical Analysis in Computer Vision, 2003;94–102
19. Trichili et al. A survey and evaluation of edge detection operators application to medical images. IEEE International Conference on Systems, Man and Cybernetics, vol 4; 2002. p. 113–8