

Upgraded RRIM2000 Rubber Clones Identification Through Image And Statistical Analysis

Mohd Azrul Aiman Bin Mahmad Azan
Faculty of Electrical Engineering, Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
E-mail: azrul_chrome@yahoo.com

Abstract— This paper studies the research work to identify the perimeter, area and radius for selected rubber seed RRIM2000 series with the aid of image processing using Sobel Edge Detection techniques. There are two groups of RRIM2000 family which are Group 2A and 2B. For this project, there are five types of rubber seed selected as a sample. Three types of group 2A which are RRIM2007, RRIM2009, RRIM2016 and the two types from Group 2B which are RRIM2012 and RRIM2025. RGB colour image for all selected seed are captured using a digital camera. Image processing involving converting RGB image to grayscale, edge detection, morphology conducted using the Matlab software to extract the shape features. 150 samples used for testing and final analysing using SPSS software to identify the clone. Data obtain from one-way ANOVA and error plot measurement shown that two of the clones series significantly different from each other in term of perimeter, area and radius classification. As a conclusion, perimeter, area and radius of rubber seed clone can be used to recognize selected RRIM2007 and RRIM2025 rubber seed clone only.

Keywords: Sobel Edge Detection, Image Processing, Rubber seed clones, RRIM2000, MATLAB.

I. INTRODUCTION

Rubber tree (*Hevea brasiliensis*) has been discovered since fifteenth century when Christopher Columbus discovered the Americas. Rubber is one of the important agricultural products of Malaysia [2]. In order to ensure continuous supply of quality latex and heveawood to the manufacturers, replanting with higher yielding rubber tree clones is hence very important [3]. Cloning produces the precise same genetic for every seed which is maintained characteristic in a plant such as taste, colour, size and resistance to diseases. Clone seed will grow at the same rate so harvesting can become streamlined. There are about 33 types of clone in the RRIM2000 Series. Each type of seed is different in term of size, weight and pattern. It is important to differentiate actual seed and the clone. Only experienced worker that can differentiate the seed through naked eye but this old method consume time and higher accuracy. Due to the advance of video camera technology, people can take a digital picture or digital video stream in any place and any time with a very simple camera or mobile phone. A digital image is not only produced by using easy and inexpensive equipment but also convenient to process by a computer system. Therefore, this research employs a simple digital camera to capture a plant seed and applies a simple image processing technique to recognize rubber seed clone [1].

A. RRIM2000 Clone Classification

RRIM2000 clone series firstly introduced in 1995 with 22 clones. Four more clones introduced in 1998 followed by seven clone the next following years. Since these clones still new and information about the quality and performance of these clones are limited, they are categorised into two groups under “LGM Planting Recommendations”[4].

- i. Group 2A
Consists of new clones that are performing well at least three years of data results in a large scale trial in various environments. This will allow the rubber planters to select new clones in Group 2 with less risk. -RRIM2007, RRIM2009, RRIM2016 [4].
- ii. Group 2B
Consists of clones recently released and potentially in small scale trial. The clone was selected based on five years of data in small scale trials in a limited environment. - RRIM2012, RRIM2025 [4].

B. Clone Characteristic

- i. RRIM2007: The size of the seed is medium, smooth, shining and brownish seed coat with a square shape. Overall growth of the clone is average and seed production is good [3,5].



Figure 1: RRIM2007

- ii. RRIM2009: The size of the seed is small, smooth and shining with light brownish seed coat. It has square to slightly rounded shape. Overall growth and seed production of this clone is good. This is the best type of clone and was recommended for both latex and timber production [3,5].



Figure 2: RRIM2009

- iii. RRIM2012: The size is small, smooth, shining with faint brownish seed coat. It has square to slightly rectangular shape. Overall growth of the clone is good and seed production is good [3,5].



Figure 3: RRIM2012

- iv. RRIM2016: The seed features of this clone are almost similar with RRIM2009 but it has a square shaped. Overall growth and seed production of the clone is very good. This clone was also recommended for latex and timber production [3,5].



