

WestminsterResearch

<http://www.westminster.ac.uk/research/westminsterresearch>

Automated Low-Cost Malaria Detection System in Thin Blood Slide Images Using Mobile Phones

Saumya Kareem Reni

Faculty of Science and Technology

This is an electronic version of a PhD thesis awarded by the University of Westminster. © The Author, 2014.

This is an exact reproduction of the paper copy held by the University of Westminster library.

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail repository@westminster.ac.uk

Automated Low-Cost Malaria Detection System in Thin Blood Slide Images Using Mobile Phones



University of Westminster

Saumya Kareem Reni

A thesis submitted in partial fulfilment of the requirements
of the University of Westminster for the degree of Doctor of
Philosophy

March 2014

ABSTRACT

Malaria, a deadly disease which according to the World Health Organisation (WHO) is responsible for the fatal illness in 200 million people around the world in 2010, is diagnosed using peripheral blood examination. The work undertaken in this research programme aims to develop an automated malaria parasite-detection system, using microscopic-image processing, that can be incorporated onto mobile phones. In this research study, the main objective is to achieve the performance equal to or better than the manual microscopy, which is the gold standard in malaria diagnosis, in order to produce a reliable automated diagnostic platform without expert intervention, for the effective treatment and eradication of the deadly disease.

The work contributed to the field of mathematical morphology by proposing a novel method called the Annular Ring Ratio transform for blood component identification. It has also proposed an automated White Blood Cell and Red Blood Cell differentiation algorithm, which when combined with ARR transform method, has wide applications not only for malaria diagnosis but also for many blood related analysis involving microscopic examination.

The research has undertaken investigations on infected cell identification which aids in the calculation of parasitemia, the measure of infection. In addition, an automated diagnostic tool to detect the sexual stage (gametocytes) of the species *P.falciparum* for post-treatment malaria diagnosis was developed. Furthermore, a parallel investigation was carried out on automated malaria diagnosis on fluorescent thin blood films and a WBC and infected cell differentiation algorithm was proposed.

Finally, a mobile phone application based on the morphological image processing algorithms proposed in this thesis was developed. A complete malaria diagnostic unit using the mobile phones attached to a portable microscope was set up which has enormous potential not only for malaria diagnosis but also for the blood parasitological field where advancement in medical diagnostics using cellular smart phone technology is widely acknowledged.

Contents

ABSTRACT.....	ii
Contents	iii
LIST OF ABBREVIATIONS	vi
LIST OF SYMBOLS	vii
LIST OF FIGURES	viii
ASSOCIATED PUBLICATIONS	xvi
ACKNOWLEDGEMENTS	xvii
Chapter 1 Introduction	1
1.1 Malaria	2
1.2 Pathology.....	2
1.3 Morphology of <i>Plasmodium</i> Species.....	4
1.4 Diagnostic Techniques.....	12
1.5 Research Aims, Methodology and Contributions	16
1.6 Thesis Outline	23
Chapter 2 Malaria Diagnosis: A Quick Review	25
2.1 Morphological Analysis of Human Blood Components in Thin Blood Films....	25
2.2 Image Analysis of Thin Blood Slides.....	27
2.3 Performance Measurements	30
2.4 Current Trends in Digital Pathology	32

2.5 Automated Diagnosis of Malaria in Thin Blood Films	33
Chapter 3 Data Analysis and Morphological Processing of Peripheral Blood Images	36
3.1 Image Acquisition and Processing	36
3.2 Analysis of Thin Blood Images	39
3.3 Improved Gray-Scale Conversion Using Contrast Adjustment	42
3.4 Morphological Filtering	49
3.5 Summary	55
Chapter 4 Blood Component Detection and Estimation	56
4.1 Annular Ring Ratio Transform Method	59
4.2 Peak Detection Algorithm.....	63
4.3 Comparison with Segmentation Methods	71
4.4 Summary.....	74
Chapter 5 White Blood Cell and Red Blood Cell Differentiation	76
5.1 Definition in terms of Cell Differentiation	76
5.2 WBC and RBC differentiation algorithm	83
5.3 Experimental Results	91
5.4 Discussion and Summary	97
Chapter 6 Malaria Parasite Identification and Life-stage Recognition	99
6.1 Stained pixel recognition	99
6.2 Identification of Infected Cells	103
6.3 Performance Analysis	119
6.4 Life stage Recognition of Plasmodium Species	138

6.5 Discussions and Conclusions	148
Chapter 7 Extensions	150
7.1 P.Falciparum Gametocyte Detection.....	150
7.2 Automated Malaria Diagnosis in Fluorescent Images	166
7.3 Conclusion.....	172
Chapter 8 Malaria Parasite Detection Using Mobile Phones.....	174
8.1 Literature Review	174
8.2 Mobile Phone Camera Enabled Malaria Diagnosis	175
8.3 Image Acquisition and Processing Using Mobile Phones	184
8.4 Discussion and Summary.....	187
Chapter 9 Conclusions.....	189
9.1 Learning Achievements	189
9.2 Learning Experience	192
9.3 Improvements and Future Work	195
REFERENCES.....	198
BIBLIOGRAPHY.....	209

LIST OF ABBREVIATIONS

ARR: Annular Ring Ratio

CCD: Charge Coupled Device

CDC: Centre for Disease Control and Prevention

FN: False Negative

FP: False Positive

GUI: Graphics User Interface

HSV: Hue Saturation Value

MM: Mathematical Morphology

NPV: Negative Predictive Value

OS: Operating System

PCR: Polymerase Chain Reaction

PPV: Positive Predictive Value

RBC: Red Blood Cell

RDT: Rapid Diagnostic Test

RGB: Red Green Blue

TN: True Negative

TP: True Positive

UoW: University of Westminster

WBC: White Blood Cell

WSI: Whole Slide Imaging

LIST OF SYMBOLS

I_δ	:	Dilated Image
I_ϵ	:	Eroded Image
n_R	:	Total number of pixels in the region R .
R	:	Circular region around the pixel at the origin.
W_R, W_B, W_G	:	Optimum weights of Red, Blue and Green components of the image respectively.
\oplus	:	Morphological Dilation
\ominus	:	Morphological Erosion
μ	:	Mean of the pixel intensity values
σ	:	Standard deviation of the pixel intensity values
σ^2	:	Variance of the pixel intensity values
θ	:	Hue angle which represents the degree to which the pixel is similar or different from its corresponding RGB representation.
\forall	:	For all.

LIST OF FIGURES

Figure 1.1: Life cycle of <i>Plasmodium</i> Parasite from Vector to Host.....	4
Figure 1.2: The life stages of <i>P. Falciparum</i> in thin blood films.....	8
Figure 1.3: Life Stages of <i>P.Vivax</i> in thin smears.....	9
Figure 1.4: Life Stages of <i>P.Ovale</i> in thin blood smears.	10
Figure 1.5: Life Stages of <i>P.Malariae</i> in thin blood smears.....	11
Figure 1.6 Automated Malaria Diagnosis -System Overview	18
Figure 2.1: Microscopic image of a Giemsa stained thin blood smear.....	26
Figure 2.2: Abnormalities in RBC cell structure.	27
Figure 2.3: Subtypes of White Blood Cells	28
Figure 2.4: Similarities in size and shape of <i>Plasmodium spp</i> life stages with other cell components.	29
Figure 3.1: Scanning pattern of peripheral blood smear for morphological evaluation.	37
Figure 3.2 Peripheral Malaria Infected Thin Blood Image.....	40
Figure 3.3 Foreground, background and stained pixel extraction.....	41
Figure 3.4: Contrast plot as a function of red and green weights for gray-scale conversion.	46
Figure 3.5: Contrast enhancement using segmented weighted averaging.	47
Figure 3.6: Contrast enhancement using segmented weighted averaging.	48

Figure 3.7: Contrast enhancement using segmented weighted averaging.	48
Figure 3.8: Morphological closing of a gray-scale thin blood image.	52
Figure 3.9: Structuring elements for morphological filtering.	53
Figure 3.10: Morphological dilation	54
Figure 3.11: Morphologically closed images.	54
Figure 4.1: Structuring elements for the ARR method	60
Figure 4.2: Illustration of ARR method with a positive image with bright objects on a lighter background.	60
Figure 4.3 Pseudo Code of ARR Transform method.	62
Figure 4.4: Structuring element which defines the search area to detect the peaks.	63
Figure 4.5 Peusdocode of Peak Detection Algorithm.	64
Figure 4.6: Demonstration of the ARR transform method on a TIFF image of size 512 x 381 pixels.	67
Figure 4.7: Demonstration of the ARR transform method on a TIFF image of size 512 x 381 pixels with a different field concentration.	68
Figure 4.8: Demonstration of the ARR transform method on a JPEG image of size 1200 x 1600 pixels with non-uniform illumination.	69
Figure 4.9: Demonstration of the ARR transform method on a JPEG image of size 1030 x 1300 pixels with a different staining concentration.	70

