

Optimizing Classification of Malaysian Plant Species through Comparative Analysis of Edge Detection Methods

Zuraini Othman^{1,*}, Muhammad Hafiz Izzuddin Mohamad Sukeri¹, Sharifah Sakinah Syed Ahmad¹, Fauziah Kasmin¹, Nur Hajar Zamah Shari², Anton Satria Prabuwono³

¹ Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

² Forestry and Environment Division, Forest Research Institute Malaysia (FRIM), 52109 Kepong, Malaysia

³ Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Rabigh 21911, Saudi Arabia

ARTICLE INFO

Article history:

Received 21 February 2025

Received in revised form 26 July 2025

Accepted 14 August 2025

Available online 1 November 2025

Keywords:

Plant species classification; Edge detection algorithms; Convolutional Neural Networks (CNNs); Image pre-processing; Malaysian Plant

ABSTRACT

The classification of plant species is a vital aspect of botanical research, biodiversity monitoring, and conservation efforts. This study aims to optimize the classification of Malaysian plant species by conducting a comprehensive comparative analysis of various edge detection methods. This study investigates the optimization of Malaysian plant species classification through a comparative analysis of edge detection methods. Focusing on five distinct species - *Murraya koenigii*, *Citrus aurantiifolia*, *Pandanus amaryllifolius*, *Polygonum minus*, and *Mentha arvensis* - the research evaluates the effectiveness of Canny, Roberts, Sobel, and Laplace edge detection algorithms across RGB, grayscale, and pre-processed leaf images. A dataset of high-resolution microscopic leaf images was developed and subjected to various image processing techniques. The study employs Convolutional Neural Networks (CNNs) for classification, assessing performance using precision, recall, and F1-score metrics. Results consistently demonstrate the superiority of the Canny edge detection method across all image types and plant species, with *Citrus aurantiifolia* exhibiting the highest classification accuracy with F1-score 93%. The research highlights the importance of species-specific considerations in edge detection and emphasizes the potential of adaptive methodologies in improving classification accuracy. These findings contribute to the advancement of automated plant identification systems, with implications for botanical research and biodiversity conservation in Malaysia.

1. Introduction

Plant species classification is a critical task in botanical research and biodiversity conservation based on leaf texture (Li et al. 2022; W. Liu et al. 2024a; X. Liu et al. 2022; Robil et al. 2021). The diversity of Malaysian flora includes various plant species with different textures feature. Studies show opportunities and challenges for effective classification (Hadi et al. 2022). There are studies that help accurate classification systems in various applications, including medicine (Pornpanomchai et al. 2011; Selda et al. 2017), agriculture (Güldenring et al. 2024; Li et al. 2022; W. Liu et al. 2024a;

* Corresponding author.

E-mail address: zuraini@utem.edu.my

<https://doi.org/10.37934/ard.146.1.244257>

X. Liu et al. 2022; Selda et al. 2017), forestry, environmental monitoring and many other fields (Abdulazeez et al. 2021; Hussein et al. 2021; Khalid & Romle 2024; Muneer & Fati 2020; Robil et al. 2021). In recent years, advances in image processing and machine learning have facilitated the development of automatic systems for plant species classification, taking advantage of the unique characteristics of leaf images, especially for herbaceous plants (Athapaththu & Piumi Ishanka 2023; Sudhakar & Swarna Priya 2023).

Related studies have shown the potential of image processing techniques in the classification of plant species. For example, classifications based on leaf shape, texture, and venation patterns have been widely studied as key features to distinguish between species. Various approaches have been proposed ranging from traditional feature extraction methods to advanced machine learning algorithms. These research have explored various techniques for the classification of plant species. Through this research it is seen that the research focus is on the analysis of leaf veins using computer vision methods. There is a study developing a mobile device using Raspberry Pi for plant identification through leaf vein image processing targeting applications for botanists, pathologists and plant breeders (Selda et al. 2017). There is also a study proposing a Confidence Refining Vein Network (CoRE-Net) to segment leaf veins to deal with challenges such as intersections and blurred boundaries that require minimal labeled samples (Li et al. 2022). In addition, there is a study that introduces DeepLabV3+ to improve the speed and accuracy of leaf vein segmentation. While there are studies using CNN architecture that improve vein shape data (Huynh et al. 2020; X. Liu et al. 2022). There are also studies that create a Grass Vein Image Quantification (GRASVIQ) framework for parallel vein segmentation (Robil et al. 2021). A subsequent study has developed a model to extract fine-grained phenotypic information, including veins (Güldenring et al. 2024).

Furthermore, researchers have developed a hierarchical leaf vein segmentation (HALVS) dataset to establish a benchmark for future studies of this field (W. Liu et al. 2024b). There are studies showing the classification of plant species using convolutional deep learning neural networks based on venation patterns (da Cruz 2020) while there are also studies using Mask R-CNN and U-Net for leaf segmentation and disease classification (Dawod & Dobre 2022). An exploration of more than 50 studies on the application of computer vision and machine learning to digitized herbarium specimens found that the importance of studies in this field in plant identification (Hussein et al. 2021). There is also the objective of studying the use of computer vision technology to identify the content and lack of plant nutrients (Sudhakar & Swarna Priya 2023). In addition, there is a classification that extracts density, color and vein texture features for classification using a Support Vector Machine (SVM) (Athapaththu & Piumi Ishanka 2023). Finally, a principal curvature-based method was also introduced for leaf vein segmentation, using Gaussian filtering and contrast enhancement to highlight vein patterns (Agarwal & Bachan 2023).