Figure 4: RRIM2016

- v. RRIM2025: The seed is large in size, rectangular, light brown and shining seed coat. Overall growth and seed production of the clone is good. The clone is recommended for both latex and timber production [6].



Figure 5: RRIM2025

II. METHODOLOGY

The process of image processing to recognize the shape of the rubber seed can be simplified as below

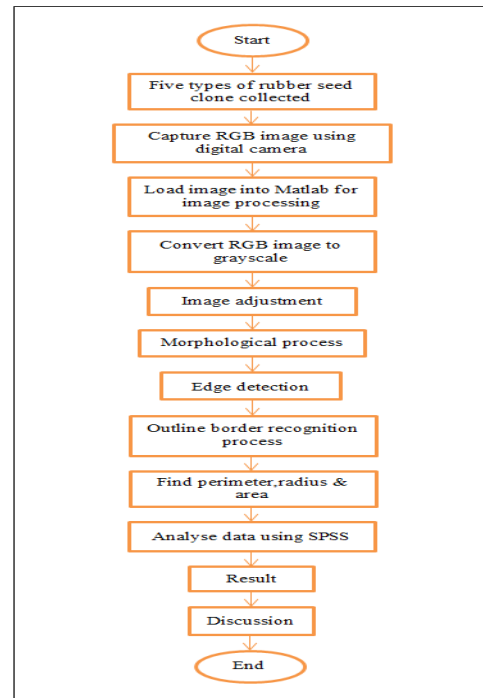


Figure 6: Methodology Process

A. Data Collection

For this project, the seed sample was collected at the Rubber Research Institute of Malaysia Sungai Buloh. They were five types of seed used which are RRIM2007, RRIM2009, RRIM2016, RRIM2012 and RRIM2025. Each sample has different quantities but for the research, only 30 seed selected from each sample because these seeds are limited and have certain season to collect.

The data collection is perform at Image Capturing Room at the Advanced Signal Processing Lab Faculty Of Electrical Engineering UiTM Shah Alam. The sample images were taken using Nikon D3100 DSLR camera attached with 18-55mm 3.5G lens which is suitable to take sharp images.

The setting for the camera is fixed and been used from the beginning of the project. Small change at the setting will affect all the data obtain. Image resolution used is 1024x683 pixels in JPEG format, RGB colour image, with the aid of Digicolor K-2500 spotlight, the ISO was set to 400 to the reduce the noise in the picture, the shutter speed is f1/20, aperture of 20 with focal length set to 50mm.

Camera and lighting setup for the image capturing process are placed at the same place for the rest of the project. Distance of each equipment is recorded, the camera was placed 180° above the seed with the distance of 9 inches. The

lighting, digicolor k-2500 was placed 45° from the sample with a piece of paper placed beside the sample to reflect the light in order to minimize the shadow around the seed as shown in Fig. 7.



Figure 7: Equipment setting for image capturing.

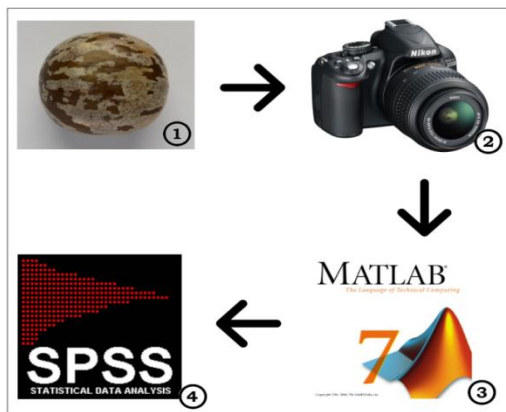


Figure 8: Image processing steps
(1) Sample seed (2) camera (3) Matlab and (d) SPSS

B. Image Processing

The most important part is image processing using Matlab which will convert the RGB image to grayscale, resizing, edge detection, outline border recognition and morphology.

