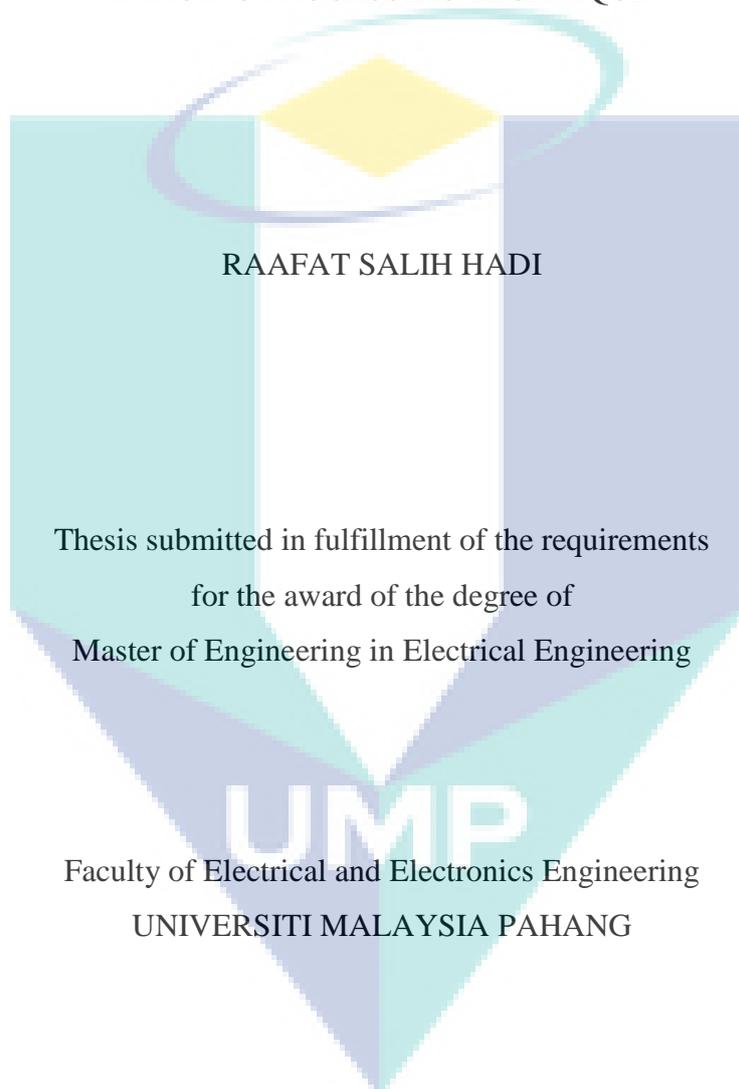


CLASSIFICATION OF HUMAN PARASITIC WORM USING MICROSCOPIC
IMAGING PROCESSING TECHNIQUE



July 2013

SUPERVISOR'S DECLARATION

We hereby declare that we have checked this thesis and in our opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Master of Engineering in (Electrical Engineering).

Signature:

Name of Supervisor:

Position:

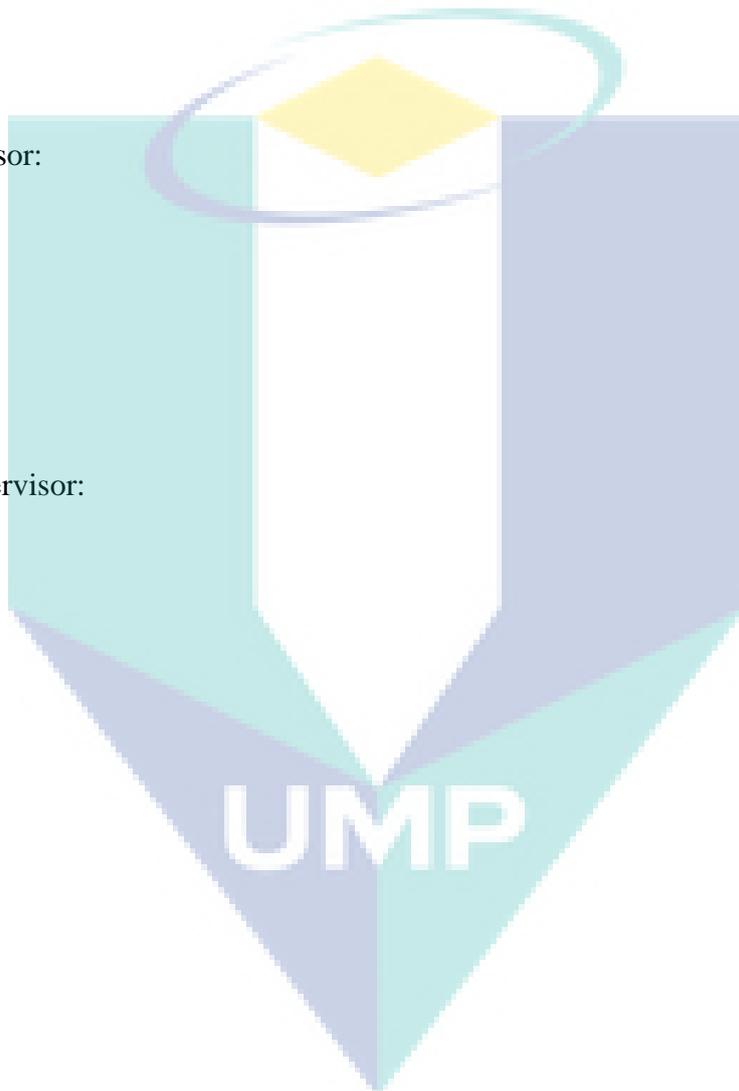
Date:

Signature:

Name of Co-supervisor:

Position:

Date:



STUDENT'S DECLARATION

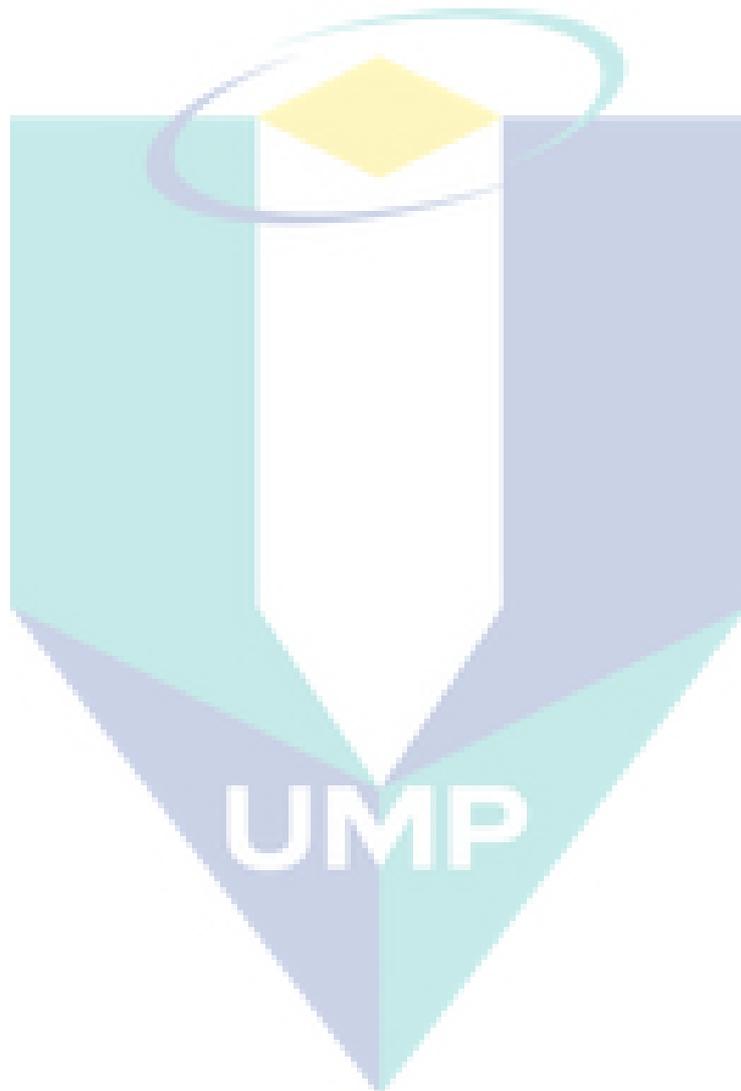
I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged. The thesis has not been accepted for any degree and is not concurrently submitted for award of other degree.

Signature:

Name:

ID Number:

Date:

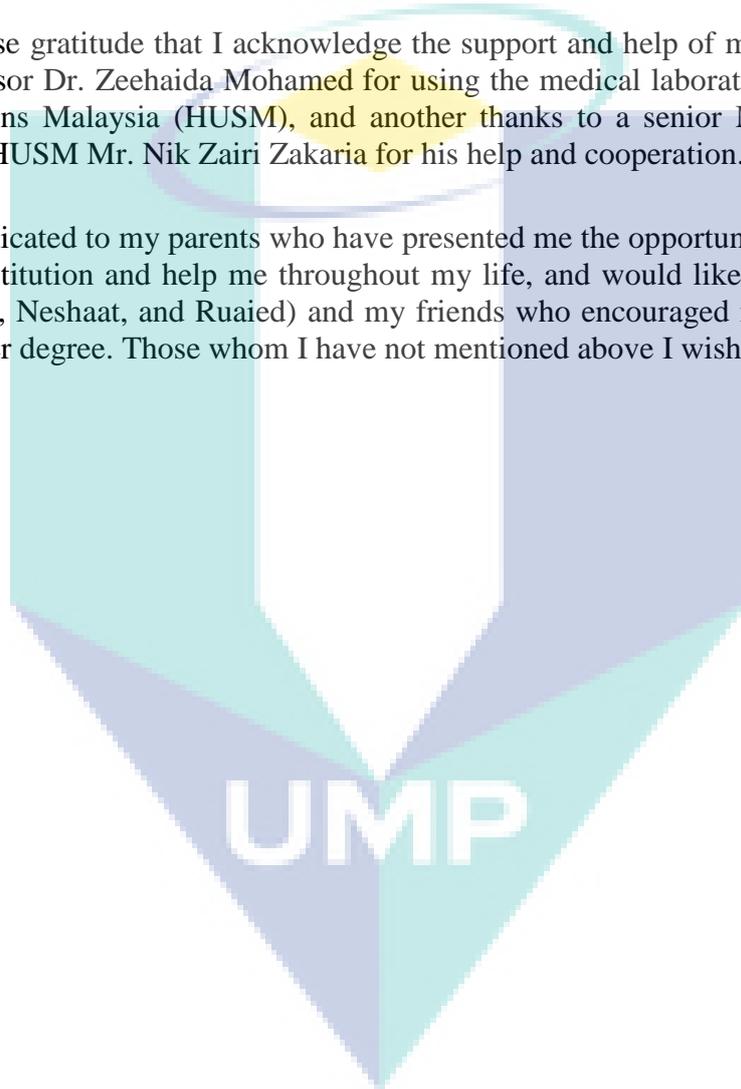


ACKNOWLEDGEMENTS

First and foremost I offer my sincerest gratitude to my supervisor, Associate Professor Dr. Kamarul Hawari bin Ghazali, who has supported me throughout my thesis with his patience and knowledge. I attribute the level of my Master degree to his encouragement and effort and without him this thesis would not have been completed or written. One simply could not wish for a better or friendlier supervisor.

It is with immense gratitude that I acknowledge the support and help of my field supervisor, Associate Professor Dr. Zeehaida Mohamed for using the medical laboratory in the Hospital of Universiti Sains Malaysia (HUSM), and another thanks to a senior Medical laboratory Technologist in HUSM Mr. Nik Zairi Zakaria for his help and cooperation.

This thesis is dedicated to my parents who have presented me the opportunity of an education from the best institution and help me throughout my life, and would like to thank my three brothers (Ahmed, Neshaat, and Ruaied) and my friends who encouraged me to continue my study for a Master degree. Those whom I have not mentioned above I wish to thank you all.



ABSTRACT

Human parasitic infection causes diseases to people whether this infection will be inside the body called endoparasites, or outside of the body called ectoparasites. Human intestinal parasite worms infected by air, food, and water are the causes of major diseases and health problems. So in this study, a technique to identify two types of parasites in human fecal, that is, the eggs of the worms is proposed. In this strategy, digital image processing methods such as noise reduction, contrast enhancement, and other morphological process are applied to extract the eggs images based on their features. The technique suggested in this study enables us to classify two different parasite eggs from their microscopic images which are roundworms (*Ascaris lumbricoides* ova, ALO) and whipworms (*Trichuris trichiura* ova, TTO). This proposed recognition method includes three stages. The first stage is a pre-processing sub-system, which is used to obtain unique features after performing noise reduction, contrast enhancement, edge enhancement, and detection. The next stage is an extraction mechanism which is based on five features of the three characteristics (shape, shell smoothness, and size). The final stage, the Filtration with Determinations Thresholds System (*F-DTS*) classifier is used to recognize the process using the ranges of feature values as a database to identify and classify the two types of parasites. The overall success rates are 93% and 94% in *Ascaris lumbricoides* and *Trichuris trichiura*, respectively.

ABSTRAK

Jangkitan parasit kepada manusia yang membawa wabak penyakit samada dari dalam badan yang dikenali sebagai endoparasites, atau dari luar badan manusia yang dikenali sebagai ectoparasites. Bahagian dalaman manusia dijangkit oleh parasit yang boleh didatangi dari udara, makanan dan air minuman adalah penyebab kepada wabak utama dan masalah kesihatan. Penyelidikan ini mengkaji satu teknik untuk mengenalpasti dua jenis parasit dalam najis manusia iaitu dengan mencadangkan kajian berkaitan telur kepada cecacing. Strategi ini adalah dengan menggunakan teknik pemprosesan gambar seperti pengurangan gangguan, peningkatan kontra, dan morfologi lain telah digunakan untuk mengekstrak telur berdasarkan sifat-sifat cecacing itu. Teknik yang dicadangkan dalam penyelidikan ini adalah untuk membolehkan kita mengenalpasti dan mengklasifikasi dua jenis telur parasit dari gambar mikroskopi iaitu roundworm (*Ascaris lumbricoides* ova, ALO) dan whipworms (*Trichuris trichiura* ova, TTO). Teknik cadangan pengenalan melibatkan tiga peringkat. Peringkat pertama adalah pra-pemprosesan sub-sistem iaitu digunakan dengan mengambil sifat-sifat unik selepas melakukan pengurangan gangguan, peningkatan kontra, peningkatan sisi, dan pengenalpastian. Peringkat seterusnya adalah mekanisma pengekstrakan iaitu berdasarkan lima sifat; tiga sifat yang dikenalpasti seperti bentuk, kelembutan cengkerang dan juga saiz. Peringkat terakhir adalah menggunakan pengelasan Filtration with Determinations Thresholds System (F-DTS) untuk mengenalpasti proses menggunakan kadar sifat tertentu sebagai pangkalan data untuk mengenalpasti dan mengklasifikasi pelbagai jenis parasit. Secara keseluruhan, kadar kejayaan adalah 93% dan 94% untuk *Ascaris lumbricoides* and *Trichuris trichiura*.

TABLE OF CONTENTS

SUPERVISOR'S DECLARATION	II
STUDENT'S DECLARATION	III
ACKNOWLEDGEMENTS	IV
ABSTRACT	V
ABSTRAK	VI
TABLE OF CONTENTS	VII
LIST OF TABLES	XI
LIST OF FIGURES	XII
CHAPTER I INTRODUCTION	1
1.1 BACKGROUND OF PARASITIC WORMS IN HUMAN	1
1.1.1 Ascaris Lumbricoides (AL)	3
1.1.2 Trichuris Trichiura (TT)	6
1.2 PROBLEM STATEMENT	8
1.3 MOTIVATION	9
1.4 OBJECTIVES OF RESEARCH	9

1.5	SCOPE OF RESEARCH	9
1.6	STRUCTURE OF THE THESIS	10
CHAPTER II	LITERATURE REVIEW	11
2.1	INTRODUCTION	11
2.2	GENERAL PARASITE DETECTION USING IMAGE PROCESSING TECHNIQUE	11
2.3	PARASITE DETECTION BASED ON FEATURES	12
2.3.1	Color and shape based detection system	12
2.3.2	Color based detection system	14
2.3.3	Shape based detection system	15
2.4	OVERVIEW OF HUMAN PARASITIC WORMS CLASSIFICATION	18
2.5	SUMMARY	22
CHAPTER III	METHODOLOGY	23
3.1	INTRODUCTION	23
3.2	DATA ACQUISITION	24
3.3	PRE-PROCESSING STAGE	27
3.3.1	Contrast enhancement	30

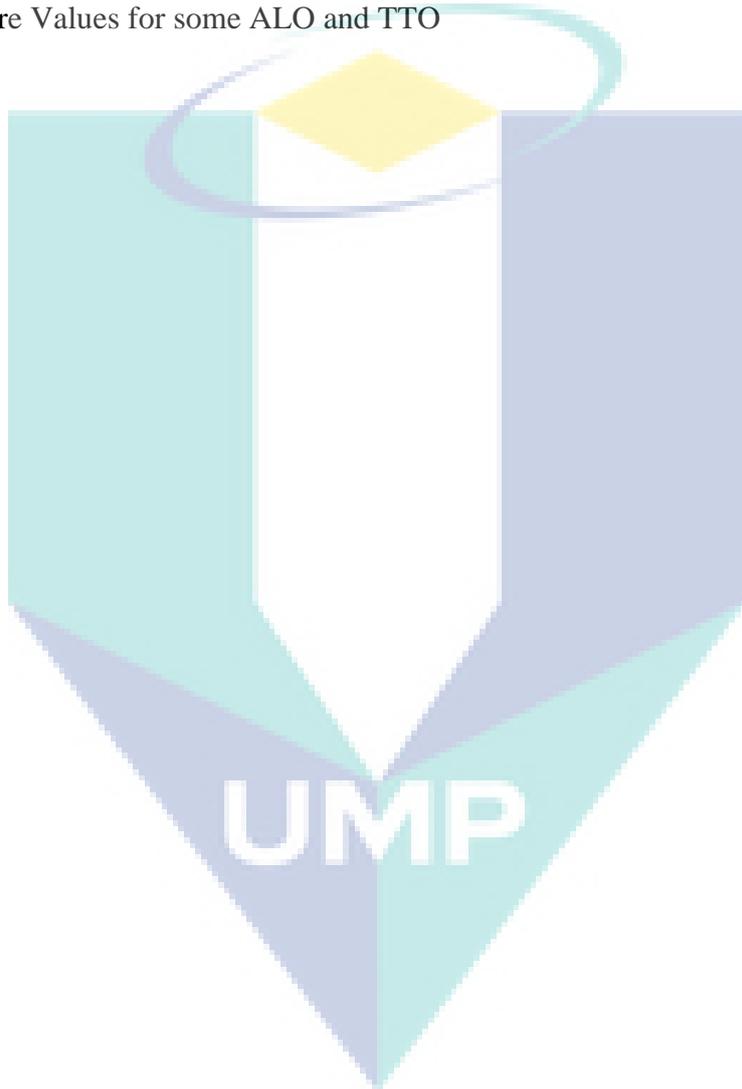
3.3.2	Edge enhancement	31
3.3.3	Filtering technique	35
3.4	ADDITIONAL PROCESSING STEPS	47
3.4.1	Circumference	48
3.4.2	Objects' bores removal	48
3.4.3	Clearing image borders	49
3.4.4	Pixel and mean values	51
3.5	FEATURE EXTRACTION	53
3.5.1	Algorithms of feature extraction	53
3.5.2	Determine threshold values of the features	55
3.6	CLASSIFICATION STAGE	61
3.7	SUMMARY	65
CHAPTER IV	RESULTS AND DISCUSSION	66
4.1	INTRODUCTION	66
4.2	CONTRAST ENHANCEMENT EVALUATION	66
4.2.1	Contrast threshold evaluation	68
4.3	EFFECT OF MEDIAN FILTERING	72

4.4	EDGE DETECTION EFFICENCY	74
4.4.1	Edge detection threshold evaluation	77
4.5	PERFORMANCE EVALUATION OF PRE-METHODS APPROACH	79
4.6	EFFECT OF VARIOUS PARAMETERS VALUES ON THE RESULTS	81
4.6.1	Comparison of contrast enhancement techniques in the results	81
4.6.2	Contrast coefficient evaluation	82
4.6.3	Edge sharpness evaluation	84
4.6.4	Median filtering evaluation	86
4.6.5	Comparison of using edge detection techniques	87
4.7	SUMMARY	90
CHAPTER V	CONCLUSION	94
5.1	FUTURE WORK	95
REFERENCES		96
APPENDICES		102
A	Program Codes	102
B	List of Publication and Awards	109

LIST OF TABLES

Table 3.1 Size of masks with various Sigma values 39

Table 3.2 Feature Values for some ALO and TTO 56



LIST OF FIGURES

Figure 1.1: The life cycle of <i>Ascaris lumbricoides</i> .	5
Figure 1.2: <i>Ascaris Lumbricoides</i> Ova (ALO).	5
Figure 1.3: <i>Trichuris Trichiura</i> Ova (TTO).	7
Figure 1.4: the life cycle of <i>Trichuris Trichiura</i> (TT).	8
Figure 2.1: Overall algorithm Chart.	13
Figure 2.2: The images show segmentation in HSV color space	15
Figure 2.3: Boundary extracted image using the Chan-Vese segmentation method	16
Figure 2.4: Flow chart of the red blood cells counting process.	17
Figure 2.5: A circle and its center determined by 3 points that are not in a straight line.	18
Figure 3.1: General diagram of the study.	24
Figure 3.2: Magnification powers of ALO and TTO parasitic eggs under microscope	26
Figure 3.3: Three different pre-processing methods	29
Figure 3.4: Using Pre-method III in Preprocessing Stage to detect ALO and TTO Parasites	29
Figure 3.5: Contrast enhancement curve.	30
Figure 3.6: Edge enhancement block diagram.	32

Figure 3.7: Unsharp masking	33
Figure 3.8: Edge enhancement with unsharp masking	34
Figure 3.9: Median Filter	36
Figure 3.10: Effect of median filter on ALO	37
Figure 3.11: Schematic of canny edge detection.	38
Figure 3.12: Curve of Gaussian G and curve of its 1st derivative of G.	40
Figure 3.13: Examples of convolution masks.	40
Figure 3.14: Gradient of image.	41
Figure 3.15: 3x3 Convolution masks of Gradient image.	42
Figure 3.16: Gradient Orientation.	43
Figure 3.17: Pixel Interpolation	43
Figure 3.18: Applying masks to an image with different sigma	46
Figure 3.19: Effect of changing in low threshold with constant high threshold = 0.7 and sigma = 1	47
Figure 3.20: Effect of changing in High threshold with constant Low threshold 0.2 and sigma = 1	47
Figure 3.21: Filling holes of objects	48
Figure 3.22: Clear boundary of input image	51

Figure 3.23: Reading a pixel value of a point in an image.	52
Figure 3.24: Showing how the five features are extracted	54
Figure 3.25: Feature values of the parasites: areas of ALO in 100 images.	57
Figure 3.26: Feature values of the parasites: areas of TTO in 100 images.	57
Figure 3.27: Feature values of the parasites: lengths of ALO in 100 images.	58
Figure 3.28: Feature values of the parasites: lengths of TTO in 100 images.	58
Figure 3.29: Feature values of parasites: widths of ALO in 100 images.	59
Figure 3.30: Feature values of parasites: widths of TTO in 100 images.	59
Figure 3.31: Feature values of parasites: boundary lengths of ALO in 100 images.	60
Figure 3.32: Feature values of parasites: boundary lengths of TTO in 100 images.	60
Figure 3.33: Feature values of parasites: roundness of ALO in 100 images.	61
Figure 3.34: Feature values of parasites: roundness of TTO in 100 images.	61
Figure 3.35: Block diagram of classification stage.	62
Figure 3.36: Classifier Tasks in Classification Stage.	62
Figure 3.37: Unwanted and confusing objects in human fecal specimens.	64
Figure 3.38: Morphological operation of parasite (<i>Ascaris lumbricoides</i> Ova (ALO))	64
Figure 3.39: Morphological operation of parasite (<i>Trichuris trichiura</i> Ova (TTO))	64

Figure 4.1: Effectiveness of three contrast enhancement techniques in image	67
Figure 4.2: Effectiveness of three contrast enhancement techniques in image	68
Figure 4.3: Effecting of gamma value in image enhancement on the ALO parasite eggs	70
Figure 4.4: Effecting of gamma value in image enhancement on the TTO parasite eggs	71
Figure 4.5: Applying of median filter to the images	73
Figure 4.6: Comparison of performance of five edge detection techniques to detect ALO	75
Figure 4.7: Comparison of performance of five edge detection techniques to detect TTO	76
Figure 4.8 : Applying various values of sigma in 'canny' edge detection and show the changing in the output of ALO parasite	77
Figure 4.9 : Applying various values of sigma in 'canny' edge detection and show the changing in the output of TTO parasite	78
Figure 4.10: Comparison of results by using three pre-methods to detect ALO.	80
Figure 4.11: Comparison of results by using three pre-methods to detect TTO.	80
Figure 4.12: Comparison of using three techniques for contrast enhancement to detect ALO and TTO.	82
Figure 4.13: Effecting of gamma on successful detection ratio for ALO.	83
Figure 4.14: Effecting of gamma on successful detection ratio for TTO.	84
Figure 4.15: Effect of alpha on edge sharpness efficiency	85
Figure 4.16: Comparison of the results of using median filter (once and twice) for	

detecting ALO 86

Figure 4.17: Comparison of the results of using median filter (once and twice) for detecting TTO 87

Figure 4.18: Comparison results of five different filters applied in ALO detection. 88

Figure 4.19: Comparison results of five different filters applied in TTO detection. 88

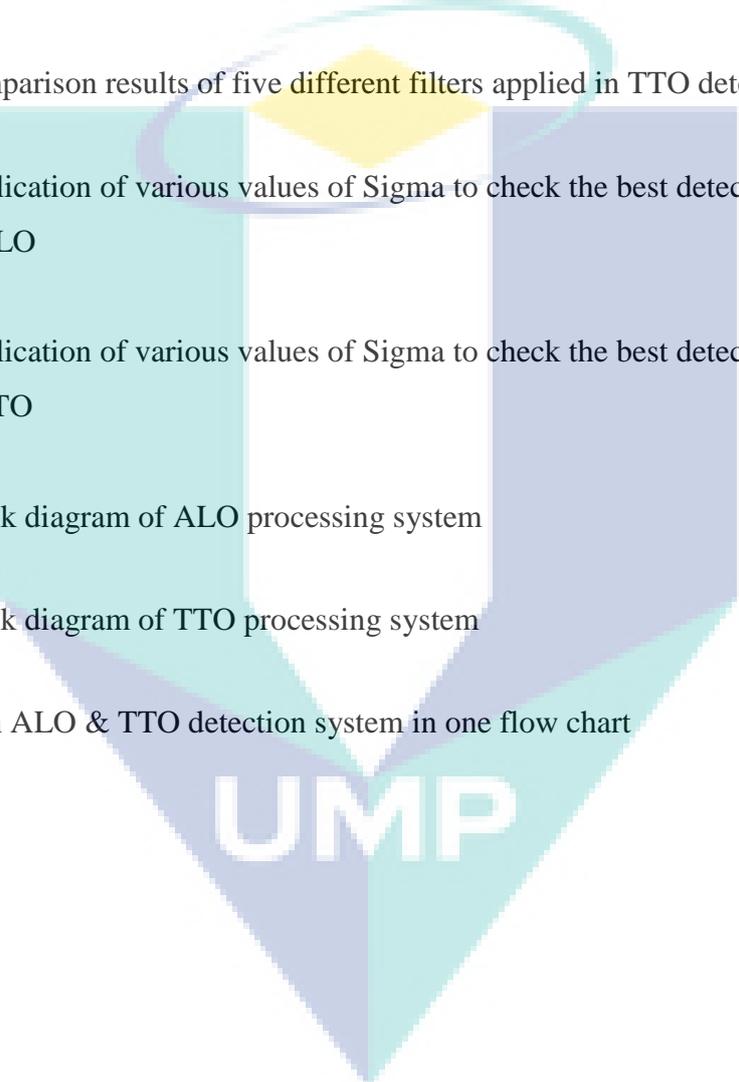
Figure 4.20: Application of various values of Sigma to check the best detection ratio in ALO 89

Figure 4.21: Application of various values of Sigma to check the best detection ratio in TTO 90

Figure 4.22: Block diagram of ALO processing system 91

Figure 4.23: Block diagram of TTO processing system 92

Figure 4.24: Both ALO & TTO detection system in one flow chart 93



UMP

THESIS CONFIDENTIAL STATUS

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THE THESIS AND COPYRIGHT

Author's full name : Raafat Salih Hadi

Date of birth : 26/04/1984

Title : Classification of Human Parasitic Worm Using Microscopic
Imaging processing Technique.

Academic Session : _____

I declared that this thesis is classified as:

CONFIDENTIAL

(Contains confidential information under the official Secret Act 1972)

RESTRICTED

(Contains restricted information as specified by the organization where research is done)

OPEN ACCESS

I agree that my thesis to be published as online open access (Full text)

I acknowledge that Universiti Malaysia Pahang reserves the right as follows:

1. The thesis is the property of Universiti Malaysia Pahang.
2. The library of Universiti Malaysia Pahang has the right to make copies for the purpose of the research only.
3. The library has the right to make copies of the thesis for academic exchange.

Certified by:

(Student's Signature)

A5946773

New IC / Passport Number
Date:

(Signature of Supervisor)

Assoc. Prof. Dr. Kamarul Hawari

Name of Supervisor
Date:

CHAPTER I

INTRODUCTION

1.1 BACKGROUND OF PARASITIC WORMS IN HUMAN

Human parasites include various protozoas and worms which may infect humans and causing parasitic diseases. Human parasites can be divided into endoparasites, which will cause infection inside the body, and ectoparasites, which will cause infection superficially within the skin.

The cysts and eggs of endoparasites may be found in fecal which offer help in the detection of parasites in the human host and also provide a means for the parasitic species to exit the host and enter other hosts (Krumhardt, 2008). Although there are different ways to contract parasitic infections, perceiving basic hygiene and cleanliness can reduce its probability (Woods & Walker, 2009).

Human intestinal parasite worms infected by air, food, and water are the causes of major diseases and health issues. The most common symptoms of worm infestation include diarrhea, stomach bloating, and digestive disorders. Other symptoms include anemia, asthma, constipation, fatigue, low immune system, nervousness, and skin rash. There are over 100 different types of parasitic worms living inside a human body. Some are microscopic in size while others can be seen quite easily. These common organisms can be found everywhere in our environment, in the air that we breathe, in the water that we drink, or in the food that we eat.

Parasites are organisms that live on or in other organism, obtaining nutrients to live. Parasites' name comes from the Greek word "para" that means beside, and "sitos", which means food. Most living parasites require a host to complete their life cycles. Besides human beings, animals can also be served as the host. The parasites vary in size from one-thousandth of one micron to whale tapeworms of one hundred feet long.

Parasites and worms can invade our bodies through food and water intake, and through transmitting agents (like a mosquito), sexual conduct or through the nose and skin. Once established, they will eat the same food you eat or worse, they will eat you.

The overseas travel and international food fads have open the door for parasitic illness to migrate from one country to other distant countries. The parasitic disorders are recognised through the identification of parasitic organisms in the blood, urine, feces, and tissues by using the appropriate recognition methods (Woods & Walker, 2009).

People with intestinal parasite infections are usually under-nourished and weak, infected with virus, fungus, or bacteria, and have various types of chemical and metal poisoning. Human intestinal parasites can be present in any kind of disease, in any kind of person, and at any age. They are responsible for many health problems because they secrete toxins and steal the vital nutrients from our bodies. They can irritate or exaggerate other health problems we may be experiencing. Indeed, everybody is under their mercy and at risk during parasitic infection. Other side cancer cases are afflicted with worms that often lump together and look like tumors (Gorbach, S.L., Bartlett, J.G., & Blacklow, N.R. 1992).

Female worms can release between 3,000 to 200,000 eggs per day depending on their type. There are 3200 varieties of parasites divided into two major groups, Protozoa and Helminths such as nematodes, trematodes, and cestodes (Krumhardt, 2008) or some sources have classified them into four major groups: Protozoa, Trematoda, Cestoda, and Nematoda. In this writing, one group has been chosen and two examples of this group will be studied. The selected group, which is the Nematodes, includes common roundworms (*Ascaris lumbricoides*), hookworms, whipworms (*Trichuris trichiura*), Pinworms, heart worms, *Strongyloides*, *Stercoralis*, *Ancylostoma*, *Caninum*, *Toxocara* worm and trichinosis. The size of these worms can vary from 2 centimeters to 35 centimeters.

Roundworm looks similar to an earthworm and can produce up to 200,000 eggs daily. Approximately 1,008 million people are infected by this type of worm, making it the most common worldwide. The most frequent symptom of a roundworm infection is

upper abdominal discomfort. Other symptoms are asthma, eye pain, insomnia, and rashes due to the secretions or waste products from the worms. Large number of worms in the body can cause blockages in the intestinal tract, hemorrhage when they penetrate the intestinal wall, appendicitis, peritonitis, abscesses in the liver, hemorrhagic pancreatitis, loss of appetite, and insufficient absorption of digested foods. The adult worm can grow up to 15 inches long.

In addition, infections from Whipworms are estimated at several hundred million worldwide. Symptoms of Whipworms infection are bloody stools, pain in the lower abdomen, weight loss, rectal prolapsed, nausea and anemia. Hemorrhage can occur when worms penetrate the intestinal wall and bacterial infections usually follow 1 to 2 inches in length.

In order to diagnose parasites infection in the human body, the most common way is by manual fecal examination. Trained experts examine the fecal specimens, search for parasitic organisms, and the eggs of helminths and cysts of protozoa. If the harmful organisms are present, they examine the sizes, shapes, numbers and sometimes colour to identify the species of parasites, the degree of infection, and appropriate therapeutic modalities. Yet, diagnosticians are reluctant to perform fecal examinations with the decreasing importance of most parasites, less material is submitted for examination, which results in an insufficiency of training material. Thus, it is difficult to maintain enough people with expertise in diagnosing parasites. Worst of all, Parasitologists are expected to maintain high standards of proficiency in the identification of rare parasites as well as common endemic ones. This study is focusing on the two most common human parasitic worms, namely:

- *Ascaris Lumbricoides* (AL)

- *Trichuris Trichiura* (TT)

1.1.1 *Ascaris Lumbricoides* (AL)

Ascaris lumbricoides, or giant roundworm, is the most common parasitic worm in a human being. According to some estimates, 25 % of humans are recorded to be infected with the disease, ascariasis, which occurs worldwide, and is mostly found in tropical and semi-tropical countries (Warren & Mahmoud, 1977). It has the highest expansion in areas of poor sanitation and where human feces are used as fertilizer. The majority of people infected with ascariasis live in Asia (up to 73%), Africa (around 12%) and South America (8% only) where some populations have an infection rate amount to 95% (Sarinas & Chitkara, 1977) (Reeder, 1988). In the United States, the expansion of infection has decreased spectacularly after the introduction of modern sanitation and loss treatment in the year 1900s (Jones, 1983). It is recorded that the current expansion of *Ascaris lumbricoides* in feces samples is approximately 2% in the United States, but it reaches more than 30% in children between 1- 5 years old, particularly in village areas in the South (Tietze & Tietze, 1991) (Jones, 1981). It is as well seen in travelers from endemic areas (Sarinas & Chitkara, 1977).

In the life cycle, *Ascaris lumbricoides* take about three months to mature. Ascariasis begins, when *Ascaris lumbricoides* eggs are accidentally swallowed. They can be obtained from dirty fingers, water or food that has been polluted with feces of an infected human. Larvae hatch from the eggs, permeate in the intestinal wall and enter the bloodstream. They stop at pulmonary arteries and then live in the lungs for two weeks. They infringe into the alveoli and travel up the respiratory system to the throat to be swallowed again. Emigration is needed for the larvae to develop into adults. Adult worms are ready to mate after they attach themselves to the intestinal wall. Adult worms survive by eating food digested by the host and live up to two years (see figure 1.1). A female can produce about 200,000 microscopic eggs per day that are passed in feces. The eggs inseminate into the infective stage in a few weeks in the best conditions in the soil and unfertilized eggs are not infective at all. The eggs are very remittent to chemicals, extreme temperatures and other rough conditions and they can survive for months. The size of adult females is 20–35 cm in length and 3–6 mm in width. Male worms are a bit smaller reaching 15–30 cm in length and 2–4 mm in diameter. The ova are oval (see figure 1.2),

have thick shell, a mammillated outer coat, and measure 45- 70 μm by 35-50 μm (Khuroo, 1996).