Several studies have been done to explore various techniques to identify Malaysian herbs through leaf analysis. A comparative study on texture features for leaf recognition, has used Histogram Oriented Gradient (HOG), Local Binary Pattern (LBP), and Robust Accelerated Features (SURF) using a multi-class Support Vector Machine (SVM) classifier (Ibrahim et al. 2018). There have also been studies investigating the use of transfer learning and fine tuning of deep learning neural networks to classify different herbaceous plant species (Khalid & Romle 2024) Low-level feature contribution in plant leaf recognition was performed using Scale Invariant Feature Transform (SIFT) for shape, color moment for color, and Segmentation-Based Fractal Texture Analysis (SFTA) for texture (Jamil et al. 2015). Their results showed that the combination of texture and shape features outperformed the combination of texture, shape and color features.

Related research works have shown vein analysis using computer vision techniques revealing significant progress and diverse approaches to plant species classification. However, there are still

critical gaps that have yet to be revealed in direct comparisons of edge detection methods specifically tailored to the unique textural characteristics of Malaysian herbaceous plant species. Although methods such as CoRE-Net and DeepLabV3+ address segmentation challenges and improve prediction accuracy, and devices as developed, offer practical field applications. However, the performance of edge detection techniques such as Canny, Roberts, Sobel and Laplace in this specific context has not been comprehensively evaluated. Moreover, most of the related previous studies emphasize vein segmentation and pattern extraction but often ignore the comparative effectiveness of different edge detection methods in improving the overall classification accuracy. Therefore, the focus of this study is on a systematic comparative analysis of edge detection methods. It can fill an important research gap and has the potential to significantly improve classification accuracy for herbaceous plant species in Malaysia. This study is expected to provide valuable insights and tools for botanists, agriculturalists and environmental scientists.

The proposed work aims to improve the classification accuracy of Malaysian plant species by conducting a comprehensive comparative analysis of various edge detection methods, including Canny, Roberts, Sobel, and Laplace. The selection of Canny, Sobel, Laplace, and Roberts edge detection algorithms for this study is based on their widespread use and proven effectiveness in various image processing applications. These methods are chosen because they have different characteristics and performance in different scenarios. Canny, known for its ability to detect weak edges and block noise, was chosen for its robustness. Sobel, on the other hand, is effective in emphasizing edges in the high-frequency region, chosen for its computational simplicity and efficiency. Laplace can detect edges by finding zero crossings after filtering the image while also being able to detect rapid intensity changes. Roberts is one of the earliest edge detection methods, has its simplicity and effectiveness in detecting edges in clear images. These algorithms represent various approaches to edge detection, from simple gradient-based methods (Roberts and Sobel) to more complex multilevel algorithms (Canny). Although other edge detection methods exist, such as the Prewitt operator, in order to provide a comprehensive comparison of basic edge detection techniques. This selection allows a comprehensive evaluation of the performance of different edge detection principles in the specific context of Malaysian plant species classification, allowing insight into which features of this algorithm are most beneficial for this specific application in various fields.

This study focuses on images of leaves from five different herbaceous plant species: Curry (*Murraya koenigii*), Lime (*Citrus aurantiifolia*), Pandan (*Pandanus amaryllifolius*), Cashew (*Polygonum minus*), and Mint (*Mentha arvensis*). This process involves several stages which are data image acquisition, image preprocessing that looks at RGB images, conversion to grayscale images, and the application of edge detection techniques to the image. The effectiveness of each method in image preprocessing will be evaluated based on classification performance metrics, such as precision, recall and F1 score, to determine the most suitable edge detection technique for this specific application. By optimizing the classification using this preprocessing technique, the results of this study developed can build a more accurate and efficient automatic system to classify plant species that contributes to progress in botanical research and biodiversity conservation.

The remainder of this paper is structured as follows: The methodology section begins with a description of the dataset development process. This is followed by an explanation of the image analysis process, detailing preprocessing techniques and the implementation of various edge detection algorithms (Canny, Sobel, Laplace, and Roberts). The methodology concludes with a discussion of the evaluation metrics used to evaluate classification performance. The results and discussion section first present the constructed dataset. It then outlines the experimental setup, including hardware specifications and parameter settings. The core of this section is a comprehensive comparative analysis of edge detection methods across different preprocessing techniques, with a

focus on species-specific performance variations. Finally, the conclusion summarizes the main findings, discusses their implications for the classification of Malaysian plant species, acknowledges the limitations of the study, and suggests directions for future research.