i. RGB to Grayscale

All the rgb colour image converted into grayscale using Matlab's RGB2GRAY function which will convert 3D images to 2D image. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Compared with the binary image, it can be considered the grayscale information of image sufficiently and described the characters of the image perfectly. In the digital image processing, the differentiating gradient operator is usually combined with the

threshold technology to detect the edge of the image [7].

ii. Sobel Edge Detection

Edge detection is a fundamental tool for image segmentation. The major property of the edge detection technique is its ability to extract the exact edge line with good orientation. The Sobel method of edge detection for image processing finds edges using the Sobel approximation to the derivative. It precedes the edges at those points where the gradient is highest. The Sobel technique performs a 2-D spatial gradient quantity on an image and so highlights regions of high spatial frequency that correspond to edges. It is a basic process detects and outlines of an object and boundaries of objects and the background in the image [8].

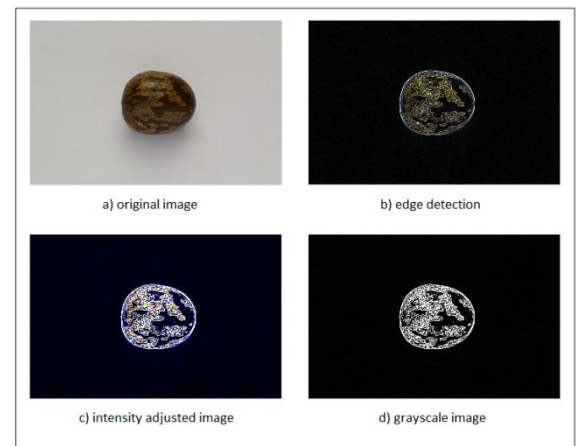


Figure 9: (a) Original image, (b) Edge detection, (c) Image adjustment, (d) grayscale image

iii. Morphological processing

Morphological image processing is a method to remove imperfections related to the form and structure of the image. This method involves opening and closing process which smoothes the contour of the object, eliminates small holes, eliminate almost all noise in the image and fill gaps in the contour.

Fig. 10 shows the outline of the seed when the morphological thresholding process is done.

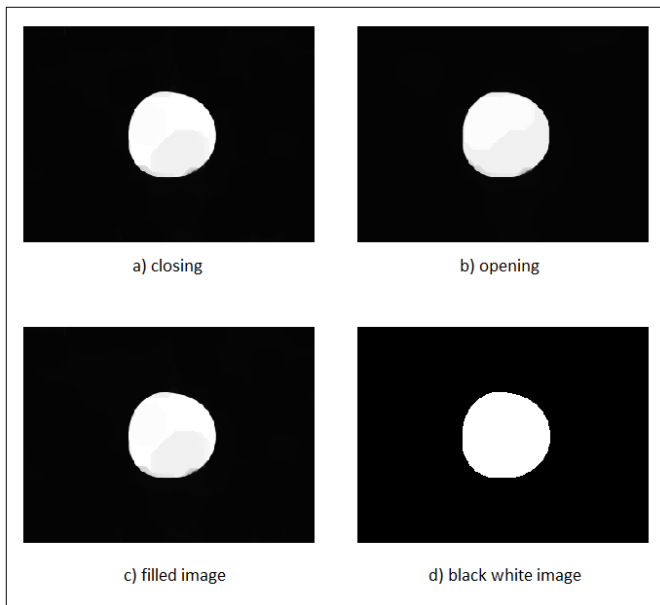


Figure 10: Resulted images of morphological process

Fig. 11 shows the image outlined border recognition process which is performed the masking between RGB image and grayscale and outlined border image that will find a good outlined border. Data from the outlined border will be used to measure the parameter of the seed.