Figure 4.10: Comparison of ARR transform method with other segmentation algorithms.	72
Figure 5.1: Application of global thresholding for blood cell extraction using intensity histogram.....	79
Figure 5.2: Effect of poor processing and illumination on histogram.	81
Figure 5.3: Effect of poor processing and illumination on histogram.	82
Figure 5.4: The RBC and WBC differentiation algorithm using localised mean and variance.	86
Figure 5.5: Disadvantages of Mean and Variance method.	87
Figure 5.6: Disadvantages of Mean and Variance method.	88
Figure 5.7: RBC and WBC differentiation algorithm Flowchart.....	90
Figure 5.8: WBC and RBC differentiation algorithm.....	92
Figure 5.9: WBC and RBC differentiation algorithm.....	93
Figure 5.10: WBC and RBC differentiation algorithm.....	94
Figure 5.11: WBC and RBC differentiation algorithm.....	95
Figure 5.12: WBC and RBC differentiation algorithm.....	96
Figure 6.1: Variable appearances of parasitic nuclei within the RBC.	100
Figure 6.2: Stained particles in a blood image.....	101
Figure 6.3: Example of images with different staining and illumination set up.	101

Figure 6.4: Example of images with different staining and illumination.	102
Figure 6.5: Example of images with different staining and illumination set up.	102
Figure 6.6: Infected cell identification using modified ARR method.....	105
Figure 6.7: Hue component representation	108
Figure 6.8: Hue- ARR algorithm..	110
Figure 6.9: Processed image in which the cells that have been identified as infected are circled.....	111
Figure 6.10: Processed image in which the cells that have been identified as infected are circled and WBC differentiated.	112
Figure 6.11: Processed image in which the cells that have been identified as infected are circled.....	112
Figure 6.12: Processed image in which the cells that have been identified as infected are circled and WBC differentiated.	113
Figure 6.13: Non-infected image with WBC differentiated.....	114
Figure 6.14: Processed image in which the cells that have been identified as infected are circled.....	114
Figure 6.15: Processed image in which the cells that have been identified as infected are circled and WBCs differentiated.....	115
Figure 6.16: Processed image in which the cells that have been identified as infected are circled and WBCs differentiated.....	116

Figure 6.17: Processed image in which the cell that has been identified as infected is circled.....	116
Figure 6.18: Processed image in which the cell that has been identified as infected is circled.....	118
Figure 6.19: Processed image in which the cells that have been identified as infected are circled.....	118
Figure 6.20: Parameters used for performance analysis.	120
Figure 6.21:Plot to estimate PPV and NPV	126
Figure 6.22: Plot to estimate the accuracy over a range of parasitemia.....	127
Figure 6.23: Plot of True negative and False positives against total number of cells analysed.....	130
Figure 6.24: Plot of True Negative and False Positive Probabilities based on the threshold.	131
Figure 6.25: Probabilities of False positive and true negative against the number of samples.....	132
Figure 6.26: Probability plot of True Positive and False Negative against the samples at 5% and 25% parasitemia	134
Figure 6.27: Probability plot of True Positive and False negative for 20 and 200 samples	134
Figure 6.28: Probability plot of True Positive and False Negative against the samples with 3% and 5% parasitemia with 5%threshold.....	136

Figure 6.29: Probability plot of True Positive and False Negative against parasitemia for 200 and 1000 samples.	136
Figure 6.30: Plot to demonstrate the impact of threshold values on the probabilities.	137
Figure 6.31: The asexual life stages in a thin blood film.	138
Figure 6.32: Life-recognition algorithm performed on the image in (a).....	140
Figure 6.33: Life-recognition algorithm performed on the image in (a).....	141
Figure 6.34: Life-recognition algorithm performed on the image in (a).....	142
Figure 6.35: GUI Demonstration of Automated Malaria Diagnostic Set-up.	144
Figure 6.36: GUI Demonstration of Automated Malaria Diagnostic Set-up.	144
Figure 6.37: GUI Demonstration of Automated Malaria Diagnostic Set-up.	145
Figure 6.38: GUI Demonstration of Automated Malaria Diagnostic Set-up. The bottom window displays the Ring trophozoite life stage marked by red circles.....	145
Figure 6.39: GUI Demonstration of Automated Malaria Diagnostic Set-up.	146
Figure 6.40: GUI Demonstration of Automated Malaria Diagnostic Set-up.	146
Figure 6.41: GUI Demonstration of Automated Malaria Diagnostic Set-up.	147
Figure 6.42: GUI Demonstration of Automated Malaria Diagnostic Set-up.	147
Figure 7.1: <i>P.falciparum</i> gametocyte morphology.	152
Figure 7.2: WBC and <i>P.falciparum</i> gametocyte in an image	153
Figure 7.3: WBC and <i>P.falciparum</i> gametocyte differentiation algorithm.....	155

Figure 7.4: P. falciparum gametocyte differentiation algorithm.....	156
Figure 7.5: WBC and P. falciparum gametocyte differentiation algorithm.....	157
Figure 7.6: P. falciparum gametocyte differentiation algorithm.....	158
Figure 7.7: P. falciparum gametocyte differentiation algorithm.....	159
Figure 7.8: P. falciparum gametocyte differentiation algorithm.....	160
Figure 7.9: P. falciparum gametocyte differentiation algorithm.....	161
Figure 7.10: P. falciparum gametocyte differentiation algorithm.....	162
Figure 7.11: P. falciparum gametocyte differentiation algorithm.....	163
Figure 7.12: WBC and P. falciparum gametocyte differentiation algorithm.....	164
Figure 7.13: Microscopic image of an infected peripheral blood smear stained with SYBR Green 1 nucleic specific dye.	167
Figure 7.14: RBC morphology in Giemsa stained (a) and Fluorescent images (b).	168
Figure 7.15: Demonstration of the similarities in the appearance of the cells in the complement of a fluorescent image in (a) and a morphologically closed image in (b).	168
Figure 7.16: Fluorescent image processing.....	169
Figure 7.17: Fluorescent image processing.....	170
Figure 7.18: Fluorescent image processing.....	171
Figure 8.1: Schematic of the objective of Malaria diagnosis using mobile phones.....	176

Figure 8.2: Screenshot of Samsung mobile phone with the application for malaria diagnosis.....	178
Figure 8.3: Screenshot of the malaria diagnosis application installed on to a Samsung mobile phone.....	179
Figure 8.4: Screenshot of the malaria diagnosis application installed on to a Samsung mobile phone.....	179
Figure 8.5: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung mobile phone.....	180
Figure 8.6: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung mobile phone.....	180
Figure 8.7: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.....	181
Figure 8.8: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.....	181
Figure 8.9: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.....	182
Figure 8.10: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.....	183
Figure 8.11: Demonstration of the mobile phone microscopy.....	185
Figure 8.12: Demonstration of mobile phone microscopy.....	186
Figure 8.13: Demonstration of mobile phone microscopy.....	186

ASSOCIATED PUBLICATIONS

1. S.Kareem, R.C.S Morling, I Kale” A Novel Method to Count the Red Blood Cells in Thin Blood Films”, IEEE International Symposium on Circuits and Systems (ISCAS), Pages 1021-1024, April 2011
2. S.Kareem, I Kale, R.C.S Morling,” Automated *P.falciparum* Detection system for Post-treatment Malaria Diagnosis using Modified Annular Ring Ratio Method”, IEEE 14th International Conference on Computer Modelling and Simulation, Cambridge, UK, March 2012.
3. S.Kareem, I Kale, R.C.S Morling,” Automated Malaria diagnosis in Fluorescent Thin Blood Image Using Annular Ring Ratio Method”, Presented at Pathology Visions Conference, Digital Pathology Association, Baltimore, Maryland, USA, October 2012.
4. S.Kareem, I Kale, R.C.S Morling,” Automated Malaria Parasite Detection in Thin Blood Films: - A Hybrid Illumination and Colour Constancy Insensitive Morphological Approach”, IEEE Asia Pacific Conference on Circuits and Systems, Kaohsiung, Taiwan, December 2012
5. S.Kareem, I .Kale, R.C.S Morling, ”A Novel Fully Automated Malaria Diagnostic Tool Using Thin Blood Films”, Pan American Health Care Exchanges (PAHCE2013), Medellin, Columbia, May 2013.
6. C.Dallet, S.Kareem, I. Kale,” Mobile Malaria Diagnosis: Real time mobile diagnostic application on Android using Java”, International Symposium for Circuits and Systems, ISCAS 2014.

Planned Submissions:

- S.Kareem , I .Kale, R.C.S Morling,” Automated Thin Blood Smear Image Analysis for Malaria Diagnosis Using Annular Ring Ratio method”, Malaria Journal, BioMed Central, to be submitted in December 2014
- S.Kareem, I .Kale, R.C.S Morling,” Mobile Phone Microscopy: Advancements in Real-time Malaria Diagnosis”, IEEE Journal of Biomedical and Health Informatics, To be submitted in February 2015

ACKNOWLEDGEMENTS

First of all I would like to thank my supervisors, Prof. Izzet Kale and Prof. Richard Charles Morling for their immense support and guidance. They have been a constant source of inspiration throughout these years and my gratitude towards them goes far beyond expression.