2. Methodology

This section provides an overview of the methodology that has been used in this study to achieve effective identification of Malaysian herbs using computer vision techniques. First, the development of the dataset is discussed, detailing the collection and preparation of leaf samples to ensure a robust and diverse representation of herb species. Next, the image analysis process that includes image acquisition, preprocessing, feature extraction, and classification techniques. Finally, the evaluation metrics used to evaluate the performance of the classification model are outlined, providing insight into the accuracy, recall and overall effectiveness of the F1 score of the proposed system.

2.1 Dataset Development

In this study, the development of a robust data set was a critical component. The dataset consists of high-resolution images of leaf samples from five different Malaysian plant species: Curry (*Murraya koenigii*), Lime (*Citrus aurantiifolia*), Pandan (*Pandanus amaryllifolius*), Kesum (*Polygonum minus*) and Pudina (*Mentha arvensis*).

This setup to obtain detailed examination of leaf vein textural features important for edge detection and classification process. The system's ability to deliver high-quality images with minimal distortion is critical in ensuring the reliability of the data used in this research. Figure 1 shows a stereo microscope equipped with a high-resolution camera used for image acquisition of leaf samples. This equipment is capable of capturing the fine details of leaf edges, which is important for edge detection analysis. Figure 2 shows the prepared leaf sample placed on the microscope stage, ready for image acquisition. Each leaf sample was placed at a different position so that many image angles could be obtained.

Images captured using this setup form the basis of our dataset, which is used to evaluate the ability of edge detection methods and subsequently classification. Each leaf image is processed to extract edge features, which are then used to accurately classify plant species.



Fig. 1. Microscope Equipment.



Fig. 2. Leaf Sample Preparation.

2.2 Image Analysis Process

In this study, the process of optimizing the classification of Malaysian plant species through a comparative analysis of edge detection methods using a detailed image analysis process consisting of three main parts, namely RGB image preparation, grayscale image conversion and edge detection with image processing. In the preparation of RGB images, each leaf sample image was resized to the desired dimensions to maintain the aspect ratio, centered and normalized by dividing the pixel value by 255 (image / 255.0). These steps ensure consistency in size and scale across all images.

Grayscale image conversion is the conversion of an RGB image to grayscale. This step simplifies the image data to reduce it to a single intensity channel. Grayscale images undergo the same pre-processing as RGB images, ensuring that they are resized and cropped to match the dimensions and central focus of the original image. Image processing is the stage where images are resized, centered, and noise reduced. Color correction is performed by scaling the RGB image to achieve an average gray scale intensity of 128. an edge detection method is applied to the processed image. Edge detection with Canny, Sobel, Laplace and Roberts edge detection methods will be used. Pixel values are normalized between 0 and 1, and all images are resized to 150x150 for standardization. These steps ensure that images are prepared consistently to allow robust comparison of different edge detection methods and their effectiveness in classifying Malaysian plant species.

2.3 Evaluations Metrics

Performance evaluation of the process used in classifying Malaysian plant species, we used several standard metrics: Precision, Recall and F1-Score. Precision measures the accuracy of the positive predictions and is calculated as the ratio of True Positives (TP) to the sum of True Positives (TP) and False Positives (FP).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

Recall indicates the ability to identify all relevant instances and is calculated as the ratio of True Positives (TP) to the sum of True Positives (TP) and False Negatives (FN).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

F1-Score provides a balance between Precision and Recall, calculated as the harmonic mean of the two.

$$\text{F1-Score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (3)$$

These evaluations provide a comprehensive classification performance, enabling a detailed comparison of different edge detection methods.

3. Results and Discussion

This section discusses the findings and interpretation of the study. Focus on three main parts, namely the constructed data set, experimental setup and comparison results. The data set constructed was analyzed to ensure its relevance for identifying Malaysian herbs. Experimental setup includes configuration of instrument parameters and image analysis process. Finally, the comparison results of different sample types and classification processes with different image processing approaches are discussed. Here, highlighting the performance of image classification and its effectiveness in identifying various herb species will be obtained. This section aims to provide a comprehensive examination of the results, offering insights and implications for future research and applications in the field of plant identification using computer vision.

3.1 Constructed dataset

The dataset constructed, as illustrated in Figure 3, consists of five different Malaysian plant species. Meanwhile, Figure 4 is an example of a sample image from each plant species that has been captured from *Murraya koenigii*, *Citrus aurantiifolia*, *Pandanus amaryllifolius*, *Polygonum minus*, and *Mentha arvensis*. This sample image dataset allows forming the basis of our comparative analysis of various image processing approaches for plant species classification. Each category of species samples in the dataset is divided into training and testing image sets, ensuring a robust framework for model development and evaluation. This structure enables thorough training of species-specific features while maintaining separate testing sets to evaluate the generalization capabilities of image processing approaches and classification algorithms.