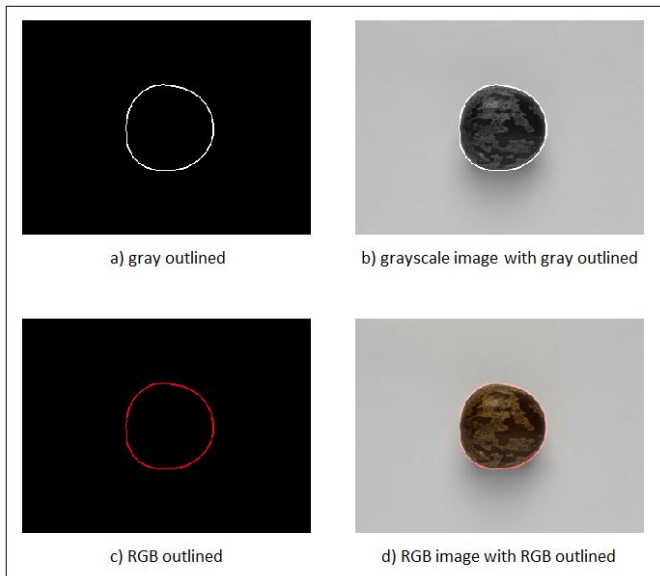


Figure 11: Outlined border recognition process to identify the perimeter.

In the final process in rubber seed image processing, radius measurement which 36 points around the outlined border from the centroid produced. Each point from the centroid to the outline are different depending on the texture

of the rubber seed. Fig. 12 shows the 36 points of radius for selected rubber seed.

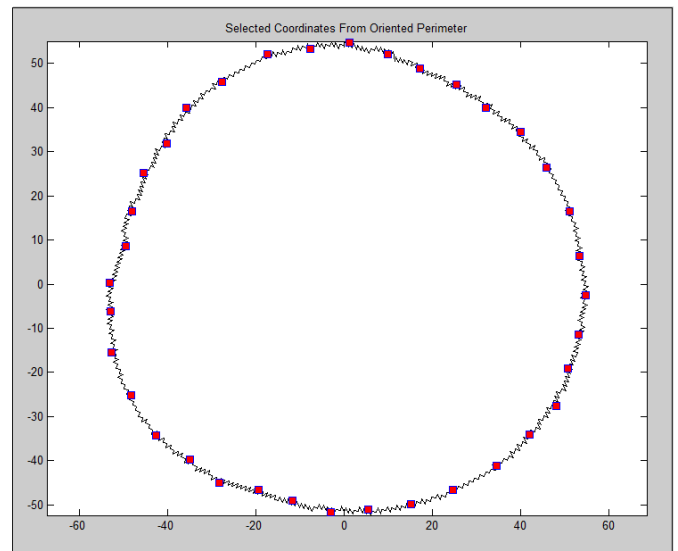


Figure 12: Radius distance of rubber seed image from centroid

III. RESULT AND DISCUSSION

All the data from Matlab are transferred to the SPSS software for the final analysis. One-way ANOVA and error plot tests conducted to prove that all the rubber seed is different from each other in term of perimeter, radius and area.

A. One-way ANOVA

Critical value, α usually set at 0.05.

H_0 = There are no significant differences in the area between the rubber seeds clones' mean

H_1 = There is a significant difference in the area between the rubber seeds clones' mean

TABLE 1 (a) : ANOVA for Area

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.057	4	.014	347.780	.000
Within Groups	.006	145	.000		
Total	.063	149			

Table 1 (a) shows there is a significant difference in area between the types of clones for all rubber seed series.

Since $p\text{-value} = 0.0000 \leq 0.05 = \alpha$, the null hypothesis (H_0) rejected and the alternative (H_1) hypothesis are accepted.

H_0 = There are no significant differences in perimeter between the rubber seeds clones' mean

H_1 = There is a significant difference in perimeter between the rubber seeds clones' mean

TABLE 1 (b) : ANOVA for Perimeter

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	4	.000	282.378	.000
Within Groups	.000	145	.000		
Total	.000	149			

Table 1 (b) shows there is a significant difference in perimeter between the types of clones for all rubber seed series. Since $p\text{-value} = 0.0000 \leq 0.05 = \alpha$, the null hypothesis (H_0) rejected and the alternative (H_1) hypothesis are accepted.