I would also like to thank Dr. Martin Parry from the Department of Life Sciences for arranging the short course on “Update in blood and tissue parasitology”. My extended thanks to Ms. Monika Kettlehut (Hospital of Tropical Diseases, London) and Ms. Claire Rogers (London School of Hygiene and Tropical Medicine) for providing the training on malaria diagnosis during the course. I would also like to thank Mr. Rick Dickinson , Mr. Keith Dunning and Mr. Mike Broderick of Newton Microscopes, Cambridge, UK for generously providing us with the portable microscope to carry out the experiments.

I would like to thank my sister Dr. Remya Kareem for going through my thesis and suggesting valuable changes. I would also like to thank my wonderful niece, Noushin Farah, for bringing smile on my face during the stressful times. I would like to thank my parents, Mr. Abdul Kareem and Mrs. Sajeetha Kareem and my mother-in-law Mrs. Chithukkutty for their colossal support and prayers which helped me to get over the difficult times. I would also like to pay respect to my late father-in-law, Mr. Abubacker whom I wish was present here among us to share our joys. My extended gratitude to all my family members including my brother-in-laws Mr. Afsal, Mr. Anoop, his wife Mrs. Noora, my sister-in-law Mrs. Jaseela, her husband Mr. Ahmed and my darling niece and nephew Fazila and Ashif respectively.

I would like to thank my friends Mr Sevket, Mr Burak, Mr Mohsen , Dr Oyku , Dr Adem, Dr Mustafa, Dr Bashar, Miss Damla, Mrs Isil and all other family friends for their constant support. Also, sincere thanks to my internal examiner Dr. Andrzej Tarczynski and all the staff of University of Westminster and Applied DSP and VLSI Research Group for the help provided.

Last, but not the least, a special thanks to my dear daughter Neha Reni for being such a sensible girl, who has adjusted with my busy schedule from the age of 4. And most of all I would like to thank my husband Reni to whom I dedicate this thesis. He has been a pillar of strength throughout these years and has held my hands when I am low and laughed with me when I was happy. Without you Reni, I would not have achieved this!

I, Saumya Kareem Reni, declare that all the material contained in this thesis is my own work.

Chapter 1 Introduction

Automated Low Cost Malaria Detection System in Thin Blood Slide Images Using Mobile Phones is an interdisciplinary research project undertaken in the field of medical diagnosis and pathological analysis, to diagnose the degree of malaria infection from a stained blood smear. The aim of the project is to develop the algorithms to provide a reliable, diagnostic platform for a quantitative analysis of malaria infection. The research has progressed through the following stages: the first being the identification of the Red Blood Cells (RBC) in which the parasites reside. A novel technique called the Annular Ring Ratio (ARR) method has been developed in order to locate the foreground objects such as RBCs and White Blood Cells in the blood. These are further differentiated using a parametric algorithm based on the size, intensity and spatial geometry of the cells and the infected cells are identified. Finally the life stages of the parasites are determined and the parasitemia, which is the ratio of infected cells to normal cells and measures the extent of the infection, is calculated. The algorithm has been implemented and tested on desktop computers and laptops. Furthermore, an android mobile application tool based on the algorithm was developed and was tested on various mobile devices. In addition to this, an automated diagnostic tool to detect the sexual stage (gametocytes) of the species *Plasmodium falciparum* for post treatment malaria diagnosis was developed. Furthermore, a parallel investigation was also carried out on automated malaria diagnosis on fluorescent thin blood films.

This chapter gives a general view on Malaria parasite and the infection, diagnostic techniques, current trends in the field of digital pathology and the objectives of the research. The chapter concludes with a brief explanation on the methodology used in the study. Sections 1.1 and 1.2 explain the disease and pathology of malaria parasite. The diagnostic techniques and current trends in the field are described in Sections 1.3 and 1.4. The final three Sections describe the research aims, methodology, contributions and the outline of the thesis. Readers who are familiar with the background and pathology of malaria parasite may prefer to skip Sections 1.1-1.4.

Automated Malaria Diagnosis Using Mobile Phones

1.1 Malaria

Malaria is a treatable infectious disease caused by a protozoan parasite of the genus *Plasmodium*, whose main victims are children under the age of five. According to the latest World Health Organization (WHO) assessment based on the data received from 104 malaria endemic countries, there were about 219 million cases of malaria in 2010 and an estimated 660,000 deaths [1]. It is most prevalent in sub-Saharan Africa in which most of the cases occur in rural areas where people do not have access to hospitals or the means to afford health care. As a consequence, the majority of cases are undocumented. An estimated US\$ 5.1 billion is needed every year between 2011 and 2020 to achieve universal access to malaria interventions, however the available funding is almost less than half of what is needed [1]. These disturbing statistics could be prevented from escalating by widespread diagnosis of the parasite at an early stage even if expert hands are not available.

The malaria transmission by mosquitoes was first discovered by British doctor, Sir Ronald Ross, while conducting researches in India, on 20th August 1897. Hence 20th August is considered to be the World Mosquito Day and Dr Ross won the Nobel prize for medicine in 1902 for his discovery and continued working for malaria prevention [2]. Many organisations and vital studies are aiming to fight against this deadly disease. The Roll Back Malaria (RBM) partnership is a global framework aimed to implement co-ordinated action against malaria and eradicate the disease by 2015 [3]. The Centre for Disease Control and Prevention (CDC) is another prominent organisation who played a critical role in reducing the disease in both developed and under developed countries [4]. More information regarding Malaria and related issues are available online and from various other sources including WHO [5], Malaria Journal [6] and annual World Malaria Reports [7], [8] and [9].

1.2 Pathology

Malaria is caused by a single-celled parasite of the genus *Plasmodium*. Mainly two types of hosts are infected by the parasite:-female Anopheles mosquitoes (which act as a vector for the transmission of the disease) and human beings, even though some cases are reported on cattle, birds and apes. The parasite enters the human blood stream through the

Automated Malaria Diagnosis Using Mobile Phones

saliva of the Anopheles Mosquito during its blood meal. Infection is relapsing and is often characterised by periodic attacks of fever, chills and anaemia. Serious infections often lead to fatal complications such as splenomegaly (enlargement of the spleen) and cerebral malaria [10]. Other complications include renal failure (Kidney failure), hypoglycaemia (low blood sugar), fluid disturbances and respiratory distress. Malaria in children is deadly whereas the disease during pregnancy leads to possible abortion and still birth due to placental infection [11].

1.2.1 Plasmodium:-Species and Life stages

The genus *Plasmodium* is a protozoan parasite and has five species that cause malaria infection in humans. These are *Plasmodium Falciparum*, *Plasmodium Vivax*, *Plasmodium Ovale*, *Plasmodium Malariae* and *Plasmodium Knowlesi*. Out of five, the first four are the most common and *P.knowlesi* was thought to infect only monkeys. But in 2008, human beings were reported to be infected by *P.knowlesi* which can easily get confused with *P.malariae* and since then rare occurrences have been registered. The most common species is *P.vivax*, but the deadliest is *P.falciparum* [12]. All the species are morphologically different but possess identical life stages in peripheral blood. The common life stages include asexual forms of ring trophozoites, trophozoites and schizonts and sexual form of gametocyte (male or female).

1.2.1.1 Lifecycle

The malaria parasite is carried by female Anopheles mosquito which acts as a vector that injects the immature form of the parasite called sporozoites into the blood stream of humans while feeding on blood. The blood carries the sporozoites to the liver where they mature through forms of schizonts to merozoites. The merozoites are released to the bloodstream where they invade the Red Blood Cells (RBCs) and eventually rupture them. The merozoites multiply asexually through stages of ring trophozoites, trophozoites to schizonts which will eventually be released as merozoites to invade more RBCs. Alternatively some of them develop to sexual forms called gametocytes.