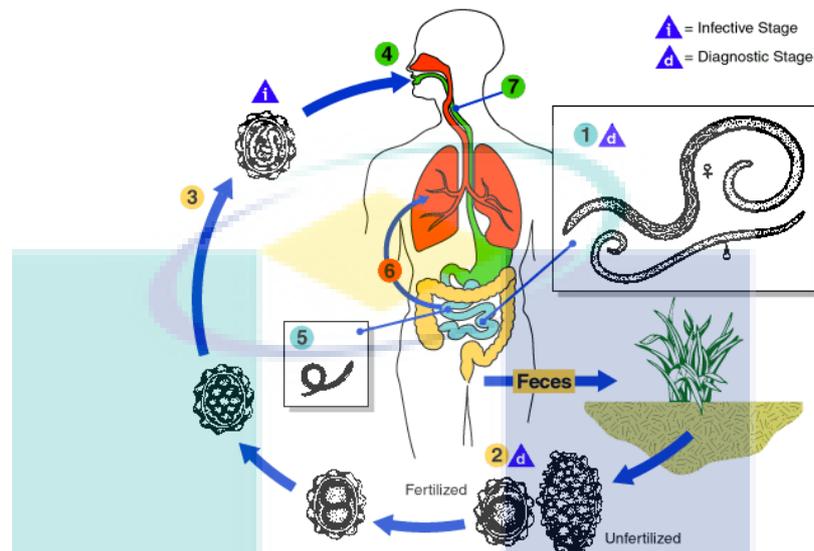


Figure 1.1: The life cycle of *Ascaris lumbricoides*. (1) Adult worms, (2) Unfertilized eggs, (3) Fertilized eggs, (4) Infective eggs are swallowed, (5) The larvae hatch, (6) emigration to the lungs, (7) The larvae mature further in the lungs to the throat and are swallowed again. (Source: DPDx, the CDC Parasitology Website)



Figure 1.2: *Ascaris Lumbricoides* Ova (ALO). (a) Fertilized egg. (b) Unfertilized egg.

The majority of infections with *Ascaris lumbricoides* are asymptomatic. However, the onus of symptomatic disease worldwide is still comparatively high because of the high prevalence of disease. The clinical disease is largely limited to individuals with a high worm load (Khuroo, 1996). If symptoms occur, they relate either to the larval migration stage or to the adult worm intestinal stage.

1.1.2 *Trichuris Trichiura* (TT)

Trichuris trichiura is a gastrointestinal nematode in the family of *Trichuridae*, gender *Trichuris* (Holland, 2006). *T. Trichiura* adult worms are characterized by long and thin anterior end that lies in a shelter in the host mucosa, and thicker end that expands into the intestinal lumen. The colour of the worms is white with length 30-50 mm (Wolfe, 1978) (Ash & Orihel, 2003). The Adult parasites are long lived; normally they live up to 5 years in their host and may remain infective for at least 1- 2 years (Ash& Orihel, 2003). At the second stage of *Trichuris trichiura* 's life cycle, the measure of larvae is roughly 260 μm x 15 μm in length (Holland, 2006). The eggs of *T. Trichiura* are barrel (or lemon) in shape with a characteristic plug at both ends, giving it a tea griddle appearance (Bundy & Cooper, 1989) (Wolfe, 1978; Holland., 2006; Stephenson, Holland & Cooper, 2000) (see figure 1.3). They are usually brown in human fecal specimens and measure roughly from 50-55 μm by 22-24 μm ; however, much larger (78 by 30 μm) eggs have been noticed in stool samples from humans infected with *T. Trichiura*, and for that reason it is somehow not easy to distinguish the eggs of *T. Trichiura* from those of other members of the *Trichuris* genus (*T. Suis* and *T. Vulpis*) by measurement alone (Wolfe, 1978) (Acha & Szyfres, 2003) (Ash & Orihel, 2003) (Yoshikawa, Yamada, Matsumoto, & Yoshida1989). Male worms are smaller than females, have a coiled posterior end (Wolfe, 1978) (Acha & Szyfres, 2003) (Stephenson, Holland, & Cooper, 2000). The female worm produces between 3,000 and 20,000 eggs per day (Bundy & Cooper, 1989). The distinctive feature of *T. Trichiura* is the stichosome, which is a glandular structure encircling the scrawny esophagus of the thin anterior, which is half the size of worm (Bundy & Cooper, 1989) (Holland, 2006).

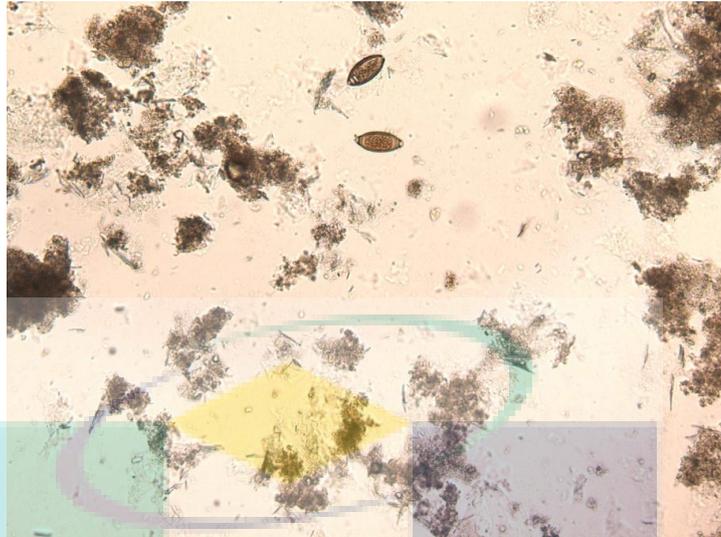


Figure 1.3: *Trichuris Trichiura* Ova (TTO)

In the life cycle of *Trichuris trichiura*, unembryonated (immature) eggs are passed with the feces (Weiss, 2001). In the soil, with the right conditions, the eggs develop into a 2-cell stage, an advanced cleavage stage, and then they embryonate; eggs become infective in 2-4 weeks. After ingestion (soil-polluted hands or food), the eggs hatch in the small intestine, and liberate larvae that mature and establish themselves as adults in the colon (Gorbach, Bartlett, & Blacklow, 1992) (Ismail, & Jayakody, 1999). The adult parasites (roughly 4 cm in length) live in the cecum and ascending colon. The adult worms are settled in that location, with the anterior parts threaded into the mucosa. The females begin to oviposit 60-70 days after infection (Mandell, Douglas, & Bennett, 1990). Female worms in the cecum produce between 3,000 - 20,000 eggs per day (Ismail & Jayakody, 1999) (Markell, Edward, John, David, Krotoski, & Wojciech, 1999). The lifespan of the adult worms is about one year (see figure 1.4).

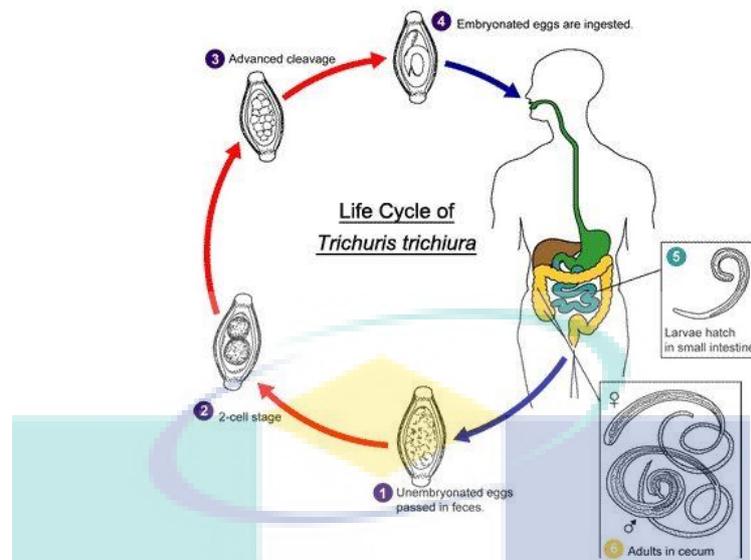


Figure 1.4: The life cycle of *Trichuris Trichiura* (TT). (1) Unembryonated egg stage, (2) 2-cell stage, (3) Advanced cleavage, (4) Embryonated egg stage, (5) The larvae hatches in the small intestine. (Source: DPDx, the CDC Parasitology Website)

Usually patients find out that they are infected with *Trichuris trichuria* when being inspected for the presence of other kind of parasites. Most cases are asymptomatic and sometimes hard to diagnose them if there are only a few worms. *Trichuris trichuria* is frequently discovered in combination with *Ascaris*, hookworm, or *Entamoeba histolytica*, which shares a hassling geographical distribution and have more severe symptoms (Doerr & Seifert, 1995).

1.2 PROBLEM STATEMENT

The problem statement of this study is listed as follows:

- The conventional methods of detecting worms in biomedical applications might carry humanity's mistakes in diagnosis of the diseases in terms of accuracy of detection (Baker GR, Norton P, 2002).
- On the other hand, these conventional methods take a time during examination of the specimen, thus with less accuracy in the result in terms of classifying and counting the number of the parasites in the slide.

1.3 MOTIVATION

To keep up with modern technologies and their facilities, we have to invest new ideas to provide the best solution to a problem. In medical laboratory applications, the new technology of equipments such as auto-focusing and auto-sliding microscope have encouraged us to search and develop new ideas to help the user for easier work in the medical field. So, writing a programme and using image processing techniques is useful to complete the whole system with such microscope that doctors need it for automatical parasite detection.

1.4 OBJECTIVES OF RESEARCH

The primary objective of the study is to classify the two categories of human intestinal parasites (ALO and TTO) using microscopic image analysis. In order to achieve this objective, the following secondary objectives have been proposed:

- Pre-processing techniques are used to develop the contrast and edge enhancements, noise reduction and edge detection.
- Processing techniques are used to develop feature extraction by clear boundary and some morphological processing.
- A classification stage is used to classify the two parasites using the threshold technique.

1.5 SCOPE OF RESEARCH

The scopes of this study are listed below:

1. This research studies the detection of most common intestinal parasitic worms in the human body which are *Ascaris Lumbricoides ova (ALO)* and *Trichuris trichiura ova (TTO)*.
2. Comparing the performances by applying a few techniques in preprocessing stage to identify and classify the eggs of two worms (ALO and TTO) through the result and adjusting the parameter values of these techniques that are used in this study to provide good results.

3. The importance of the sequence of the steps in pre-processing and processing stages, is one of the scopes that we try to approve it for better results.
4. Extract specific features of the ALO and TTO parasites to achieve better parasite recognition.

1.6 STRUCTURE OF THE THESIS

This thesis is divided into five chapters:

In Chapter I, the general introduction is presented which includes the background of human parasitic worms and types of microscopes and digital image processing in medical application. Then, the motivation, objectives and scope of this thesis are presented. The next chapter is the literature review. Here the theory related to the issues in all areas such as blood, urine and fecal with the solution are lighted. Each topic is described, and the approaches of different studies are addressed. This is followed by the methodology chapter. The proposed solutions are discussed and presented. Then, every step associated with the methodology is carefully explained, including the description of the approach and implementation details. In Chapter IV, a sequence of experiments to test the performance of the solutions is given. The goal and characteristics of the experiments are described. Then, the results of these experiments are shown and discussed. The last chapter is on conclusions and future work. The conclusions gained from the discussion of the results are presented. Recommendations for future work are also given.

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

Object detection is one of the most challenging problems in computer vision. It is difficult due to the significant amount of variation between images belonging to the same object category. Other factors, such as changes in viewpoint and scale, illumination, partial occlusions and multiple instances further complicate the problem of object detection.

Several laboratory methods, imaging techniques and endoscopy in addition to clinical picture and geographic location, are the ways to diagnose parasitic diseases. Parasitic diseases may be presented with a wide assortment of clinical semblances according to the tissue infested. Microscopy is a direct way of detecting the parasite by examination of different specimens (stool, urine, blood, CSF and tissue biopsies). Recently, nanotechnology can be used as diagnostic procedures utilizing nanodevices. Control and prevention of parasitic diseases depend on the interactions among many factors such as the environment, human behavior, and socio-cultural factors that determine transmission and stability of parasites.

2.2 GENERAL PARASITE DETECTION USING IMAGE PROCESSING TECHNIQUE

In previous studies, microscopic imaging analysis has been used for recognition and detection of parasites (Daugochies, Imarom, & Bollwahn, 1999). The aim of the study is to classify several parasite types occurring in human beings by using morphological features extracted in interactive image tools (Joachim, Dulmer, & Daugochies, 1999).

These morphological features were subsequently included in the computation of classification indices. The parasite types were then recognised by using the calculated indices (Sommer, 1996). In this chapter, we will focus on the techniques that were used to detect the parasites in human beings whether these parasites are in blood, urine, CSF, or feces, in order to explore few techniques in image processing that might help us improve the work and the results of the best detection approach. The most important thing is to study the techniques in image processing for same purpose. Generally, just a few techniques used image processing for medical applications to detect parasites automatically and these techniques are based on parasite characteristics.

2.3 PARASITE DETECTION BASED ON FEATURES

In order to present some techniques, we have grouped them into three main groups: a) colour and shape based detection, b) colour based detection and c) shape based detection. In this work, we are using one of these groups as a baseline to detect the two parasitic worms, these groups are described below.

2.3.1 Colour and shape based detection system

In some researches, the detection of parasites was based on the colour and the shape of these objects as features to recognise them in the images. We present some related works in detecting parasites using techniques which rely on colours and geometric features in the classification in the blood, urine, or fecal. We are focusing on the techniques and the features generally. Some of these works are:

Suwalka and Sanadhya have proposed an innovative digital technique detecting parasitic protozoa of the genus Plasmodium. It contained collection of a blood smear, its staining with Romanowsky stains and examination of the Red Blood Cells (RBCs) for intracellular malarial parasites (Suwalka and Sanadhya, 2012). In blood samples, if the red corpuscles of vertebrates are infected by malarial parasites, they will have a specific shape which we can identify their presence. Recent research has mentioned that the shape of the affected red blood cells can be detected by using the 2D moments of the image of the

infected cell. In their work, an algorithm was performed to identify the type of parasite by their colour and shape. The idea to identify the shape was based on equations of perimeter and area of different geometric structures in the image, see Figure 2.1.

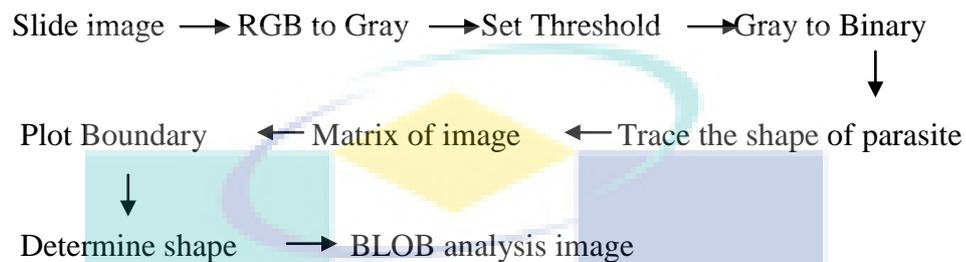


Figure 2.1: Overall algorithm chart

Kareem, Kale and Morling explained a modified Annular Ring Ratio (ARR) method which successfully located and differentiated gametocytes of *P.falciparum* species in the thin blood smear images. ARR transform method (Kareem, Morling, Kale, 2012) aimed at locating the centre of each cell existing in the image. It also sources the location information of the cells. Firstly, the method performed morphological dilation using a concentric ring structuring element and then erosion using a disk shape structuring element to remove the artifacts as well as other noises existing in the image. This can help removing the holes or bright patch in the centre of RBCs. It then used the ratio of transform (outer to inner average intensity ratio) similar to the top hat transformation (Jackway, 2000). The method identified the existence of any white blood cells (WBCs) in the image, and neglected other artifacts and non-infected cells. It used the information based on the structure, colour, and the geometry of the cells and did not require any segmentation or non-illumination correction techniques that are commonly used for cell detection.

Calva and García have proposed an interesting way that uses classification of the microscopic urine components through the study of their complex internal structure (Calva, Landa, & Lehman, 2004), which is a continuation of previous work. This would

be equal to analyzing the texture of the sediment contained in each sample by using the fractal geometrical methods (Mandelbrot, 1982) (Falconer, 1997).

Combined with neural networks, their study used different methods for image processing, separating the colour channels in the corresponding base RGB. With a simple neural network (perceptron multi-layer and back propagation), they obtained very interesting results for the classification of some sediments and parasites (Calva, García1, Martínez, Salgado, & Lehman, 2009).

2.3.2 Colour based detection system

Basically, in this group the detection is using the colour as a feature to classify the parasites and we are presenting some works that used the techniques which rely on colour feature based method, such as:

Makkapati has proposed a method in the detection of malaria parasites in stained blood smears which is critical for treatment of the disease using image processing (Makkapati, 2009). A scheme based on HSV colour space that segments Red Blood Cells and parasites by detecting dominant hue range and by calculating optimal saturation thresholds was presented in his work.

Ray (2010) has proposed a computationally efficient method for segmentation and characterization of malaria parasites from peripheral blood smear images using image processing. Normalized Cut (Ncut) Algorithm is a method in which they cut a graph into two components so as to express the cost of the cut as a small fraction of the total affinity within a group (Forsyth & Ponce, 1997). The algorithm was applied in various colour spaces to find its optimal performance for microscopic blood smear images. They tested the efficacy of results in RGB, YCbCr, HSV and NTSC using the Rand's Index. The results indicated that the performance of the Ncut algorithm was best in HSV color space, see Figure 2.2.

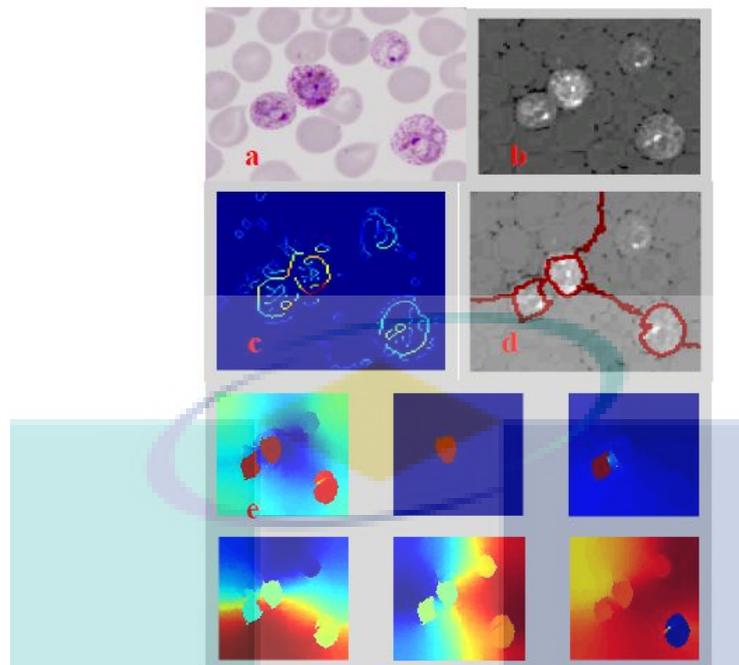


Figure 2.2: The images show segmentation in HSV colour space using nb segment = 6 (a) Image grabbed from microscope, (b) Image brightness (c) Displaying edges (d) Displaying the segmentation (e) Displaying the eigenvectors.

2.3.3 Shape based detection system

In this study, the methodology of our work is using the geometrical features of the parasites in detection and classification. Most related works, especially for fecal parasite detection, are using the same ideological way, in other words, they used shape features of the parasites as a baseline to detect the parasitic worms. Generally, this group of detection is widely used, and here some of these approaches that are related to this group and showing the techniques of which kind of features is used, such as:

Halim, Bretschneider and Li (2006) proposed a technique in their study to estimate parasitaemia (the percentage of infected erythrocytes which is used to measure progress of experimental Plasmodium infection in infected hosts) from blood smear images by extracting healthy and parasite infected red blood cells. The developed approach accounts for uncertain imaging conditions due to microscope settings (Halim, Bretschneider & Li, 2006). The technique was used based on a multi-stage estimation process with minimal

prior knowledge starting from a model representation of red blood cells (RBCs). This technique is based on outside characteristics of the RBCs from shape and size. Based on a pattern matching with parameter optimization with cross-validation against the expected biological characteristics, red blood cells are found. In the final stage, the parasitaemia measure was performed by partitioning the uninfected and infected cells using an overlooked and in comparison a training-based technique.

Purwar and Shah (2011) introduced a method based on digital image processing of Giemsa-stained thin smear image to facilitate the diagnostic process. The most important step in image segmentation was to extract meaningful regions, or in other words, recognise objects from the background. The way described in this case is to use edge detection algorithms (Gonzalez, 2001) but it uses gradient information followed by morphological boundary closing. A method was required to overcome its problems. Chan, and Vese (2001) and Jampana (2010) have proposed energy minimization of the image to detect edges of objects within an image. The aim of using Chan-Vese based boundary detection algorithm to segment an image into meaningful regions, in this case to separate RBC and artifacts from the background, is shown in Figure 2.3.

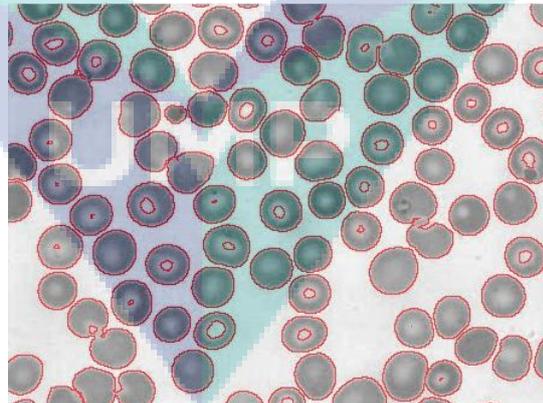


Figure 2.3: Boundary extracted image using the Chan-Vese segmentation method

Mahmood and Mansor (2012) have proposed a computer vision system that could detect and estimate the number of red blood cells in the blood sample image. Morphological part is a very powerful tool in image processing, and it has been used to segment and extract the red blood cells from the background and other cells, see Figure

2.4. The algorithm used cell features such as shape of red blood cells for counting process by using the Hough transform technique. The result presented was based on images with normal blood cells.

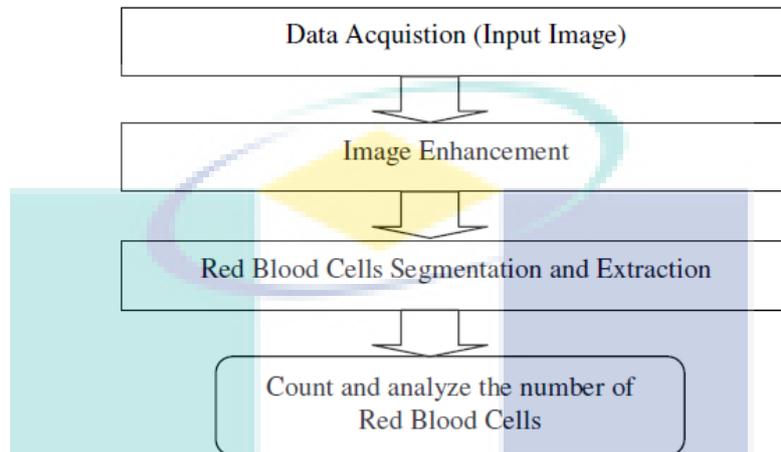


Figure 2.4: Flow chart of the red blood cells counting process

Ghate and Jadhav (2012) have proposed a method to explore the possibility of a computerized diagnosis of malaria and to develop a novel image processing algorithm to dependably detect the existence of malaria parasite in thin smears of peripheral blood sample. They achieved this goal by using Image Segmentation and Thinning techniques to detect malaria parasites in images obtained from Giemsa stained peripheral blood samples.

Soni and Mishra (2011) have proposed a technique that took benefits of morphological operation and thresholding at the appropriate position in the whole process to maximize the productivity of algorithm and distinguish between the simple RBC and malaria parasite. With the tested test algorithms, 'SUSAN edge detection technique' gave good localization of edges but formed a thick border making cell separation difficult. If the staining of RBC was not properly done than the edge of parasite affected RBC, it could be easily detected by using SUSAN algorithm, which was an important advantage of that particular algorithm.

Cao and Cai Zhong have proposed a method to detect red blood cells in urine. Their study focused on the detection of red blood cells in urine image captured from the microscope by image processing. After the urine image was pre-processed by improved Sobel operator, red blood cells are localized by using Hough Transform and their features, which are extracted and selected by Principal Component Analysis (PCA), which can be classified with LDA (Linear Discriminant Analysis) (Cao & Zhong, 2009). Classical Hough transform of circle detection has disadvantages of large parameter space with low efficiency of detection. In order to overcome these flaws, their method of using the geometrical features to detect the circle centre in the image is shown in Figure 2.5. That is to say, it could determine a circle and the centre by any three points that are not in a straight line. Overall, this study provided a reliable and convenient detective method for medical research.

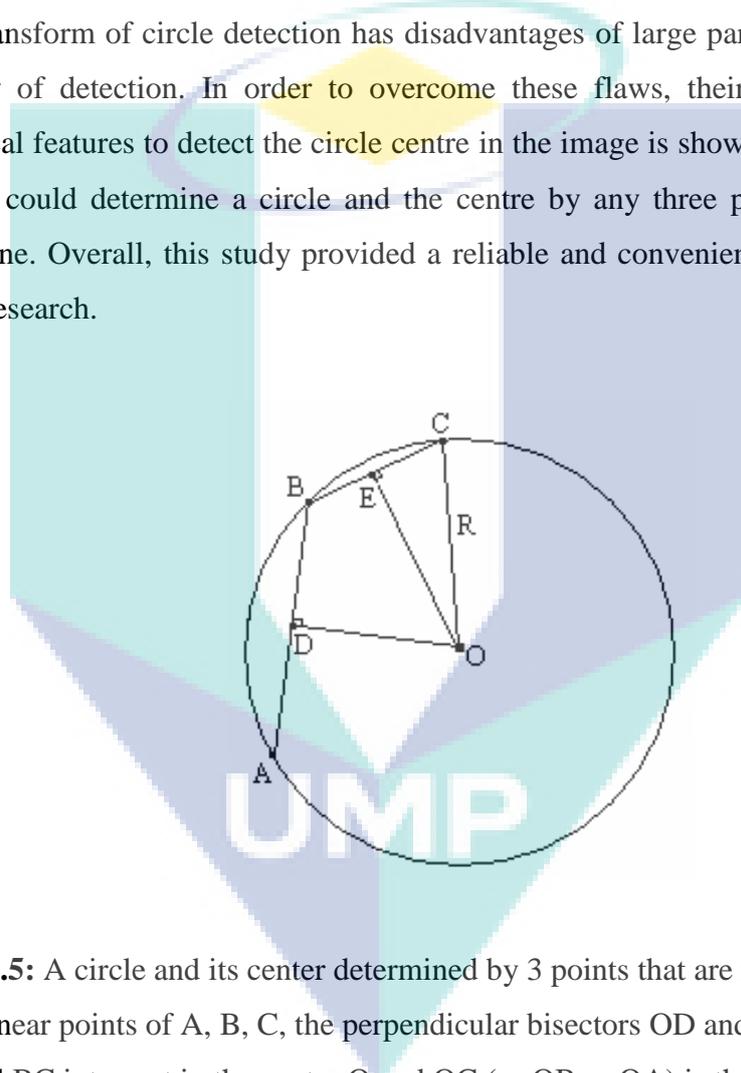


Figure 2.5: A circle and its center determined by 3 points that are not in a straight line. 3 un-collinear points of A, B, C, the perpendicular bisectors OD and OE of line segments AB and BC intersect in the centre O and OC (or OB or OA) is the length of the radius

2.4 OVERVIEW OF HUMAN PARASITIC WORMS CLASSIFICATION

Some approaches for human parasitic worm detection are presented. In this study, to detect the parasites in fecal we are using third group based detection as we mentioned previously which is the shape based detection method. The parasitic eggs are identified

based on their geometrical features such as size and overall shape category and most of researchers use such detection method. The reason for not using 'colour based detection' method in fecal parasite detection is the colour of specimens will change the type of the stain which is used for the illumination of the microscope with the stain's effects. In previous studies, techniques based on shape are performed for the same purpose in detecting parasites in human fecal and these techniques are useful to find the best way in our study. Some of these studies with these interesting techniques are mentioned below.

Yang et al. has proposed a method to detect common helminth eggs automatically in microscopic fecal specimen images. The seven typical species of human helminth were chosen as follows (Ash and T. C, 1990): *ascaris lumbricoides*, *trichuris trichiura*, *capillaria philippinensis*, *clonorchis sinensis*, *paragonimus westermani*, *diphyllobothrium latum*, and *taenia*.). For that purpose, the digital image processing techniques and artificial neural network (ANN) classifier were used. ANN classification was performed by two sub-classifiers (Yang, Park, Kim, Choi, & Chai, 2001). Most ANNs classify given patterns according to their geometric groups and their characteristics in the pattern space. . In his study, four features were selected based on three morphological characteristics: representing shape, shell smoothness, and size. The first stage (ANN-1) of the proposed ANN classification system separated the eggs from confusing artifacts, while the second stage (ANN-2) classified eggs by species. The performance of the ANN was evaluated using a cross - validation method to avoid the dependency on the selection of training samples.

The targets in his study for the developed algorithm were seven typical human helminth eggs. For the identification of more parasitic species than a detailed analysis of materials inside the eggs (Gonzalez, 1993), more features were used in this study and even the best features, in terms of maximum variances and minimum redundancy, would be able to deduce from them using dimensionality reduction methods like Karhunen–Loève transformation (Freeman and Skapura, 1992).

Sengur and Turkoglu have proposed a method to recognise parasite eggs using pattern recognition techniques. In their study, the feature extraction mechanism based on

the invariant moments and artificial neural network classifier was used. Pattern recognition is divided into a series of stages, starting with feature extraction from the occurring patterns, which is the conversion of patterns of features that are noticed as a condensed representation, ideally having all the necessary information (Sengur & Turkoglu, 2004). At the next stage, the feature selection step, a smaller number of significant features that best represented the given pattern without redundancy were found. Finally, classification was performed: a specific pattern was specified for a specific class according to its characteristic features, selected for it (Avci & Akpolat, 2006). The techniques utilized to pattern recognition use artificial intelligence approaches (Bishop, 1996).

Dogantekin & Yilmaz have proposed a robust technique based on invariant moments – adaptive network based fuzzy inference system (ANFIS) for recognition of human parasite eggs in microscopic images. The authors presented a digital image processing method used for feature extraction stage of IM-ANFIS (Bishop, 1996). Recently, the pattern recognition principles have come into significance. The technique presented in their study enabled us to classify 16 different parasite eggs from the microscopic images. This proposed recognition method has four stages.

In the first stage, a pre-processing sub-system was realized for gaining unique features from the same group of patterns using some digital image processing operations, which are noise reduction, contrast enhancement, segmentation by using thresholds, morphological process, and logical process. In the second stage, a feature extraction mechanism using invariant moments was applied. Seven numbers Hu invariant moments of all of these pre-processed microscopic images of human parasite eggs were calculated. In the third stage, an adaptive network based on fuzzy inference system (ANFIS) classifier was used for the recognition process.

Finally, the testing stage is whereby the performance of the proposed recognition system is tested with the developed computer simulations on MATLAB. In their study, the Hu invariant moments were used for feature extraction. There are some advantages of this invariant moment method such as invariant under translation, rotation, and scaling

operations. Therefore, the proposed human parasite eggs, namely, Fertilized *Ascaris Lumbricoides*, Unfertilized *Ascaris Lumbricoides*, *Diphyllobotrium Latum*, *Enterobius Vermicularis*, *Fasciola Hepatica*, *Giardia Lamblia*, *Hymenolepsis Diminuta*, *Hymenolepsis Nana*, Hookworm, *Kellicoti*, *Paragonimus Westermani*, *Schistosoma Haematobium*, *Schistosoma Japonicum*, *Schistosoma Mansoni*, *Taenia Saginata*, and *Trichuris Trichura*, could be recognised from their images even they are translated, rotated, or scaled.

Avci and Asaf (2009) have proposed an expert diagnosis system to classify human parasite eggs based on multi-class SVM, which was based on invariant moments and multi-class support vector machine (MCSVM) for classification of human parasite eggs in microscopic images. This method consists of four stages, that is, pre-processing, feature extraction, classification, and testing. For this purpose, the proposed approach was tested by using test data to detect 16 kinds of parasites.

In the pre-processing stage, their study contained four digital image processing operations. These include a median filtering for noise reduction, contrast enhancement, thresholding, and morphological and logic processes. In image processing, moments were applied as feature vectors for classification as well as for image texture properties and shape descriptors of objects in the images.

In the feature extraction stage, they used the method of Hu's invariant moments (Hu, 1962). They derived a set of seven functions that made use of the central moments of pre-processed image for characterizing the shape of each parasite egg type. The output of these invariant moments was independent of any translation, rotation or mirror image of a particular blob. Furthermore, these invariant moments could be used in coupling with both the blob image itself and the edge processed contour image.

In the classification stage, the multi-class support vector machine (MCSVM) classifier was performed for classification of features extracted. In the testing stage which is the last stage in their study, the rest of the database was applied for testing the proposed

expert diagnosis system which is based on MCSVM for classification of human parasite eggs in microscopic images.

Badawi and Sonni have proposed a software to recognise the parasite in microscope images whether the object is Taenia egg or not and calculate the error percentage. Basically, the programme used a number of Matlab techniques in image processing to enable us to detect the parasite (Badawi and Sonni, 2008). In their study, the feature measured was the radius by using a number of equations and statistical analysis to estimate the radius of Taenia. They did not rely on the colour variation of the egg, which is not a problem but the problem lies with the noise. They have mentioned that using a number of algorithms and techniques could not remove the noise completely, but it could reduce its level because some noises were deeply affected in the image and this is the only reason for an error to appear in false analysis.

2.5 SUMMARY

In this chapter, we can conclude that the parasites in the human body found in blood, urine, or fecal are detected based on three approaches which rely on their features, which are: shape and colour based system, colour based system, and shape based system. The usual approach used to detect and classify the parasites in fecal is the shape based system which relies on geometric features of the parasites rather than the other two strategies.

This research aims at improving and finding the solutions of some errors that might find in such work. This study presents a method to identify two types of parasites in fecal which can directly cause diseases in human beings.

CHAPTER III

METHODOLOGY

3.1 INTRODUCTION

In this study, fecal parasite detection technique based on Filtration and Determination Threshold System (*F-DTS*) was proposed. This can remove the unwanted and confusing objects through using a threshold of features' values. The recognition method includes three stages. The first stage is a pre-processing sub-system the purpose of which is to obtain unique features after performing contrast enhancement, noise reduction, edge enhancement and detection, segmentation and other morphological process. These are applied to the feature extraction stage. The next stage is a feature extraction mechanism which is based on five features of the three characteristics (shape, shell smoothness, and size). The final stage is the Filtration and Determination Threshold System (*F-DTS*) classifier which is used to recognise the process using the ranges of feature values as a database to identify and classify the two types of parasite. This technique enables us to classify two different parasite eggs from the microscopic images, namely, roundworms (*Ascaris lumbricoides* ova, ALO) and whipworms (*Trichuris trichiura* ova, TTO). The block diagram of this approach is shown in Figure 3.1. This approach starts with data acquisition, then pre-processing stage, processing stage, feature extraction and finally classification stage.