H_0 = There are no significant differences in radius between the rubber seeds clones' mean

H_1 = There is a significant difference in radius between the rubber seeds clones' mean

TABLE 1 (c) : ANOVA for Radius

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3698.021	4	924.505	265.777	.000
Within Groups	504.382	145	3.478		
Total	4202.403	149			

Table 1 (c) shows there is a significant difference in radius between the types of clones for all rubber seed series. Since $p\text{-value} = 0.0000 \leq 0.05 = \alpha$, the null hypothesis (H_0) rejected and the alternative (H_1) hypothesis are accepted.

B. Multiple Comparisons

Table 2 (a) below show the comparison between mean of area and significant values for all series clones to each other. The result shows that only RRIM2007 and RRIM2025 have the entire mean significant below than 0.05. So, it can conclude that only RRIM2007 and RRIM2025 are significantly different between each other. Whereas others series clones are no significantly different between each other.

TABLE 2 (a) : Multiple Comparisons for Area

(I) clone	(J) clone	Mean Difference (I-J)	Std. Error	Sig.
RRIM2007	RRIM2009	.02090366*	.00164887	.000
	RRIM2012	.01550952*	.00164887	.000
	RRIM2016	.01909194*	.00164887	.000
	RRIM2025	-.03111355*	.00164887	.000
RRIM2009	RRIM2007	-.02090366*	.00164887	.000
	RRIM2012	-.00539414*	.00164887	.012
	RRIM2016	-.00181172	.00164887	.807
	RRIM2025	-.05201722*	.00164887	.000
RRIM2012	RRIM2007	-.01550952*	.00164887	.000
	RRIM2009	.00539414*	.00164887	.012
	RRIM2016	.00358242	.00164887	.196
	RRIM2025	-.04662308*	.00164887	.000
RRIM2016	RRIM2007	-.01909194*	.00164887	.000
	RRIM2009	.00181172	.00164887	.807
	RRIM2012	-.00358242	.00164887	.196
	RRIM2025	-.05020549*	.00164887	.000
RRIM2025	RRIM2007	.03111355*	.00164887	.000
	RRIM2009	.05201722*	.00164887	.000
	RRIM2012	.04662308*	.00164887	.000
	RRIM2016	.05020549*	.00164887	.000

*. The mean difference is significant at the 0.05 level.

Table 2 (b) showed the comparison between mean of perimeter and significant values for all series clones to each other. The result shows that only RRIM2007 and RRIM2025 significantly different between each other.

TABLE 2 (b) : Multiple comparison for Perimeter

(I) clone	(J) clone	Mean Difference (I-J)	Std. Error	Sig.
RRIM2007	RRIM2009	.00075311*	.00005977	.000
	RRIM2012	.00048498*	.00005977	.000
	RRIM2016	.00063040*	.00005977	.000
	RRIM2025	-.00098645*	.00005977	.000
RRIM2009	RRIM2007	-.00075311*	.00005977	.000
	RRIM2012	-.00026813*	.00005977	.000
	RRIM2016	-.00012271	.00005977	.246
	RRIM2025	-.00173956*	.00005977	.000
RRIM2012	RRIM2007	-.00048498*	.00005977	.000
	RRIM2009	.00026813*	.00005977	.000
	RRIM2016	.00014542	.00005977	.112
	RRIM2025	-.00147143*	.00005977	.000
RRIM2016	RRIM2007	-.00063040*	.00005977	.000
	RRIM2009	.00012271	.00005977	.246
	RRIM2012	-.00014542	.00005977	.112
	RRIM2025	-.00161685*	.00005977	.000
RRIM2025	RRIM2007	.00098645*	.00005977	.000
	RRIM2009	.00173956*	.00005977	.000
	RRIM2012	.00147143*	.00005977	.000
	RRIM2016	.00161685*	.00005977	.000

*. The mean difference is significant at the 0.05 level.