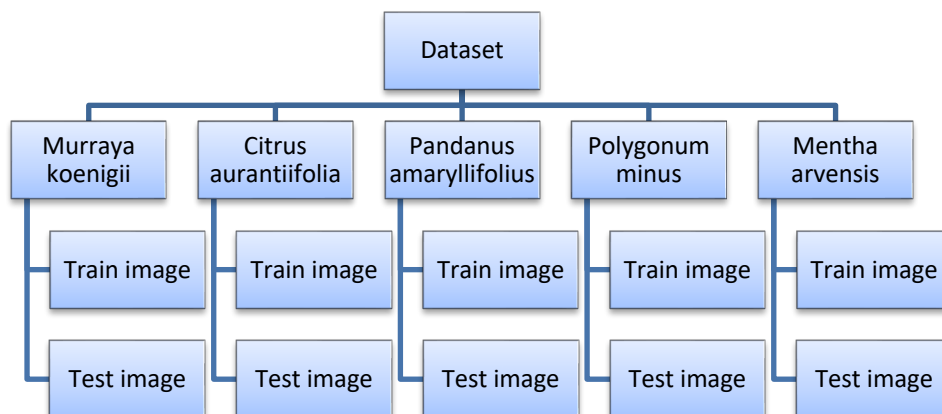


Fig. 3. Structure of the constructed dataset for Malaysian plant species classification

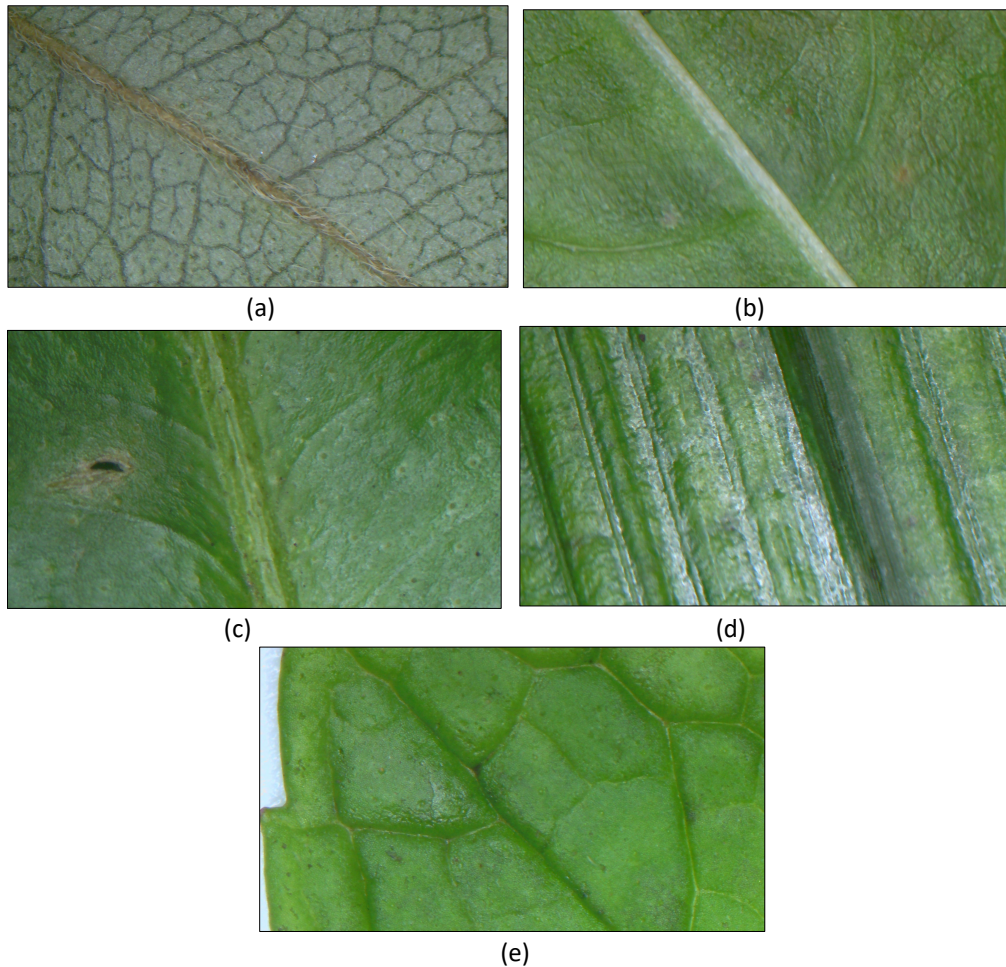


Fig. 4. Sample images had been captured from (a) *murraya koenigii*, (b) *citrus aurantifolia*, (c) *pandanus amaryllifolius*, (d) *polygonum minus*, and (e) *mentha arvensis*.

The selection of these five diverse plant species, each native or commonly found in Malaysia as a cooking ingredient and its uses for various fields. Provide a representative sample to evaluate the effectiveness of various edge detection techniques in different processing approaches in classifying the unique textural features of Malaysian flora. The development of this dataset allowed us to investigate the performance of image processing approaches with different edge detection methods across a wide range of leaf structures, textures and shapes typical of the region's plant biodiversity. By systematically applying and comparing algorithms on these data sets, it is further possible to determine the most effective method to highlight the distinctive characteristics of each species. This comparative analysis is important to optimize the classification accuracy of Malaysian plant species, potentially leading to improved automatic identification systems for botanical research, biodiversity monitoring and conservation efforts in the region.