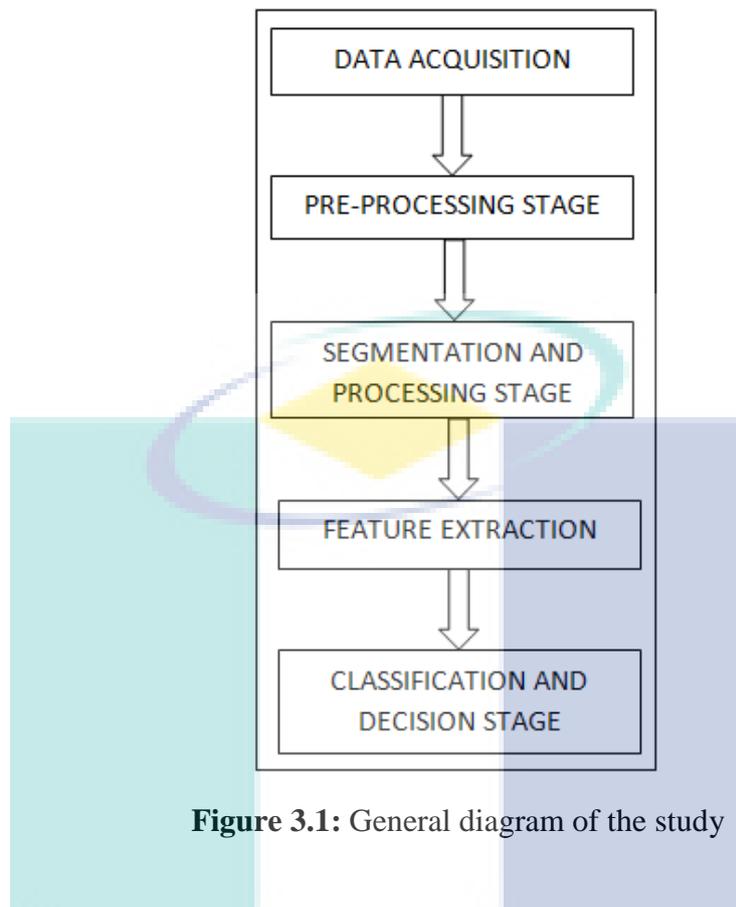


Figure 3.1: General diagram of the study

3.2 DATA ACQUISITION

In this research, parasite images were collected as an input data from Hospital Universiti Sains Malaysia (HUSM), Kelantan, Malaysia. For both types of parasites, *Ascaris Lumbricoides ova* (ALO) and *Trichuris trichiura ova* (TTO), 200 images were collected and 100 for each parasite, by using a digital microscope which is attached to a digital camera, connected to a PC and runs with special software. The tools and their features are listed as follows:

Camera: Olympus XC50, U-CMAD3, Japan.

Microscope: Model- BX41TF-FL_CCD, Olympus, serial number: BG22578.

Software: AnalySIS docu, copyright 1986-2007, Olympus Soft Imaging Solutions GmbH.

In a conventional method, to prepare a slide of fecal for examination, there are a few steps that we need to follow, which are: (Dr. Zeehaida, field supervisor)

1. Place a drop of saline at the centre of a slide.

2. With an applicator stick, select 1-2 mg of feces carefully avoiding the non-fecal elemental.
3. Mix the feces with saline until the suspension is even.
4. Remove coarse particles from the suspension and cover with 22 x 22 mm cover slip.
5. Examine the entire preparation systematically under microscope using 10 x and 40 x objectives.
6. If suspected objects are seen, prepare another smear using Lugol's iodine or add in Lugol's iodine to the saline smear from the side of the cover slip.
7. Repeat examination under light microscopy.

Since we are trying to let a system take all the responsibility in the examination and detect the parasites, we can short the steps to:

1. Do the same steps 1-4 from the above, and
2. Add Lugol's iodine to the saline smear from the side of the cover slip, then it is ready to get examined.

In a microscope, the magnification powers might have many objectives to enlarge objects like 10x - 20x. In this study, the input data or image acquisitions are under the magnification power 10x and the values of feature extraction belong to the magnification power 10x only. The main reasons to choose magnification power 10x rather than 20x and 40x are:

1. In 10x, the general characteristics of parasites are clear and enough to determine the features of these parasites.
2. The field of view under 10x is wider and can recognise the eggs among the other surrounding objects.
3. In the lab, both 20x and 40x are mostly used to study for an internal egg descriptions, while the features of this study focus on the descriptions of an external egg.
4. To decrease the number of images which are taken from the whole slide to be examined, since 10x has less magnification power and wide field of view.

Figure 3.2 clearly shows the difference between images which are taken from the same microscope for both parasites ALO and TTO respectively under three different magnification powers of 10x, 20x and 40x.

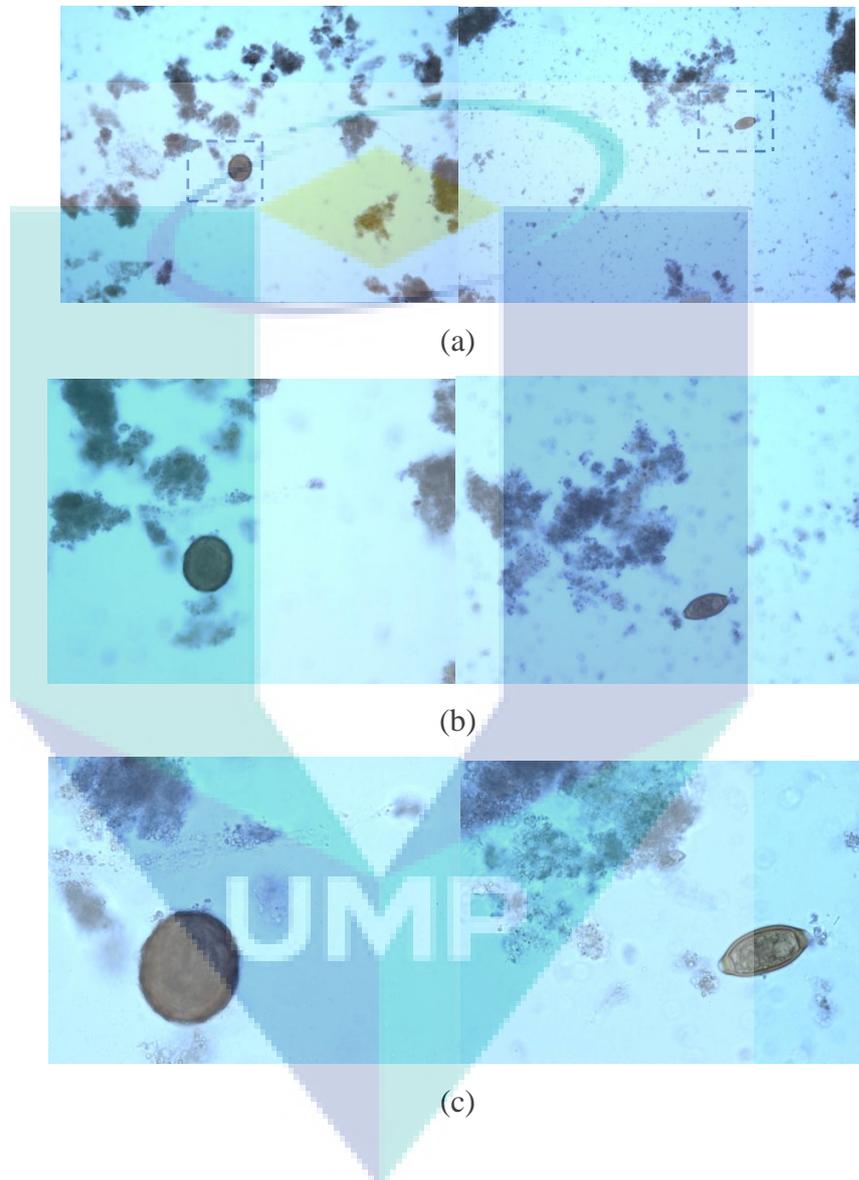


Figure 3.2: Magnification powers of ALO and TTO parasitic eggs under microscope from left to right, respectively: (a) Under magnification 10x; (b) Under magnification 20x; (c) Under magnification 40x.

3.3 PRE-PROCESSING STAGE

In the preprocessing stage, the theory of techniques used in this stage and the advantages of using them will be discussed. Firstly, to avoid the errors faced by the user of this system, we provide an additional step and name it “pre-using step”, the function of which is to resize all input images to uniform size. In data acquisition, purposely we capture images with different sizes to see if the variety of these sizes will bring us any problem, especially when this study depends on the shape and size characteristics of the objects in the images. The sizes of image acquisitions are two sizes are 966x3864 and 1932x7728, and these two sizes are not readable in the system because the images are too big. That was the reason of applying this step to resize the image to a reasonable size to be read directly and without any problem. So, the size that has been applied in this programme is 483x644 which is considered half or a quarter of the original size.

The sequences of applying the techniques in the preprocessing stage are very important and later on in this chapter we will show a few methods that are applied during the experiments, and these methods are different in sequences in using some other techniques.

Based on ideas, we have performed a method in the preprocessing stage, and we have tried two other experimental methods to find the fit one that can work perfectly in our system in order to detect ALO and TTO parasites in the images. In related works, most studies used the same procedure in the preprocessing stage such as noise reduction, contrast enhancement, and edge detection techniques. In this study, we keep using the same ideas but the sequences of applying these steps are surprising in the following three methods, as shown in Figure 3.3:

1. Preprocessing method I (Pre-Method I): contains noise reduction using median filter and image enhancement by contrast enhancement and adjustment with gray thresholds.
2. Preprocessing method II (Pre-Method II): contains noise reduction using median filter, contrast enhancement, and then edge detection using ‘canny’ filter.

3. Preprocessing III (Pre-Method III): contains an image enhancement using contrast enhancement, edge enhancement using an unsafe, then noise reduction using median filter and an edge detection using 'canny' filter.

Among these methods, only one is the most suitable due to keeping the data of the image away from being lost. For example, in pre-method I, the sequences of this method is noise reduction (median filter), contrast enhancement using converting images from gray images to binary images, and threshold images.

Additionally, pre-method II ignores the converted image to prevent the loss of details regarding input data, but it uses edge detection with 'canny' filter. That is one way to detect the edge of hard sharp objects. Since the two kinds of parasite eggs in this study have very clear borders, using edge detection with canny filter is one of the main steps of the preprocessing stage. However, using pre-method II is not a solution for all images that have been tested because some eggs have hard background artifacts, which means their borders would not be very clear and sharp.

While in pre-method III, most problems in pre-methods I & II have been solved for better output. The differences between pre-method II and III are the sequences of the steps inside this stage besides adding an edge enhancement (sharpness) technique. At the beginning, contrast enhancement was performed, so the brightest and the darkest pixel values are estimated. Sharpness method is used to sharpen the outlines of all objects. Then two times of median filter are used to reduce artifacts and remove as much noise as it can.

Based on the experimental results, pre-method III was the best method. Pre-method III removed much noise since the median filter was used twice. Then using edge detection can clarify the edge of all objects in the image within specific thresholds of low-level and high-level pixel intensity and specific value of Gaussian's sigma.

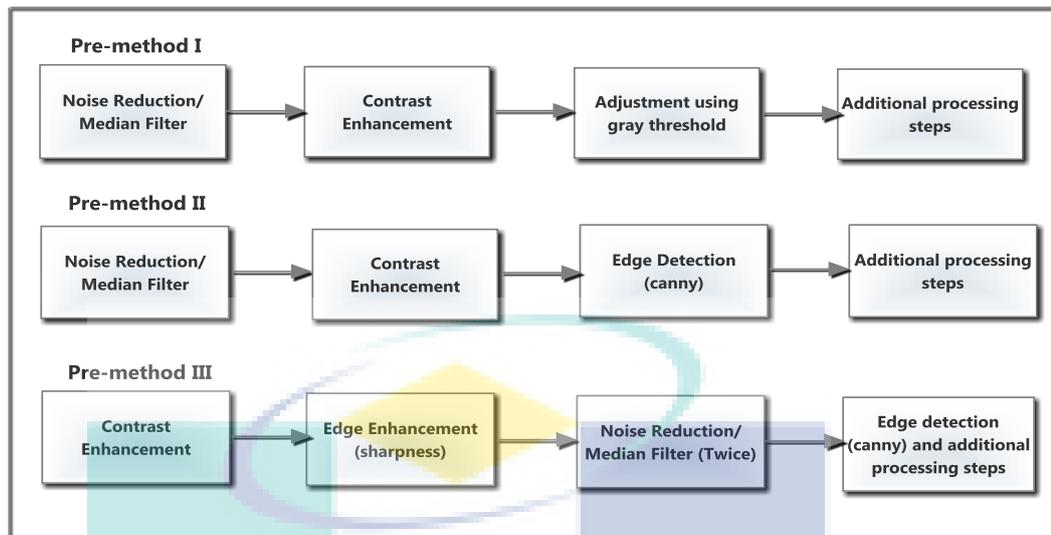


Figure 3.3: Three different pre-processing methods

Finally, we could find the most suitable method that we can rely on it to get a better performance in the preprocessing stage. Thus, details of pre-method III are explained in this chapter step by step and sequentially as a part of the whole approach in this study, as shown in the block diagram in Figure 3.4.

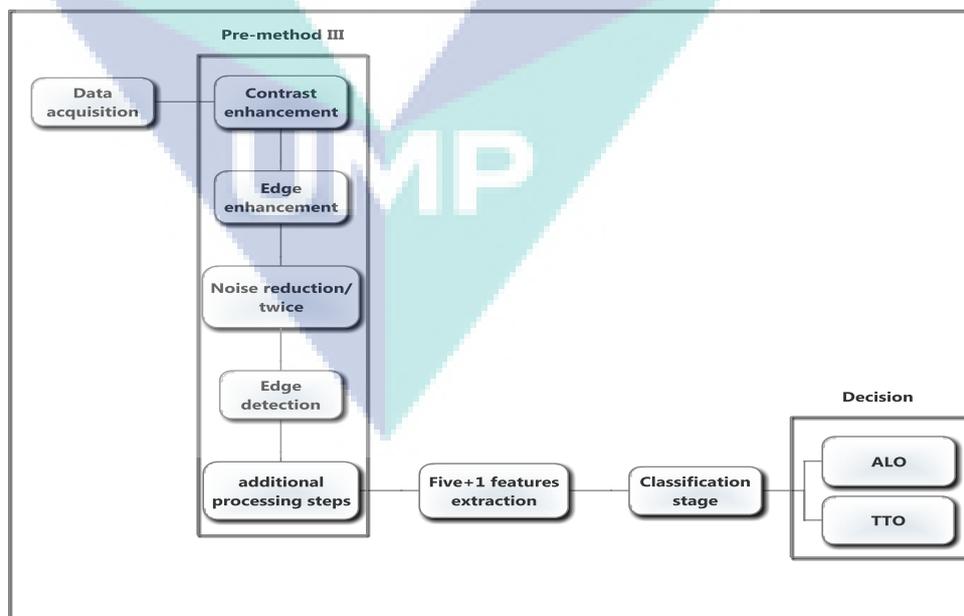


Figure 3.4: Using Pre-method III in the Pre-processing Stage to detect ALO and TTO Parasites

3.3.1 Contrast enhancement

The aim of using the contrast enhancement is to make the image brighter than the original because in fecal samples there are many artifacts that have dark backgrounds and can increase the error of detecting the eggs of parasites, and a lot of image information will be lost. Image contrast enhancement is the process of adjusting digital images, so that the results would be more suitable for display or for analytical purpose. For example, it can remove the noise or brighten an image, making it easier to identify key features. An image will lose its contrast when there are no sharp differences in the image between black and white. Brightness are based on the overall lightness or darkness of an image. To change the contrast or brightness of an image, the Adjust Contrast technique performs *contrast stretching*. In this process, pixel values below a specified value are displayed as black, pixel values above a specified value are displayed as white, and pixel values in between these two values are displayed as shades of gray. The result is a linear mapping of a subset of pixel values to the entire range of grays, from black to white, producing an image of higher contrast (see figure 3.5).

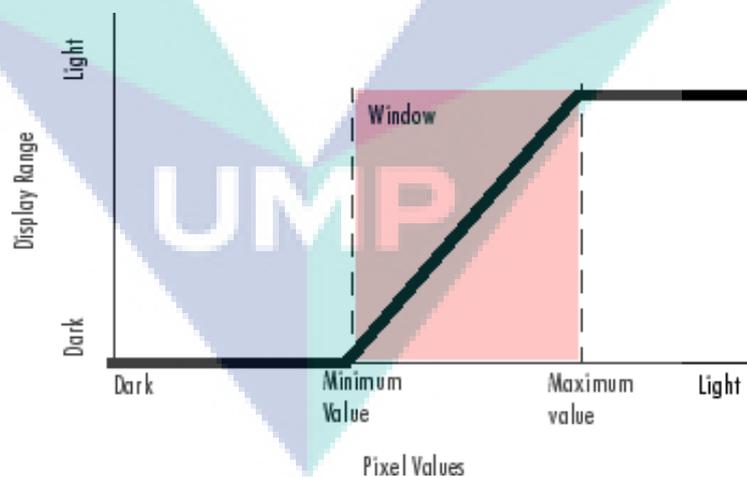


Figure 3.5: Contrast enhancement curve

There are three techniques of contrast enhancement, namely (Gonzalez, 2002):

1. Adjust image within intensity values.
2. Histogram equalization.
3. Adaptive histogram equalization.

After applying and testing these techniques, the best performance is the first technique. We will discuss more about the reason of choosing this technique rather than the other two techniques in the next chapter.

3.3.2 Edge enhancement

After we finished enhancing the contrast of images, we have to adjust and fix some characteristics of the objects in the images by edge enhancement. The major aim of edge enhancement is to enhance the appearance of an image to make it visually more attractive or to modify the visibility of certain features (Lewis, 1990). The edge enhancement technique to improve all high spatial frequency details in an image includes edges, lines, and points of high gradients (Richards, 1986). Thus, to detect the eggs of the parasites we need to enhance the edges of the objects in the images after they got enhanced in the contrast of overall image contents.

One of the techniques usually used for edge enhancement is unsharp masking, and in the next chapter we will give the reason of choosing this technique. There are different ways to perform unsharp mask. One of them is to subtract the blurred or low-pass filtered image from the original image which contains the most high-quality components. Most image editing tools, such as Photoshop, is an adaptive way to use the three parameters and the radius threshold. More information about the use of these parameters and the clarity of the image can be found (Pratt, Wiley & Sons, 2001). Another approach to improve the region's image is by applying proper design of the high-pass filter or mask to the image. Edges have important information about the image. Therefore, if unsharp mask is applied and then binaries the output, the result of binary image containing most edge information using a 2-D filter, will be used to enhance the edges and apply directly to enhance the edges inside the image. In this part, the image smoothed is subtracted from the original image and the result of this approach goes towards a sharper content of the image. The procedure of unsharp masking is shown below:

1. Apply blur filter to the original image.
2. Subtract the output result from step (1) from the original image.
3. Multiply the output result in step (2) by some weighting fraction.

4. The last step is to add the output result in step (3) to the original image.

The above steps are shown in Figure 3.6.

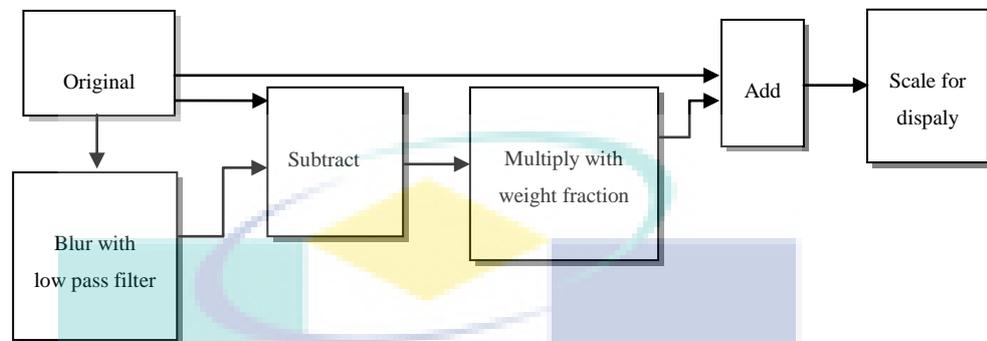


Figure 3.6: Edge enhancement block diagram

Mathematically, the unsharp masking operation is shown below:

$$f'(m, n) = f(m, n) + \alpha [f(m, n) - \bar{f}(m, n)] \quad (3.1)$$

where $f(m, n)$ is the original image, $\bar{f}(m, n)$ is the blurred version of the original image, α is the weight fraction, and $f'(m, n)$ is the sharpened result.

In command terms, the filtering and subtracting can be performed in one command using the linearity of the filter which is 3×3 filter:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which is the identity filter and the unsharp masking can be performed by a filter of the form:

$$f = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{k} \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

where k is constantly chosen for the best result.

So subtracting a blur from a scaled version of the original image is in effect and the scaling factor may be splinter between the identity and blurring filters (Arce 2005), see Figure 3.7.

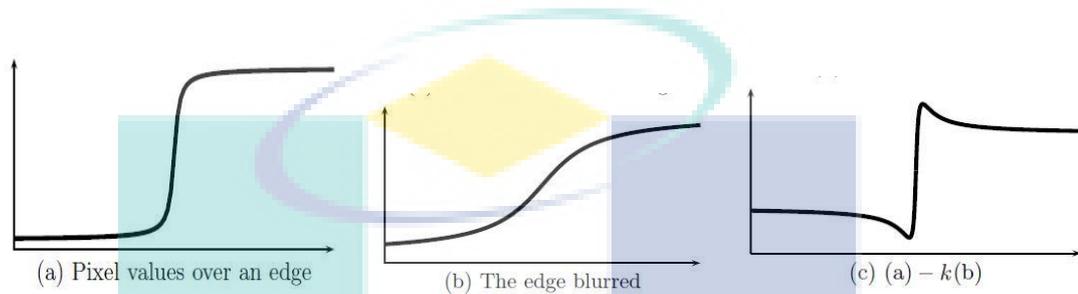


Figure 3.7: Unsharp masking

The unsharp option of 2-D filter performs such filters; the filter created has the form

$$\frac{1}{\alpha + 1} \begin{bmatrix} -\alpha & \alpha - 1 & -\alpha \\ \alpha - 1 & \alpha + 5 & \alpha - 1 \\ -\alpha & \alpha - 1 & -\alpha \end{bmatrix}$$

where α is an optional parameter, its default is 0.2. If $\alpha = 0.5$ the filter is:

$$\frac{1}{3} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 11 & -1 \\ -1 & -1 & -1 \end{bmatrix} = 4 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - 3 \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

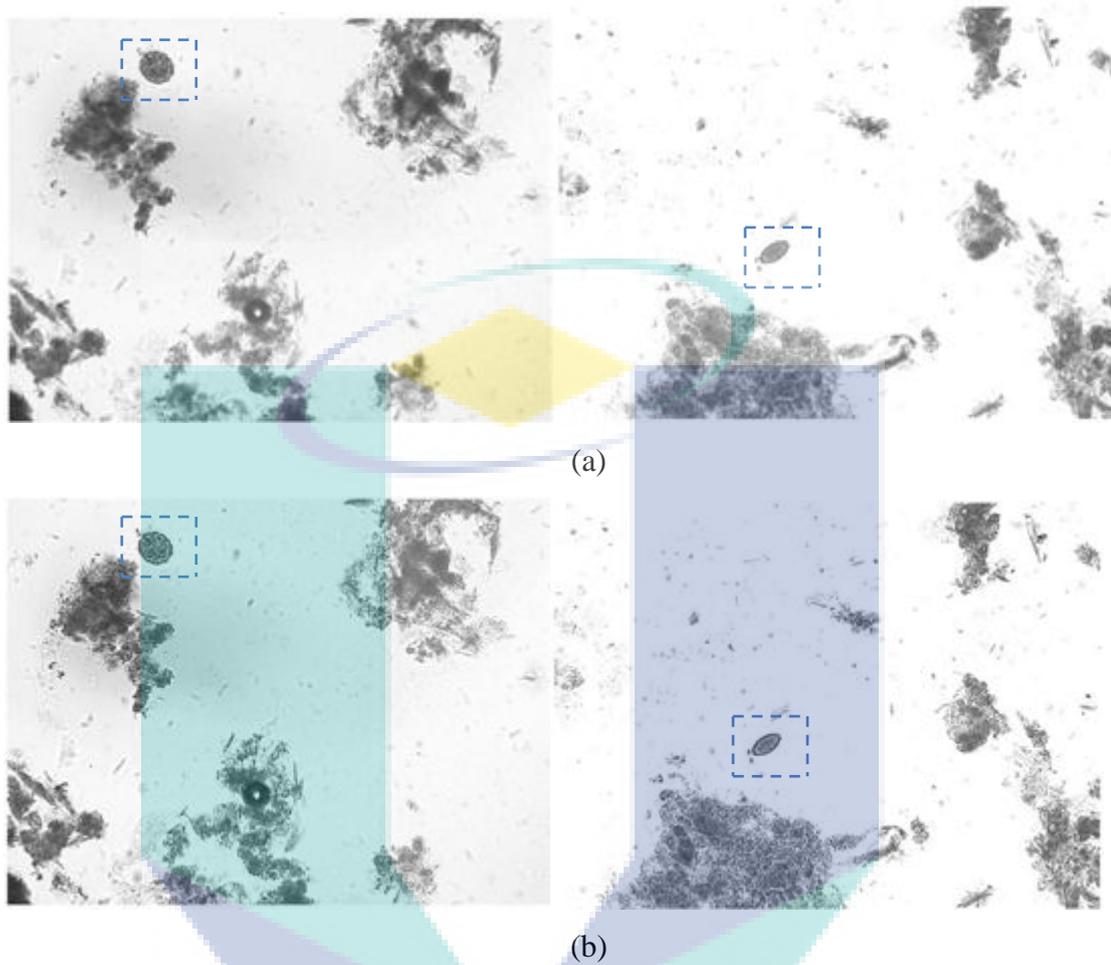


Figure 3.8: Edge enhancement with unsharp masking. (a) Image after contrast enhancement and before unsharp masking; (b) Image after applying unsharp masking.

Figure 3.8 shows the effect of unsharp masking on an input image. We can see clearly the objects in the original image in Figure 3.8.a, which becomes sharper after applying the unsharp masking, as shown in Figure 3.8.b, by using ‘2-D filter’. Thus, the result will be:

$$\begin{array}{r}
 \begin{matrix} -0.1667 & -0.6667 & -0.1667 \\ -0.6667 & 4.3333 & -0.6667 \\ -0.1667 & -0.6667 & -0.1667 \end{matrix} \\
 \text{2-D filter('unsharp')} =
 \end{array}$$

3.3.3 Filtering technique

When an image is snapped by a camera or other imaging system, often the intended vision system has been unable to use it directly. Sometimes the image is corrupted by random variations in intensity, variations in illumination, or weak contrast that should be dealt with in the early stages of vision processing. The image of detecting parasite passed through the edge enhancement makes all objects sharper and clearer outside the border. In the output image, we found some bores at the external edge of the parasites and in order to avoid any lose of this data, we had to enclose these edges and remove some kinds of noises in the images. The strategy to make this happened is to use suitable filter technique.

Related researches had proved that a non-linear filter, such as median, is better than linear which is based on a mean to remove noises, such as salt and pepper, and to give a better edge enhancement. Because of this, we considered using a median filter in this study. It is a non-linear operation which is mostly used in the image processing to reduce "salt and pepper" noise. This is more effective than convolution if the aim is to reduce noise and preserve edges. One kind of smoothing technique is median filtering, as is linear Gaussian filtering. All smoothing techniques are effective on removing noise in smooth patches or the smooth regions of a signal, but negatively affect edges. Often though, at the same time while reducing the noise in a signal, it is very important to preserve the edges. The edges are very important to the visual appearance of images. For small to moderate levels of (Gaussian) noise, the median filter is clearly better than Gaussian blur at removing noise while preserving edges for a given, fixed window size (Arias-Castro and Donoho, 2009).

However, its performance not much better than Gaussian blur in high levels of noise, whereas, in speckle noise and salt and pepper noise (impulsive noise), it is most effective (Arce, 2005). Because of this, median filtering is more widely used in digital image processing. The main idea of the median filter is simple which is to run through the signal entry by entry, and replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which runs, entry by entry, over

the whole entire signal. For 1D signal, the most obvious window is only the first few preceding and following entries, whereas for 2D or higher-dimensional signals such as images, more complex window patterns are possible like "box" or "cross" patterns.

Median filter is considered to be the best known order-statistics filter, which replaces the pixel value by the median of the gray levels in the neighborhood of that pixel. The original pixel value is included in the computation of the median. Median filter is quite popular because, for certain types of random noise, it provides excellent noise reduction capability, with considerably less blurring than linear smoothing filter of similar size (Yang, Lin, Gabbouj, Astola, and Neuvo, 1995) (Ng and Ma, 2006).

$$\hat{f}(x, y) = \text{median}_{(s,t) \in S_{xy}} \{g(s, t)\} \quad (3.2)$$

That S_{xy} represents the set of coordinates in a rectangular sub image window of size $m \times n$ which is centered at the point (x, y) . The arithmetic median filtering process computes the median value of the corrupted image $g(x, y)$ in the area defined by S_{xy} . A standard median operation is performed by sliding a window of odd size (e.g. 3x3 window) over an image. At each window position the sampled values of a signal or image are isolated, and the median value of the samples has replaced the sample in the center of the window, as shown in Figure 3.9.

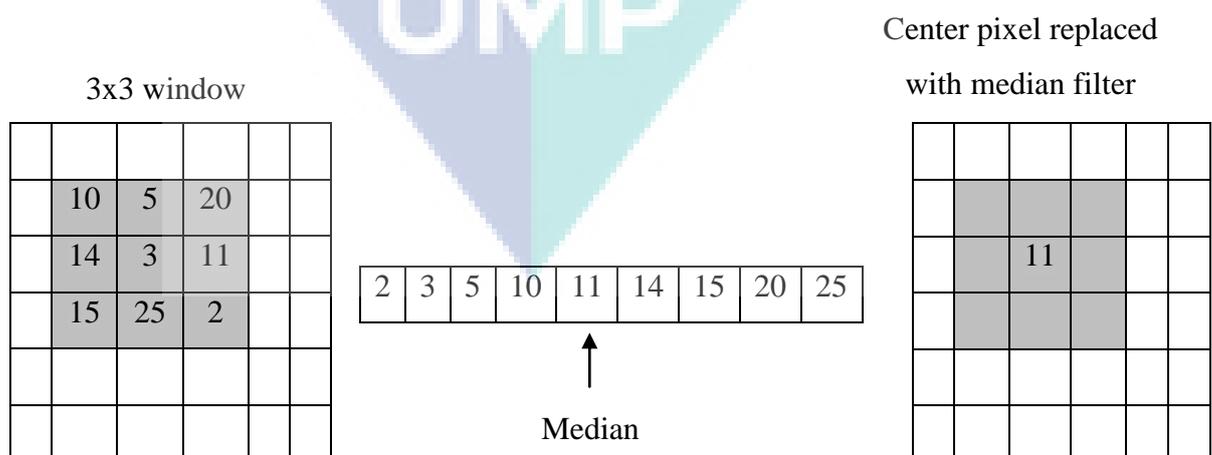


Figure 3.9: Median Filter

For a real example, Figure 3.10 provides a median filter effecting in an input image of ALO and TTO and shows the output of this process. This example is one of the data that is used in this study and the same process applies to all input images. In this figure, obviously we can see how the median filter enclosed the edge of the parasite and fill up some bores inside the objects and remove the noise that surrounds the parasites and some of them are attached with the parasites.

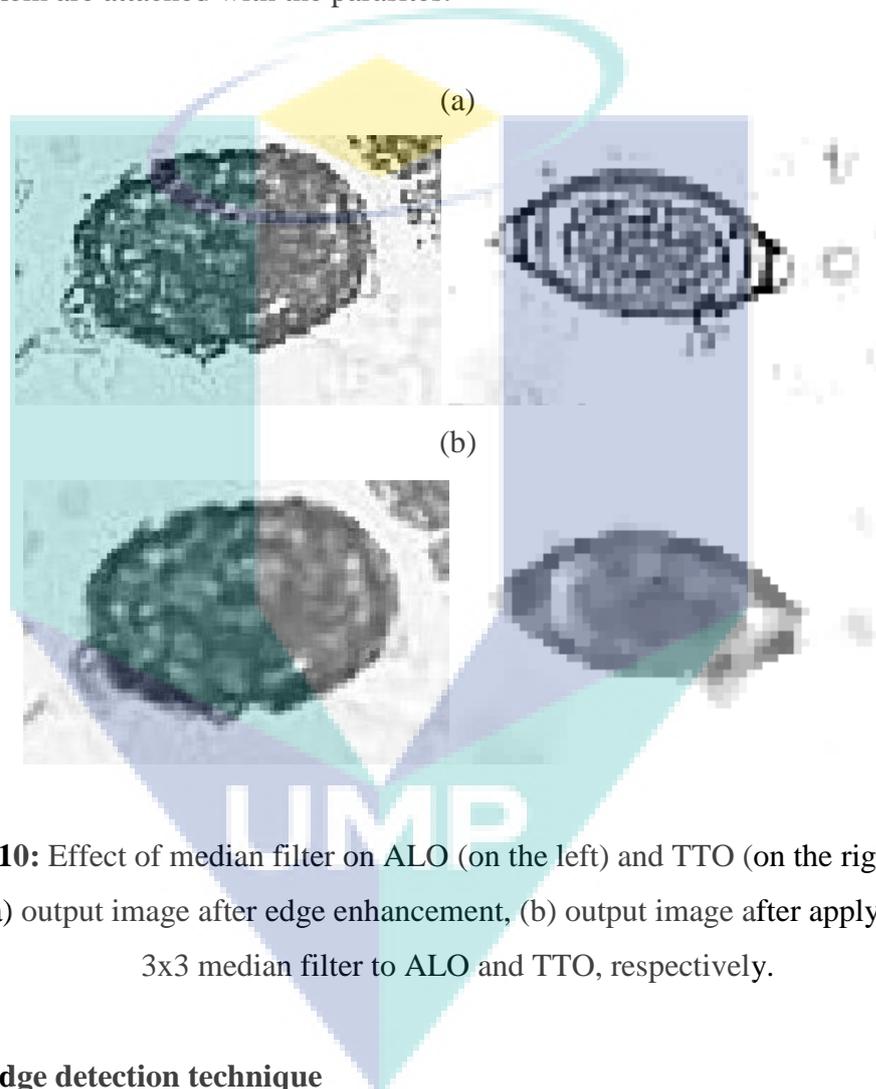


Figure 3.10: Effect of median filter on ALO (on the left) and TTO (on the right) grayscale image, (a) output image after edge enhancement, (b) output image after applying 2x2 and 3x3 median filter to ALO and TTO, respectively.

3.3.4 Edge detection technique

The output of the previous step, which is the filtering technique, is ready to go through the last step or pre-processing stage. This step allows the existence of the parasite because it is going to detect the edge of the objects within specific threshold of intensity that is related to detect our parasite. The edge in an image is a curve which follows a path of fast change in intensity of the image. Edges often relate to the boundaries of objects in the scenery. Edge detection is a way to identify the edges of objects in an image.

The “edge” function is used to find edges (Burger & Burge, 2007). This function seeks place in the image where the intensity rapidly changes, by using one of these criteria:

- For places where the 1st derivative of the intensity is larger than some thresholds in magnitude.
- For places where the 2nd derivative of the intensity has a zero crossing.

Edge can provide a number of derivative estimators (Canny, 1986), each of which implements one of the definitions above. For some of these estimators, the operation can be specified whether it should be sensitive to horizontal edges, vertical edges, or both edges. The ‘edge’ appears a binary image containing 1’s where edges are found and 0’s elsewhere. The most powerful method for edge-detection that ‘edge’ provides is the ‘Canny’ method. The Canny method differs from the other edge-detection methods in that it uses two different thresholds, one to detect strong edge and the other for weak edge, and the output will include the weak edges only if they are connected to strong edges. This method of edge-detection is therefore less likely than the others to be fooled by noises, and more likely to detect true weak edges in the image. The Canny edge detector has given a better edge detection for biological images, which are usually noisy and have less well-defined edges (see Figure 3.11).

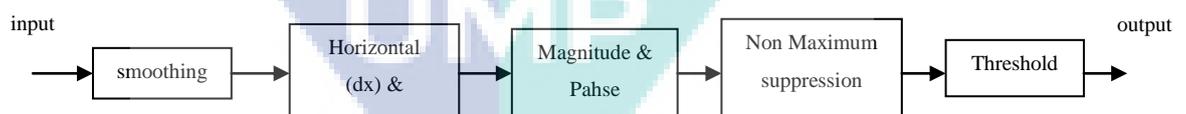


Figure 3.11: Schematic structure of Canny edge detection

- Smoothing

The Gaussian smoothing operator is a 2-D convolution operator that is used to blur images and remove details and noise. In 2-D, an isotropic (i.e. circularly symmetric) Gaussian has the form of:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (3.3)$$

where σ is the standard deviation of the distribution. The effect of Gaussian convolution is to blur an image and the degree of smoothing is determined by the standard deviation of the Gaussian. To generate the masks, the first step is the computation of the mask size. The size of the mask should not be so large compared to the mask lobes, otherwise the result would be in unnecessary computational overhead during convolution. At the same time, the mask size should not be too small to lose the primary lobe characteristics. Choice of the mask size is based on analyzing the Gaussian and applying a threshold T . Figure 3.12 explains the idea.