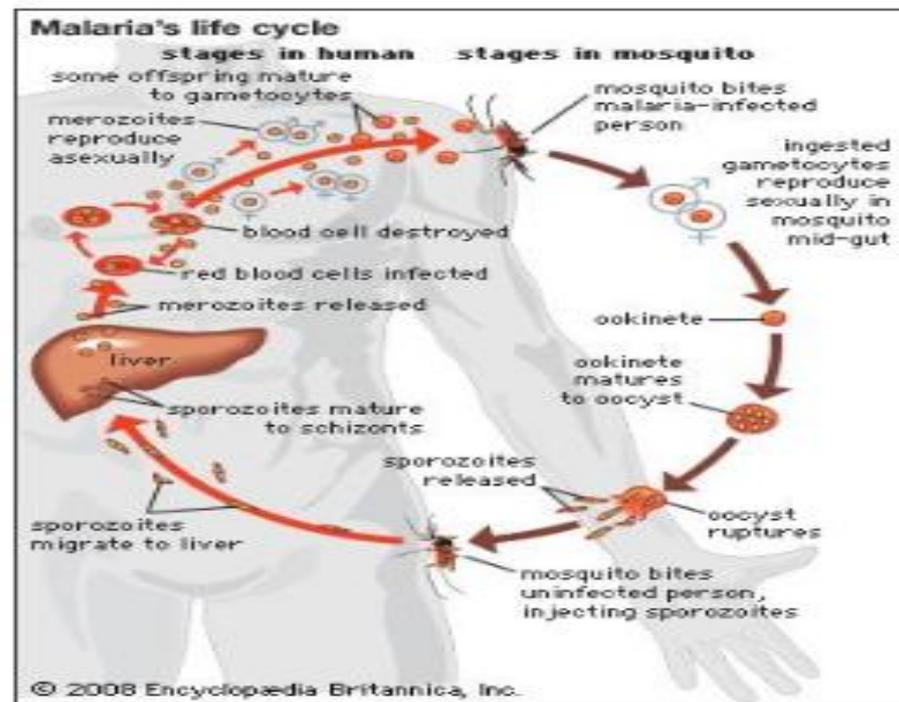


Figure 1.1: Life cycle of *Plasmodium* Parasite from Vector to Host.

Reproduced from [12].

Once a non-infected mosquito has a blood meal from an infected person, the gametocyte enters the mosquito gut where they mate and produce sporozoites which in turn will cause further transmission of the disease. Figure 1.1 shows the entire lifecycle of the parasite from vector to host and back again. The intervals between the release of each generation of merozoite to the blood stream to invade RBCs is usually 48-72 hours and is responsible for the recurrent symptomatic attacks of chills and fever in the host [12].

1.3 Morphology of *Plasmodium* Species

The morphology of each species and its effect on the RBCs depends on the species. Each parasite impacts on the size and appearance of the erythrocytes (RBC) differently, which makes them easy to distinguish between themselves. However, all of them go through the same life stages in peripheral human blood. Sometimes it is hard to discriminate between species in the same life stages; especially the early stages (Trophozoite ring stages). Tables 1.2, 1.3, 1.4 and 1.5 demonstrate the key features of each species.

Automated Malaria Diagnosis Using Mobile Phones

Table 1.1: Key Features of *P.falciparum* life stages.

Reproduced with permission from [13]

<i>Plasmodium</i> species	Stages found in blood	Appearance of Erythrocyte (RBC)	Appearance of Parasite
<i>P. falciparum</i>	Ring	normal; multiple infection of RBC more common than in other species; Maurer's clefts (under certain staining conditions)	delicate cytoplasm; 1 to 2 small chromatin dots; occasional appliqué (accolé) forms
	Trophozoite	normal; rarely, Maurer's clefts (under certain staining conditions)	seldom seen in peripheral blood; compact cytoplasm; dark pigment
	Schizont	normal; rarely, Maurer's clefts (under certain staining conditions)	seldom seen in peripheral blood; mature = 8 to 24 small merozoites; dark pigment, clumped in one mass
	Gametocyte	distorted by parasite	crescent or sausage shape; chromatin in a single mass (macrogametocyte) or diffuse (microgametocyte); dark pigment mass

Table 1.2: Key Features of *P.Vivax* life stages.

Reproduced with permission from [13]

<i>Plasmodium</i> species	Stages found in blood	Appearance of Erythrocyte (RBC)	Appearance of Parasite
<i>P. vivax</i>	Ring	normal to 1.25×, round; occasionally fine Schüffner's dots; multiple infection of RBC not uncommon	large cytoplasm with occasional pseudopods; large chromatin dot
	Trophozoite	enlarged 1.5 to 2×; may be distorted; fine Schüffner's dots	large amoeboid cytoplasm; large chromatin; fine, yellowish-brown pigment
	Schizont	enlarged 1.5 to 2×; may be distorted; fine Schüffner's dots	large, may almost fill RBC; mature = 12 to 24 merozoites; yellowish-brown, coalesced pigment
	Gametocyte	enlarged 1.5 to 2×; may be distorted; fine Schüffner's dots	round to oval; compact; may almost fill RBC; chromatin compact, eccentric (macrogametocyte) or diffuse (microgametocyte); scattered brown pigment

Automated Malaria Diagnosis Using Mobile Phones

Table 1.3: Key features of *P.Ovale* life stages.
Reproduced with permission from [13]

<i>Plasmodium</i> species	Stages found in blood	Appearance of Erythrocyte (RBC)	Appearance of Parasite
<i>P. ovale</i>	Ring	normal to 1.25×, round to oval; occasionally Schüffner's dots; occasionally fimbriated; multiple infection of RBC not uncommon	sturdy cytoplasm; large chromatin
	Trophozoite	normal to 1.25×; round to oval; some fimbriated; Schüffner's dots	compact with large chromatin; dark-brown pigment
	Schizont	normal to 1.25×, round to oval, some fimbriated, Schüffner's dots	mature = 6 to 14 merozoites with large nuclei, clustered around mass of dark-brown pigment
	Gametocyte	normal to 1.25×; round to oval, some fimbriated; Schüffner's dots	round to oval; compact; may almost fill RBC; chromatin compact, eccentric (macrogametocyte) or more diffuse (microgametocyte); scattered brown pigment

Table 1.4: Key features of *P.Malariae* life stages.
Reproduced with permission from [13]

<i>Plasmodium</i> species	Stages found in blood	Appearance of Erythrocyte (RBC)	Appearance of Parasite
<i>P. malariae</i>	Ring	normal to 0.75×	sturdy cytoplasm; large chromatin
	Trophozoite	normal to 0.75×; rarely, Ziemann's stippling (under certain staining conditions)	compact cytoplasm; large chromatin; occasional band forms; coarse, dark-brown pigment
	Schizont	normal to 0.75×; rarely, Ziemann's stippling (under certain staining conditions)	mature = 6 to 12 merozoites with large nuclei, clustered around mass of coarse, dark-brown pigment; occasional rosettes
	Gametocyte	normal to 0.75×; rarely, Ziemann's stippling (under certain staining conditions)	round to oval; compact; may almost fill RBC; chromatin compact, eccentric (macrogametocyte) or more diffuse (microgametocyte); scattered brown pigment

Automated Malaria Diagnosis Using Mobile Phones

The two key attributes utilised by the microscopists are the appearance of the erythrocyte and appearance of the parasite, the latter being also used for life-stage recognition and species identification. The similarities in the feature between *P.vivax* and *P.ovale* often make the diagnosis difficult. It is also difficult to differentiate the gametocytes of *P.vivax*, *P.ovale* and *P.malariae* due to similar morphology. However the gametocytes are asymptomatic and apart from enhancing the transmission, they are not fatal if appears in the blood. The gametocytes of *P.falciparum* possess distinctive crescent or sausage shape which makes the diagnosis easier. Figures 1.2, 1.3, 1.4 and 1.5 illustrate the morphology of *P.falciparum*, *P.vivax*, *P.ovale* and *P.malariae* respectively. All the images and data were reproduced from the CDC's malaria website with written permission [13].