TABLE 2 (c) : Multiple comparison for Radius

(I) clone	(J) clone	Mean Difference (I-J)	Std. Error	Sig.
RRIM2007	RRIM2009	5.64929096*	.48155991	.000
	RRIM2012	3.93234363*	.48155991	.000
	RRIM2016	5.24615679*	.48155991	.000
	RRIM2025	-7.65819785*	.48155991	.000
RRIM2009	RRIM2007	-5.64929096*	.48155991	.000
	RRIM2012	-1.71694733*	.48155991	.004
	RRIM2016	-.40313417	.48155991	.919
	RRIM2025	-1.33074888E1	.48155991	.000
RRIM2012	RRIM2007	-3.93234363*	.48155991	.000
	RRIM2009	1.71694733*	.48155991	.004
	RRIM2016	1.31381316	.48155991	.055
	RRIM2025	-1.15905415E1	.48155991	.000
RRIM2016	RRIM2007	-5.24615679*	.48155991	.000
	RRIM2009	.40313417	.48155991	.919
	RRIM2012	-1.31381316	.48155991	.055
	RRIM2025	-1.29043546E1	.48155991	.000
RRIM2025	RRIM2007	7.65819785*	.48155991	.000
	RRIM2009	13.30748881*	.48155991	.000
	RRIM2012	11.59054148*	.48155991	.000
	RRIM2016	12.90435464*	.48155991	.000

*. The mean difference is significant at the 0.05 level.

Table 2 (c) shows the comparison between mean of radius and significant values for all series clones to each other. The result shows that only RRIM2007 and RRIM2025 have all the mean difference below than 0.05 same as the analysis for mean of area. So, it can be assume only RRIM2002 and RRIM2015 are significantly different between each other.

C. Error Plots

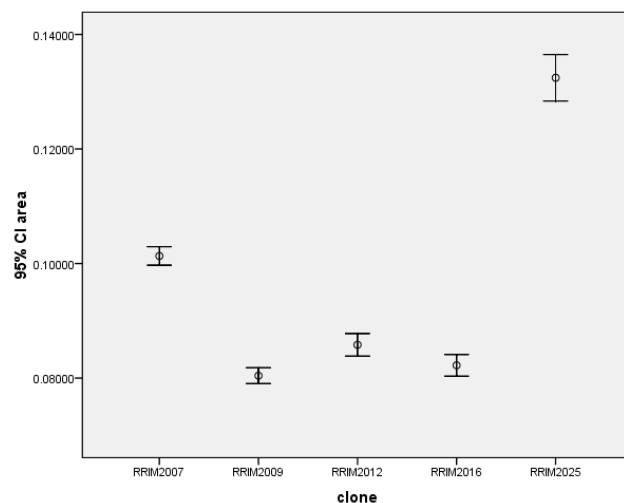


Figure 3 (a): Error plot for area comparison

Fig. 3 (a) shows the error plots for area of rubber seeds. From the figure, there is very small significantly different between RRIM 2009 and RRIM2012 but for RRIM2016, the plot overlaps with RRIM2012.

For RRIM2007 and RRIM2025, there is a huge gap between the seed so it is significantly different and not similar to each other. As a conclusion, rubber seed clone recognition using area can be applied to RRIM2007 and RRIM2025.

Table 3 (a): Descriptive Statistics of Area

	N	Mean	Minimum	Maximum
RRIM2007	30	.1013190	.09009	.10966
RRIM2009	30	.0804154	.07113	.08681
RRIM2012	30	.0858095	.07331	.09851
RRIM2016	30	.0822271	.07205	.08976
RRIM2025	30	.1324326	.11476	.15015
Total	150	.0964407	.07113	.15015

Table 3 (a) shows the Descriptive Statistics area due to minimum, maximum and mean value for all type of clone.