T represents a real-number between 0 and 1 and width of the mask basically is derived on T as follows: First of all, the size of half mask, s -Half is computed by finding the point on the curve where the Gaussian value drops below T , i.e. then:

$$e^{(-x^2/(2\sigma^2))} = T \quad s - half = \text{round}\sqrt{-2\log(T) \cdot \sigma^2}$$

$$\text{Mask size} = 2(s - half) + 1$$

To combine both positive and negative sides of the mask. The lowest limit of sigma is 0.5, below this the mask size will be less than 3x3, which is not reasonable for finding F_x and F_y . The values of mask size in this method for various sigma values are shown in Table 3.1:

Table 3.1 Size of masks with various Sigma values

Sigma (σ)	Size of Mask
0.5	3x3
1	5x5
2	9x9
3	13x13
4	19x19

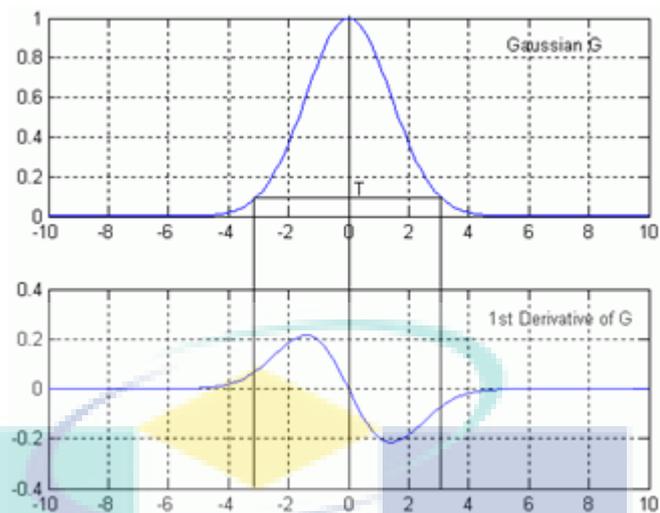


Figure 3.12: Curve of Gaussian G and curve of its 1st derivative of G

The masks are applied to the input images by using convolution. The result then will be scaled down by the same factor which is used to scale the masks up. To write output to the image files, the min and max values will be scaled to 0 and 255 respectively, as shown in Figure 3.13.

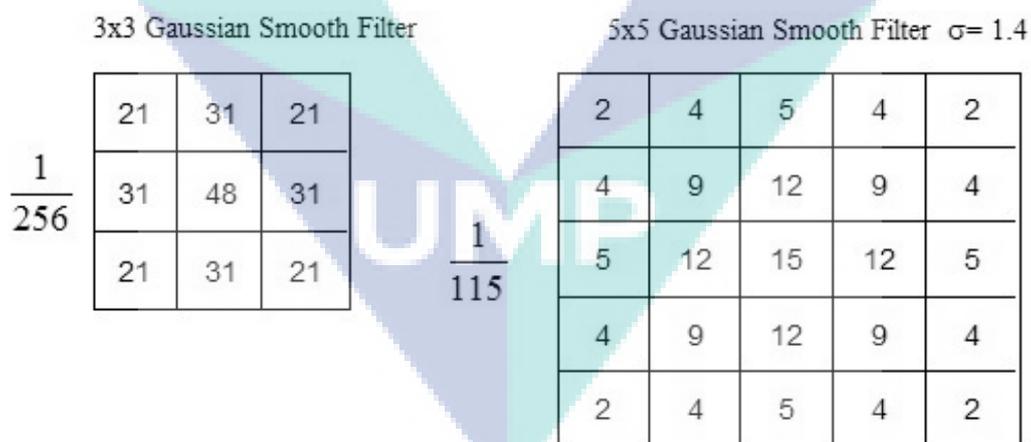


Figure 3.13: Examples of convolution masks

- Gradient Calculation

After smoothing the image by eliminating the noise, the next move is to find the edge strength by applying the gradient of the image. Most edge-detection methods work

on the presumption that an edge occurs when there is a discontinuity in the intensity function or a very sharp intensity gradient in the image, as shown in Figure 3.14. Most edge-detecting operators based their concepts on gradient-calculators. Since the gradient is a continuous function notion and images are discrete functions, so it has to be approximated. Gradient calculations are most often done by using convolution since derivatives are linear and shift invariant. Numerous kernels have been proposed for finding edges, some of them are: Roberts Kernel, Kirsch Compass Kernel, Prewitt Kernel, and Sobel Kernel.

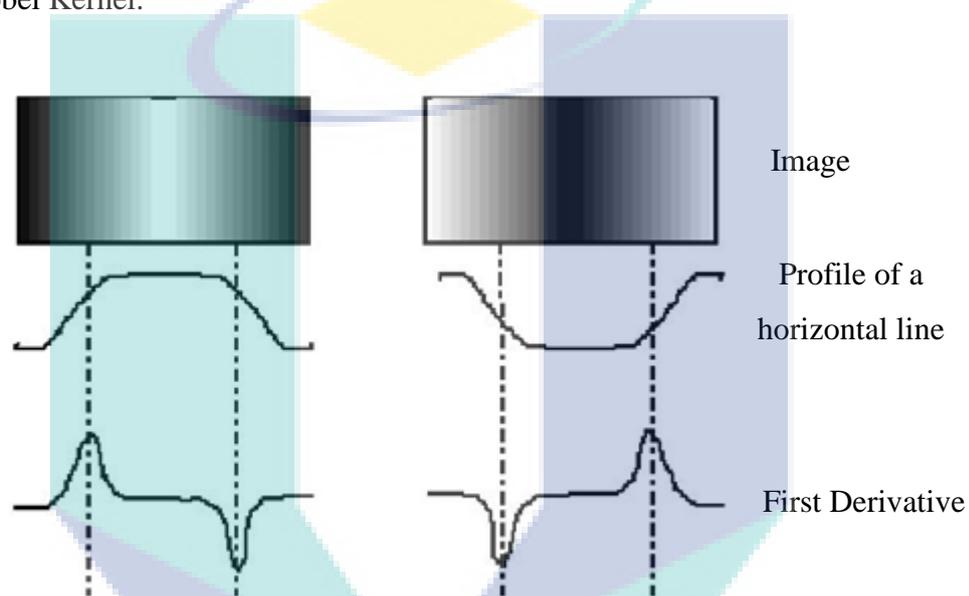


Figure 3.14: Gradient of image

The Prewitt kernels are based on the simple idea of the central difference between rows for horizontal gradient and difference between columns for vertical gradient. In image processing, the 1st derivative is the gradient, defined as the vector:

$$\frac{\partial I}{\partial x} \approx \frac{I(x+1,y)-I(x-1,y)}{2}, \text{ and } \frac{\partial I}{\partial y} \approx \frac{I(x,y+1)-I(x,y-1)}{2} \quad (3.4)$$

The convolution masks are derived from equations which is shown in Figure 3.15.

Horizontal Convolution		
0	0	0
-1	0	1
0	0	0

Vertical Convolution		
0	-1	0
0	0	0
0	1	0

Figure 3.15: 3x3 Convolution masks of Gradient image

These convolutions are used to calculate the horizontal and vertical gradients.

- Magnitude and Phase

The convolution of the image with horizontal and vertical gradients produces horizontal gradient (∂x) and vertical gradient (∂y) respectively. The absolute gradient magnitude ($|G|$) is calculated by the mean square root of the horizontal (∂x) and vertical (∂y) gradients. That is,

$$|G| = \sqrt{(\partial x)^2 + (\partial y)^2} \quad (3.5)$$

The direction of the gradient (θ) is calculated by the arc-tangent of the vertical gradient to the horizontal gradient, see Figure 3.16:

$$\theta = \arctan(\partial y / \partial x) \quad (3.6)$$

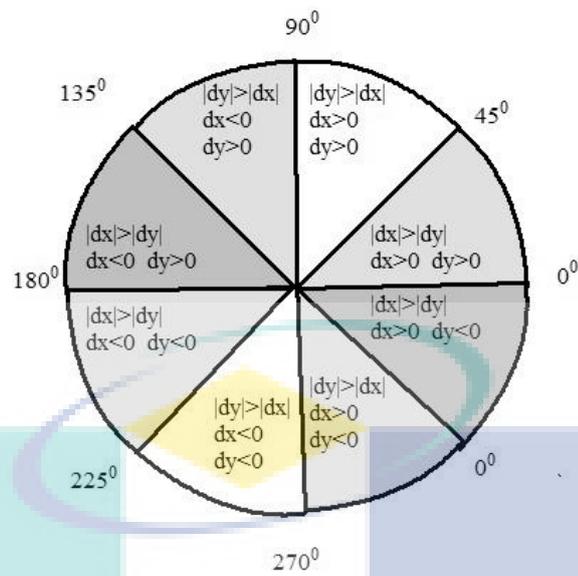


Figure 3.16: Gradient Orientation

- Non-Maximum Suppression

Once the direction of the gradient is known, the values of the pixels found in the neighborhood of the pixel under analysis are interpolated. The pixel that has no local maximum gradient magnitude is eliminated. The comparison is made between the actual pixel and its neighbors, along the direction of the gradient. For example, if the approximate direction of the gradient is between 0° and 45° , the magnitude of the gradient at $P_{x,y}$ is compared with the magnitude of the gradient at adjacent points, as shown in Figure 3.17.

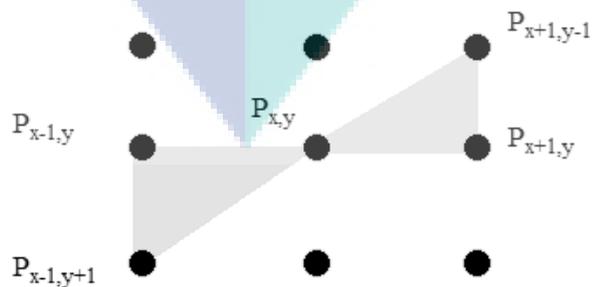


Figure 3.17: Pixel Interpolation

where,

$$P_{x,y} = |dx_{x,y}| + |dy_{x,y}| \quad (3.7)$$

The values of the gradient at the point Pa and Pb are defined as follows:

$$P_a = \frac{P_{x+1,y-1} + P_{x+1,y}}{2} \quad (3.8)$$

where $P_{x+1,y-1} = |dx_{x+1,y-1}| + |dy_{x+1,y-1}|$, and $P_{x+1,y} = |dx_{x+1,y}| + |dy_{x+1,y}|$

$$P_b = \frac{P_{x-1,y+1} + P_{x,y+1}}{2} \quad (3.9)$$

where $P_{x-1,y+1} = |dx_{x-1,y+1}| + |dy_{x-1,y+1}|$, and $P_{x,y+1} = |dx_{x,y+1}| + |dy_{x,y+1}|$

The center pixel $P_{x,y}$ is considered as an edge, if $P_{x,y} > P_a$ and $P_{x,y} > P_b$. If neither condition is not satisfied, then the center pixel is eliminated.

- Threshold

The output image of the non-maximum suppression stage may consist of broken edge contours, single edge points which contribute to noise. This can be eliminated by threshold with *hysteresis*. Two thresholds are considered for hysteresis, one high and the other low. If any edge response is above a high threshold, the pixels will constitute a definite edge output of the detector for a particular scale. Individual weak responses usually correspond to noise, but if these points are connected to any of the pixels with high threshold, they are more likely to be actual edges in the image. Such connected pixels are treated as edge pixels if their response is above the, low threshold.

To get thin edges, two thresholds [high threshold (HT) and low threshold (LT)] are used. If the gradient of the edge pixel is above the HT, it is considered as an edge pixel. If the gradient of the edge pixel is below the LT then it is unconditionally set to zero. If the gradient is between these two, then it is set to zero unless there is a path from this pixel to a pixel with a gradient above the HT; the path must be entirely through pixels with gradients of at least LT with sigma. SIGMA is the standard deviation of a Gaussian filter which is applied to an input image prior to edge detection.

Canny's Edge Detector has the ability to generate single-pixel thick continuous edges that the programme has four inputs: input image, smoothing value parameter (sigma), high threshold HT, and low threshold LT. An input requires the value of sigma and the output will generate x- and y-derivative masks. In Canny's edge detector method (Deriche, 1987), the masks used 1st derivative of a Gaussian in x- and y-directions.

Clearly, the effect of increasing sigma is dramatically highlighted in these 5 examples in Figure 3.18. When the sigma is increasing, the details of object's edge will be less till it reaches to the smoothest edge and some small objects will lose the details of their edges.

Finally in Canny's edge detection algorithm, two thresholds are applied. Since the edges would follow recursively by looking at the neighbours, at first the border pixels are made zero, so that finding neighbours would not go out of bounds of the image. Then, the image will be scanned from left to right, top to bottom. The first pixel in the non-maxima repressed magnitude image which is above the threshold, HT, is certified an edge and all its neighbours will be recursively followed, and those above the threshold, LT, are marked as an edge. A visited map is maintained as well, so that the recursion will not loop infinitely. Thus there are really two stopping conditions, they are: if a neighbour is below the LT, it would not recurse on it; and, if a neighbour has already been visited, then it would not recurse on it too, see Figure 3.21 & Figure 3.22.

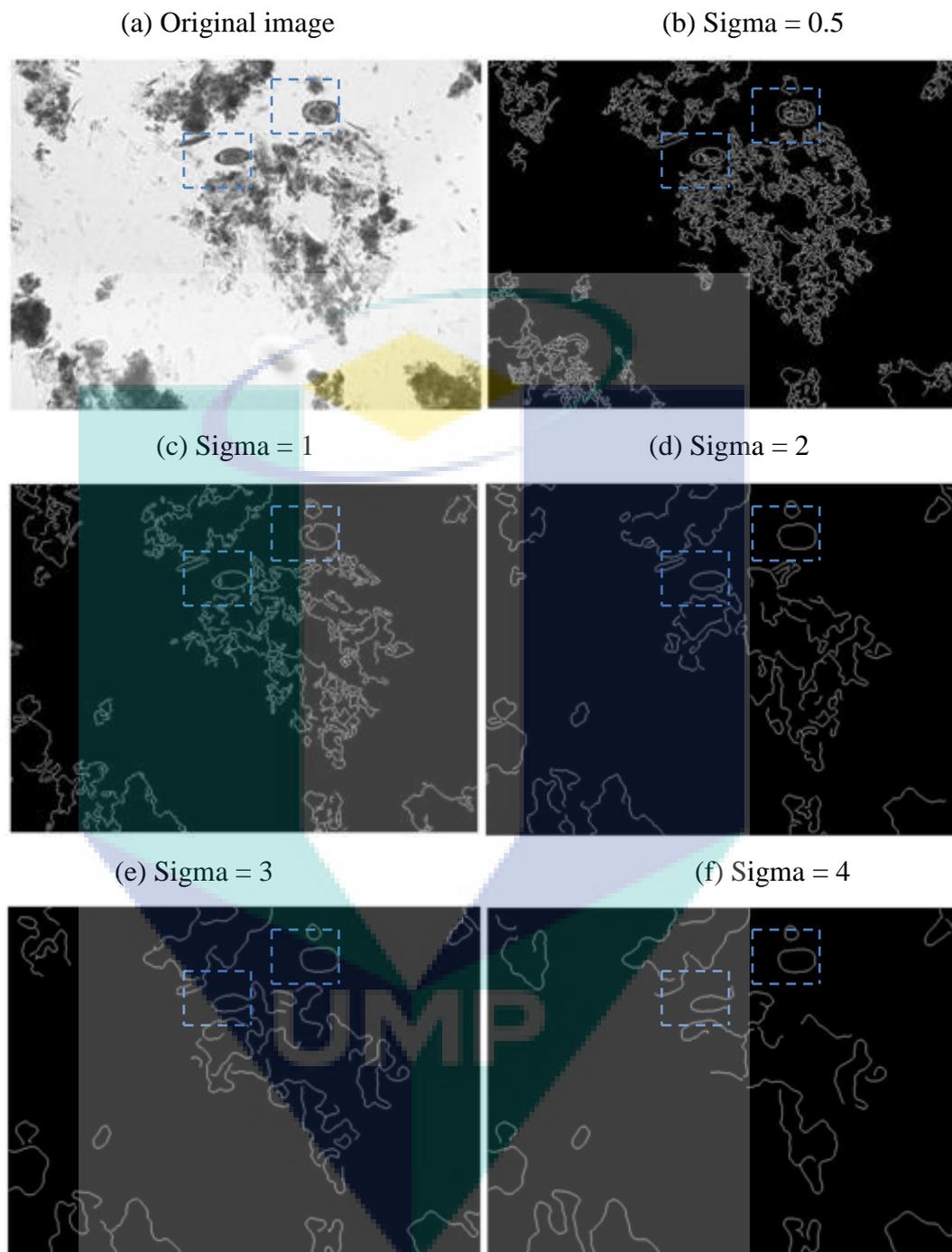


Figure 3.18: Applying masks to an image with different sigma. (a) Original image, (b) mask with sigma = 0.5, (c) mask with sigma = 1, (d) Mask with sigma = 2, (e) mask with sigma = 3, and (f) mask with sigma = 4.

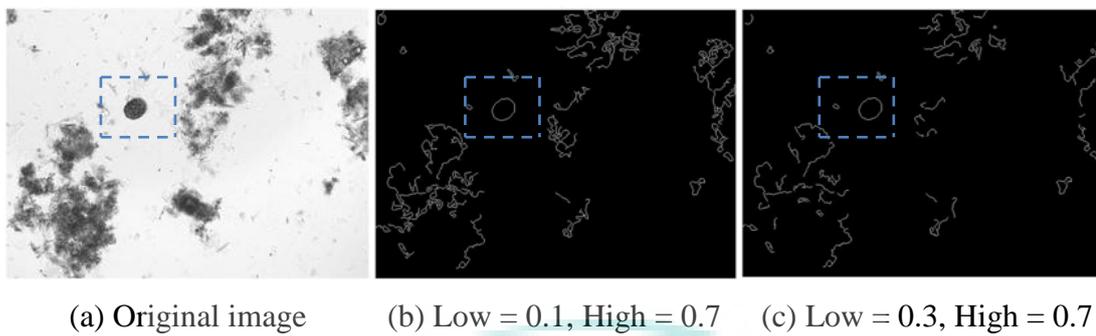


Figure 3.19: Effect of changing in low threshold with constant high threshold = 0.7 and sigma = 1 in image (a) Original image, (b) Low threshold = 0.1, (c) Low threshold = 0.3

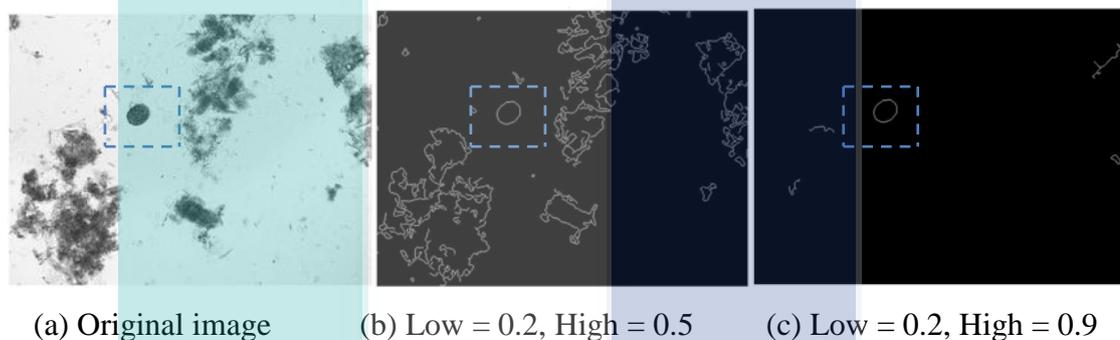


Figure 3.20: Effect of changing in High threshold with constant Low threshold 0.2 and sigma = 1 in image (a) Original image, (b) Low threshold = 0.5, (c) Low threshold = 0.9

3.4 ADDITIONAL PROCESSING STEPS

In this chapter, a few steps are needed to be processed in our work in order to prepare the data to feature extraction stage. Some of these steps are essential to extract a feature of the parasites such as circumference and pixel values and the other step was a part of morphological image processing such as object's bores removal by filling up the holes and this was needed to remove interrupted objects with the borders of an image. These steps are explained in detail as follows:

3.4.1 Circumference

Cell array is a useful means of holding various kinds and sizes of information. When manipulations or computations need to be performed on all or a subset of values in a cell array, apply a function to each cell in the cell array is a way to obtain the characteristic that is useful and can be utilized. This feature is the length of an outside object's boundary.

Circumference of an object = summation of pixels that make the outside's edge of that object.

3.4.2 Objects' bores removal

After detecting the boundary of the objects using edge detection, it needs to fill up the holes in all objects. It is a step which must be taken before going further to extract the important features of the objects inside the image, especially when these features depend on the shape of the object.

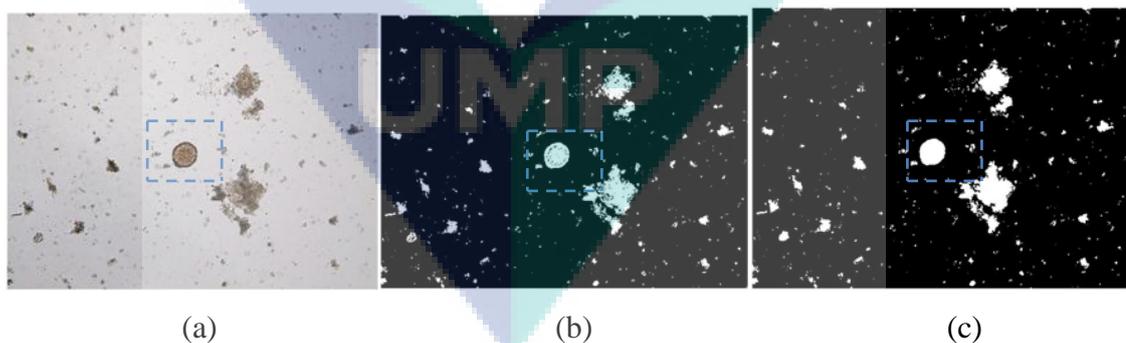


Figure 3.21: Filling holes of objects (a) original image, (b) output image in the system before applying filling hole, (c) output image in the system after applying filling hole.

In this grayscale image, a hole is represented as an area of dark pixels surrounded by lighter pixels. Same function, it is used to fill holes in the binary image. While this hole is a set of background pixels that cannot be achieved by filling in the background from the edge of the image (Soille, 1999). In Figure 3.23, we can see how the bores are

filled up, and how to reduce the errors that might happen in the further stage in feature extraction.

3.4.3 Clearing image borders

An input image can be a grayscale or binary image and the output image is grayscale or binary, respectively. For two dimensions, the default connectivity is 8, while 26 for three dimensions and for higher dimensions there is a special function for connectivity definition. For grayscale images, function of clear border tends to reduce the overall intensity level in addition to suppressing border structures (Soille, 1999).

While for same function with connectivity, the input image and output image are the same which can be either grayscale or binary but the difference is that the conductivity is selected which can have any of the following scalar values:

- For two dimensional connectivity: 4 or 8 connected neighborhood.
- For three dimensional connectivity: 6, 18 or 26 connected neighborhood.

The input image can be a numeric or logical array of any dimension, and it must be non-scattered and real. The output image has the same class as the input image. A simple binary image is used in examples to illustrate the effect of the clear border function with different connectivity.

Output =

```

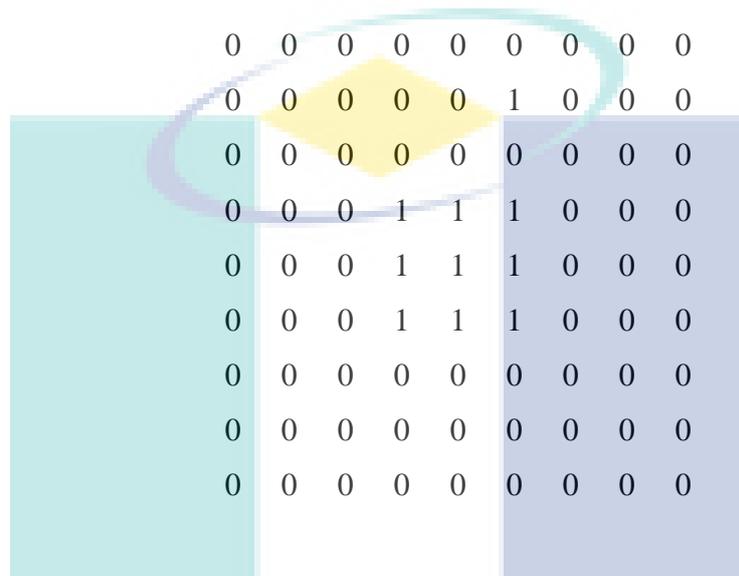
0 0 0 0 0 0 1 0 0
0 0 0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 0 0 0
0 0 0 1 1 1 0 0 0
0 0 0 1 1 1 0 1 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0

```

By using a 4-connected neighborhood, the pixel at (2,6) is not considered connected to the border pixel (1,7), so it is not clear.

Output = Clearlearborder [Input,4]

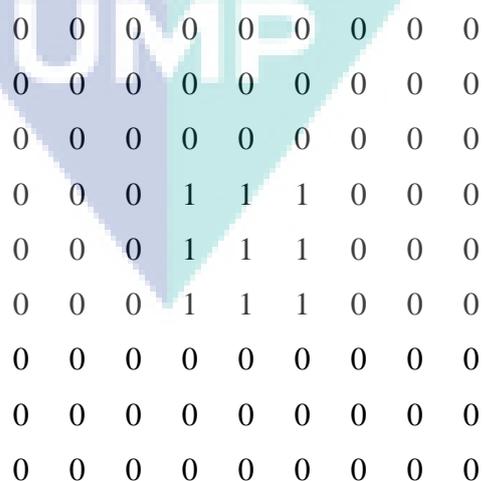
Output =



While using an 8-connected neighborhood, the pixel at (2,6) is considered connected to a pixel (1,7) so both are cleared.

Output = Clearlearborder [Input,8]

Output =



It is necessary to remove the unwanted objects attach to the border of the image to reduce the number of objects that are going to be under study, see Figure 3.22.

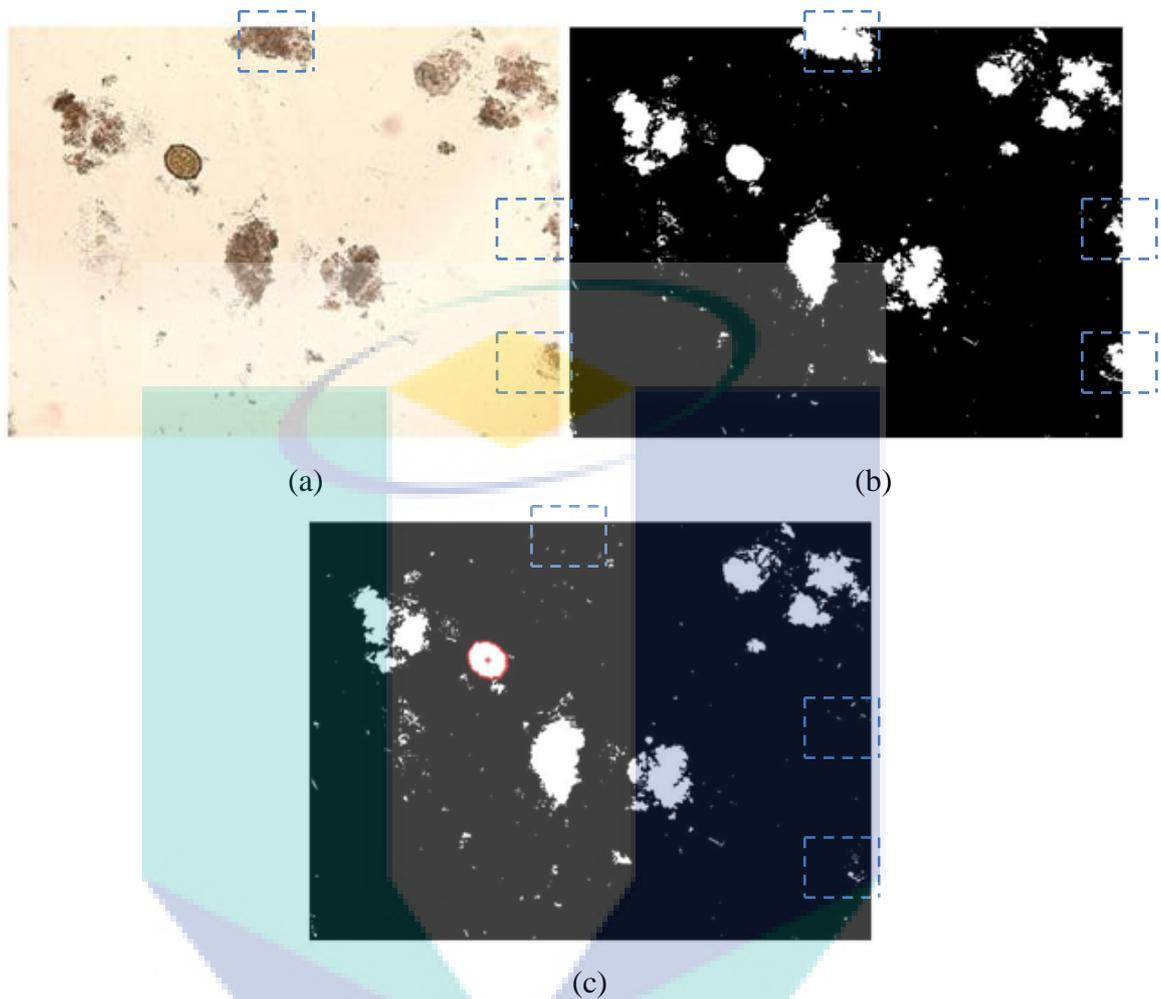


Figure 3.22: Clear boundary of input image (a) original image, (b) Before clear border, (c) After clear border.

3.4.4 Pixel and mean values

In this part of processing, we find three values of the pixel which are located in the center of each object and then determine the mean of these three values to be used later on as an additional step of feature extraction.

- Pixel values

Pixel color values include the red, green, and blue color values of specified image pixels. To determine the values of one or more pixels in an image and return the values in a variable, use the pixel color value function. The pixels can be specified by passing their coordinates as input arguments and find their values, see Figure 3.23.

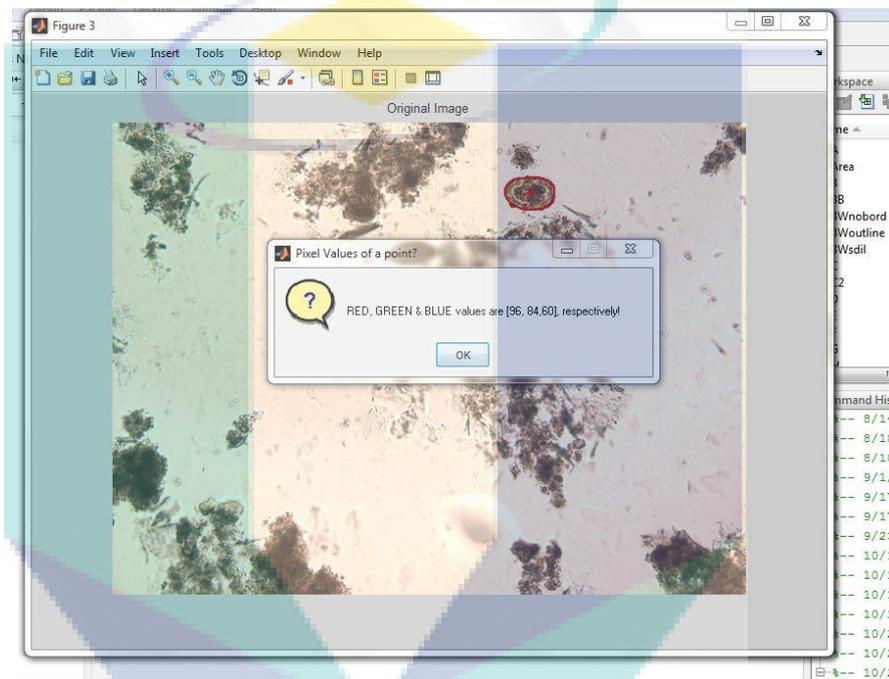


Figure 3.23: Reading a pixel value of a point in an image

And then, use a formula which converts RGB values into gray scale values by forming a weighted sum of the R , G , and B components:

$$0.2989 * R + 0.5870 * G + 0.1140 * B$$

The summation of this question gives one value in gray level of each coordinated point. In this study, pixel value in gray scale will be used as a way to remove a certain kind of objects in terms of features which are very near to the wanted objects.

- Average or Mean value

The mean values of the elements comes along different dimensions of an array. When an input image is a vector, the mean of the input-image returns. When an input image is a matrix, the mean of the input image will treat the columns of input image as vectors, returning a row vector of mean values. When an input image is a multi-dimensional array, the mean of the input-image will treat the values along the first non-single to 'n' dimension as vectors, returning an array of mean values. The mean with dimension will return the mean values for elements along the dimension of input image specified by scalar dimension. For matrices, 'mean' of an input image with dimension is a column vector containing the mean value of each row.

3.5 FEATURE EXTRACTION

When the input algorithm is too large to be addressed and suggests notoriously redundant (for example, the same measurement in both feet and meters), the input data will be transformed with a reduced representation set of feature vectors. An input conversion feature set is named feature extraction. If the extracted features are carefully selected, it is expected that the property will extract relevant information from incoming data to perform the desired task using this reduced representation instead of entering the full size of the input. In this work, we applied some techniques to extract the features that our task is performed perfectly to detect the eggs from the input images, based on these features.

3.5.1 Algorithms of feature extraction

Using 'Measure properties of image regions' in this study to extract some of necessary properties of objects in the image to become main features which help to catch the wanted objects (Gonzalez, Woods, and Eddins 2004), these properties are:

1. Calculating the number of pixels of an object--the Area (represents feature num.1).
2. Measuring the length of an object--MajorAxisLength (represents feature num.2).
3. Measuring the width of an object--MinorAxisLength (represents feature num.3).