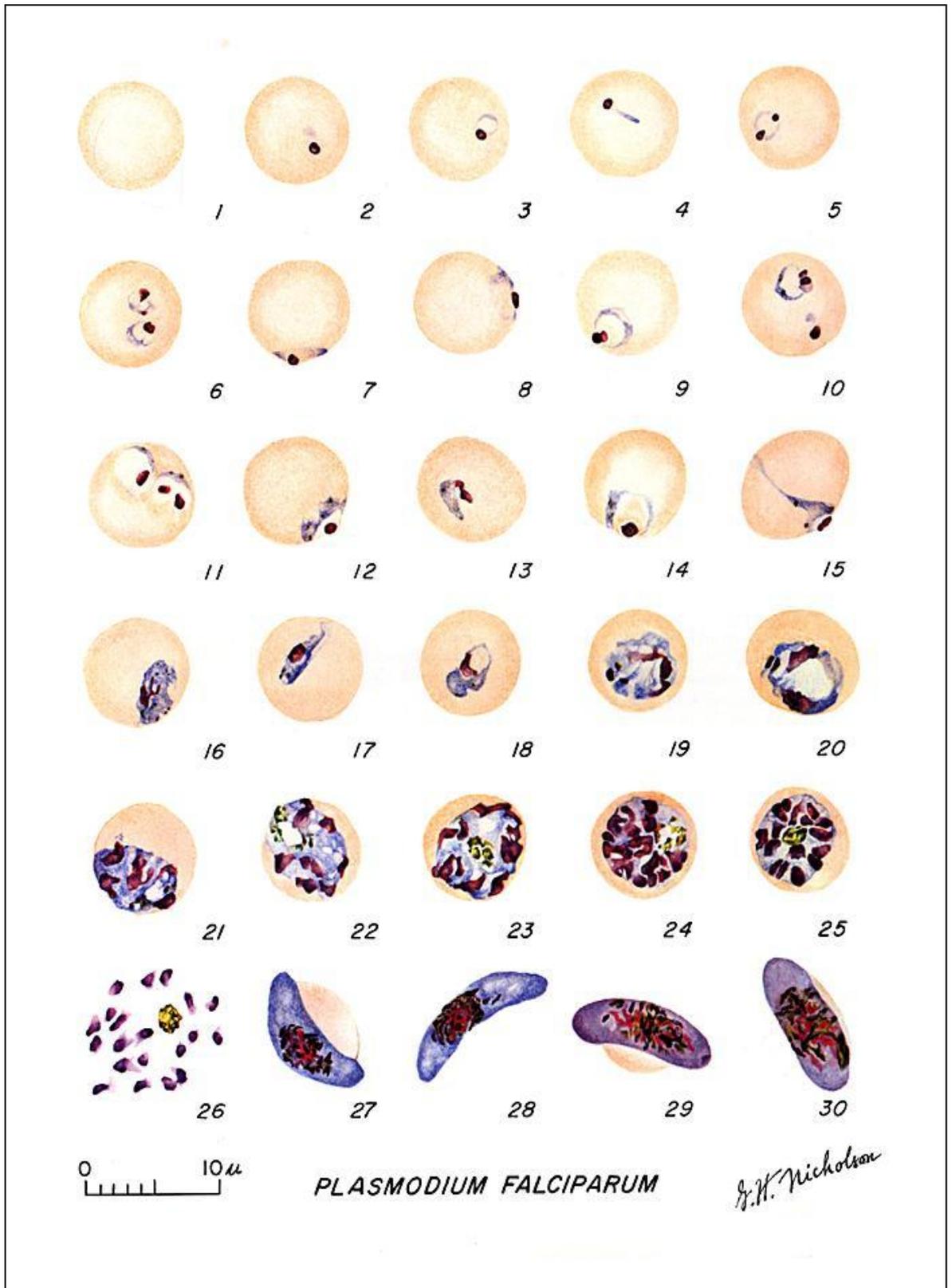


Figure 1.2: The life stages of *P. Falciparum* in thin blood films.

1: Normal red cell; 2-18: Trophozoites (among these, 2-10 correspond to ring-stage trophozoites); 19-26: Schizonts (26 is a ruptured schizont); 27, 28: Mature macrogametocytes (female); 29-30: microgametocytes (male). Note the distinctive shape of the gametocytes. Reproduced with permission from [13]

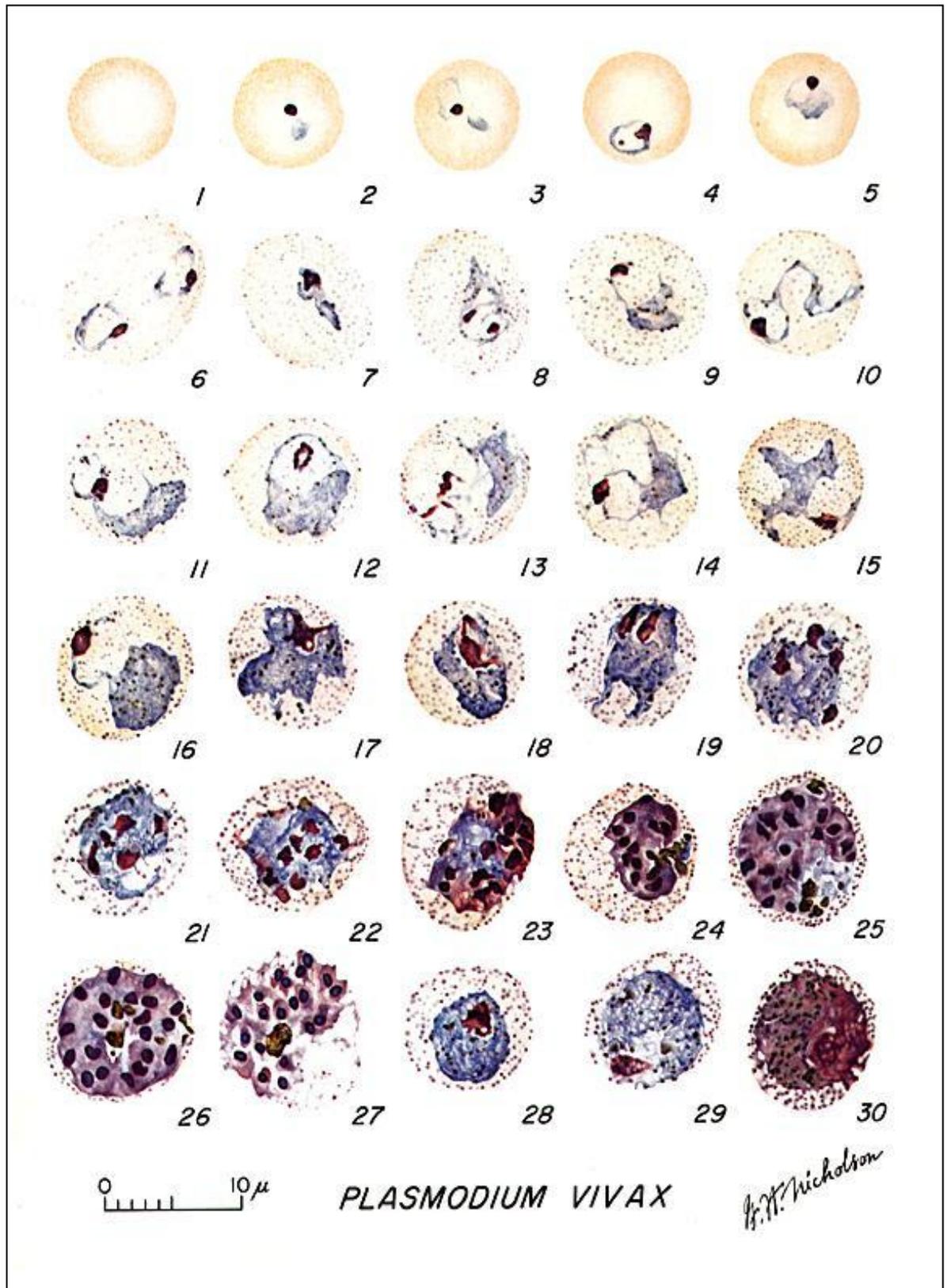


Figure 1.3: Life Stages of *P. Vivax* in thin smears.

1: Normal red cell; 2-6: Young trophozoites (ring stage parasites); 7-18: Trophozoites; 19-27: Schizonts; 28 and 29: Macrogametocytes (female); 30: Microgametocyte (male). Reproduced with permission from [13]