3.2 Experimental Setup

For the image acquisition process, a high-precision trinocular stereo zoom microscope was used to capture detailed images of leaf veins from the identified species. This microscope has a 7X-45X zoom magnification capability, able to provide clear and sharp images essential for the accurate classification of plant species. It is equipped with a 56-LED ring light that offers adjustable illumination to ensure optimal lighting conditions and reduce glare during image acquisition. The instrument has quality optics with a 360-degree rotatable binocular head tilted at 45 degrees for ergonomic viewing.

A working distance of 100mm allows sufficient space to manipulate the specimen under observation. The microscope camera has a resolution of 1600X1200 (UXGA) and supports various video output formats such as 720P and 1280x1024, enabling high-resolution image capture. Additional camera features include digital noise reduction, automatic white balance, and the ability to compare captured images with dynamic video, improving the clarity and accuracy of the images obtained.

This study focused on classifying Malaysian herbal plants using a data set of leaf images captured through a microscope. The dataset includes various species such as *Murraya koenigii*, *Citrus aurantifolia*, *Polygonum minus*, and *Mentha arvensis*. These leaves are obtained from the market and self-grown plants. Images were preprocessed to standardize the dataset, including resizing to 150x150 pixels, conversion to grayscale and normalization of pixel values between 0 and 1. This preprocessing aims to reduce noise and create a uniform basis for subsequent analysis.

The experimental framework uses several edge detection algorithms, namely Canny, Sobel, Laplace, and Roberts, to extract texture features from leaf images. This algorithm was chosen for its insensitivity to noise and its ability to accurately represent plant structure. OpenCV was used for image manipulation tasks and implementation of edge detection algorithms. The extracted edge features, along with additional features such as texture, color and shape, are used to train the classification model. For the classification task, a convolutional neural network (CNN) was implemented using TensorFlow and Keras. This project explores the use of transfer learning with pre-trained models to improve classification accuracy. The performance of different combinations of edge detection algorithms and classification models was evaluated using metrics such as accuracy, precision, recall and F1 score. This comparative analysis aims to identify the optimal approach for the classification of Malaysian plant species, focusing on the effectiveness of each edge detection method in capturing the unique textural features of the selected plants.

3.2 Comparative Results

This section shows a detailed analysis of the comparative results obtained from the study. Focusing on the effectiveness of different image processing techniques to optimize the classification of Malaysian plant species. The comparative results section is divided into three main parts of different edge detection approaches using RGB images, grayscale images and image processing. Each section discusses the methodology used, the results achieved, and their impact on the overall classification accuracy. This analysis highlights the potential of each technique offering insights into the most effective methods for improving plant species identification through edge detection.

Table 1
Precision Table for RGB

Plant Species	Sobel	Laplace	Canny	Roberts
<i>Murraya koenigii</i>	0.81	0.8	0.82	0.79
<i>Citrus aurantiifolia</i>	0.88	0.89	0.9	0.86
<i>Pandanus amaryllifolius</i>	0.72	0.73	0.75	0.71
<i>Polygonum minus</i>	0.84	0.85	0.87	0.83
<i>Mentha arvensis</i>	0.76	0.77	0.79	0.75

Table 2
Recall Table for RGB

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.83	0.86	0.88	0.81
Citrus aurantiifolia	0.92	0.93	0.94	0.9
Pandanus amaryllifolius	0.78	0.79	0.81	0.75
Polygonum minus	0.9	0.91	0.91	0.89
Mentha arvensis	0.82	0.83	0.83	0.81

Table 3
F1-Score Table for RGB

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.82	0.83	0.85	0.8
Citrus aurantiifolia	0.9	0.91	0.92	0.88
Pandanus amaryllifolius	0.75	0.76	0.78	0.73
Polygonum minus	0.87	0.88	0.89	0.86
Mentha arvensis	0.79	0.8	0.81	0.78

Comparison results with selected Malaysian plant species and edge detection algorithms for RGB parts are presented in Table 1 (Precision for RGB), Table 2 (Recall for RGB), and Table 3 (F1-Score for RGB). Here, Citrus aurantiifolia exhibits the highest overall performance across all algorithms for precision, recall and F1-Score in RGB, suggesting its distinctive features are well captured by the edge detection method. In contrast, Pandanus amaryllifolius showed the lowest scores across the board in RGB, indicating it may be more challenging to classify accurately using this edge detection technique.

These results highlight the effectiveness of the Canny edge detection algorithm for RGB images in this Malaysian plant classification task. The varying performance across selected plant samples in RGB emphasizes the importance of choosing an edge detection method based on the specific characteristics of each plant type when optimizing a classification system.