4. Determining the coordinates of the object's center--Centroid.

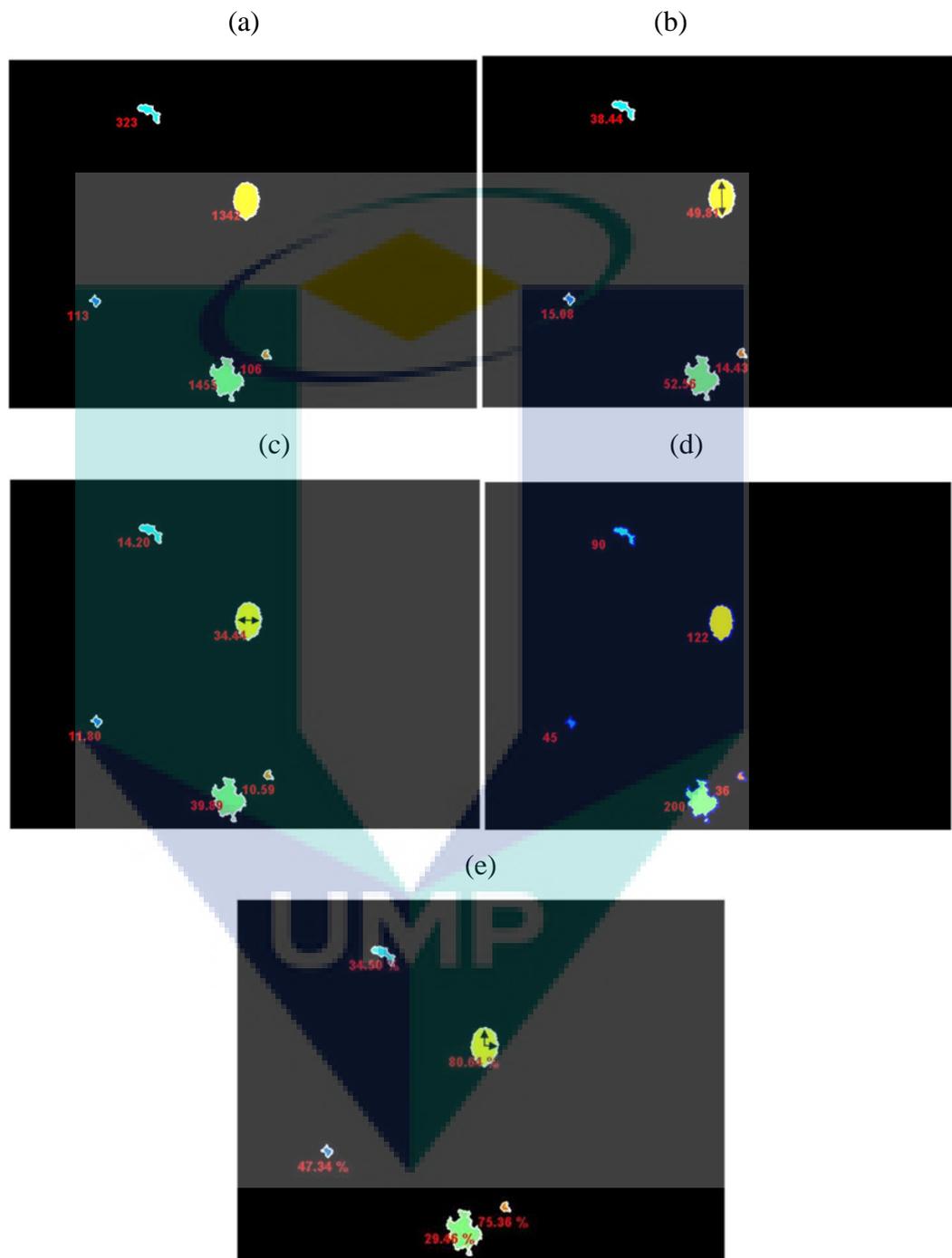


Figure 3.24: Showing how the five features are extracted, (a) calculating the object's Area in the image, (b) calculating the object's length in the image, (c) calculating the object's width in the image, (d) calculating the object's boundary length in the image, (e) calculating the object's roundness in the image.

Feature number four is the circumference of the parasite that can extract from a function as mentioned before in this chapter (in 3.3.1). while feature number five is the roundness of the parasites that can be determined as follows:

1. Obtain (X,Y) boundary coordinates corresponding to the number of objects.
2. By squaring the value of differences between adjusting element in boundary obtained from (1), using this function: $\text{delta_sq} = \text{diff}(\text{boundary}).^2$
3. Find the perimeter which is a summation of the square value of summation delta_sq and 2, as shown: $\text{perimeter} = \text{sum}(\text{sqrt}(\text{sum}(\text{delta_sq}, 2)))$.
4. Obtain the area calculated corresponding to a certain object.
5. Compute the roundness metric as shown in the following equation: $\text{metric} = 4 * \pi * \text{area} / \text{perimeter}^2$;

Now we have five features of each object in the input image. The following example shows the extraction of these features (Figure 3.24).

3.5.2 Determine threshold values of the features

The features have to be extracted by using digital image processing techniques for localizing and classifying the parasite eggs in a given image. To develop a totally automated system and image segmentation, the algorithm for feature extraction and classification should all be automated. In addition, dedicated hardware equipment is necessary to scan complete images of each specimen under the microscope. In this study, a method to detect common helminth eggs in microscopic fecal specimen images and to identify the correct species using digital image processing technique is proposed.

In the feature extraction stage, five features have been selected based on three characteristics. Those characteristics represent shape, shell smoothness, and size (section area, object's length, object's width, length of boundary, and roundness). The five parameters are the major features that are able to detect the parasite eggs and each parameter has limited maximum and minimum values, which is called range. These ranges

are recorded by experimenting and testing until the best results are obtained. Practically speaking, each object will have five values of five features, and each value of five features should be within the range of feature values. Otherwise, one of these feature values will fall out of the range, indicating the object is not a parasite. In other words, every object in the image will not be qualified for parasite suspicion if any feature's value of the five parameters is not within the range. Table 3.2 shows feature values records of some parasites.

Table 3.2 Feature Values for some ALO and TTO.

Image No.	ALO Feature Value (FV)	TTO Feature Value (FV)
	[area, length, width, boundary length, roundness]	[area, length, width, boundary length, roundness]
FV1	[1239, 42.31, 37.35, 111, 91]	[551, 35.73, 19.68, 78, 84]
FV2	[1797, 55.88, 41.19, 139, 85]	[536, 37.27, 18.37, 77, 78]
FV3	[1234, 43.78, 35.98, 138, 81]	[529, 36.10, 18.70, 79, 79]
FV4	[1647, 54.70, 38.78, 121, 86]	[564, 37.99, 19.43, 85, 76]
FV5	[1466, 45.68, 40.95, 121, 86]	[532, 36.45, 18.83, 76, 80]
:	:	:
FV96	[1810, 57.72, 40.18, 139, 80]	[544, 36.47, 18.62, 79, 77]
FV97	[1622, 52.87, 39.16, 127, 86]	[549, 36.67, 18.88, 80, 80]
FV98	[1375, 48.40, 36.28, 122, 85]	[558, 35.06, 19.93, 78, 85]
FV99	[1363, 45.64, 30.09, 118, 91]	[591, 36.43, 20.37, 80, 83]
FV100	[1591, 51.02, 39.84, 131, 87]	[558, 35.78, 19.73, 79, 82]

Experimentally, the five feature extractions and we can call them 'the five parameters' which are extracted based on the values of these parameters. The following charts show and clarify these values referring to 100 images of both kinds of diseases ALO and TTO and based on the values, the range of the five parameters are selected. Figure 3.25 presents the value of the section area of every single parasite within 100 images of ALO and same goes to TTO as shown in Figure 3.26. We can see that the

threshold of extracting the area of ALO and based on experiment, the pixels is between 1127 and 1908 and for TTO between 470 and 647.

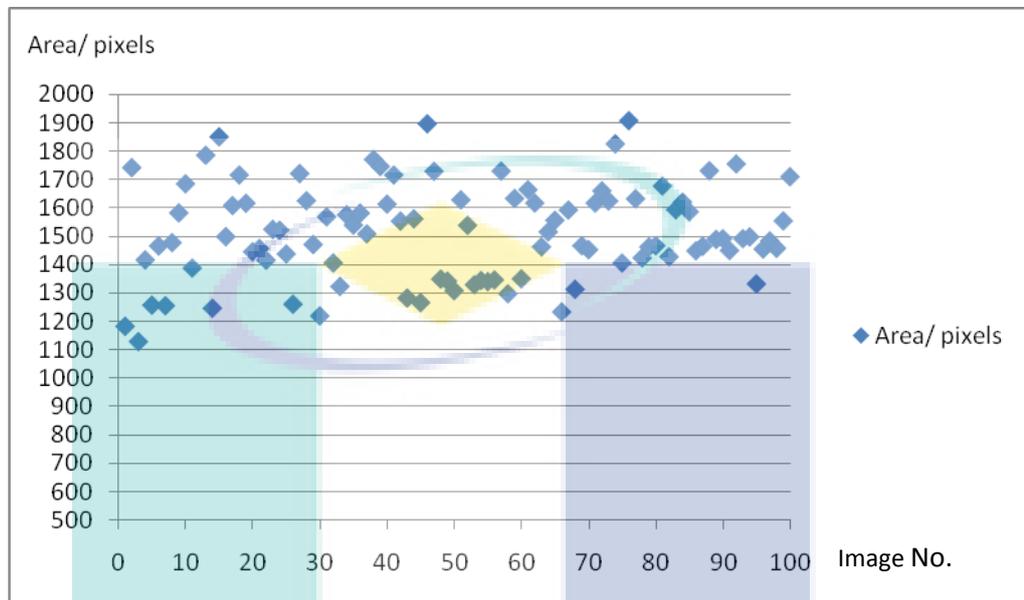


Figure 3.25: Feature values of the parasites: areas of ALO in 100 images

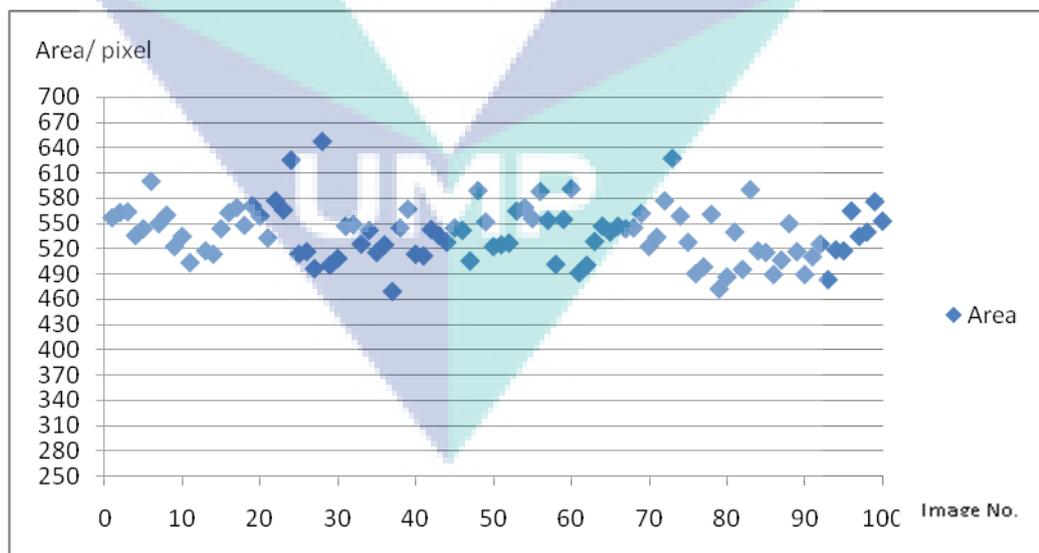


Figure 3.26: Feature values of the parasites: areas of TTO in 100 images

In Figure 3.27, the chart presents the value of the length of every single parasite in image for 100 images of ALO and same goes to TTO as shown in Figure 3.28. We can see

that the threshold of ALO length to be extracted is between 41 and 59 pixels; for TTO length it is between 31 and 41 pixels.

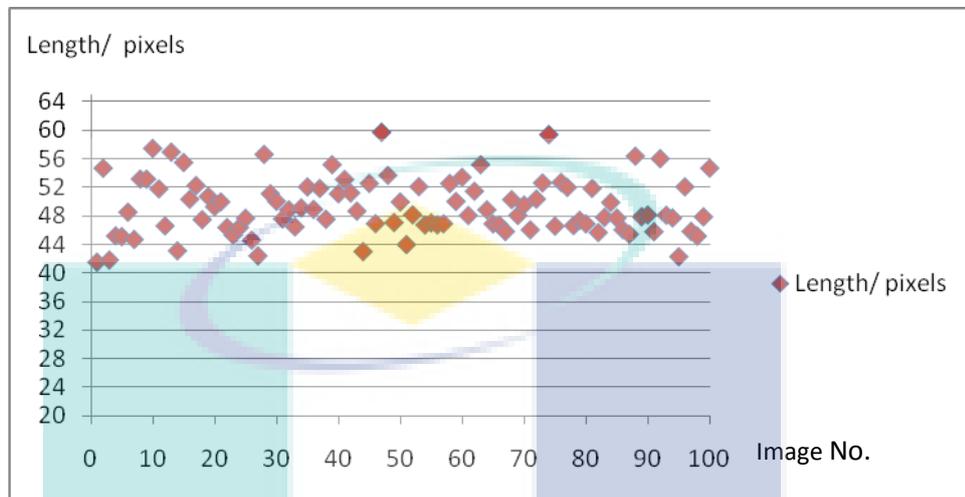


Figure 3.27: Feature values of the parasites: lengths of ALO in 100 images

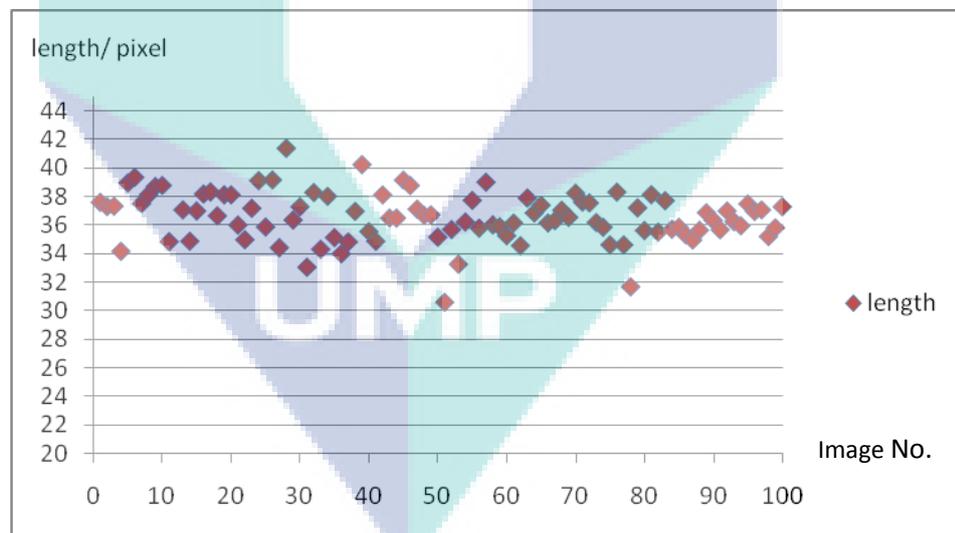


Figure 3.28: Feature values of the parasites: lengths of TTO in 100 images

The chart in Figure 3.29 presents the value of the width of every single parasite in image for 100 images of ALO and same goes to TTO, as shown in Figure 3.30. These values show the width of ALO is between 32 and 46 pixels, while for TTO it is between 16 and 21 and these values represent the thresholds of the extracted objects.

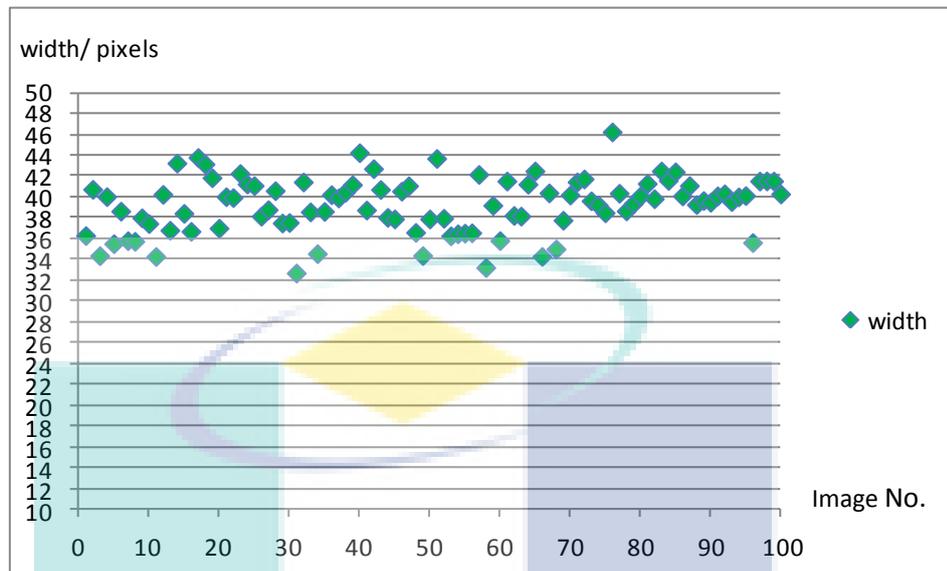


Figure 3.29: Feature values of parasites: widths of ALO in 100 images

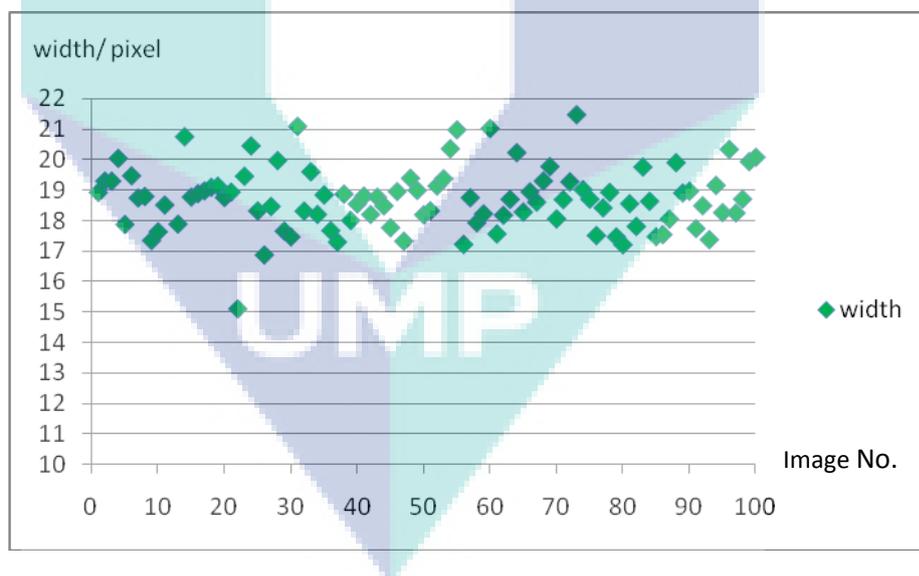


Figure 3.30: Feature values of parasites: widths of TTO in 100 images

In Figure 3.31, the chart presents the value of the boundary length of every single parasite in image for 100 images of ALO and same goes to TTO as shown in Figure 3.32. As previous charts, the thresholds of ALO and TTO can be extracted here for boundary lengths which are between 106 to 146 and 74 to 90 for ALO and TTO, respectively.

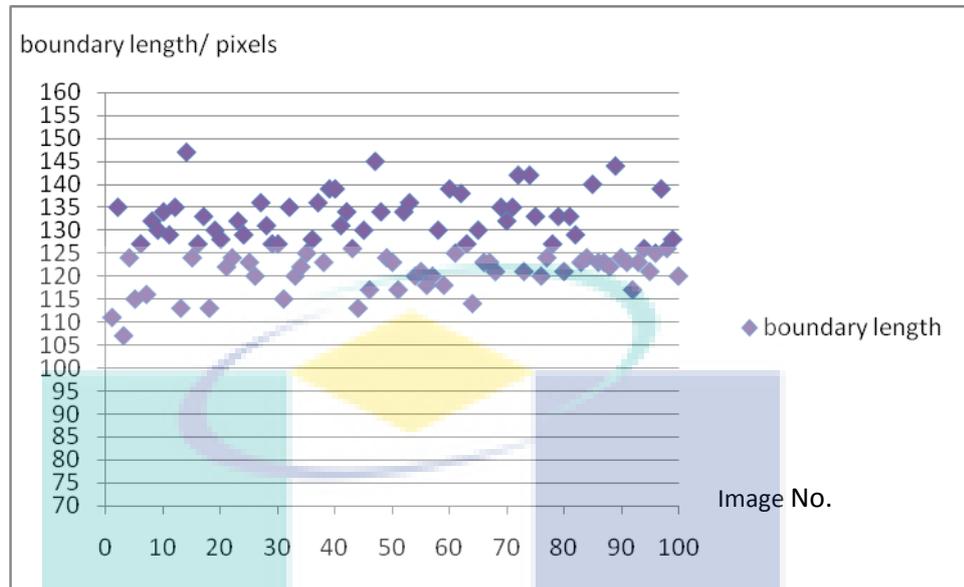


Figure 3.31: Feature values of parasites: boundary lengths of ALO in 100 images.

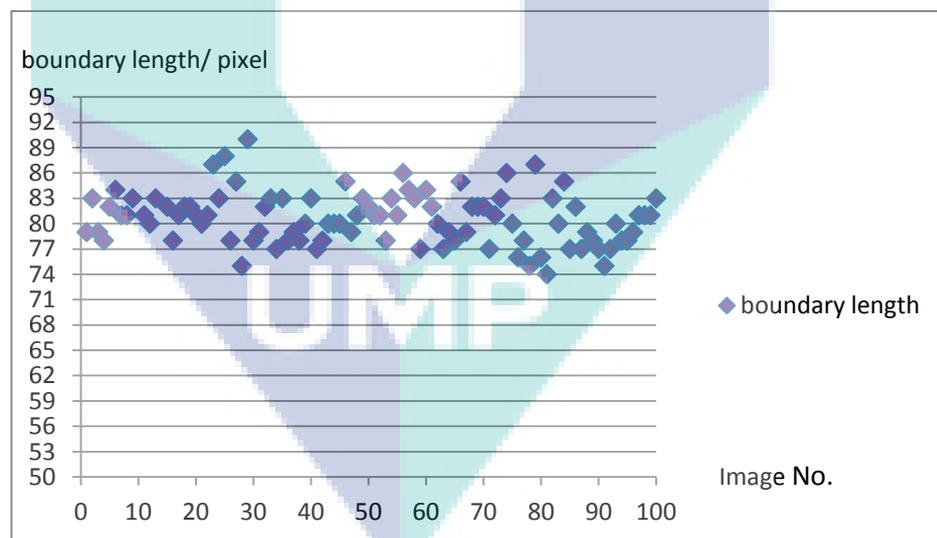


Figure 3.32: Feature values of parasites: boundary lengths of TTO in 100 images.

Figure 3.33 presents the value of the roundness of every single parasite in image for 100 images of ALO and same goes to TTO, as shown in Figure 3.34. In these two charts, we see that the limited values of the object's roundness for ALO and TTO are 75% to 94% and 73% to 87% respectively.

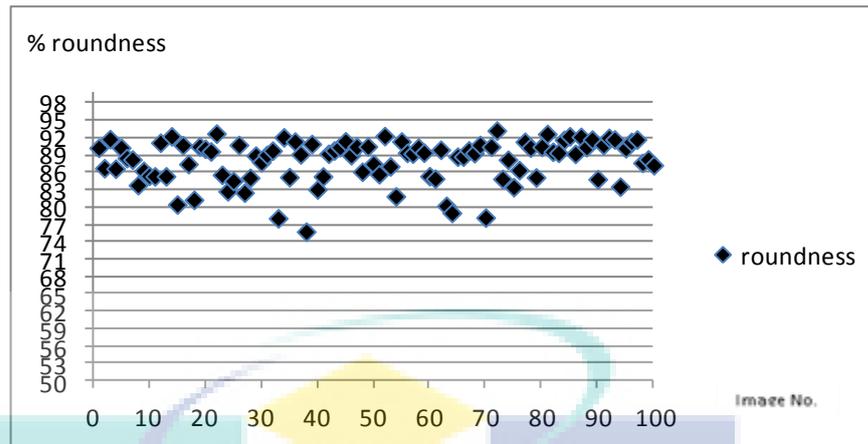


Figure 3.33: Feature values of parasites: roundness of ALO in 100 images

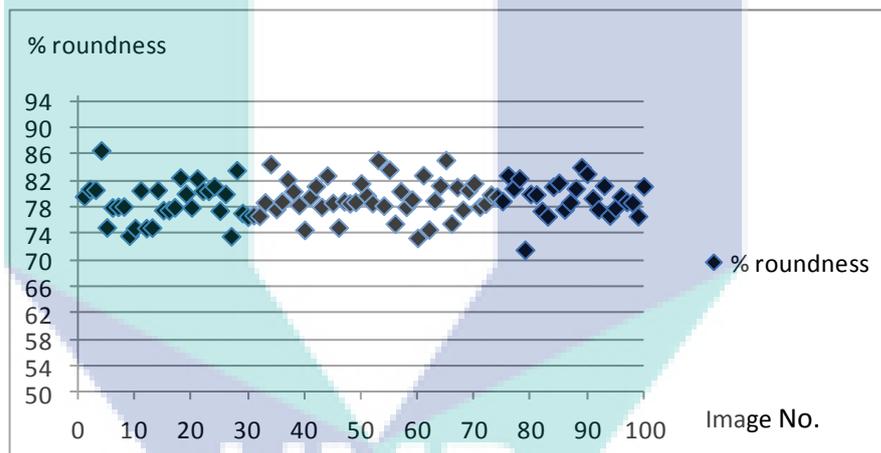


Figure 3.34: Feature values of parasites: roundness of TTO in 100 images

3.6 CLASSIFICATION STAGE

Classification stage is the most important stage since it classifies the objects detected in the images to check if they are parasites or artifacts and then makes a decision about it, see Figure 3.35.

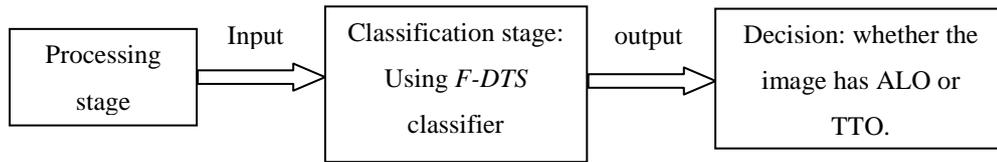


Figure 3.35: Block diagram of classification stage

In detail, the next diagram in Figure 3.36 will clarify how the classification stage works with Filtration and Determination Threshold System (*F-DTS*) classifier, which is based on the values of the features, as shown in Table 3.1.

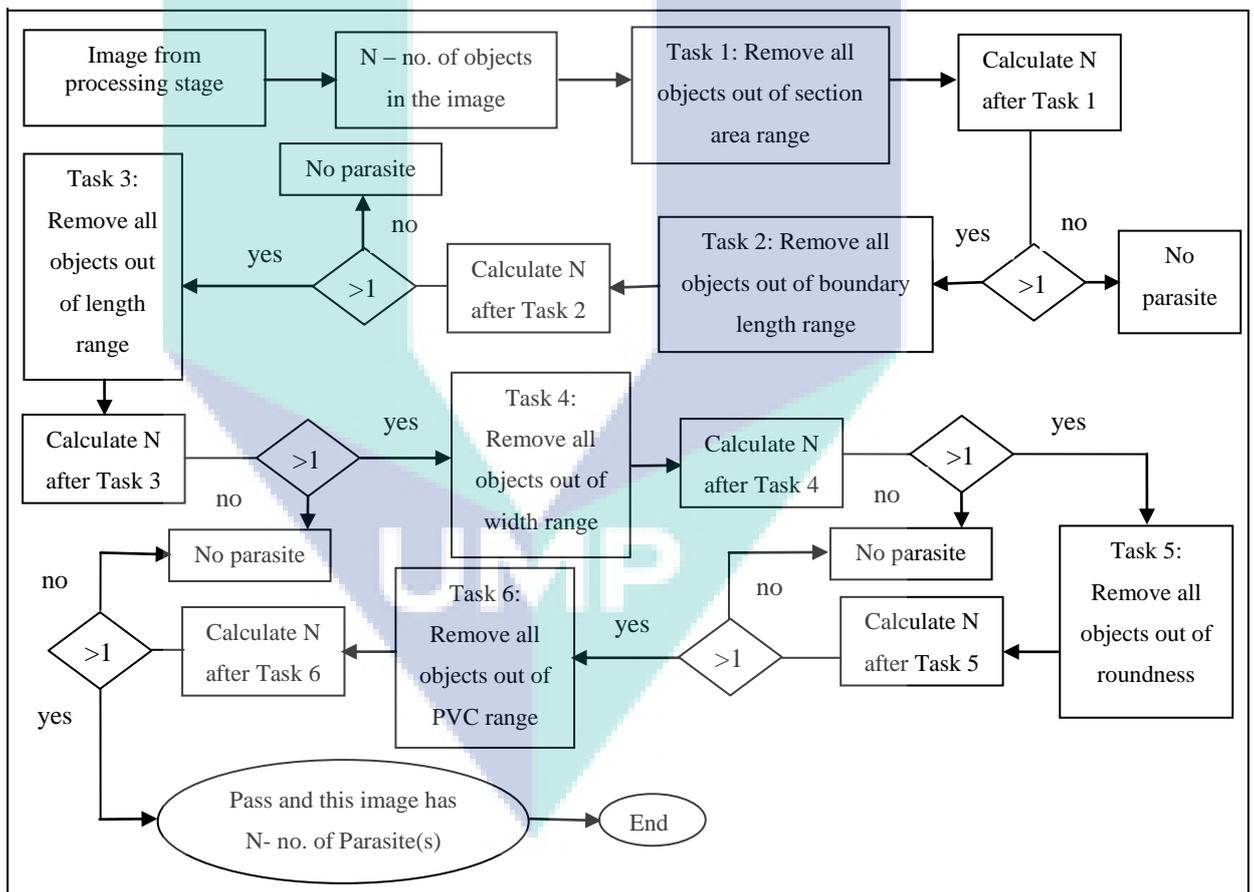


Figure 3.36: Classifier Tasks in Classification Stage

This information is for both parasites ALO and TTO, as well as others when both of them are studied at the same time. After every single task of six tasks, the programme will recalculate N. N is the number of objects remaining inside the image, so as to test if

the image captured objects. When there is one object or more than one, the system will move to the next task but if there is no object remained which means the image has no parasite detected, the system will stop and request for the next image and so on.

After the processing stage, the system calculates the number of objects inside the input image – N, and then moves to:

1. Task -1 which deletes all objects out of the area range ($850 < \text{area} < 1900$ for ALO and $350 < \text{area} < 750$ for TTO; note, that all measurements are in unit pixel). The system will calculate N again (N- number of remaining objects). Here the logic of this system works by giving 0 or 1 (if $N = 0$ then the logic will give 0, while $N > 0$ the logic will give 1). The system passes to the next task (Task-2) only if the logic of previous task is 1. Otherwise, if the logic is 0, the system will stop and give 'no parasite' as a result of this image and will request for the next image.
2. Task-2 is about removing all objects out of boundary length range ($100 > \text{boundary length} > 161$ for ALO and $69 < \text{boundary length} < 94$) and the same as the previous procedure that recalculating N and if the logic is 1, will move to Task-3; if the logic gives 0, it will stop and request for the next image.
3. Task-3 is about deleting all objects out of the length range ($41 < \text{length} < 60$ for ALO, $34 < \text{length} < 44$ for TTO) and again after calculating N, if the logic is 0, the system stops and requests new image but if the logic is 1, Task-4 will be the next task.
4. Task -4 which is about removing all objects out of a width range ($32 < \text{width} < 46.5$ for ALO, $15 < \text{width} < 21.6$ for TTO) and N will be calculated, so if the logic is 1, it will go on to Task-5 but if logic is 0, the system requests the next image.
5. Task-5 is about deleting all objects out of roundness range ($79 < \text{roundness} < 98$ for ALO, $74 < \text{roundness} < 91$ for TTO) and same as before recalculating N. Task-6 is the last task of this system only if the logic is 1 to give the last decision whether the image has N-number of parasite(s) (ALO, TTO or both) or if this image has no parasite, it will request a new image. Task-6 is an additional task and practically it helps removing special kind of confusing objects appeared in most fecal specimens, see Figure 3.37.
6. Task-6 is about removing all objects out of PVC range (PVC is the Pixel Value of Centroid in an object at the grayscale level which is $20 < \text{PVC} < 200$ for both ALO and TTO), but if the logic is 0, the system will stop and ask for the next image.

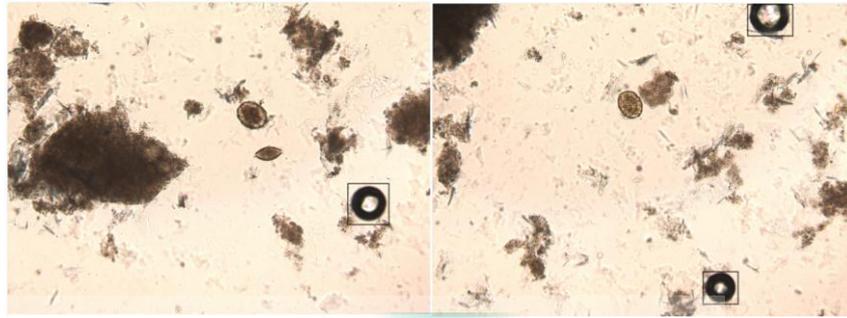


Figure 3.37: Unwanted and confusing objects in human fecal specimens

After extracting the five features and determining their values, the next step is to use them to detect the parasites in the images by using *F-DTS* classifier. The results of processing steps roughly are shown in terms of output images in Figures 3.38 & 3.39.

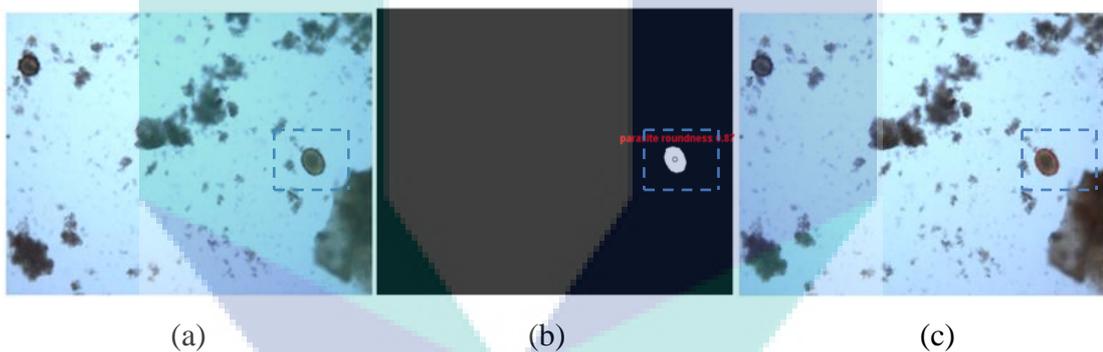


Figure 3.38: Morphological operation of parasite (*Ascaris lumbricoides* Ova (ALO)). (a) Original image, (b) Parasite detection image, (c) Outlined the parasitic body in the original image

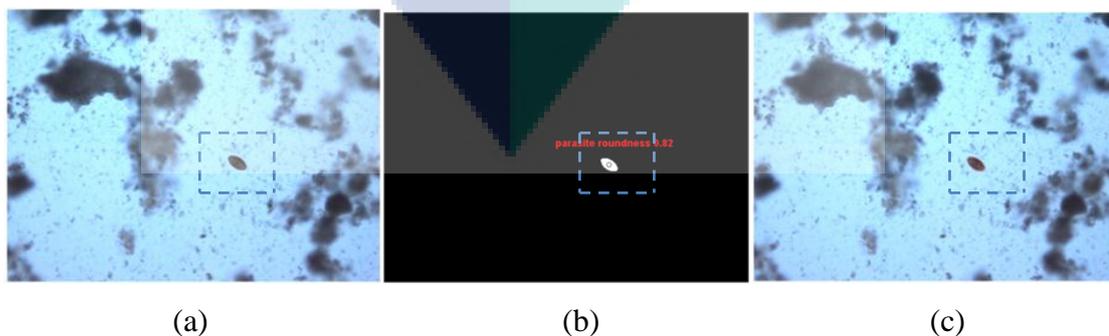
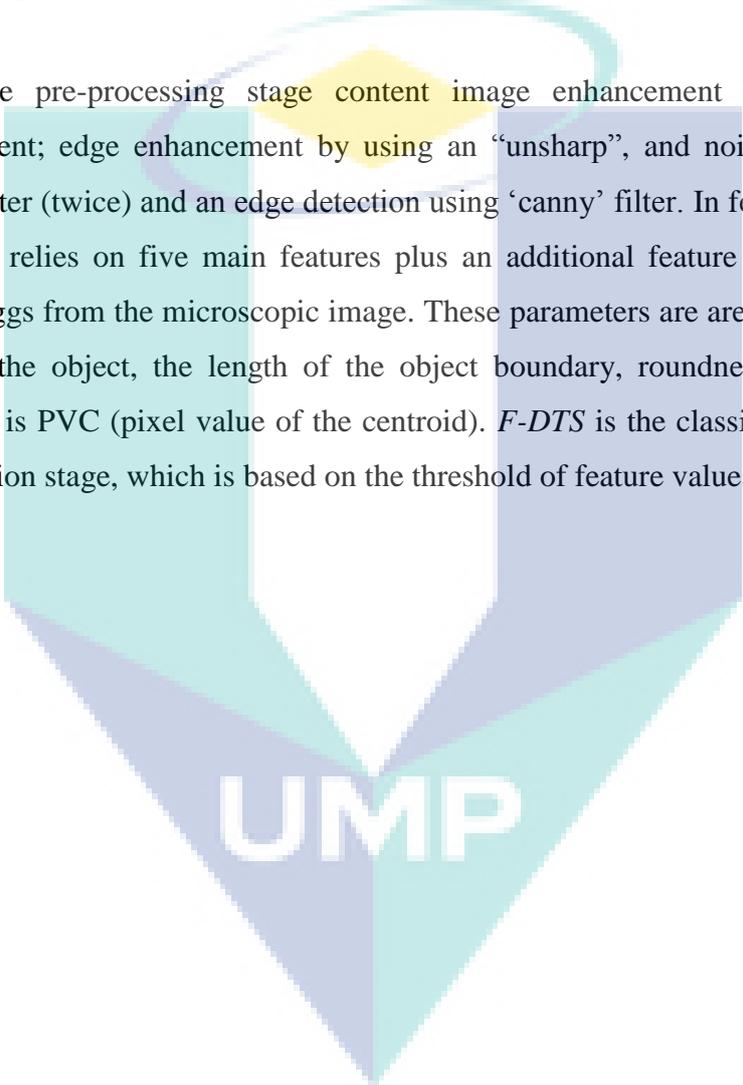


Figure 3.39: Morphological operation of parasite (*Trichuris trichiura* Ova (TTO)). (a) Original image, (b) Parasite detection image, (c) Outlined the parasitic body in the original image.