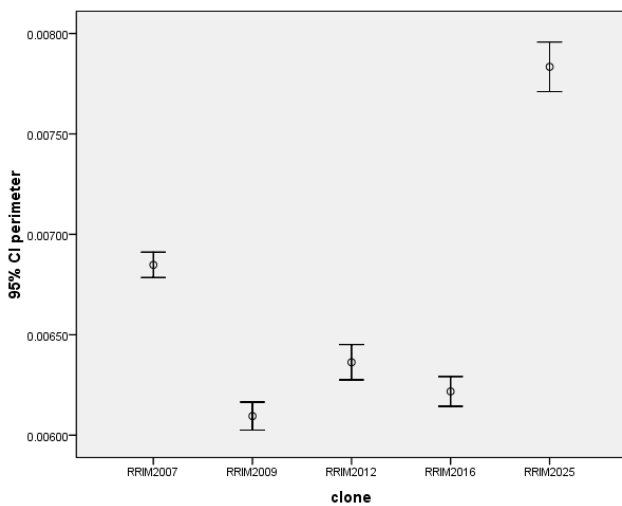


Figure 3 (b): Error plot for perimeter comparison

Fig. 3 (b) shows the error plots for perimeter of rubber seeds. From observation, RRIM2009 and RRIM2012 have small significant different and RRIM2016 overlap with RRIM2012. But for RRIM2007 and RRIM2025, there is significantly different between the seeds. As a conclusion, RRIM2007 and RRIM2025 can be recognized using perimeter method.

Table 3 (b): Descriptive Statistics of Perimeter

	N	Mean	Minimum	Maximum
RRIM2007	30	.0068480	.00642	.00719
RRIM2009	30	.0060949	.00565	.00649
RRIM2012	30	.0063630	.00582	.00705
RRIM2016	30	.0062176	.00580	.00659
RRIM2025	30	.0078344	.00725	.00835
Total	150	.0066716	.00565	.00835

Table 3 (b) shows the Descriptive Statistics of perimeter due to minimum, maximum and mean value for all types of clones.

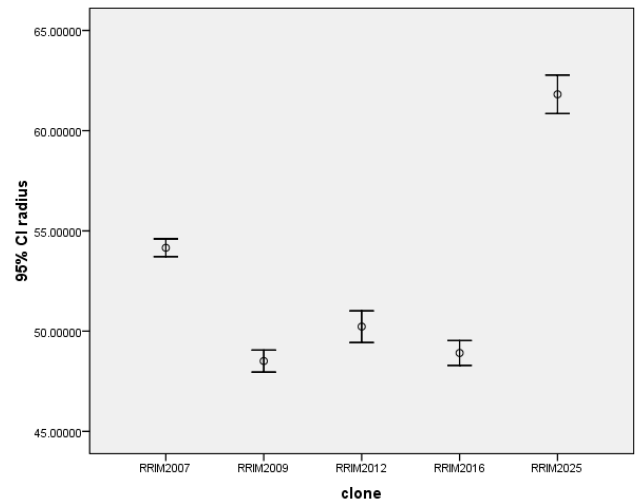


Figure 3 (c): Error plot for radius comparison

Fig. 3 (c) shows the error plots for a radius of rubber seeds. There is very small significantly different for RRIM2009 and RRIM2012 but for RRIM2016, its overlap with RRIM2012. But it is significantly different with RRIM2007 and RRIM2025 clone series. As a conclusion, only rubber seed series for 2007 and 2025 can be recognized by using a radius of seed.

Table 3 (c): Descriptive Statistics of Radius

	N	Mean	Minimum	Maximum
RRIM2007	30	54.1563806	51.13673	56.29217
RRIM2009	30	48.5070896	45.40122	51.51462
RRIM2012	30	50.2240370	46.14490	58.55208
RRIM2016	30	48.9102238	45.65325	51.83544
RRIM2025	30	61.8145784	57.61976	65.81466
Total	150	52.7224619	45.40122	65.81466

Table 3 (c) shows the Descriptive Statistics of radius due to minimum, maximum and mean value for all types of clones.