Table 4
Precision Table for Greyscale

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.76	0.77	0.79	0.75
Citrus aurantiifolia	0.85	0.86	0.87	0.84
Pandanus amaryllifolius	0.69	0.7	0.72	0.68
Polygonum minus	0.82	0.82	0.83	0.8
Mentha arvensis	0.73	0.74	0.76	0.72

Table 5
Recall Table for Greyscale

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.79	0.8	0.82	0.78
Citrus aurantiifolia	0.88	0.9	0.91	0.88
Pandanus amaryllifolius	0.74	0.75	0.77	0.73
Polygonum minus	0.86	0.87	0.88	0.85
Mentha arvensis	0.77	0.78	0.8	0.76

Table 6
F1-Score Table for Greyscale

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.78	0.79	0.81	0.77
Citrus aurantiifolia	0.87	0.88	0.89	0.86
Pandanus amaryllifolius	0.72	0.73	0.75	0.71
Polygonum minus	0.84	0.85	0.86	0.83
Mentha arvensis	0.75	0.76	0.78	0.74

Comparison results for grayscale images across different plant species for this study and edge detection algorithms in the grayscale section are presented in Table 4 (Precision), Table 5 (Recall), and Table 6 (F1-Score). Notably, Citrus aurantiifolia exhibited the highest overall performance across all algorithms for precision, recall and F1-Score in grayscale images. These results suggest that the distinctive features are well captured by the edge detection method even without color information. In contrast, Pandanus amaryllifolius and Mentha arvensis showed the lowest scores across the board, indicating they may be more challenging to classify accurately using this edge detection technique in grayscale.

These results highlight the effectiveness of the Canny edge detection algorithm for grayscale images in the Malaysian plant classification task selected for this study. Canny's consistent superior performance across all metrics and plant species underscores its robustness in capturing relevant edge features for classification. The varying performance across species emphasizes the importance of considering species-specific features when choosing an edge detection method to optimize a classification system, even in the context of grayscale images.

Table 7
Precision Table for Pre-processing Images

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.82	0.83	0.85	0.81
Citrus aurantiifolia	0.89	0.9	0.91	0.88
Pandanus amaryllifolius	0.75	0.76	0.78	0.74
Polygonum minus	0.86	0.87	0.88	0.81
Mentha arvensis	0.79	0.8	0.81	0.79

Table 8
Recall Table for Pre-processing Images

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.86	0.87	0.89	0.85
Citrus aurantiifolia	0.93	0.94	0.95	0.92
Pandanus amaryllifolius	0.79	0.8	0.82	0.78
Polygonum minus	0.9	0.91	0.92	0.85
Mentha arvensis	0.83	0.84	0.85	0.92

Table 9
F1-Score Table for Pre-processing Images

Plant Species	Sobel	Laplace	Canny	Roberts
Murraya koenigii	0.84	0.85	0.87	0.83
Citrus aurantiifolia	0.91	0.92	0.93	0.9
Pandanus amaryllifolius	0.77	0.78	0.8	0.76
Polygonum minus	0.88	0.89	0.9	0.83
Mentha arvensis	0.81	0.82	0.83	0.72

Comparison results for pre-processed images on different plant species in this study and edge detection algorithms for the pre-processing part of the image are presented in Table 7 (Accuracy), Table 8 (Recall), and Table 9 (F1-Score). Notably, *Citrus aurantiifolia* exhibits the highest overall performance across all algorithms for precision, recall and F1-Score in pre-processed images suggesting its distinctive features are well preserved and enhanced by pre-processing techniques. On the other hand, *Pandanus amaryllifolius* showed the lowest score overall. This suggests this species may be more challenging to classify accurately even after pre-processing.

Figure 5 graph shows the F1 score of the Canny edge detector applied to images pre-processed using RGB, grayscale and custom pre-processing methods. The x-axis represents five different plant species (*Murraya koenigii*, *Citrus aurantiifolia*, *Pandanus amaryllifolius*, *Polygonum minus*, *Mentha arvensis*), and the y-axis represents the F1 score, a metric that combines precision and recall.

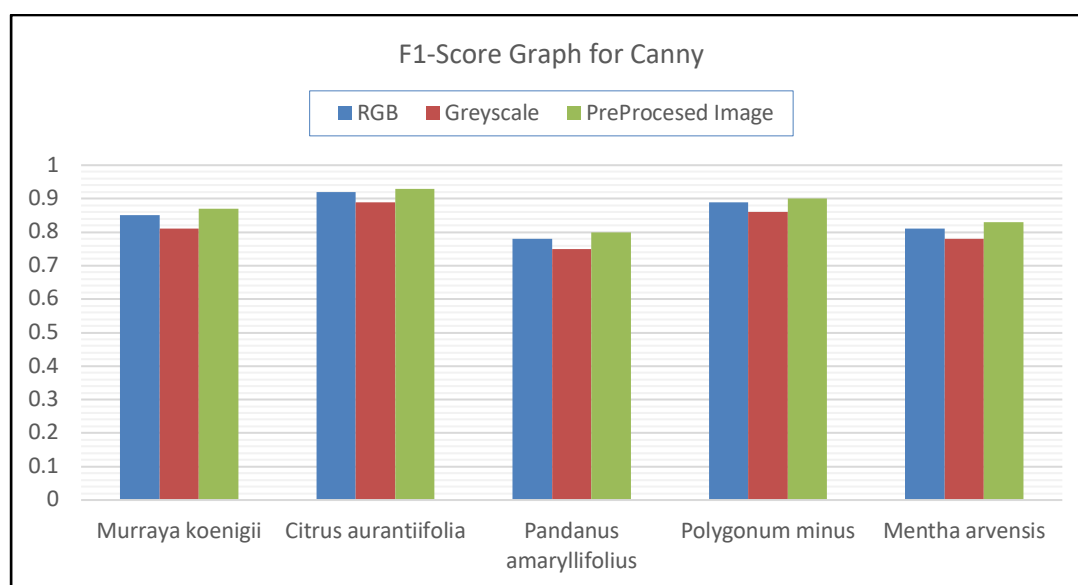


Fig. 5. F1-Score Comparison of Canny Edge Detector on RGB, Greyscale and Pre-processed Image.