3.7 SUMMARY

In short, this chapter discusses the techniques in the preprocessing stage which could help filtrate and identify the wanted objects (parasite eggs) among many surrounded objects and artifacts. The sequences of the preprocessing stage are very important in terms of not losing the main data of the input image.

The pre-processing stage content image enhancement is by using contrast enhancement; edge enhancement by using an “unsharp”, and noise reduction by using median filter (twice) and an edge detection using ‘canny’ filter. In feature extraction stage, this study relies on five main features plus an additional feature (*Five+1*) in detecting parasite eggs from the microscopic image. These parameters are area, length of the object, width of the object, the length of the object boundary, roundness, and the additional parameter is PVC (pixel value of the centroid). *F-DTS* is the classifier that is used in the classification stage, which is based on the threshold of feature values.

The logo of UMP (Universiti Malaysia Perlis) is a shield-shaped emblem. It features a central white shield with a blue border. The shield is divided into four quadrants by a white cross. The top-left and bottom-right quadrants are light blue, while the top-right and bottom-left quadrants are light purple. The letters 'UMP' are written in white, bold, sans-serif font across the center of the shield.

UMP

CHAPTER IV

RESULTS AND DISCUSSION

4.1 INTRODUCTION

In this chapter, we are going to evaluate techniques that have been used in the pre-processing stage such as contrast enhancement, contrast threshold, edge enhancement, edge threshold, median filtering and performance of three pre-processing methods. We will do some comparisons of performing a few techniques to get the best results and then analyze the effect of using parameters in the results. All these studies will come sequentially with overall results.

4.2 CONTRAST ENHANCEMENT EVALUATION

In Figure 4.1, we can see a real example with images showing the effectiveness of three techniques that are tested for best contrast enhancement in order to detect the parasite eggs. As mentioned in the previous chapter, there are three contrast enhancement techniques, namely:

1. Adjust image within intensity values.
2. Histogram equalization.
3. Adaptive histogram equalization.

From the images, if we apply the ‘histogram equalization’ technique, the output image will be darker than the original image; the ‘adaptive histogram equalization’ technique will give an output image with a bright background; whereas the ‘adjust the image within intensity values’ technique will yield a brighter background in the output image than the ‘adaptive histogram equalization technique’. Getting the output image

brighter with less big dark spot objects can help remove the artifacts with unwanted objects and focusing on the egg features which in the end can extract them easily.

Based on that analysis, the ‘adjust the image within intensity values’ technique is the one that gives perfect contrast and helps the system detect both diseases of ALO and TTO. The transfer function which is used in ‘adjust image within intensity values’ technique is the gamma contrast adjustment.

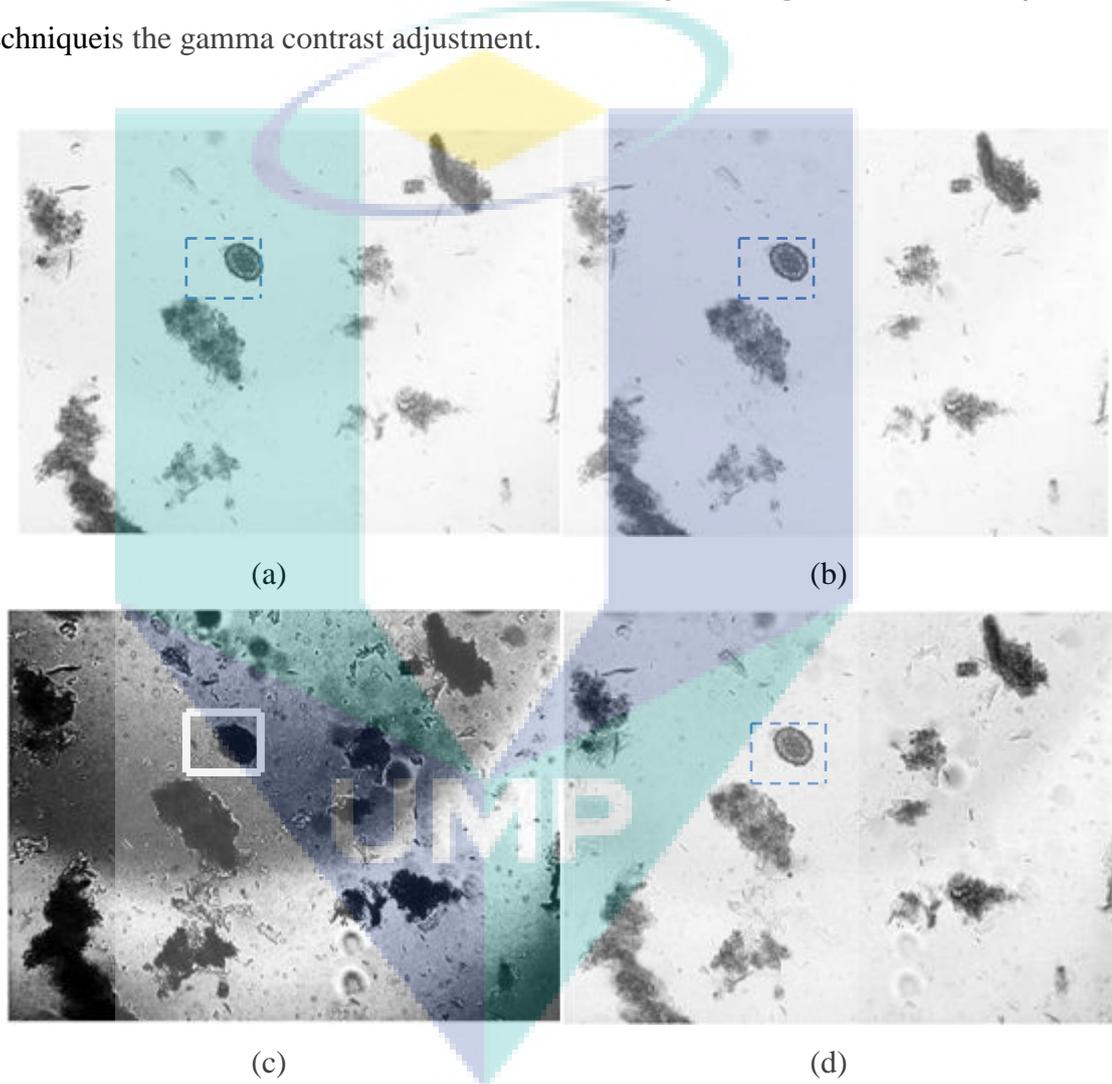


Figure 4.1: Effectiveness of three contrast enhancement techniques in image to detect ALO. (a) Grayscale of an original image, (b) using adjust image within intensity values, (c) using histogram equalization, and (d) using adaptive histogram equalization

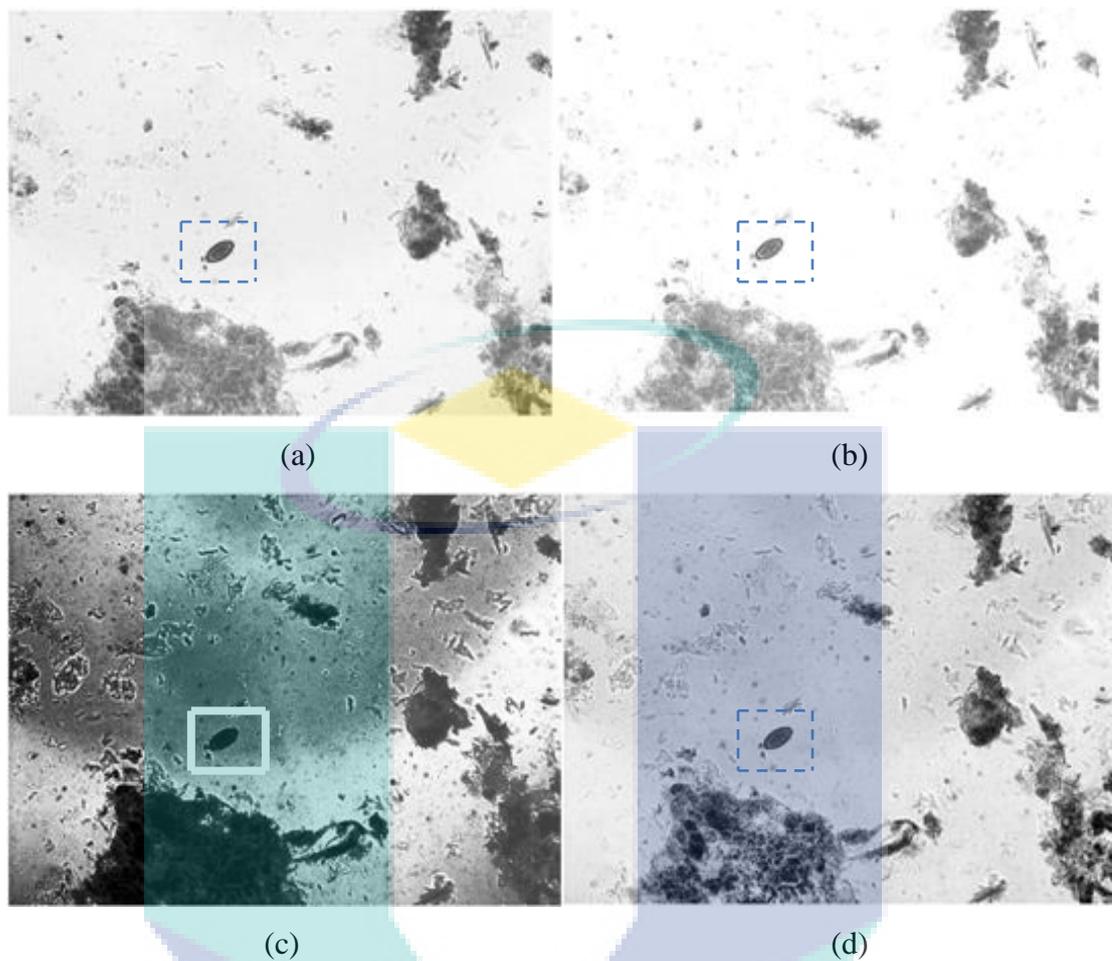


Figure 4.2: Effectiveness of three contrast enhancement techniques in image to detect TTO. (a) Grayscale of an original image, (b) using adjust image within intensity values, (c) using histogram equalization, and (d) using adaptive histogram equalization

4.2.1 Contrast threshold evaluation

Basically, the external borders of two kinds of parasites are clear and thick. We have used some steps to reduce the artifacts and bright the borders for easy removing of the surrounding objects.

In chapter 3, contrast enhancement was a part of treatment to enhance the image to become better in characteristics with a good border sharpening for easy distinction from the noises of the image. The technique used to enhance the contrast and adjust image intensity values is the adjust image within intensity values and this technique has `low_in` and

high_in, low out and high out values of image pixels with gamma value. The low_in and high_in are the intensities in the input image which are replaced to low_out and high_out in the output image. For low_in, high in, low_ out and high_out, all values are default values. By experimenting, we found out that using the default values of low_in, high_in, low_out and high_out (= [0.0 1]) is better to keep the image information safe from losing.

Gamma is a very important factor and it is the main player in this step. In details, we have to talk about the algorithm of this factor. Also we use Logarithmic-law and Power-law transformations to map a narrow range of intensity values in the input image into a wider range of intensities in the output image. The logarithmic transformation is given by:

$$s = c \log(1 + r)$$

where c is constant and $r \geq 0$. The log transformation is often used to compress dynamic range. As with logarithmic transformation, we use power-law transformation to expand a certain range of intensity values. The power-law transformation is shown below:

$$s = cr^\gamma$$

The convention of the exponent of the power-law equation is called gamma. When gamma (γ) is greater than 1, the range of darker intensity values of the image will be expanded; if gamma is less than 1, the range of brighter intensity values of the image will be expanded. Definitely, when gamma = $c = 1$, we have a petty identity mapping.

Meanwhile, as the exponent is called gamma, the operation used to correct the power-law response phenomenon is called gamma correction. ‘Adjust image within intensity values’ can be taken as an additional argument which specifies the gamma correction factor. Referring to the value of gamma, the mapping between values in the input and output images may be non-linear. For example, the value halfway which is between low and high may map to a value either greater than or less than the value halfway, that is between the bottom and top.

Gamma can be any value between 0 and infinity. If gamma is 1 (the default value), the mapping is linear. If gamma is less than 1, the mapping is weighted towards higher (brighter) output values. If gamma is greater than 1, the mapping is weighted towards lower (darker) output values. Experimentally, various values of gamma are tested to see the effect of changing that parameter value on the performance of the output image in contrast enhancement stage.

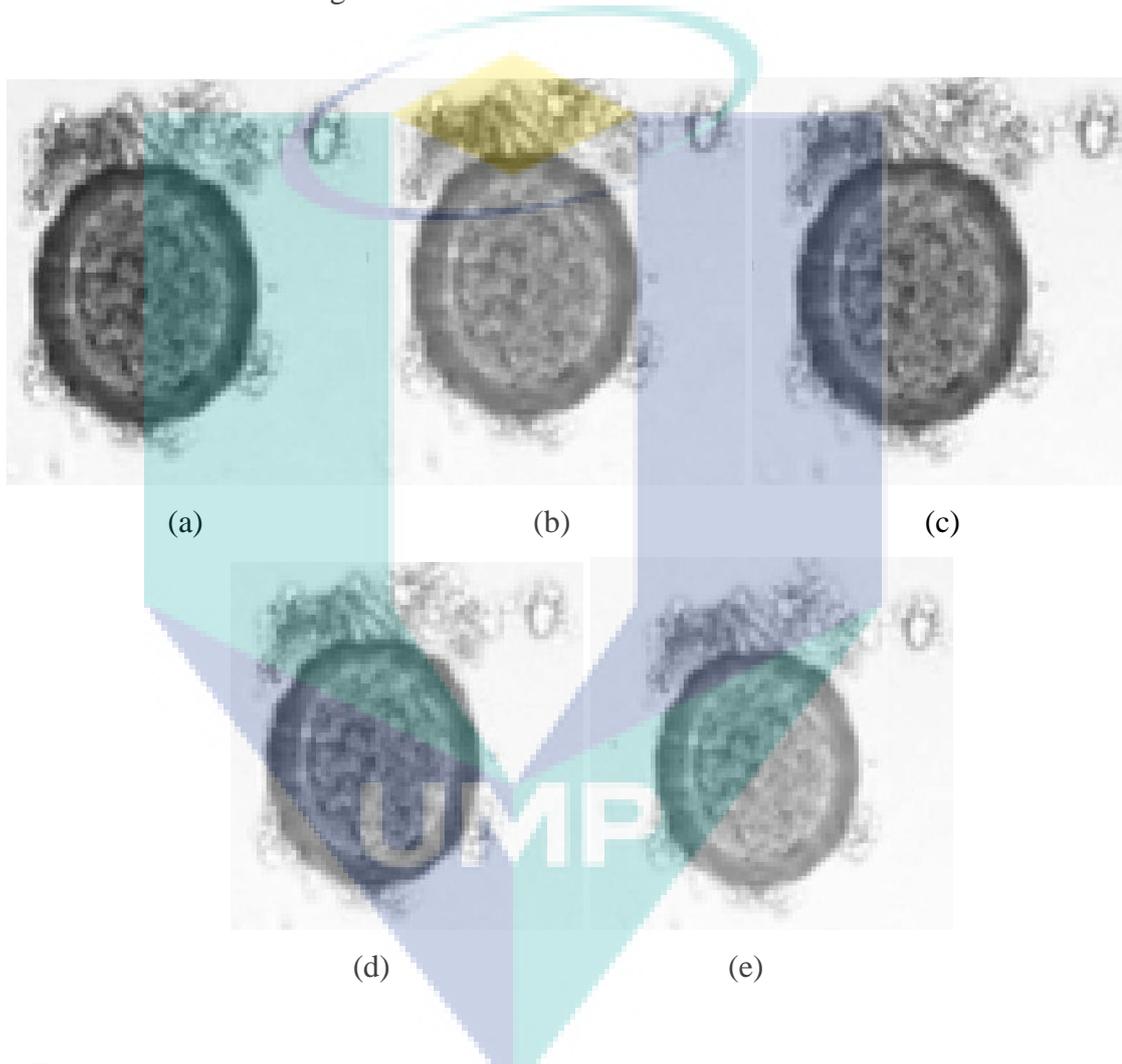


Figure 4.3: Effect of gamma value in image enhancement on the ALO parasite eggs for best feature extracted among the artifacts and the unwanted objects. (a) Image with gamma value = 1, (b) image with gamma value = 0.9, (c) image with gamma value = 0.8, (d) image with gamma value = 0.7, and (e) image with gamma value = 0.6

In figure 4.3, some differences can be seen in terms of appearance of the parasite with surrounding artifacts and objects which are sometimes attached to the eggs. The most

important step is to reduce the darkness of the data and keep the image data away from losing its major characteristics.

Obviously, in image of Figure 4.3a, we can see contact region between the egg and the artifact which is tight because the gamma is equal to 1 and it gives an output with dark spots, and this contact region has reduced the dark spots of the image whenever the gamma value is applied which is less than the previous one; such as 0.9, 0.8, 0.7 and 0.6 for ALO, as shown in Figures 4.3b, 4.3c, 4.3d and 4.3e, respectively. Similarly, it can be seen in Figure 4.4 that when we talk about TTO after applying gamma, the values will be 0.5, 0.6, 0.7, 0.8, and 0.9.

These varieties of gamma values for some input images are not affected but for other images that have parasites with contacting artifacts, they will be affected. Therefore, the image has to be enhanced using the noise reduction technique.

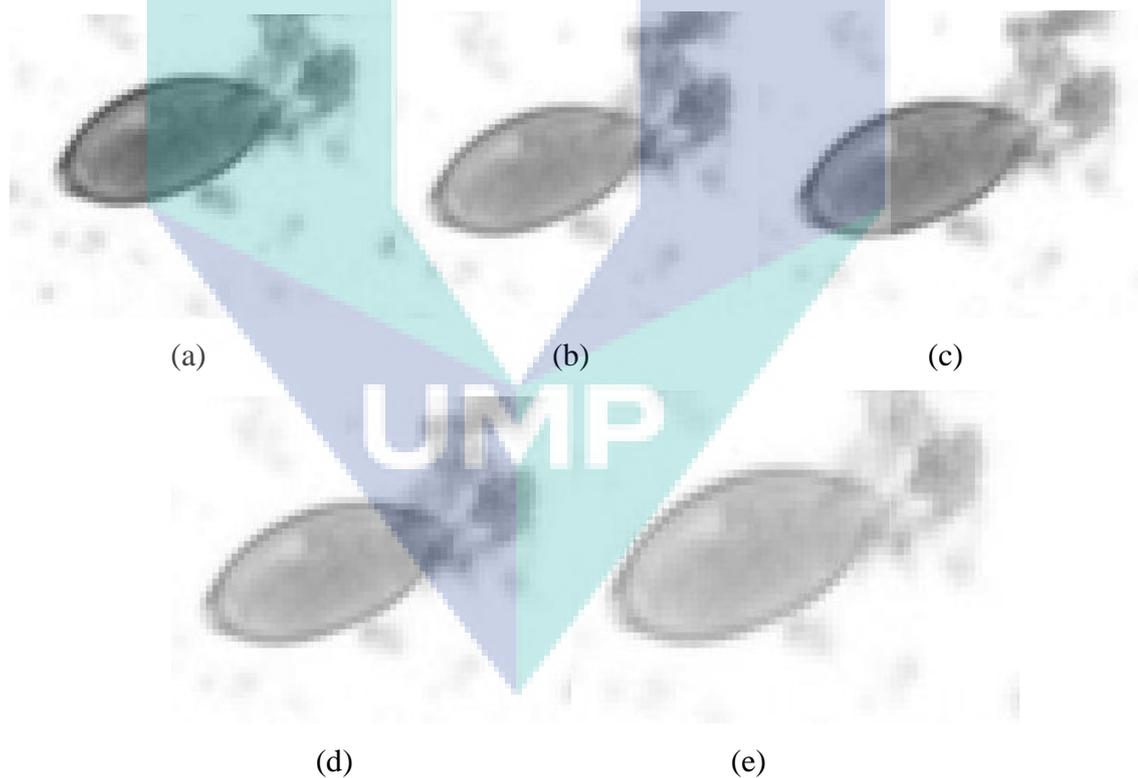


Figure 4.4: Effect of gamma value in image enhancement on the TTO parasite eggs for best feature extracted among the artifacts and the unwanted objects. (a) image with gamma value = 0.9, (b) image with gamma value = 0.8, (c) image with gamma value = 0.7, (d) image with gamma value = 0.6, and (e) image with gamma value = 0.5

While checking the performance of various gamma values on most images that have contacted artifacts, we can see the performance of gamma value equal to 0.8 is much better in detecting of ALO, while in TTO, we can consider that gamma value equal to 0.7 is the suitable value of enhancement and keep the input data safe from missing and the brightness is quite reasonable. Later on in this chapter we will show the comparison of applying these values to the image enhancement technique after setting other steps' parameters.

4.3 EFFECT OF MEDIAN FILTERING

Median filtering is one of the suitable filters for noise reduction and has been used in this work. During our work, few images were difficult to delete or reduce the noises, especially the small ones attached to our objects (eggs). Then, an idea has popped out. That idea could solve the problem of removing some of the small noises and give a higher percentage of successful detection for both kinds of parasites ALO and TTO.

The idea is to use a median filter twice with the same mask size rather than using it once only and we have proved that using a median filter twice can give slightly a better result than using it once.

At the preprocessing stage, median filtering with a mask size 2x2 for ALO detection was applied twice to reduce the error of ALO detecting. In Figure 4.5, we can see the differences of using one median filter or twice to the image after the edge enhancement step. The median filtering with a mask size 3x3 for TTO detection was proceeded twice.

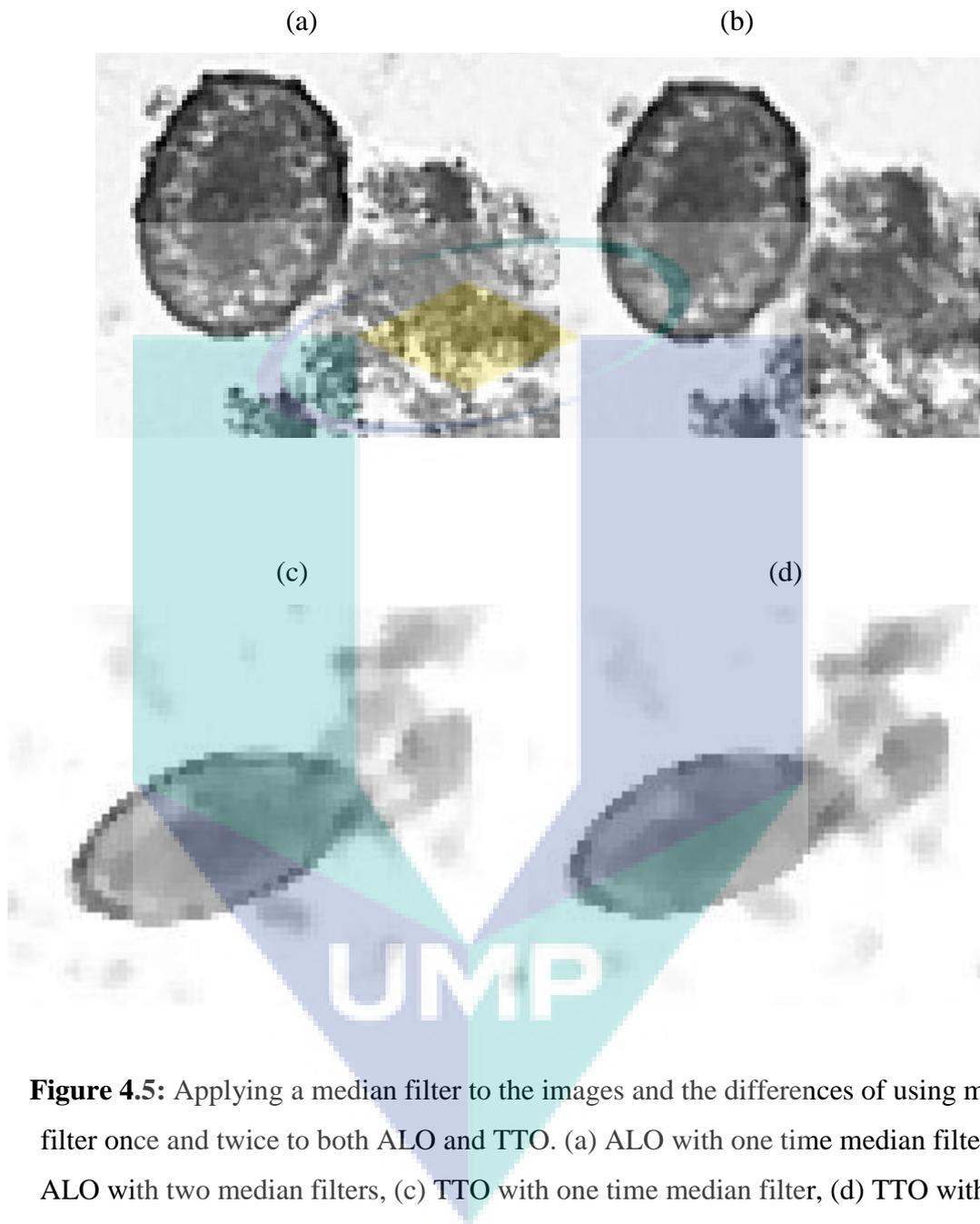


Figure 4.5: Applying a median filter to the images and the differences of using median filter once and twice to both ALO and TTO. (a) ALO with one time median filter, (b) ALO with two median filters, (c) TTO with one time median filter, (d) TTO with two median filters

Using the median filter twice can reduce the noises of the image generally and the artifacts that are contacted to the parasites specially, also it can minimize the resolution of contact region between the parasites and the artifacts. Basically, with facts and results, we could approve the benefit of using median filter twice even though the difference is very slight but in terms of accurate detection it is still needed and this is considered an

improvement. Later on in this chapter we will show the effect by using median filter twice rather than one.

4.4 EDGE DETECTION EFFICIENCY

The efficiency of edge detection comes with trying various techniques of edge detection and applying different values of factors until we reach to the best result with high efficiency. In this study, we have tried five techniques of edge detection to find out which is a suitable one for better detection with good results. Parameters are very important because they are the keys to eliminate unwanted objects in the image and focus on the threshold and the characteristics of the wanted objects.

The five edge detection techniques are 'sobel', 'prewitt', 'log', 'roberts' and 'canny'. These techniques are performed and obtained certain results in terms of success in parasite detection. In Figures 4.6 and 4.7, we can see the performance of five edge detection techniques in the same input image of ALO and TTO in order to find the differences between them and to see which one is the best to detect the parasites without losing their features and to see which technique can give good edge detection. Basically, when we have a look at the output of the five techniques and make comparison between them, we will find out that 'canny' technique is very suitable in giving the desired edge detect.

In the other techniques, we can see that the tiny objects can be clearly detected and those surrounded the parasites can get in contact with the out border of the eggs and that thing will increase the errors and change some features of the eggs that the system is relying on the classification stage. While in the 'canny' technique, the edge is very clear and the small details will be ignored in order not to get involved with the major features of the parasites.

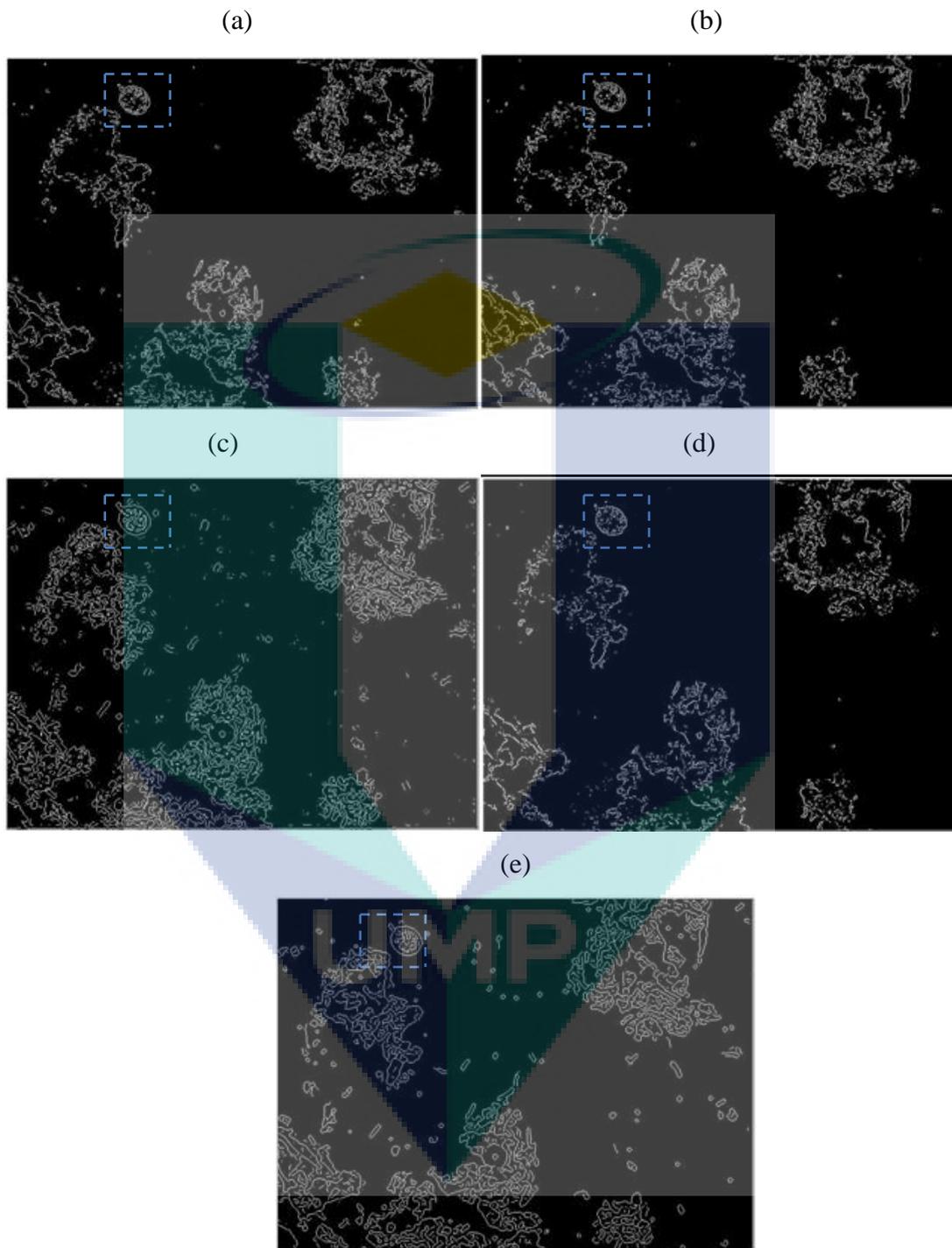


Figure 4.6: Comparison of performance of five edge detection techniques to detect ALO, (a) sobel technique, (b) prewitt technique, (c) log technique, (d) roberts technique, and (e) canny technique.

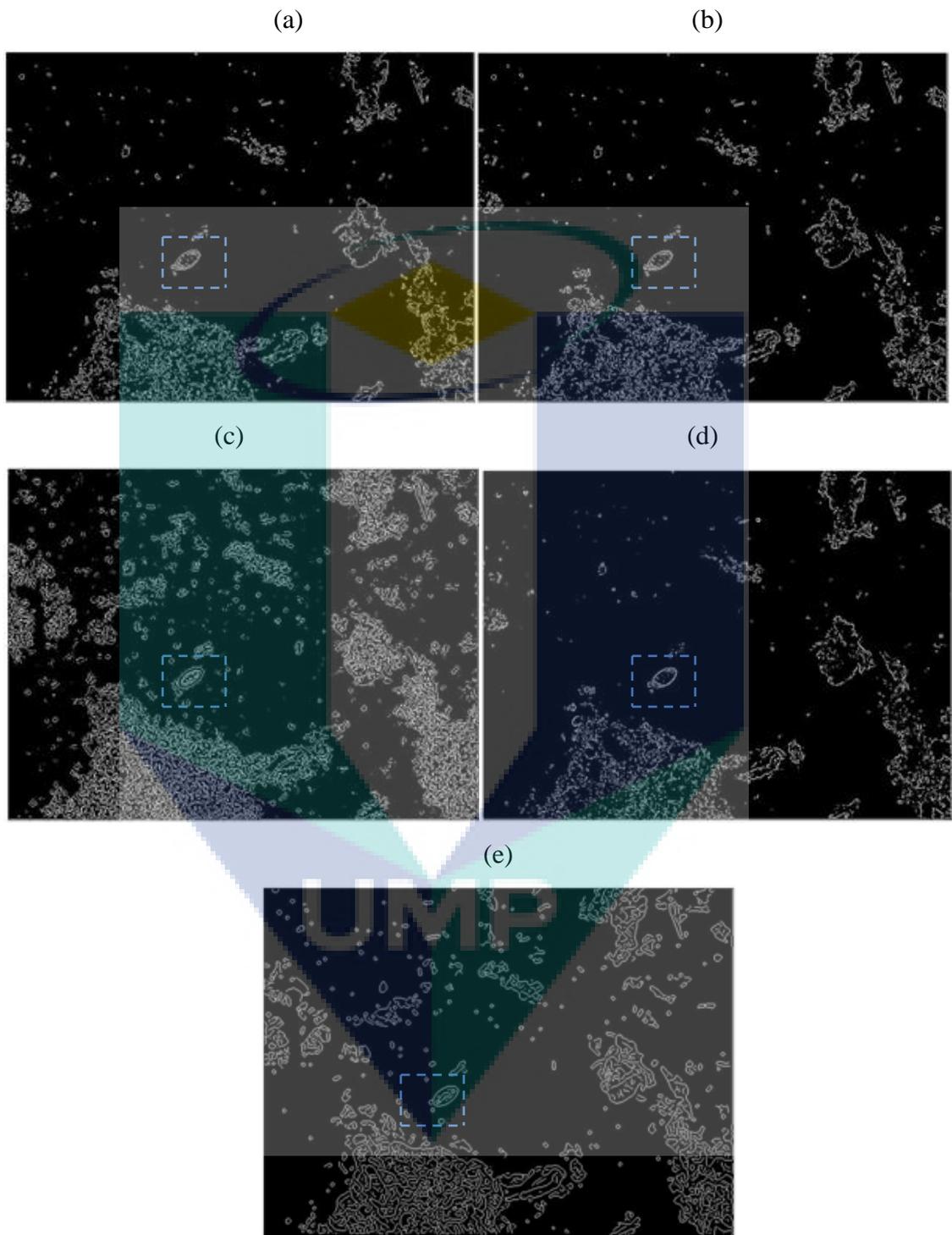


Figure 4.7: Comparison of performance of five edge detection techniques to detect TTO, (a) sobel technique, (b) prewitt technique, (c) log technique, (d) roberts technique, and (e) canny technique

4.4.1 Edge detection threshold evaluation

After the suitable technique for edge detection is selected, now we need to choose the most suitable parameter values of that technique. In the pre-processing stage, “canny” filter is selected as an edge detection technique and run with thresholds and coefficient. In this study, in order to reach the best results to detect the parasites in the input image and among many artifacts, experimentally we have tried various coefficients of canny filter which is SIGMA and it represents the standard deviation of a Gaussian filter that applies to an input image prior to edge detection. The results of applied coefficients and their effect on the images of ALO and TTO detection are shown in Figures 4.8 and 4.9 respectively.