IV. CONCLUSION AND RECOMMENDATION

A. Conclusion

From this research, image processing contributing a huge help to rubber industry in Malaysia by recognizing and differentiate the best series clones. Parameter used in research like perimeter, area and radius are useful in order to differentiate between clones. With these parameters, it can help to produce a model system for rubber seed clones classification. With the aid of Matlab and SPSS, it can

conclude that RRIM 2007 and RRIM2025 can be recognized using perimeter, area and radius because of the shape of the seed itself that is very different from other seeds. For RRIM2009 and RRIM2012, there is a significant difference between both types but the data cannot be used to identify the seed because of the small significant differences between both type and for RRIM2016, the texture of the seed almost identical to RRIM2012 and cannot be recognized using all the parameters. This is due to the condition of the seed that's been collected a year ago affected the shape of the seed. To recognize all the seed, it needs a high accuracy of information from digital image to the image processing in order to differentiate each type of rubber seed clones.

B. Recommendation

From the result obtained, the accuracy in the area and radius is not high enough in order to recognize different types of rubber tree seed clones. Sobel edge detection can perform better under the circumstances of obvious edge and low noise. So many noises in a digital image will affect the data obtain. To improve the problem, other edge detection methods such as canny edge detection can be use and 3D computer generated images can be used to determine the perimeter, area and radius for rubber seeds clone.

ACKNOWLEDGEMENT

The author would like to acknowledge Mr Fairul Nazmie, project supervisor for his advice, ideas and critical guidance throughout the preparation of this project. Also acknowledgement to Rubber Research Institute of Malaysia (RRIM) for providing all the data and seed and sharing their knowledge and not to forget to Image Capture Room at the Advanced Signal Processing Lab Faculty Of Electrical Engineering UiTM Shah Alam for providing the equipment to conduct this research.

REFERENCES

- [1] Benjamaporn Lursthut and Chomtip Pornpanomchai, " *Plant Seed Image Recognition System (PSIRS)*": IACSIT International Journal of Engineering and Technology, Vol. 3, No. 6, December 2011.
- [2] Shafar Jefri Mokhtatar, Noordin Wan Daud and Nazera Arbain Department of Crop Science, Faculty of Agriculture, University Putra Malaysia,; " *Performance of Hevea brasiliensis on Haplic Ferralsol as Affected by Different Water Regimes*".
- [3] M. Z. A. Aziz, D. R. Othman, D. M. Benong, and D. O. S. Huat, *RRIM2000 SERIES Clones: Characteristic and Description*: Rubber Research Institute Of Malaysia, 1997".

- [4] Masahuling Benong, Ramli Othman, Ong Chin Wei & Nurmi Rahayu Abdul Hamid Unit Pembaikan dan Perlindungan Tanaman: " *Klon-Klon Siri RRIM 2000*".
- [5] H. M. Salleh,; " *An Automated Shape Recognition for Rubber Seed Clones Through Imaging Techniques*", in Faculty of Electrical Engineering, vol. Bsc.Eng (Hons.) Electrical
- [6] M. F. Muhammad,; " *An Automated Color Recognition for Rubber Tree Seed Clones Using MCS 600 Spectrometer*", in Faculty of Electrical Engineering, vol. Bsc.Eng (Hons.) Electrical. Shah Alam: Universiti Teknologi Mara, 2009.
- [7] Linan Fan, Yong Wen, Xinhe Xu , " *Research on Edge Detection of Gray-scale Image Corrupted by Noise Based on Multi-structuring Elements*", School of Information Science & Engineering Northeastern University, Shenyang
- [8] Muthukrishnan and M.Radha,; " *Edge Detection Techniques for Image Ssegmentation*", International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 6, Dec 2011.
- [9] M. Thaler, H. Hochreutener, " *Image processing basics using MATLAB*", February 2008, c ZHAW
- [10] R. C.Gonzalez, R. E.Woods, and S. L.Eddins, " *Digital Image Processing Using MATLAB*", Prentice Hall, 2004.