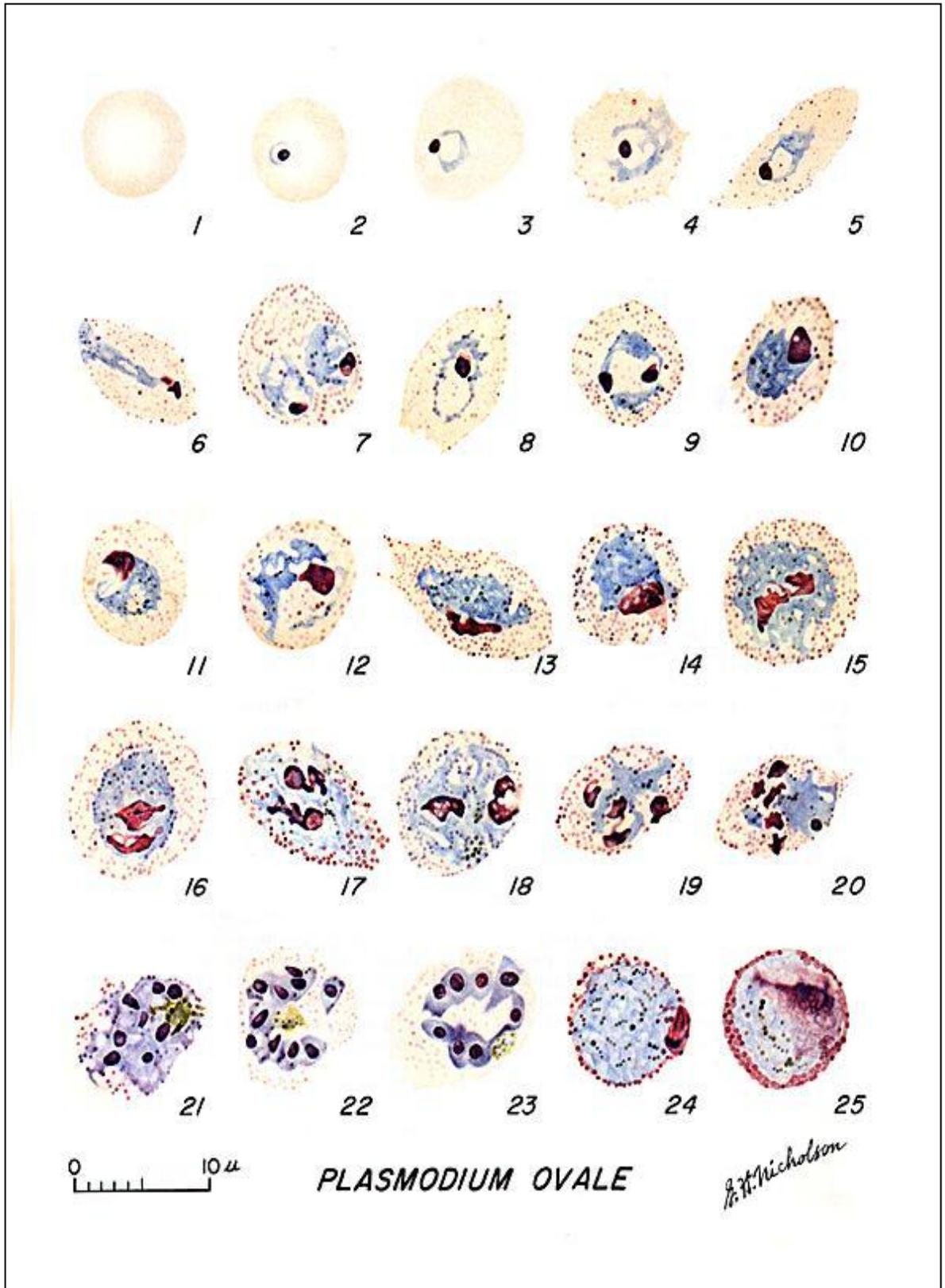


Figure 1.4: Life Stages of *P.Ovale* in thin blood smears.

1: Normal red cell; 2-5: Young trophozoites (Rings); 6-15: Trophozoites; 16-23: Schizonts; 24: Macrogametocytes (female); 25: Microgametocyte (male). Reproduced with permission from [13]

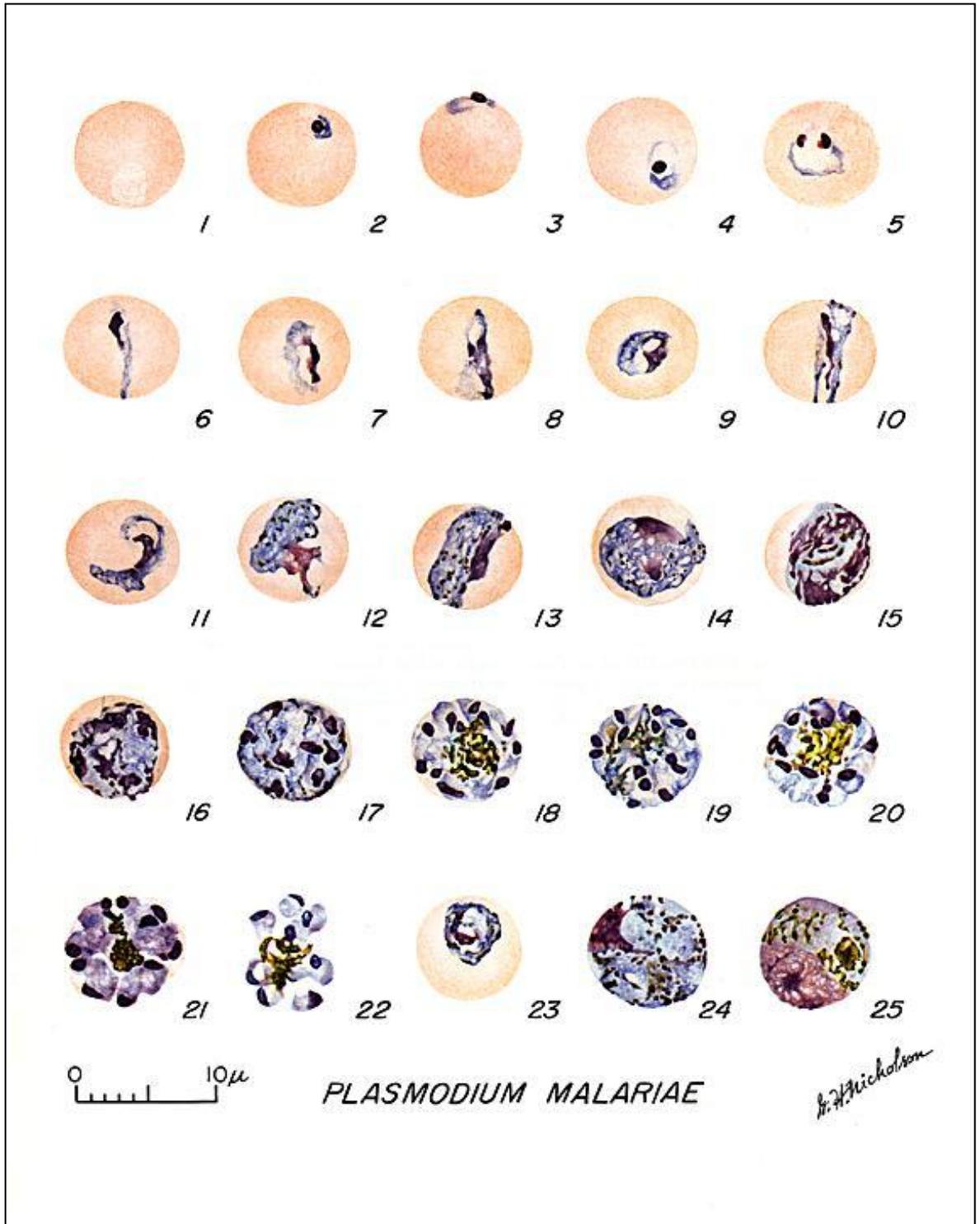


Figure 1.5: Life Stages of *P. Malariae* in thin blood smears.

1: Normal red cell; 2-5: Young trophozoites (rings); 6-13: Trophozoites; 14-22: Schizonts; 23: Developing gametocyte; 24: Macrogametocyte (female); 25: Microgametocyte (male). Reproduced with permission from [13]

Automated Malaria Diagnosis Using Mobile Phones

1.4 Diagnostic Techniques

The accuracy and reliability of diagnostic procedures plays a vital part in the treatment and subsequent elimination of deadly diseases like malaria. Recent developments in medical diagnosis, combined with traditional methods, provide the following list of diagnostic techniques to detect the parasite.

- Manual Diagnosis
- Molecular Diagnosis
- Rapid Diagnostic Tests
- Automated Diagnosis using digital slide analysis

Detailed descriptions of all the above methods are given below.

1.4.1 Manual diagnosis

Manual diagnosis remains the most standard procedure to diagnose malaria. However, the technique is laborious, time consuming and need skilled expertise. The two classical techniques for manual diagnosis are clinical diagnosis and microscopic diagnosis.

Clinical diagnosis is the least expensive and widely practised method. It is based on physical findings at examination derived from patient's signs and symptoms. There are certain clinical algorithms to be followed for diagnosis but the method requires only a minimally trained examiner. However the method shows very low specificity (probability of negative result for negative samples) and can lead to over diagnosis due to the non-specific nature of signs and symptoms. For example, *P.falciparum* infection is more fatal with severe illnesses whereas the other species show acute illness but with a very low mortality rate. In some cases the parasite remains dormant in the liver leading to the person being an asymptomatic carrier. Even though the clinical diagnosis often shows around 100% success rate, the overlapping of malaria symptoms with other tropical diseases impair the diagnostic specificity and hence the accuracy can be reduced. Therefore it is always advisable to combine this method of diagnosis with some kind of laboratory diagnostic measurements [14].

Automated Malaria Diagnosis Using Mobile Phones

Microscopic examination of malaria is conducted on stained blood films using Giemsa, Wrights or Field stains, out of which microscopic detection and identification of the parasite on Giemsa stained thick or thin blood films is considered to be the gold standard for laboratory diagnosis [14]. A comparative description of thick and thin blood smear is provided in Table 1.5.

Table 1.5: Comparisons of Thick and Thin Peripheral Blood Smears

Characteristics	Stained Thick Blood Film	Stained Thin Blood Film
Preparation	3 drops of blood in 1 cm ² area. Forms multiple layers of cells	1 drop of blood in 1 cm ² area. Forms single layer of cells
Screening	RBCs are dead and parasites are more concentrated. Screened for parasites and mixed infections.	RBCs are preserved with methanol. Rapid screening is possible.
Observable Components	WBCs and Parasites	RBCs, WBCs, Parasites, Platelets and Artefacts
Parasite Identification	Highly sensitive to parasite identification due to high concentration in a small area.	For low level infection, parasite identification can be difficult. Useful for detailed analysed following a positive thick smear examination
Species Identification	Species identification not possible due to destruction of RBCs.	Useful for species identification.