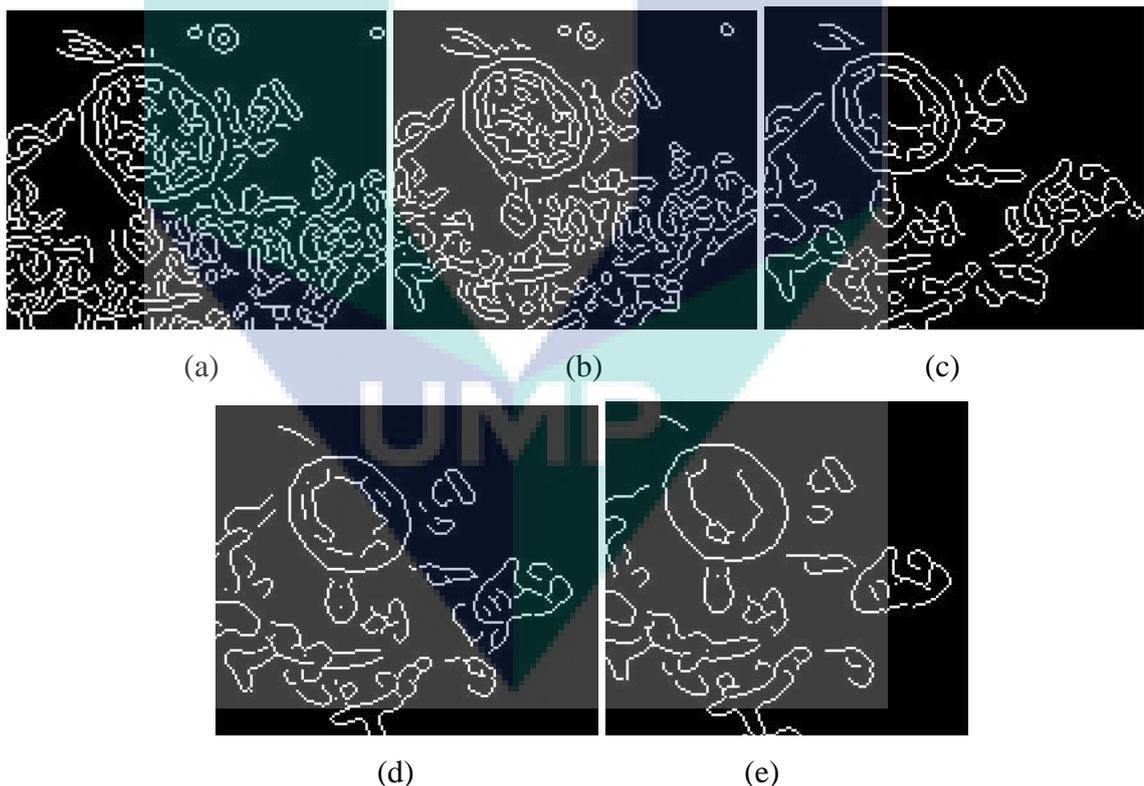


Figure 4.8 : Applying various values of sigma in ‘canny’ edge detection and show the changing in the output of ALO parasite, (a) image with sigma = 0.8, (b) image with sigma = 1, (c) image with sigma = 1.4, (d) image with sigma = 1.8, and (e) image with sigma = 2

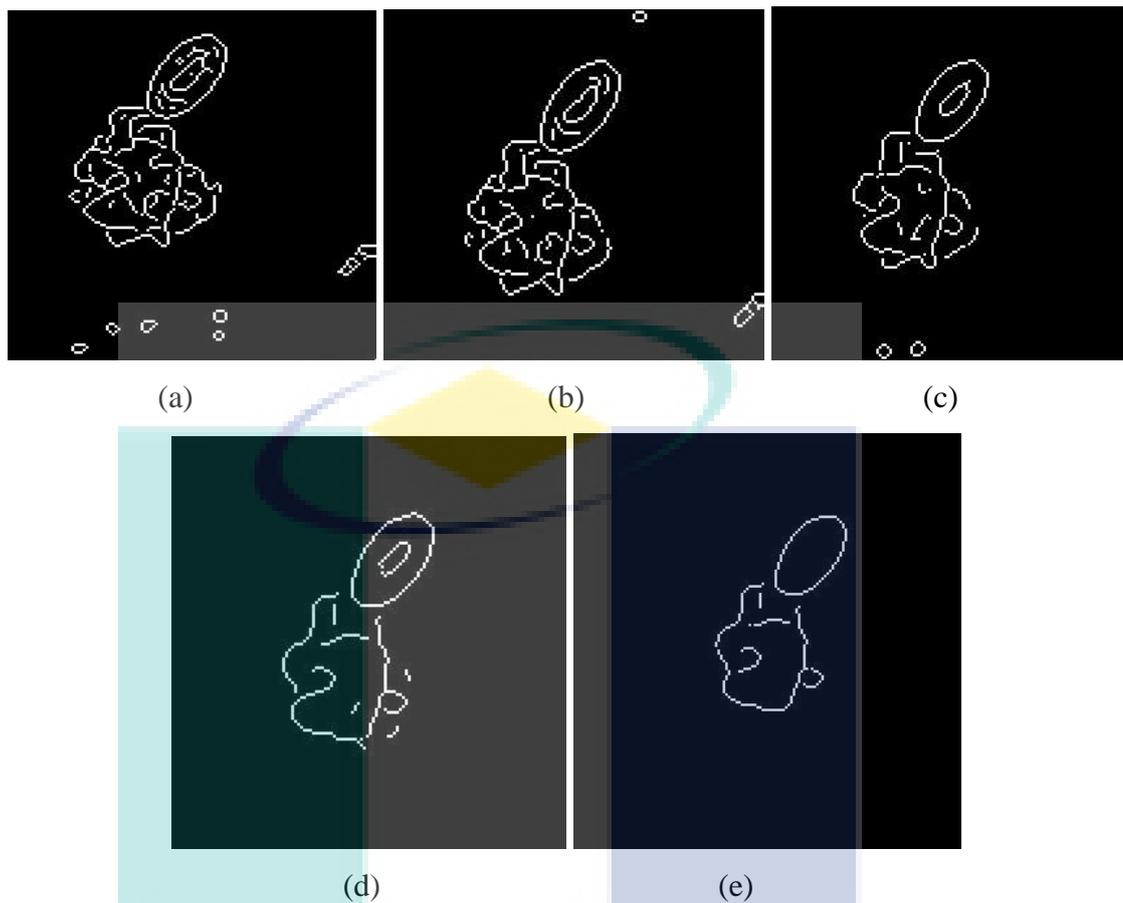


Figure 4.9 : Applying various values of sigma in ‘canny’ edge detection and show the changing in the output of TTO parasite, (a) image with sigma = 0.8, (b) image with sigma = 1, (c) image with sigma = 1.4, (d) image with sigma = 1.8, and (e) image with sigma = 2

Clearly, we can see that by applying different values of sigma we get different output details. Figures 4.8 and 4.9 are examples of performing 100 images and after analyzing these output images, we could reach to a decision that using sigma equal to 1.4 was the best because when we compare the output details of the images we can see that when sigma is below 1.4, the image will have much tiny and small objects involved in the parasites and when apply value above 1.4, we can lose data and features of the parasites, so both cases cause the system to make errors and therefore difficult to detect the images. Maybe using other sigma values will give the same results to most images, To avoid some other images that have a lot of artifacts from getting involved with parasites, we have to consider 1.4 as a coefficient value in ‘canny’ edge detection technique with default thresholds.

4.5 PERFORMANCE EVALUATION OF PRE-METHODS APPROACH

As mentioned in chapter 3 in the pre-processing stage, three different pre-processing methods are tested and each method has various steps of processing. These methods are:

1. Pre-processing method I (Pre-Method I): contains noise reduction using median filter and image enhancement by contrast enhancement, adjustment with gray thresholds and then additional processing steps with thresholds such as clear boundary of the image, fill up the holes of the objects in the image.
2. Pre-processing method II (Pre-Method II): contains noise reduction using median filter, edge detection using 'canny' filter and then additional processing steps with thresholds such as clear boundary of the image, fill up the holes of the objects in the image.
3. Pre-processing III (Pre-Method III): contains image enhancement by contrast enhancement, edge enhancement using unsharp, then noise reduction using median filter (twice) and an edge detection using 'canny' filter, and then additional processing steps with thresholds such as clear boundary of the image, fill up the holes of the objects in the image.

When we applied these methods separately in our system, we have seen the difference in the results in terms of success rate as shown in Figures 4.10 and 4.11 .

The chart in Figure 4.10 shows the comparison in performance of three pre-processing methods for ALO detection. When we use pre-method I, it yields 73% success rate of detecting ALO by processing 100 images of parasite type. This ratio came out with big errors because we did not use any edge enhancement or any filter for edge detection. Pre-method II gives a successful detection ratio which reaches to 81% for ALO, and this result still has errors due to using edge detection technique without using image contrast and edge enhancement. While in pre-method III, the success rate is higher compared to the previous pre-methods in I and II which reaches to 93%. This is due to the use of image

contrast, edge enhancement, median filter (twice), and an edge detection technique which is “canny” filter. With selected parameter values, the results comparison of three methods is shown in Figure 4.10 .

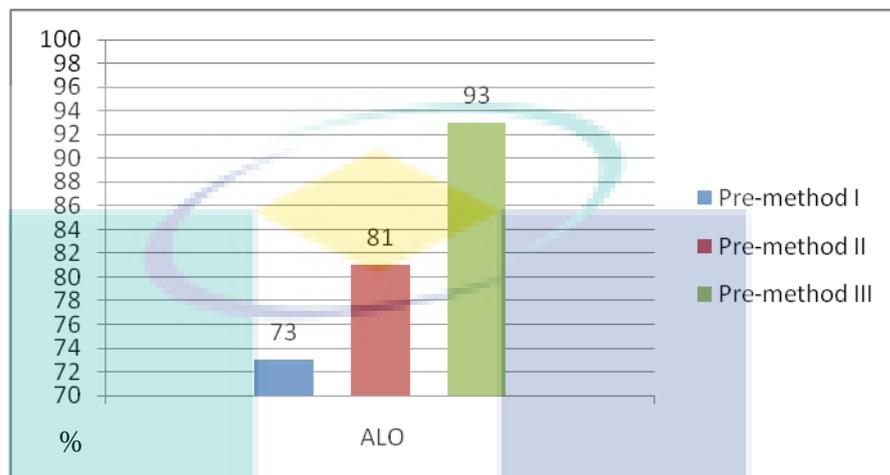


Figure 4.10: Comparison of results by using three pre-methods to detect ALO



Figure 4.11: Comparison of results by using three pre-methods to detect TTO

The chart in Figure 4.11 shows the comparison in performance of three pre-processing methods for TTO detection. When we use pre-method I, the success rate of detecting TTO by training 100 images of parasite type is 75%. This ratio came out with big errors because we did not use any edge enhancement or any filter for edge detection.

Pre-method II gives a successful detection ratio which reaches to 84% for TTO, and this result still has errors due to using edge detection technique without using image contrast and edge enhancement. While in pre-method III, the success rate is higher compared to the previous pre-methods in I and II which reaches to 94% of detecting TTO. This is because the image contrast and edge enhancement with an edge detection technique, which is the “canny” filter, is used. With specific parameter values, the results comparison of three methods is shown in Figure 4.11 .

After testing the three experimental methods in the pre-processing stage, fortunately we reach to the right decision by using pre-method III in our system to detect both ALO and TTO eggs. The main reason that pre-method III is succeeded because this method has a focus on edge detection since the eggs of our parasites have a clear and sharp outside edge.

4.6 EFFECT OF VARIOUS PARAMETERS VALUES ON THE RESULTS

Earlier in this chapter, we analyzed every step of pre-processing stage and discussed the purpose of choosing such technique. Some of these techniques have parameters and values, so in this chapter we will prove in terms of overall results how the parameter values are chosen for good results.

4.6.1 Comparison of contrast enhancement techniques in the results

Basically, three techniques are used to test contrast enhancement. Here we prove in terms of numbers of results of the system by examining 100 images of each parasite from ALO and TTO.

Figure 4.12 shows the comparison of the effectiveness of the above three techniques We can see in this chart that the ‘adjust image within intensity values’ technique, is better than ‘histogram equalization’ and ‘adaptive histogram equalization’. For the first technique, the ratio of success reached to 93% and 94% for ALO and TTO,

respectively. The second technique gives 31% and 15% of success for ALO and TTO. The third technique gives 52% and 43% of success for ALO and TTO.

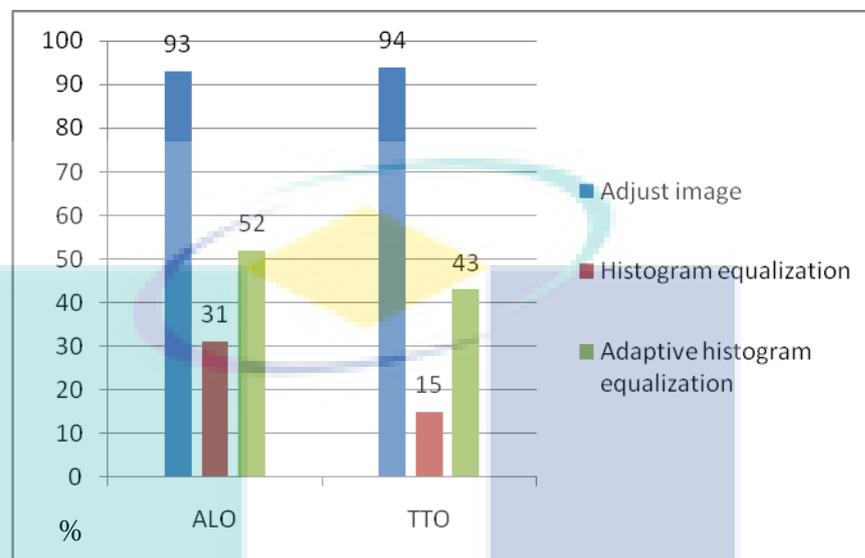


Figure 4.12: Comparison of using three techniques for contrast enhancement to detect ALO and TTO

4.6.2 Contrast coefficient evaluation

To evaluate the edge of parasite, we used various values of gamma to find out how these values affect the threshold of the object's edge, see Figures 4.13 and 4.14.

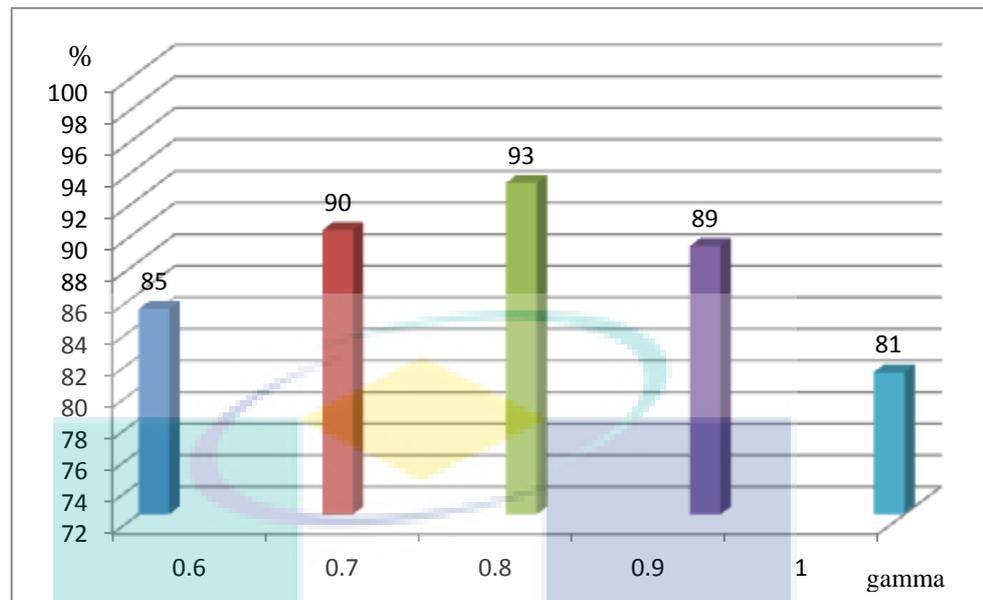


Figure 4.13: Effect of gamma on successful detection ratio for ALO

In Figure 4.13, we can see the effect of applying various values of gamma on the results of detecting ALO parasites from the images. We applied a few values to find out which one is the best to clarify the image and adjust the edge of objects in the image for better detection. We started with a certain value of gamma and then we went slightly up to test the result of the next value with 5 different values of gamma. The first value of gamma is 0.6, the ratio of success is 85%. With gamma equals to 0.7, the ratio is 90%; when gamma equals to 0.8, the ratio reached 93%; when it is 0.9 and 1, the ratios are 89% and 81% respectively. From these results, we can see that when it is 0.8, the ratio is higher than others. Based on that we have considered, we can use this gamma value to detect the parasite since this value could read the threshold of the edge in ALO eggs.

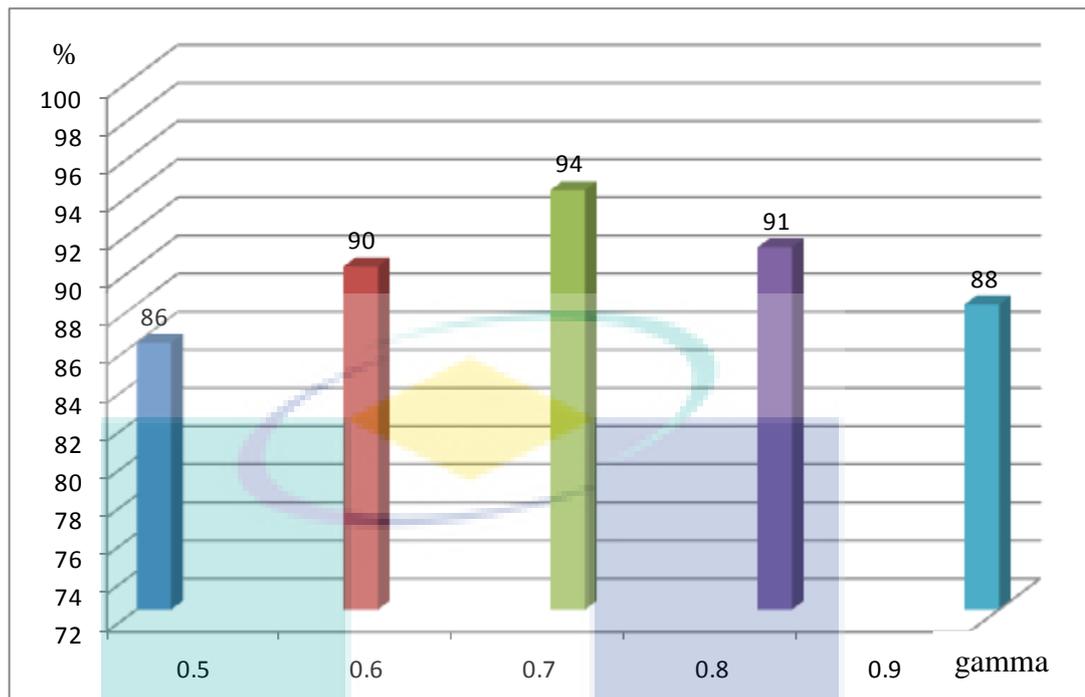


Figure 4.14: Effect of gamma on successful detection ratio for TTO

In Figure 4.14, which is similar to Figure 4.13, we can see the effect of applying various values of gamma on the results of detecting TTO parasites from the images. Here, we applied some values to find out which one is the best to clarify the image and adjust the edge of objects in the images for better detection. When the gamma equals to 0.5, the ratio of success is 86%. While gamma equals to 0.6, the ratio is 90%; when it equals to 0.7, the ratio reaches 94%; the ratios are 91% and 88% respectively when they are in 0.8 and 0.9. . From these results we can see that in 0.8, the ratio of detection is higher than others and based on that we consider, this gamma value in our system can detect the eggs of TTO from the input images.

4.6.3 Edge sharpness evaluation

The perception of sharpness is related to the clarity of detail and edge definition of an image. Sharpness has a large influence on the perceived image quality, and many image capture and display systems offer automated sharpness control, customizable sharpness settings, and adaptive sharpness enhancement as competitive features.

Sharpness metric by itself can also be used as a control variable in sharpness enhancement algorithms in high-quality digital video, or as a quality indicator for situations in which the quality is sufficiently high and other factors remain constant. Combined with other metrics, sharpness can be used to compute an overall quality.

The unsharp option of special 2-D filter performs such filters; the filter created has the form:

$$\frac{1}{\alpha + 1} \begin{bmatrix} -\alpha & \alpha - 1 & -\alpha \\ \alpha - 1 & \alpha + 5 & \alpha - 1 \\ -\alpha & \alpha - 1 & -\alpha \end{bmatrix}$$

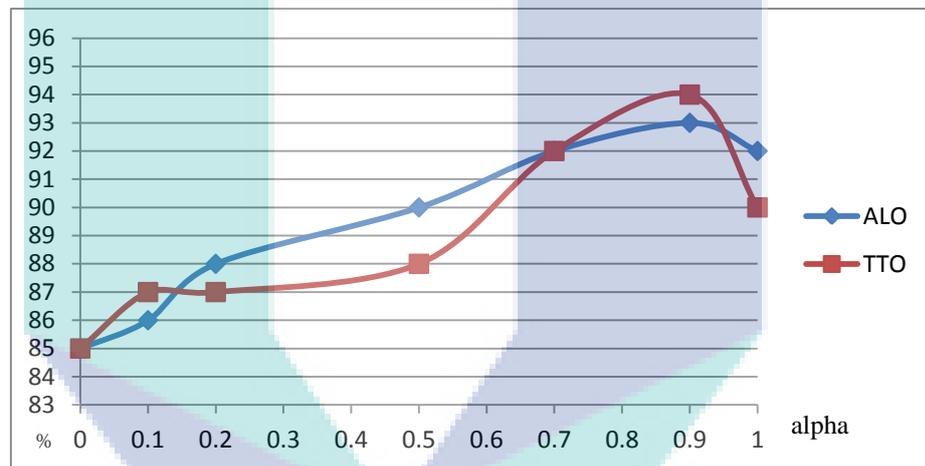


Figure 4.15: Effect of alpha on edge sharpness efficiency for both ALO and TTO results

In Figure 4.15, we can see the curve of ALO goes slightly up when the value of alpha increases from 0.0 to 1.0 and we recorded few points on the curve when alpha equals to 0.0, 0.1, 0.2, 0.5, 0.7, 0.9 and 1.0. The results of success in detecting ALO in terms of percentage are 85%, 86%, 88%, 90%, 92%, 93% and 92%, respectively. Obviously, at value 0.9, it recorded the highest ratio of success which is 93%. This alpha value has been considered in the system. After applying the unsharp masking as mentioned in chapter 3 by using ‘2-D filter’ function with $\alpha = 0.9$, the filter is:

$$\begin{array}{r}
 -0.4737 \quad -0.0526 \quad -0.4737 \\
 2\text{-D filter('unsharp')} = \quad -0.0526 \quad 3.1053 \quad -0.0526 \\
 -0.4737 \quad -0.0526 \quad -0.4737
 \end{array}$$

While the other curve in the same figure, which represents TTO, has shown an unstable truck when we apply the values of alpha from 0.0 – 1.0. The values of alpha that applied are 0.0, 0.1, 0.2, 0.5, 0.7, 0.9 and 1.0. The ratios of success in detecting TTO are 85%, 87%, 87%, 88%, 92%, 94% and 90%, respectively. From these results, we can see the highest rate of success is 94% when alpha equals to 0.9.

4.6.4 Median filtering evaluation

Experimentally, we find that using the median filter twice in a pre-processing stage can reduce the error of detecting ALO parasite. The ratio of error will be reduced to 6% in ALO, which means that this is better than using the median filter once, see Figure 4.16. When we use a median filter one time only, it gives 87% of success rate for ALO. While applying median filter twice, it gives a better result since it reaches 93% for ALO.

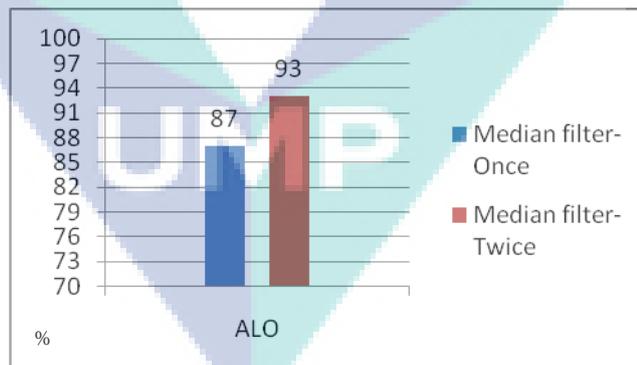


Figure 4.16: Comparison of the results of using median filter (once and twice) for detecting ALO

The median filtering with a mask size 3x3 for TTO detection was proceeded twice. Experimentally, we find this can reduce the error of detecting ALO parasite in the pre-processing stage. The record of error is reduced to 4% in ALO, which means using the

median filter twice is better than it once, see Figure 4.17. When we use a median filter one time only, it gives 90% success rate for TTO; if we apply twice, it reaches 94%, which is better.

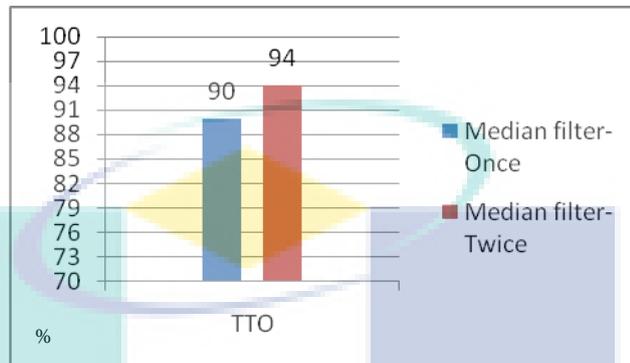


Figure 4.17: Comparison of the results of using median filter (once and twice) for detecting TTO

4.6.5 Comparison of using edge detection techniques

The differences in performance by using five techniques are seen in Figure 4.18: the sobel technique is 42% only for ALO detection; the prewitt technique is 48% ; log technique produces 56%;, the roberts technique reports 20% only. the canny technique has the best percentage and the ratio is 93%. Figure 4.18 illuminates the performance of the five techniques in terms of output image after using these techniques and we can see that in the Canny technique has eliminated many unwanted small details and it focuses on high threshold cells based on parameters' values.

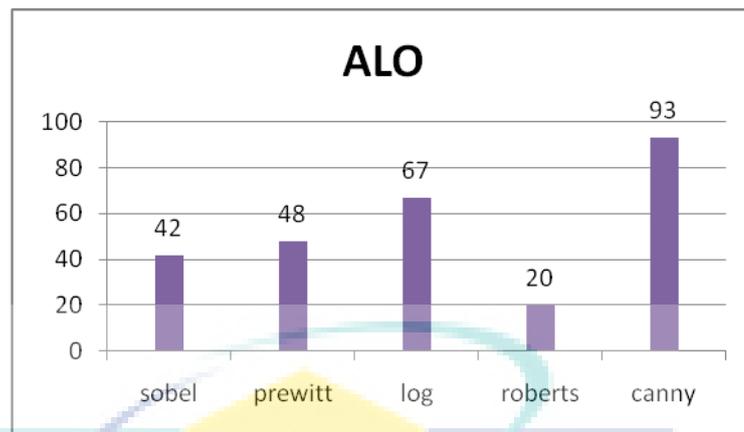


Figure 4.18: Comparison results of five different filters applied in ALO detection

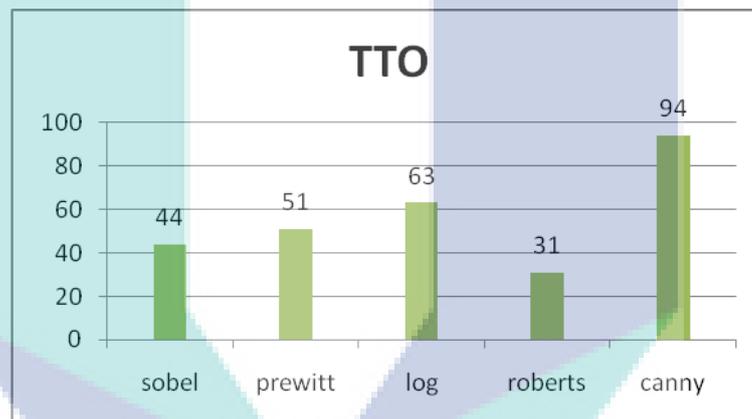


Figure 4.19: Comparison results of five different filters applied in TTO detection

In Figure 4.19, to detect TTO, the five techniques are performed and the ratios of successful detection in sobel, prewitt, log, roberts, and canny techniques are 44%, 51%, 60% 31% and 94%, respectively. So canny filter is the most suitable technique than the other four. Figure 4.11 shows the performance of the five edge detection techniques in terms of output image after using these techniques and we can see in canny technique that it focuses on high threshold cells based on parameters' values and it eliminates the small details with low thresholds inside the images.

From Figures 4.18 and 4.19, the nearest success ratio to canny technique ratio is the 'log' technique. This is because both techniques are required to detect the specific

threshold objects' edges. So when any technique is relying on factors, then that technique will be powerful. The factor's value is very important to reach the main purpose of using that technique.

Various values of the factor 'sigma' are applied to choose the perfect one with good results as we mentioned earlier in this chapter, but here we will show the differences in the results due to using various values of sigma. The results of these applied factors and their effects on the images in ALO detection are shown in Figure 4.20, and the same goes to TTO in Figure 4.21.

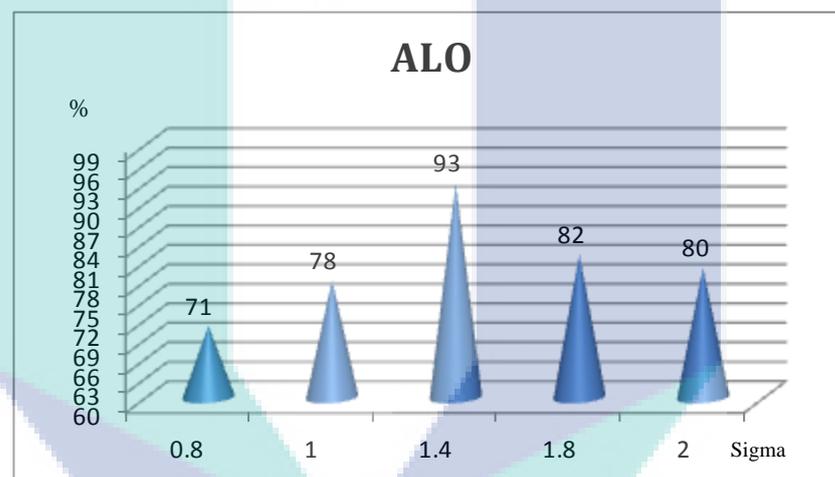


Figure 4.20: Application of various values of sigma to check the best detection ratio in ALO

Figure 4.20 shows that when sigma is equal to 1.4, it gives the best result in a percentage of successful detection for ALO parasite, that is, 93% compared with the results of other values of sigma such as 0.8, 1, 1.8 and 2. The percentages of success are less which are 71%, 78%, 82% and 80%, respectively. So, based on experimental work we started applying 'canny' filter with 0.8 of sigma values slightly up until 2, and we found that the best result with fewer errors is 1.4 of sigma values.

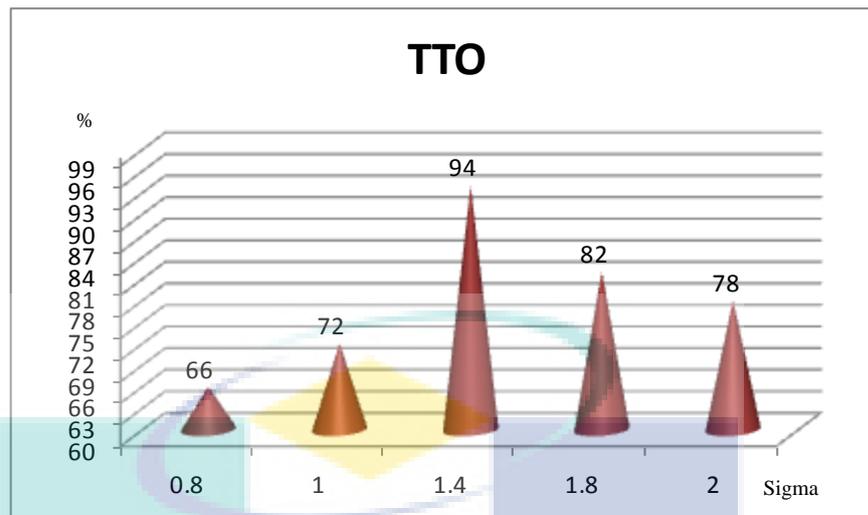


Figure 4.21: Application of various values of Sigma to check the best detection ratio in TTO

In Figure 4.21, we tried to find the best result to detect TTO parasite by applying sigma values slightly up from 0.8 to 2., The chart shows that when sigma values at 0.8, 1, 1.4, 1.8 and 2 are applied, it gives 66%, 72%, 94%, 82% and 78% of success ratio, respectively. That means when sigma equals to 1.4, we can see the best result compared with other values.

4.7 SUMMARY

The goal of this study is to detect eggs of parasitic worms for two kinds of diseases which are *Ascaris Lumbricoides ova (ALO)* and *Trichuris trichiura ova (TTO)*. The results and the success of this study came from logical ideas and experiments. We have experimented several methods and techniques to achieve the objective which is to detect these diseases and we made comparisons between these methods and techniques in terms of efficiency of results. Furthermore, a user friendly environment is provided, which helps to show the results and simultaneously consuming less manpower and time, which reaches 1-3 seconds per image as an average processing time for all 200 images as compared to conventional methods

Based on the results, we could make a decision by choosing the most suitable method and use better techniques to solve the issues that we have faced in the system to get the highest quality results with fewer errors. We can summarize the whole system procedure with their parameters and factor values in one block diagram, as shown in Figures 4.22 and 4.23 for ALO and TTO processing system, respectively.

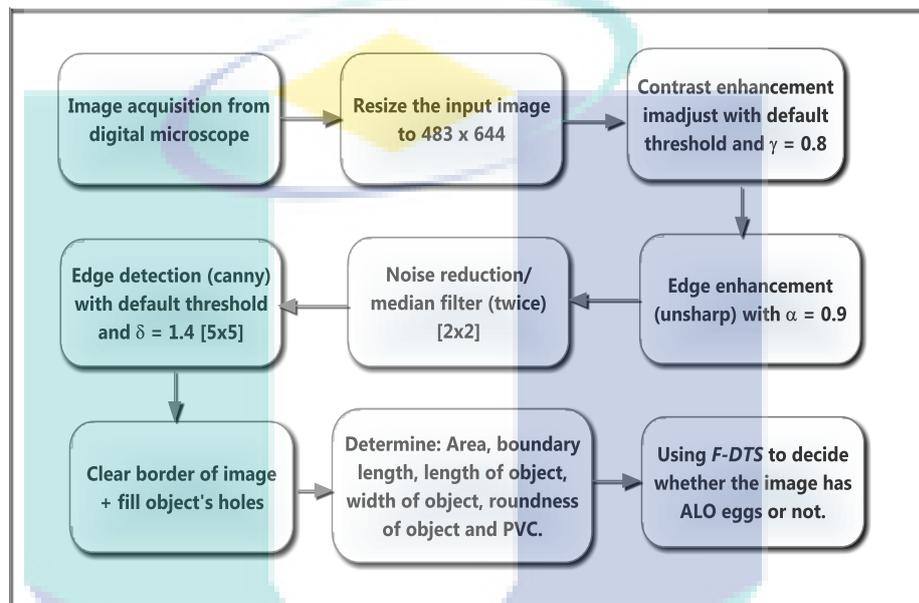


Figure 4.22: Block diagram of ALO processing method

As we can see in Figure 4.22 that the images obtained from digital microscope passed through the image resize from any image size to 483x644. Then, contrast enhancement is the next item to adjust and enhance the images with $\gamma = 0.8$ to get a better vision for the next step, which is edge enhancement that sharpens all object's borders to be sharper and clearer under alpha 0.9, and prepare to remove the noises in the images using median filter twice with a mask of size 2x2.