Overall, the custom pre-processing method achieved the highest F1 scores across all plant species. This shows its effectiveness in improving the performance of edge detection methods. Grayscale images performed slightly lower but still reasonable, while RGB images exhibited the lowest F1 scores. This suggests that color information may not be very important for edge detection in the context of this study.

These results highlight the effectiveness of the Canny edge detection algorithm for pre-processed images in the Malaysian plant classification task selected in this study. Canny's consistent performance across all metrics and plant species underlines its robustness in capturing relevant edge features for classification, even after a pre-processing step. Adding an image pre-processing process improves its overall classification accuracy compared to grayscale image and RGB image results. The varying performance across plant species emphasizes the importance of looking at species-specific characteristics when choosing edge detection methods and pre-processing techniques to optimize a classification system.

Performance varied across different plant species, with *Citrus aurantiifolia* consistently scoring the highest and *Pandanus amaryllifolius* the lowest. This emphasizes the importance of species-specific considerations in classification approaches. The consistent high performance in classifying *Citrus aurantiifolia* across all methods may be due to its distinctive leaf structure, characterized by

clear, distinct edges and a relatively simple venation pattern. On the other hand, the lower accuracy observed with *Pandanus amaryllifolius* can be attributed to its more complex leaf structure. It shows parallel venation and possible finer edge transitions that are challenging for some algorithms to detect well.

Additionally, the impact of color information and pre-processing techniques on classification accuracy highlights the need for a multifaceted approach in developing effective plant identification systems. The custom pre-processing method that has been developed through this study can reduce the noise. This approach has been adapted to address the specific challenges posed by selected Malaysian plant leaf images, such as varying lighting conditions and complex textures. The advantage of this method lies in its ability to increase the visibility of edges while preserving critical texture details. This results in better performance across all edge detection algorithms than standard RGB or grayscale processing.

This finding has significant implications for optimizing the classification of Malaysian plant species. The consistent effectiveness of the Canny algorithm suggests it should be a key consideration in edge detection-based classification systems. However, the varying performance across species and image types suggests that a one-size-fits-all approach may not be sufficient. Instead, the results show the potential benefits of adaptive or hybrid methodologies that can take advantage of the strengths of different edge detection algorithms and image pre-processing techniques based on specific plant characteristics. Furthermore, the challenges in classifying certain species such as *Pandanus amaryllifolius* for all conditions indicate that complementary feature extraction methods may be needed to develop a more comprehensive and robust classification system for this species. Overall, this study not only demonstrates the effectiveness of edge detection methods in plant classification but also opens up space for future research in optimizing this technique for diverse plant species.

4. Conclusions

In conclusion, this study has shown the effectiveness of the edge detection algorithm, especially the Canny method, in the classification of Malaysian plant species. Through a comprehensive analysis of various edge detection techniques applied to RGB, grayscale and pre-processed images. The main factors influencing classification accuracy across different plant species can be derived. The consistently superior performance of the Canny algorithm across all image types and plant species underscores its robustness and suitability for capturing textural features important for accurate plant identification.

The findings of this study highlight the importance of species-specific considerations in developing an optimized classification system. The varying performance of different plant species shows high accuracy for *Citrus aurantiifolia*. Whereas, challenges are shown by *Pandanus amaryllifolius*, emphasizing the need for a tailored approach in edge detection and feature extraction. Furthermore, the effect of color information and pre-processing techniques on classification accuracy shows that a multifaceted approach by combining insights from different image representations can lead to more robust and accurate classification.

This research provides a basis for future work in optimizing plant species classification systems. The results show the potential to leverage the strengths of different edge detection algorithms based on specific plant characteristics. Additionally, this study opens the door to investigate feature extraction methods to address the limitations of edge detection in classifying certain challenging species. Therefore, future research can contribute to the development of a more sophisticated and accurate plant identification system that ultimately improves our ability to study and conserve plant biodiversity in Malaysia.

Acknowledgement

The deepest gratitude and thank you to the Ministry of Education (MOE) and Universiti Teknikal Malaysia Melaka (UTeM) for the financial supports through FRGS. Grant No: FRGS/1/2020/FTMK-CACT/F00462.