The conventional microscopic malaria diagnosis starts with a routine examination of thick blood films. According to the WHO learner's guide for microscopic examination of malaria [8], an examination of up to 200 fields, each field having approximately 200 WBCs(0.025µl of blood) , would be sufficient to provide a conclusion. For a negative conclusion 100 fields are recommended to be observed and a further 100 fields would be sufficient to examine if a parasite is observed, in order to calculate the parasitemia. The thick film examination is followed by thin film examination where 25 fields (each with an average of 200 RBC/field) are necessary to examine to calculate the parasitemia. This is based on an objective of 100x magnification with each field consisting of 200-250 RBCs.

Automated Malaria Diagnosis Using Mobile Phones

Parasitemia is calculated in order to enumerate the parasites. Parasitemia is a measure of infection quantified against blood elements such as RBCs and WBCs [15]. For thin blood films parasitemia is calculated as total number of parasites per RBCs or WBCs and for thick films, it is against WBCs.

As per the standards recommended by the Clinical and Laboratory Standards Institute (CLSI), the total number of parasites per 100 WBCs could be counted to estimate the parasitemia on either thick or thin blood films. The alternative is to calculate the number of parasites per 1000 RBCs in thin films. At least 10 fields with 100 RBCs should be counted to get the percentage [16].

$$\text{Parasitemia in thin blood films} = \frac{\text{Infected RBCs}}{\text{Total number of RBCs}} \quad (1.1)$$

For low level infection up to 2000 RBCs and for high level infection 500 RBCs has to be examined to measure the parasitemia [14].

$$\text{Parasitemia in thick blood films} = \frac{\text{Number of Parasites}}{\text{Number of WBCs}} \times \text{WBC count per microlitre} \quad (1.2)$$

For thick smears the numbers of either the parasites or the WBCs are tallied until a count of 500 for parasites or 1000 WBCs for quantification is achieved [14].

Apart from thin and thick smear examination, microscopic diagnosis has been recently extended to fluorescent dyed slides. Stained nucleic acids of parasites using fluorescent dyes are used for detection with either a fluorescent microscope or a light microscope with interference filter system [14]. Visual recognition is comparatively easier with fluorescent microscopy but training of the microscopist and appropriately chosen dyes are critical factors to enhance accuracy.

Microscopic malaria diagnosis is by far considered to be the most effective diagnostic method, but it is highly time-consuming, labour intensive and requires special training. The efficiency can be affected by poor skills maintenance, slide preparation, quality of the laboratory facilities and supply and the condition of the microscope. Moreover, the accuracy of the system solely depends on the expertise of the microscopist [17].

Automated Malaria Diagnosis Using Mobile Phones

1.4.2 Molecular Diagnosis

Molecular diagnosis is done through Polymerase Chain Reaction (PCR) which is a confirmatory test for malaria. The method is used to confirm the presence of parasites and for species recognition by extracting the genomic DNA of *Plasmodium Spp* and amplify them. The method undertakes two-step nested PCR by first amplifying the DNA with genus specific primers (DNA sequences to target and amplify) followed by species specific primers for species identification [14]. Molecular diagnosis is more sensitive than microscopic examination for low level and mixed infections. However, the procedure is expensive, including access to reagents and cost of labour and hence the implementation is difficult in laboratories located in poorer areas where malaria is endemic [18].

1.4.3 Rapid Diagnostic Tests

Rapid diagnostic tests (RDT) are immunochromatographic (separation of antigens or antibodies) methods based on detection of malarial antigens, where particular antibodies are used to detect the antigens in peripheral blood. Both asexual and sexual form of *plasmodium Spp* produces parasite lactase dehydrogenase (pLDH) antigen which is detected during the tests. The method is useful for discriminating *P.falciparum* from other species but cannot identify others specifically [14]. These tests are mainly used for research studies and even though they give results faster than microscopic examination, the detection rate is impaired for low level infection and hence is not very reliable. Several factors including manufacturing process and laboratory and other environmental conditions may affect the performance of the tests. It is usually recommended in situations such as outbreaks or occupational exposure where microscopic examinational capabilities exceed [17].

1.4.4 Automated Diagnosis Using Digital Slide Analysis

Automated malaria diagnosis is the latest addition to the several identification techniques for malaria. The method works on the principle of digitising the whole microscopic examination process with minimum human interference. This is, in one way, imitating all the procedures and techniques of manual microscopy and incorporating it in to a machine-vision platform. The diagnostic system consists of hardware and software components. The

Automated Malaria Diagnosis Using Mobile Phones

hardware requires imaging equipments such as a slide scanner, camera enabled or digital microscope that has provisions for auto-focusing, automated or semi-automated field selection and slide movements and processing devices such as a computer, mobile phone or tablet. The software necessitates robust image processing algorithms for pre-processing and overall diagnosis of the infection. Apart from the initial set-up cost, the automated system needs only basic training and possesses consistent timing with accuracy reaching up to that of manual diagnosis. These algorithms need to achieve diagnostic accuracy approaching that of an experienced pathologist yet be incorporated into a platform used by technicians with only basic training. Furthermore, the time taken should not be longer (and preferably shorter) than that taken using manual methods. The methods can be used to for quantitative analysis, research, peripheral blood screening, post treatment diagnosis as well as an alternative or secondary option for manual microscopy.

1.5 Research Aims, Methodology and Contributions

This research is of interdisciplinary nature and hence the proposed methods in the study contribute to both medical and image processing areas. The primary aim of the research was to develop a diagnostic tool which can contribute significant improvements in clinical pathology. This has been achieved by investigating a set of problems involved in the processing of the blood image, finding the solutions by proposing novel algorithms and finally evaluating them.

Automated microscopic diagnosis of malaria had been a subject of several research efforts over the past decade. Three major research projects have been conducted at the University of Westminster between 2000 and 2009. Digital thin film examinations were first studied by Cecilia Di Ruberto [19], [38] and later by K.N.R Mohano Rao [20], [39], [40]. Their work on blood image processing provided the basis for parasite detection. Ruberto identified the stained objects by their colour characteristics and distinguished them as parasitic and non-parasitic objects. The research also addressed the issues of stained White Blood Cells (WBCs) and platelets. Rao considered all the stained objects as parasites and identified the foreground objects first and then differentiated the stained particles. These approaches however will tag all the stained platelets and other artefacts as parasites and thereby impair the specificity of the automated diagnosis.

Automated Malaria Diagnosis Using Mobile Phones

The recent and most remarkable work in this field is done by Boray Tek whose contributions remain in the highly referenced category [21], [42], [54]. Tek developed a computerised diagnostic tool for malaria which performs minimum area watershed segmentation and Radon transform [51] based marker extraction algorithms for blood cell segmentation. The process first estimates the stained classes using non-parametric histogram method and uses features such as Hu-moments and colour autocorrelation to develop a K-nearest Neighbour classification rule [52] and other pattern recognition algorithms for parasite and life stage recognition. The method requires pre-processing algorithms for colour normalisation, which includes transforming the original image to match a reference image colour characteristic and perform further transformation of the foreground regions.

The work described in this thesis is different from the above studies performed at the University of Westminster. Since it is ultimately aiming for usage in mobile phones, it intends to reduce the computational complexity and thereby keep the implementation cost to a minimum. The segmentation technique used in the research is a novel algorithm performed using morphological filters and structuring elements. The research also avoids complex pattern recognition algorithms and neural network architectures. Instead, extensive use of morphological operations is the key feature of the work. An overview of the diagnostic procedure is given in Figure 1.6.