After removing the noises, edge detection using 'canny' filter is the best choice to detect the object's borders as we mentioned previously in this chapter, compared with other techniques that experimentally we have tested them to get the best edge detection. The value of sigma was very important to detect the parasite edge under default thresholds due not to lose the data of images, and the value of δ is 1.4 by 5x5 mask. Some

segmentation steps are applied preparation to determine the five features plus one (five+1) that are going to be used in *F-DTS* classifier as thresholds to detect the ALO parasites in the images.

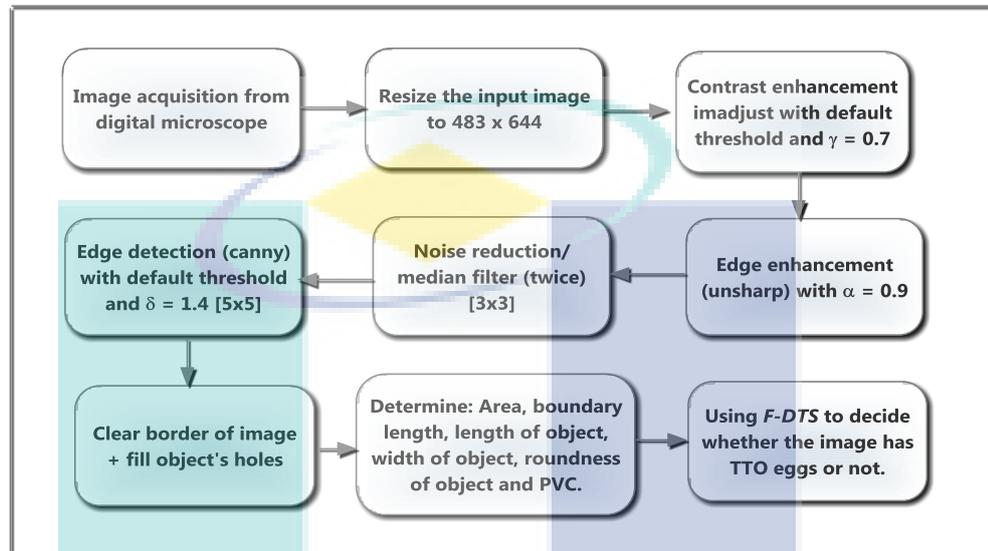


Figure 4.23: Block diagram of TTO processing method

In Figure 4.23, the scenario of detecting TTO is exactly the same as ALO but with different factor values of certain steps. For example, in contrast enhancement $\gamma = 0.7$, and in noise reduction the mask of median filter is 2×2 . While other steps are exactly the same in factors and sequences, like in the edge enhancement, the alpha is 0.9, the same in the ALO system; and in edge detection, the filter type, sigma value and the mask size are the same as in the ALO system.

Indeed, we can summarize the system of detecting ALO and TTO in one flow chart as shown in Figure 4.24.

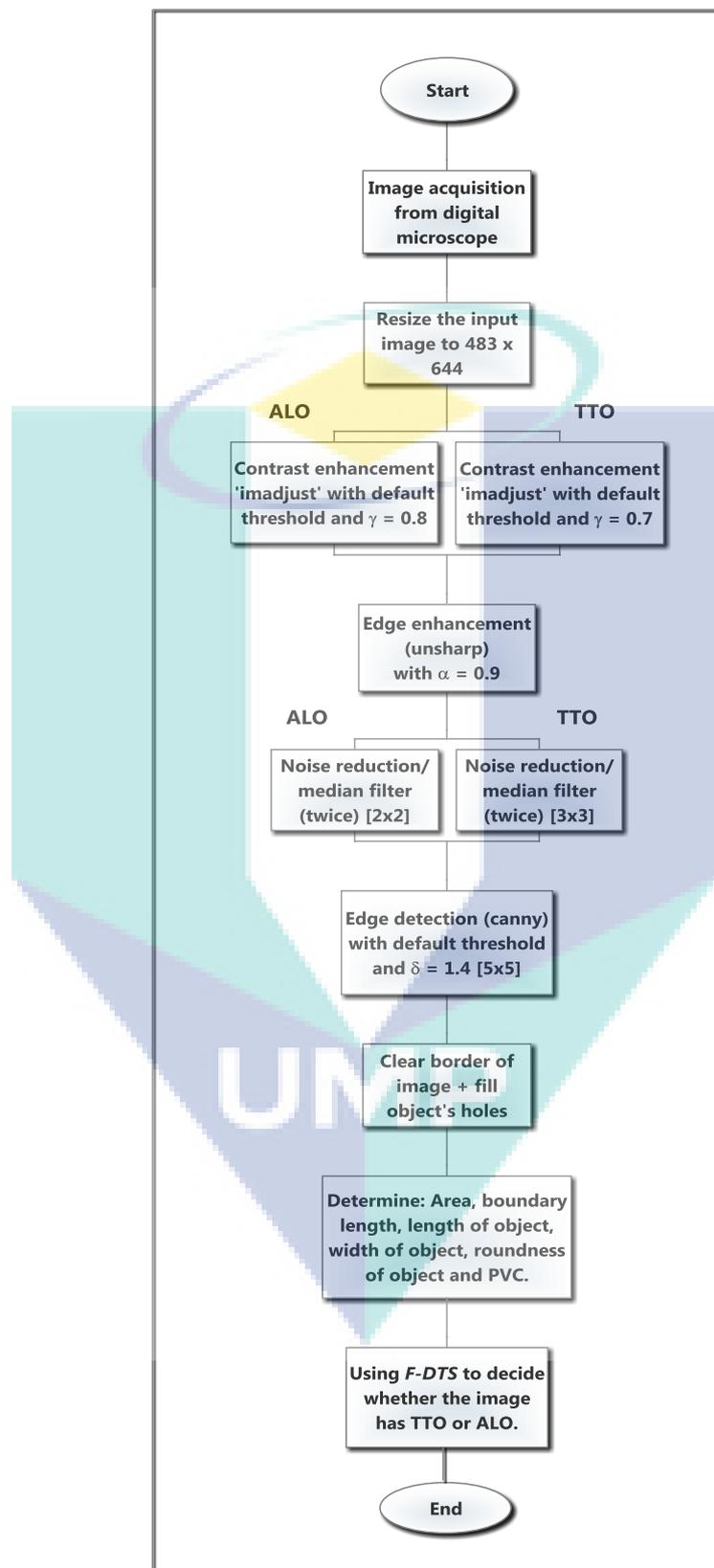


Figure 4.24: Both ALO & TTO detection method in one flow chart.

CHAPTER V

CONCLUSION

This work has achieved the objective of classifying human parasitic worms (ALO and TTO) using microscopic imaging processing technique. After the review and analysis of previous literature and comparison of many related systems, depending on the shape based detection method to identify the parasites from the microscopic images, *F-DTS* classifier was proposed for easy diagnosis with some improvement in terms of enhancement and trying few ideas for a better result.

An image processing technique is proposed to diagnose two types of human parasitic worms, *Ascaris Lumbricoides ova* (ALO) and *Trichuris trichiura ova* (TTO). The algorithm presented in this study is limited to the basic diagnosis of these kinds of parasitic worms. Basically, we have experimented three methods in the preprocessing stage to achieve the best one that gives good results in the end. One of these three methods has simple and strong technique performed in this study. This technique is capable of detecting the existence of ALO and TTO parasites within a few seconds per image; thus, this method can replace the conventional methods of detecting worms in biomedical applications. This software programme can count the number of parasites which have been detected for each single patient. Furthermore, a user friendly environment is provided, which helps to show the results and simultaneously consuming less manpower and time, which reaches 1-3 seconds per image as an average processing time for all 200 images as compared to conventional methods.

The highlighted points of this study are listed as follows:

1. The noise reduction technique used the medial filter twice which in turn can reduce the noise, especially those in touch with parasite eggs.

2. The two kinds of parasites (ALO & TTO) have a very clear and externally thick boundary. The 'canny' filter is a powerful technique to detect the edge since it has parameters.

3. In this study, five parameters are extracted from the parasites which are the main features to detect these parasites automatically plus one additional parameter in order to remove the object that is very close in the features to ALO parasite and that could increase the error, which will confuse the detection system.

4. Filtration and Determination Threshold System *F-DTS* is used in the classification stage. These thresholds are experimentally recorded and they are used for detecting the ALO and TTO in the system.

5.1 FUTURE WORK

This work has focused on improving certain parts of the pre-processing stage and on extracting more features that could help detect the parasite eggs with high accuracy. Using *F-SDTS*, which relies on thresholds, was not a new novel but a limitation of the parameters and works with fix environments gave to this technique specialty to be applied in detecting ALO and TTO parasites.

For future work, one of the suggestions is to try chain technique in edge detection that is quite interesting in drawing the outline shape of the objects and then can separate the objects which are in touch and this is the best way to reduce the errors.

Other work can be done is that by using color as a feature in the detection system. This idea has never been used in detecting parasites in fecal and everyone knows there is an obstruction to use different stain in the effects of microscope illumination, but it can be done when we use certain stain with same illumination environments of the microscope during the capture of the images from the fecal specimen.

REFERENCES

- Acha, P. N., & Szyfres, B. 2003. Trichuriasis of Animal Origin. *Zoonoses and Communicable Diseases Common to Man and Animals* 3rd ed. *Pathogen Regulation Directorate, Public Health Agency of Canada*. 302-304.
- Ajoy, K. R. 2010. Segmentation of Blood Smear Images using Normalized Cuts for Detection of Malarial Parasites. *IEEE Indian Conference*, 1-4.
- Al Badawi A. & Al Sonni, M.S. 2008. Detecting and Analysis of Parasite (Taenia) Using Matlab. *Springer link*, 21: 552-554
- Arce, G.R. 2005. *Nonlinear Signal Processing: A Statistical Approach*, Wiley and Sons.
- Arias-Castro, E. & Donoho, D.L. 2009. Does median filtering truly preserve edges better than linear filtering? *Annals of Statistics*, 37(3):1172.
- Ash L. R. & T. C, 1990. *Orihel, Atlas of Human Parasitology*, 3rd ed. Chicago, IL: ASCP.
- Ash, L. R., & Orihel, T. C. 2003. Intestinal Helminths. In P. R. Murray (Ed.), *Manual of Clinical Microbiology*, 8th Ed., 2031-2060.
- Avcı & Akpolat, 2006. Speech recognition using a wavelet packet adaptive network based fuzzy inference system. *Expert Systems with Applications*, 31(3):495–503.
- Avcı, D. a& Varol, A. 2009. An expert diagnosis system for classification of human parasite eggs based on multi-class SVM. *Expert Systems with Applications*. 36: 43–48.
- Baker GR, Norton P, 2002. Patient safety and healthcare error in the Canadian healthcare system. Ottawa, Canada: Health Canada, 2:1–167.
- Bishop, C. M. 1996. *Neural networks for pattern recognition*. Oxford University.
- Bundy, D. A., & Cooper, E. S. 1989. Trichuris and trichuriasis in humans. *Advances in Parasitology*, 28:107-173.
- Burger, W., & Burge, M.J. 2007. Principles of Digital Image Processing. *Fundamental Techniques*, 164.

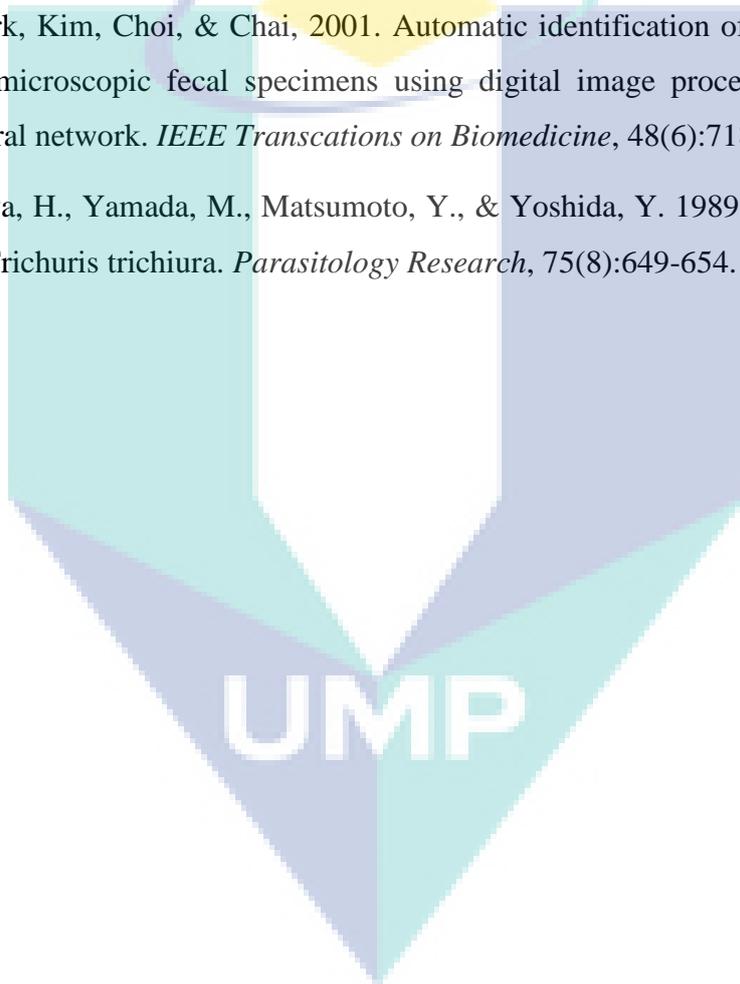
- Calva et al., 2009. Urine and Copro Recognition with Generalized Entropy and Neural Networks. *IJCSNS International Journal of Computer Science and Network Security*, 9(4): 173-179 .
- Calva Méndez D., Landa Quezada A., & Lehman, M.2004. A Neural Network Approach to Fractal Geometry. *WSEAS Int. Conf. on Mathematical Biology and Ecology*. 3(5): 1238- 1241.
- Canny, J. A. 1986. Computational Approach To Edge Detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8 (6): 679–698.
- Cao, C.Z. 2009. Detection of Red Blood Cell in Urine Micrograph. *ICBBE IEEE*. 136-141.
- Chan T.F., & Vese, L.A. 2011. Active Contours Without Edges. *IEEE Trans on Image Process*, 10:266-277.
- Daugochies, Imarom, & Bollwahn, 1999. Differentiation of porcine imeriaspp. By morphologicalgorithms. *Veterinary Parasitology*, 81:201–210.
- Deriche, R. 1987. Using Canny's criteria to derive a recursively implemented optimal edge detector. *Int. J. Computer Vision*, 1:167–187.
- Doerr, W. & Seifert, G. 1995. *Tropical Pathology*, 2nd edn. Springer, Berlin Heidelberg New York, 871-875.
- Dogantekin, E. & Yilmaz, M. 2008. A robust technique based on invariant moments – ANFIS for recognition of human parasite eggs in microscopic images. *Expert Systems with Applications: An International Journa*, 35(3): 1367-1378.
- Falconer, K. 1997. *Techniques in Fractal Geometry*, John and Wiley Sons, Chichester.
- Forsyth, D.A. & Ponce J., 1997. *Computer Vision A Modern Approach*. Pearson Education.
- Freeman, J. A. & Skapura, D. M. 1992. *Neural Networks: Algorithms, Applications, and Programming Techniques*. New York: Addison-Wesley.
- Ghate, D.A. & Prof. Jadhav, C., 2012. *Automatic Detection of Malaria Parasite from Blood Images*, Department of Computer, D. Y. Patil, College of Engineering, Pimpri Pune, Maharashtra India, TIJCSA.

- Gonzalez, R. C. 1993. *Digital Image Processing*. New York: Addison- Wesley.
- Gonzalez, R.C. 2002. *Digital Image Processing Using MATLAB*. Second Edition, University of Tennessee.
- Gonzalez, R.C., Woods, R.E., & Eddins, S.L. 2004. *Digital Image Processing using MATLAB*, Upper Saddle River, Prentice Hal.
- Gorbach, S.L., Bartlett, J.G., & Blacklow, N.R. 1992. *Infectious Diseases*. Saunders, Philadelphia. 1646-1654.
- Halim, S. et al., 2006. Estimating Malaria Parasitaemia from Blood Smear Images. *ICARCV IEEE*. 1-6.
- Holland, C.V.2006. *Gastrointestinal nematodes - Ascaris, hookworm, Trichuris, and Enterobius*. Wiley-Blackwell. 713-736.
- Hu, M.K. 1962. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2): 179- 187.
- Ismail, M.M. & Jayakody, R. L. 1999. Efficacy of albendazole and its combinations with ivermectin or diethylcarbamazine (DEC) in the treatment of *Trichuris trichiura* infections in Sri Lanka. *Annals of Tropical Medicine and Parasitology*, 93(5):501-504.
- Jackway, P.T, 2000. Improved morphological tophat. *Electronics Letters*, 36(14):1194-1195.
- Jampana, P.V. 2010. *Computer Vision based Sensors for Chemical Process*.
- Jiraamonninit. C., & Wongkamchai, S. 2006. The Prevalence of Intestinal Parasitic Infections among Schoolchildren with Annual Anthelmintic Treatment in Narathiwat Province, Thailand. *J Trop Med Parasitol*. 29:45-50.
- Joachim, Dulmer, and Dauschies, 1999. Differentiation of two *Oesophagostomum* spp. from pigs, *O. dentatum* and *O. quadrispinulatum*, by computer-assisted image analysis of fourthstage larvae. *Parasitology International*, 48: 63–71.
- Jones J.E. 1981. The royal roundworm: *Ascaris Lumbricoides*. *J fam Pract*; 13:271.
- Jones J.E. 1983. Parasites in Kentucky: the past seven decades. *J ky Med Assoc* 81: 621.

- Kareem S., Kale I., & Morling R.C.S, 2012. Automated P.falciparum Detection System for Post- treatment Malaria Diagnosis Using Modified Annular Ring Ratio Method. *14th International Conference on Computer Modelling and Simulation IEEE*, 430-436.
- Kareem, S., Morling, R.C.S, & Kale, I. 2012. A Novel Method to Count the Red Blood Cells in Thin Blood Films. *IEEE International Symposium on Circuits and Systems (ISCAS)*, 1021-1024.
- Khuroo, M.S. 1996. Ascariasis. *Gastroenterol Clin North Am.* 25:553.
- Krumhardt, B., Ph.D., 2008. Science Group Leader, Biology Instructor, Urban Campus, Des Moines, Iowa, USA 50314.
- Lewis, R. 1990. *Practical Digital Image Processing*. Ellis Horwood Limited, Simon and Schuster International Group.
- Mahmood, N.H. & Mansor, M.A., 2012. Red blood cells estimation using Hough Transform Technique. *SIPIJ*, 3(2): 53-64.
- Mandelbrot, B.B. 1982. *The Fractal Geometry of Nature*. Freeman, San Francisco.
- Mandell, G.L, Douglas, R.G. Jr. & Bennett. 1990. *Principles and Practice of Infectious Diseases*, 3rd edn. Churchill Livingstone, New York.
- Markell, Edward K., John, David T., Krotoski, & Wojciech A. 1999. *Markell and Voge's Medical Parasitology*. Eight Ed. W.B. Saunders Co. New York.
- Ng, P.E. & Ma, K.K. 2006. A Switching Median Filter with BDND for Extremely Corrupted Images. *IEEE Trans Image Processing*. 15(6): 1506-1516.
- Pratt, W.K., Wiley, J. & Sons. 2001. *Digital Image Processing: PIKS Inside*. Third Edition, 288. John Wiley and Sons.
- Purwar, Y., Sirish L., & Shah, S.L. 2011. Automated and unsupervised detection of malarial parasites in microscopic images. *Purwar et al. Malaria Journal*. 10:364
- Reeder, M.M. 1988. The radiological and ultrasound evaluation of ascariasis of the gastrointestinal, biliary, and respiratory tracts. *Semin Roentgenol*, 33:57.

- Richards, J.A. 1986. *Remote Sensing Digital Image Analysis: An introduction*, Springer-Verlag.
- Sarinas, P.S., & Chitkara, R.K. 1977. Ascariasis and hookworm. *Semin Respir Infect.* 12:130.
- Sengur A., & Turkoglu I., 2004. Parasite egg cell classification using invariant moments. *Proceedings of 4th international symposium on intelligent manufacturing systems.* 98–106.
- Soille P. 1999. *Morphological Image Analysis: Principles and Applications*, 164-165. Springer.
- Soille, P. 1999. *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, 173-174.
- Sommer. 1996. Digital image analysis and identification of eggs from bovine parasitic nematodes. *Journal of Helminthology*, 70: 143–151.
- Soni, J. & Mishra, N. 2011. Automatic differentiation between RBC and Malarial parasites based on Morphology with first order features using image processing. *International Journal of Advances in Engineering & Technology IJAET*, 1(5): 290-297.
- Stephenson, L. S., Holland, C. V., & Cooper E. S. 2000. The public health significance of *Trichuris trichiura*. *Parasitology*, 121: 73-95.
- Suwalka, I. & Sanadhya, A. 2012. Identify Malaria Parasite Using Pattern Recognition Technique. *International Conference on Computing, Communication and Applications (ICCCA)*, 1-4.
- Tietze, P.E., & Tietze, P.H. 1991. The roundworm, *Ascaris Lumbricoides*. *Prim Care*, 18-25.
- Vishnu, V., & Makkapati, 2009. Segmentation of Malaria Parasites in Periphreal Blood Smear Images. *ICASSP IEEE*. 1361-1364.
- Warren, K.S., & Mahmoud, A.A. 1977. Algorithms in diagnosis and management of exotic diseases. Xxii. Ascariasis and tococariasis. *J Infect*, 135- 868.

- Wolfe, M. S. 1978. Oxyuris, trichostrongylus and trichuris. *Clinics in Gastroenterology*, 7(1): 201-217.
- Woods, G.L., & Walker, D.H. 2009. Detection of infection or infectious agents by the use of cytologic and histologic stains. *Clin Microbiol Rev.* 9(3):382-404.
- Yang R., Lin L., Gabbouj M., Astola J., & Neuvo Y. 1995. Optimal Weighted Median Filters Under Structural Constraints. *IEEE Trans. Signal Processing.* 43:591-604.
- Yang, Park, Kim, Choi, & Chai, 2001. Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network. *IEEE Transactions on Biomedicine*, 48(6):718–730.
- Yoshikawa, H., Yamada, M., Matsumoto, Y., & Yoshida, Y. 1989. Variations in egg size of *Trichuris trichiura*. *Parasitology Research*, 75(8):649-654.



Appendix A

Programme Codes

```

tic; % Start timer.
clear all
clc
    ims = 16;
image_num = 16;
RL = 0;
OR = 0;
imr = image_num - ims + 1;
text='D:\ParaImage\ALO\';
NM = 0;
NM1 = 0;
r = 0;
for im= ims:image_num

    filename=strcat(text,int2str(im));
    filename=strcat(filename, '.jpg');
    imageaa=imread(filename);

    RGB = imresize(imageaa, [483 644]);
    II = rgb2gray(RGB);

    V= imadjust(II, [], [], [0.8]);

    Hf = fspecial('unsharp', 0.9);
    sh1 = imfilter(V,Hf, 'replicate');

    s = medfilt2(sh1, [2 2]);
    K = medfilt2(s, [2 2]);

    [BW,thresh] = edge(K, 'canny', [], 1.4);

    I2=BW(:,:)-1;
    I3= I2* -1;
    no = imclearborder(I3, 4);
    wb = imfill(no, 'holes');
    bw1 = bwareaopen(wb, 850);
    [A,L] = bwboundaries(bw1, 'noholes');

    cc = bwconncomp(L);
    st = regionprops(cc, 'Area', 'Centroid');
    M = ismember(L, find([st.Area]< 2000));
    [AM,LM] = bwboundaries(M, 'noholes');
    D = cellfun('length',AM);

    % boundary which is between 101 and 160...
    S = ismember(LM, find(D >= 101 & D <= 160));
    [F,V1] = bwboundaries(S);
    H = cellfun('length',F);

```

```

% Showing up dimensions (length and width) of the object.
if H ~= 0
for i = 1: length(F)
R = bwlabel(V1);
cc = bwconncomp(R);
stats = regionprops(cc, 'All');
L1 = stats(i).MajorAxisLength;
W1 = stats(i).MinorAxisLength;
A1 = stats(i).Area;
end
III1 = ismember(V1, find([stats(:).MajorAxisLength] > 41.0000));
[C2,E2] = bwboundaries(III1);
III1= ismember(E2, find([stats(:).MajorAxisLength] < 61.0000));
[C0,E0] = bwboundaries(III1);
II2 = ismember(E0, find([stats(:).MinorAxisLength] > 32.0000));
[C,E] = bwboundaries(II2);

c3 = bwconncomp(E);
stat3 = regionprops(c3, 'All');

III2 = ismember(E, find([stats(:).MinorAxisLength] < 46.0000));

[T,P] = bwboundaries(III2);
BB = cellfun('length',T);
if BB ~= 0

c2 = bwconncomp(P);
stat0 = regionprops(c2, 'All');

% loop over the boundaries
for k = 1:length(T)

% obtain (X,Y) boundary coordinates corresponding to label 'k'
boundary = T{k};

% computes a simple estimate of the object's perimeter
delta_sq = diff(boundary).^2;
perimeter = sum(sqrt(sum(delta_sq,2)));

% obtain the area calculation corresponding to label 'k'
area = stat0(k).Area;

% computes the roundness metric
metric(k) = 4*pi*area/perimeter^2;
end

G= ismember(P, find(metric(:) >= 0.7400));
% figure,imshow(G), title('G')
U= ismember(G, find(metric(:) < 0.9800));

[T12,P12] = bwboundaries(U);

c2 = bwconncomp(P12);
stat010 = regionprops(c2, 'All');

```

```

if metric(k) < 0.8000
    im;
    metric(k);
    RL = RL+1;
elseif metric(k) >= 0.8000
    OR = OR+ 1;
end

```

```

H3 = cellfun('length',T12);
if H3 ~= 0
c11 = bwconncomp(P12);
sta1 = regionprops(c11, 'All');

```

```

for k = 1: length(T12)
cente = sta1(k).Centroid;
Peri = sta1(k).Perimeter;
Plist = sta1(k).PixelList;
ww1 = sta1(k).BoundingBox;
a = ww1(1);
b = ww1(2);
c = ww1(3);
d = ww1(4);
i = cente(1);
j = cente(2);
UX = impixel(IMG,i,j);
mme3(k) = mean(UX);
end

```

```

ZX = ismember(P12, find(mme3(:) < 200));
ZZ = ismember(ZX, find(mme3(:) > 15));
[Te,Pe] = bwboundaries(ZZ);
Hy = cellfun('length',Te);
if Hy ~= 0
WE = 1;

```

```

else WE = 0;
    [Ty,Pe] = bwboundaries(ZZ);
end
else WE = 0;
    [Ty,Pe] = bwboundaries(P12);
end
else WE = 0;
    [Ty,Pe] = bwboundaries(P);
end
else WE = 0;
    [Ty,Pe] = bwboundaries(V1);
end

```

```

if WE == 1;
NM = NM + length(Te);
else NM = NM + 0;
end

```

```

%-----TTO-----
---
```

```

V9= imadjust(II,[0.01 .9],[],[0.7]);

Hf = fspecial('unsharp',0.9);
sh19 = imfilter(V9,Hf,'replicate');

s9 = medfilt2(sh19,[3 3]);
K9 = medfilt2(s9,[3 3]);

[BW,thresh] = edge(K9,'canny',[],1.4);

I2=BW(:,:)-1;
I3= I2* -1;
no = imclearborder(I3, 4);
wb = imfill(no,'holes');
bw1 = bwareaopen(wb,380);
[A,L] = bwboundaries(bw1,'noholes');

c9 = bwconncomp(L);
sta9 = regionprops(c9,'All');
for k = 1:length(A)

    % obtain (X,Y) boundary coordinates corresponding to label 'k'
    boundary = A{k};

    % compute a simple estimate of the object's perimeter
    delta_sq = diff(boundary).^2;
    perimeter = sum(sqrt(sum(delta_sq,2)));

    % obtain the area calculation corresponding to label 'k'
    area = sta9(k).Area;

    % compute the roundness metric
    met(k) = 4*pi*area/perimeter^2;
end

cc = bwconncomp(L);
st = regionprops(cc,'Area','Centroid');
M = ismember(L, find([st.Area]< 750));

BWn = imclearborder(M, 4);
[AM,LM] = bwboundaries(BWn,'noholes');
cc = bwconncomp(LM);
st = regionprops(cc,'All');
for k =1:length(AM)
Ar = st(k).Area;
L = st(k).MajorAxisLength;
W = st(k).MinorAxisLength;
end
D = cellfun('length',AM);

% boundary which is between 101 and 160...
S = ismember(LM, find(D >= 69 & D <= 93));
[F,V1] = bwboundaries(S);
H = cellfun('length',F);
% Showing up dimensions (length and width) of the object.

```

```

if H ~= 0
for i = 1: length(F)
R = bwlabel(V1);
cc = bwconncomp(R);
stats = regionprops(cc, 'All');
L1 = stats(i).MajorAxisLength;
W1 = stats(i).MinorAxisLength;
A1 = stats(i).Area;
end
    III1 = ismember(V1, find([stats(:).MajorAxisLength] > 34.0000));
    [C2,E2] = bwboundaries(III1);
    III1= ismember(E2, find([stats(:).MajorAxisLength] < 44.0000));
% imshow(III1), title('2')
    II2 = ismember(E0, find([stats(:).MinorAxisLength] > 15.0000));
    [C,E] = bwboundaries(II2);

c3 = bwconncomp(E);
stat3 = regionprops(c3, 'All');

    III2 = ismember(E, find([stats(:).MinorAxisLength] < 22.0000));

[T,P] = bwboundaries(III2);
BB = cellfun('length', T);
if BB ~= 0

c2 = bwconncomp(P);
stat0 = regionprops(c2, 'All');

% loop over the boundaries
for k = 1:length(T)

    % obtain (X,Y) boundary coordinates corresponding to label 'k'
    boundary = T{k};

    % compute a simple estimate of the object's perimeter
    delta_sq = diff(boundary).^2;
    perimeter = sum(sqrt(sum(delta_sq, 2)));

    % obtain the area calculation corresponding to label 'k'
    area = stat0(k).Area;

    % compute the roundness metric
    metric(k) = 4*pi*area/perimeter^2;
end

G= ismember(P, find(metric(:) >= 0.7300));
U= ismember(G, find(metric(:) < 0.9400));

[T12,P12] = bwboundaries(U);

c2 = bwconncomp(P12);
stat010 = regionprops(c2, 'All');

for k = 1:length(T12)

    % obtain (X,Y) boundary coordinates corresponding to label 'k'

```

```

boundary = T12{k};

% compute a simple estimate of the object's perimeter
delta_sq = diff(boundary).^2;
perimeter = sum(sqrt(sum(delta_sq,2)));

% obtain the area calculation corresponding to label 'k'
area = stat010(k).Area;

% compute the roundness metric
met(k) = 4*pi*area/perimeter^2;
end

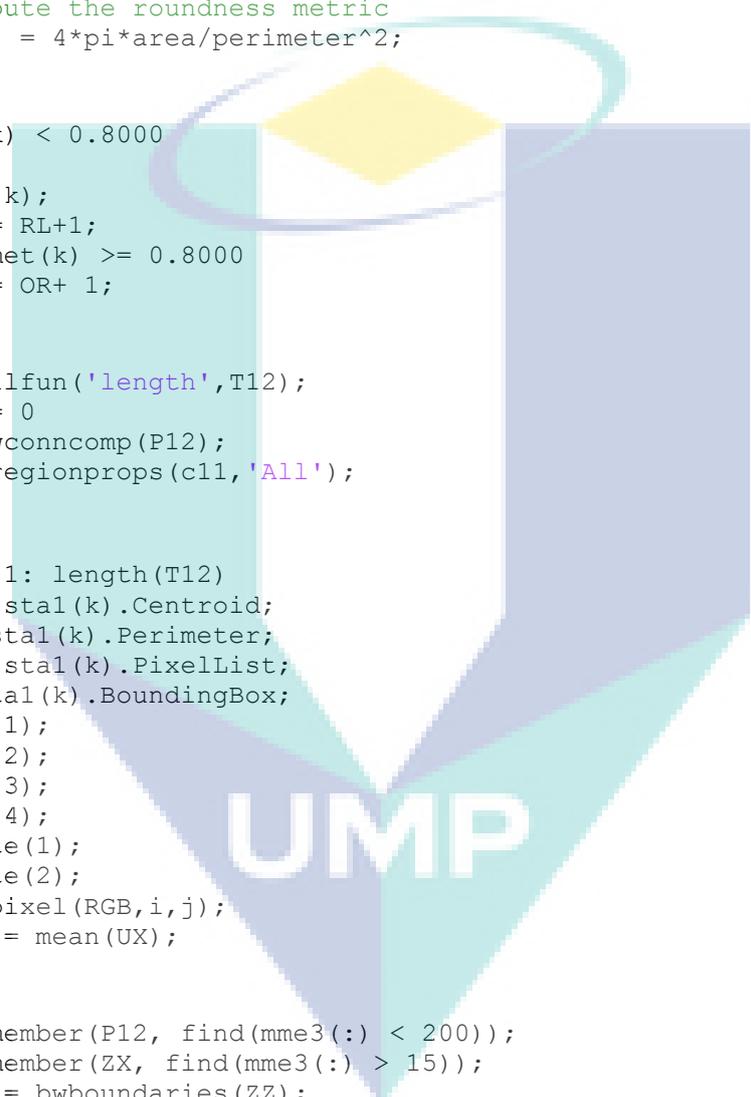
if met(k) < 0.8000
    im;
    met(k);
    RL = RL+1;
elseif met(k) >= 0.8000
    OR = OR+ 1;
end

H3 = cellfun('length',T12);
if H3 ~= 0
c11 = bwconncomp(P12);
sta1 = regionprops(c11, 'All');

for k = 1: length(T12)
cente = sta1(k).Centroid;
Peri = sta1(k).Perimeter;
Plist = sta1(k).PixelList;
ww1 = sta1(k).BoundingBox;
a = ww1(1);
b = ww1(2);
c = ww1(3);
d = ww1(4);
i = cente(1);
j = cente(2);
UX = impixel( RGB,i,j);
mme3(k) = mean(UX);
end

ZX = ismember(P12, find(mme3(:) < 200));
ZZ = ismember(ZX, find(mme3(:) > 15));
[Ty,Py] = bwboundaries(ZZ);
Hy = cellfun('length',Ty);
    if Hy ~= 0
        [aa ss] = size(Ty);
WE = 1;
    if aa >= 2
        im;
        aa;
        hi = hi+1;
    else xix = 1;
    end
else WE = 0;
        [Ty,Py] = bwboundaries(ZZ);

```



```

end
else WE = 0;
    [Ty,Py] = bwboundaries(P12);
end
else WE = 0;
    [Ty,Py] = bwboundaries(P);
end
else WE = 0;
    [Ty,Py] = bwboundaries(V1);
end

if WE == 1;
NM1 = NM1 + length(Ty);
else NM1 = NM1 + 0;
end
ALL = uint8(Pe) - uint8(zeros(size(V)));
% for Bb
ALLL = uint8(Py) - uint8(zeros(size(V1)));
LLL = ALL + ALLL;

[LT,LP] = bwboundaries(LLL);

%-----

elapsedTime = toc;
res = elapsedTime/ 60;
so = rem(elapsedTime,60);
message = sprintf('Got (%.0f) ALO Ova, and\n\n (%.0f) TTO Ova in this
image within %.0f min: %.0f sec.', NM, NM1, res, so);
reply = questdlg(message, 'Results!', 'Ok', 'Ok');
end

```

APPENDIX B

List of Publication and Awards

PUBLICATION

- Certificate of participation at MCEECT 2012 Conference in University Malaysia Pahang/ Malaysia.
- Certificate of participation in IEEE Symposium on Computer Application and Industrial Electronics 2012-Conference in Kota Kinabalu, Sabah, Malaysia. And published in ISCAIE2012/ pp. 142-147.
- Journal publication in American Journal of Scientific Research (AJSR), index: ISI - published, Issue 87.
- Journal publication in Modern Applied Science (MAS), Index: Scopus – published, Vol. 7, No. 5.

AWARDS

- Silver Medal Award from Citrex exhibition 2012 in University Malaysia Pahang/ Malaysia.
- Bronze Medal Award from Bio-Malaysia exhibition 2012 in KLCC, Kuala Lumpur, Malaysia.