References

- [1] Abdulazeez, A.M., Zeebaree, D.Q., Zebari, D.A. & Hameed, T.H. 2021. Leaf identification based on shape, color, texture and vines using probabilistic neural network. *Computacion y Sistemas* 25(3): 1–15.
- [2] Agarwal, D. & Bachan, P. 2023. Segmentation of leaf veins using principal curvature and morphological methods. In *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*. Vol. 3 pp. 2–7.
- [3] Athapaththu, S.W. & Piumi Ishanka, U.A. 2023. Plant Leaf Recognition Using Texture, Colour, and Vein Density Features. In *ICARC 2023 - 3rd International Conference on Advanced Research in Computing: Digital Transformation for Sustainable Development*. pp. 102–107.
- [4] da Cruz, J.M.V. 2020. Plant Species Identification Through Leaf Venation Extraction and CNNs. *PQDT - Global*. Universidade de Lisboa (Portugal) PP - Portugal, Portugal <https://libproxy.utem.edu.my/login?url=https://www.proquest.com/dissertations-theses/plant-species-identification-through-leaf/docview/2652591531/se-2?accountid=34984>.
- [5] Dawod, R.G. & Dobre, C. 2022. Upper and Lower Leaf Side Detection with Machine Learning Methods. *Sensors* 22(7).
- [6] Gldenring, R., Andersen, R.E. & Nalpantidis, L. 2024. Zoom in on the Plant : Fine-Grained Analysis of Leaf, Stem, and Vein Instances. *IEEE ROBOTICS AND AUTOMATION LETTERS* 9(2): 1588–1595.
- [7] Hadi, N.A., Halim, S.A., Lazim, N.S.M. & Alias, N. 2022. Performance of CPU-GPU Parallel Architecture on Segmentation and Geometrical Features Extraction of Malaysian Herb Leaves. *Malaysian Journal of Mathematical Sciences* 16(2): 363–377.
- [8] Hussein, B.R., Malik, O.A., Ong, W.-H. & Slik, J.W.F. 2021. Application of Computer Vision and Machine Learning for Digitized Herbarium Specimens: {A} Systematic Literature Review. *CoRR abs/2104.0*.
- [9] Huynh, H.X., Truong, B.Q., Nguyen Thanh, K.T. & Truong, Di.Q. 2020. Plant Identification Using New Architecture Convolutional Neural Networks Combine with Replacing the Red of Color Channel Image by Vein Morphology Leaf. *Vietnam Journal of Computer Science* 7(2): 197–208.
- [10] Ibrahim, Z., Sabri, N. & Mangshor, N.N.A. 2018. Leaf recognition using texture features for herbal plant identification. *Indonesian Journal of Electrical Engineering and Computer Science* 9(1): 152–156.
- [11] Jamil, N., Hussin, N.A.C., Nordin, S. & Awang, K. 2015. Automatic Plant Identification: Is Shape the Key Feature? *Procedia Computer Science* 76(Iris): 436–442.
- [12] Khalid, F. & Romle, A.A. 2024. Herbal Plant Image Classification using Transfer Learning and Fine-Tuning Deep Learning Model. *Journal of Advanced Research in Applied Sciences and Engineering Technology* 35(1): 16–25.
- [13] Li, L., Hu, W., Lu, J. & Zhang, C. 2022. Leaf vein segmentation with self-supervision. *Computers and Electronics in Agriculture* 203(September 2021): 107352.
- [14] Liu, W., Li, A., Wu, Z., Li, Y., Ge, B., Lan, G., Chen, S., Li, M., Liu, Y., Yuan, X. & Dong, N. 2024a. Revealing Hierarchical Structure of Leaf Venations in Plant Science via Label-Efficient Segmentation: Dataset and Method.
- [15] Liu, W., Li, A., Wu, Z., Li, Y., Ge, B., Lan, G., Chen, S., Li, M., Liu, Y., Yuan, X. & Dong, N. 2024b. Revealing Hierarchical Structure of Leaf Venations in Plant Science via Label-Efficient Segmentation: Dataset and Method. *CoRR abs/2405.1*.
- [16] Liu, X., Xu, B., Gu, W., Yin, Y. & Wang, H. 2022. Plant leaf veins coupling feature representation and measurement method based on DeepLabV3+. *Frontiers in Plant Science* 13(November): 1–20.
- [17] Muneer, A. & Fati, S.M. 2020. Efficient and automated herbs classification approach based on shape and texture features using deep learning. *IEEE Access* 8: 196747–196764.
- [18] Pornpanomchai, C., Rimdusit, S., Tanasap, P. & Chaityod, C. 2011. Thai herb leaf image recognition system (THLIRS). *Kasetsart Journal - Natural Science* 45(3): 551–562.
- [19] Robil, J.M., Gao, K., Neighbors, C.M., Boeding, M., Carland, F.M., Bunyak, F. & Mcsteen, P. 2021. GRASVIQ : an image analysis framework for automatically quantifying vein number and morphology in grass leaves 629–648.
- [20] Selda, J.D.S., Ellera, R.M.R., Cajayon, L.C. & Linsangan, N.B. 2017. Plant identification by image processing of leaf veins. *ACM International Conference Proceeding Series Part F1313*: 40–44.
- [21] Sudhakar, M. & Swarna Priya, R.M. 2023. Computer Vision Based Machine Learning and Deep Learning Approaches for Identification of Nutrient Deficiency in Crops: A Survey. *Nature Environment and Pollution Technology*.