Automated Malaria Diagnosis Using Mobile Phones

1.5.1 System Overview

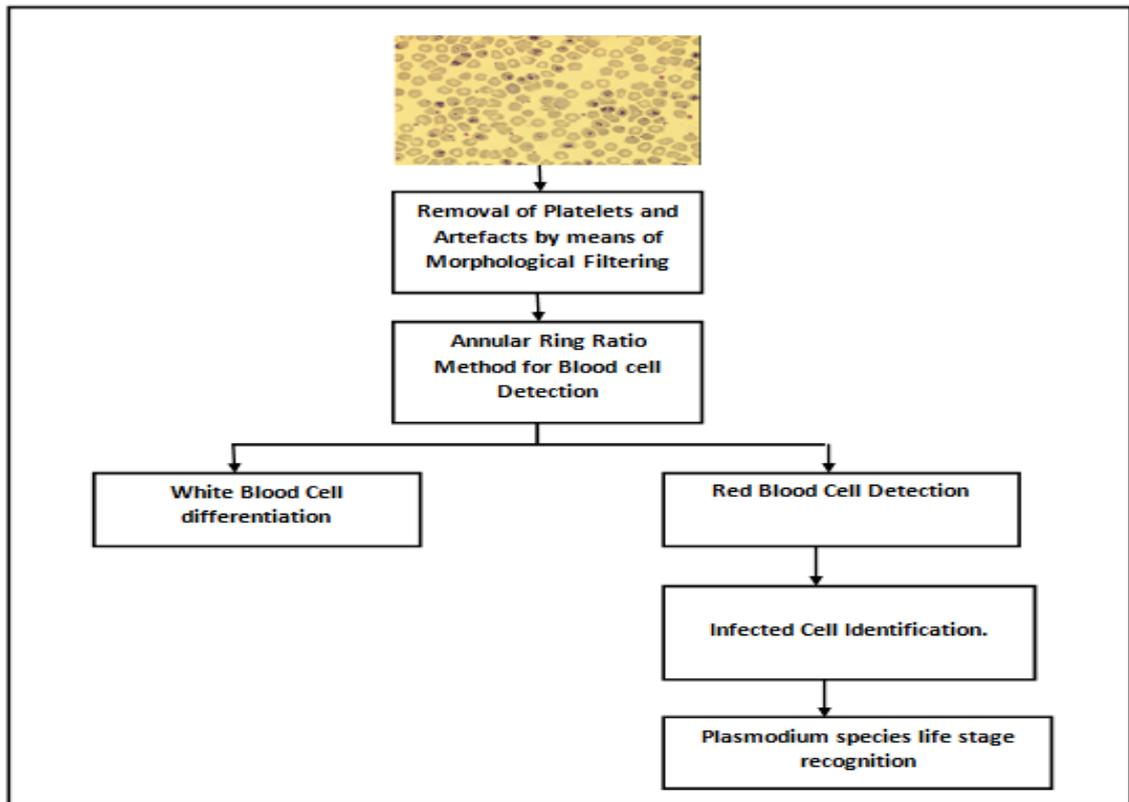


Figure 1.6 Automated Malaria Diagnosis -System Overview

This research is concerned with automated slide analysis described in Section 1.4.4. It has concentrated mainly on the software part of the diagnostic system, in developing novel algorithms for parasite detection to be incorporated onto a mobile phone. The following outcomes were achieved:

- Morphological pre-filtering of the image to facilitate cell isolation with poor quality images;
- Identification and location of both RBCs and WBCs;
- Detection of infection within the RBCs;
- *Plasmodium* species life stage recognition –sexual and asexual life stages;
- Gametocyte detection tool for post treatment diagnosis;
- Fluorescent image analysis and parasite diagnosis;
- Design of Android based mobile application tool for malaria diagnosis;

Automated Malaria Diagnosis Using Mobile Phones

The outcomes were attained through analysis and processing of peripheral thin blood images. In the early stages of the research, the experiments were conducted on images that were already available in the Applied DSP and VLSI Research Group (ADVVRG) database. These included images provided by the National Institute of Medical Research (NIMR), UK and some taken directly using a Canon A360 camera mounted on an Olympus microscope. As the research progressed, a variety of images were collected from various sources including Hospital for Tropical Diseases (HTD), Pondicherry Institute of Medical Science (PIMS) and Centre for Medical Electronics, Anna University, India. In addition, a set of images were acquired from thin blood slides which belongs to the National External Quality Assessment Service (NEQAS) for Microbiology using mobile phone cameras placed on the eyepiece of a microscope. These slides were provided by the specialists of London School of Hygiene and Tropical Medicine (LSHTM) and HTD. The images were acquired using HTC-1, Samsung Galaxy S2, S3 and Apple Iphone-4. Later, a complete diagnostic arrangement was set up, using a portable Newton Nm1 series microscope [103] incorporated on to the mobile phone, which enabled on-the-spot diagnosis.

Mathematical morphological operations are the quintessential paradigm used throughout the line of study. The method of diagnosis avoids pattern recognition algorithms and neural network architectures that are commonly used in these modes of investigations. The research has tackled several sub-problems of the diagnosis, and a number of image processing concepts were utilised to overcome them. The research is only concerned about the problems associated with the machine vision point of view and tried to achieve comparable parasite detection ability to manual diagnosis. In addition to the main outcomes, parallel investigations were carried out on *P.falciparum* gametocyte detection and malaria diagnosis in fluorescent images.

1.5.2 Contributions

As previously explained, this research focused on malaria diagnosis by processing the image of a thin blood film. During the course of this study, a number of image processing algorithms were developed which has contributed to the interdisciplinary field of imaging and medicine. The main contributions to the knowledge and related publications are explained below:

Automated Malaria Diagnosis Using Mobile Phones

- A morphological pre-processing and filtering algorithm was developed during the research, which removes noise and artifacts present in the image and provide a uniform intensity distribution across the foreground components. The processing of blood images is often difficult because of non-uniform intensity distribution across the foreground components such as RBCs and WBCs, the light centre patch within the RBCs, as well as noise present due to poor slide preparation and impurities in the microscope lenses. Conventional image processing algorithms tackle these issues separately using complex pre-processing procedures which involves filtering using median filters, subtraction techniques or morphological closing or opening, contrast enhancement or colour normalisation for improved visibility and uniform illumination of foreground components and techniques such as ‘hole filling’ and connected component extraction to remove the light centre pallor within the foreground components [35], [37], [42], [53]. The morphological filtering algorithm proposed in the thesis successfully addresses the above issues, thereby reduces the computational cost and contributes to the field as an effective pre-processing tool, not only for blood images, but also for images with a non-uniform background with medium to high contrast between foreground and background objects such as geological and histopathological images. The method is described in detail in Section 3.4.
- A novel method called the Annular Ring Ratio (ARR) transform for blood component identification was developed which separates the foreground component (cells such as RBCs and WBCs) from the background based on local variation of intensities. This method advances the field of mathematical morphology because of its insensitivity towards non-uniform illumination and uneven film distribution. The method is fundamentally different from conventional segmentation technique since it locates and estimates the centre of the cells rather than regionalises them. The method is not specific to malaria diagnosis and hence can also be used as a morphological image processing tool for other applications such as normal blood image analysis, structure analysis of GPS photographic data and other segmentation techniques. The method along with the pre-processing technique mentioned in the previous paragraph is an effective morphological object identification tool and can be used for quantitative estimation. Chapter 4 of this thesis describes the ARR

Automated Malaria Diagnosis Using Mobile Phones

algorithm, its advantages and experimental results in detail. The work was presented at the IEEE International Symposium on Circuits and Systems (ISCAS) in 2011 [56].

- The research also proposed an object identification algorithm based on size and average intensity distribution combined with morphological closing. The algorithm can be applied to differentiate objects with different size and eccentricity present in the image and could be modified according to the application. In this research the algorithm is used to differentiate the foreground objects as WBCs and RBCs and can be used as a WBC detection tool not only for malaria diagnosis but also for normal blood analysis. A detailed explanation with experimental results is provided in Chapter 5 of this thesis.
- The Hue-ARR algorithm explained in Chapter 6 is another major contribution of this research to knowledge. It explains a novel image processing algorithm for the identification and subsequent estimation of objects in the image using the parameters obtained from the hue component of the image. This algorithm can be mainly used for images with diverse object properties with a high level of pixel intensity variation and non-uniform background. This algorithm is a simple alternative to conventional pattern recognition algorithms and does not require any feature selection or extraction procedures. The method also estimates parasitemia in the blood and was published at Pan American Health Care Exchange (PAHCE) conference in 2013 [69].
- The work also contributed to the field of clinical pathology, a *P.falciparum* gametocytes detection tool for post treatment malaria diagnosis which aids in controlling the transmission of the disease. The algorithm is based on the mean, variance and geometry of the stained pixels. The proposed method can also be used as a WBC detection tool and the diagnosis of other peripheral parasites with unusual geometry such as trypanosome and babesia [63]. This work was presented at the IEEE Computer Modelling and Simulation Conference (UKSIM), 2012 [79].
- The experiments performed on fluorescent images contributed a novel algorithm for malaria diagnosis based on the ARR transform method. The algorithm utilise modified ARR transform method and has extended applications to other fluorescent blood images as well and has scope for future work on fluorescent microscopy. The work was presented at the Pathology Visions Conference in 2012 [85].

Automated Malaria Diagnosis Using Mobile Phones

In addition, a comparative study on two morphological methods used for *Plasmodium* detection and life stage recognition was conducted and results were presented at the IEEE Asia Pacific Conference on Circuits and Systems (APCCAS) in 2012 [68]. The detailed description of these experiments are described in Chapter 6.

All the algorithms developed in this research are implemented and tested in Matlab [111]. During the final stages of the research, further investigations were carried out to implement these algorithms on to mobile phones and an Android application for malaria diagnosis using these algorithms were developed as part of a collaborative work [112]. A significant number of images used for the experiments were acquired using the camera of the mobile phone. A comprehensive journal paper on automated malaria diagnosis is currently under preparation.

1.5.3 Testing and Validation

As mentioned in Section 1.5.1, the images used in this study are obtained from diverse sources and has been acquired under different imaging, hardware and laboratory conditions. Hence the experiments conducted to test the efficiency and robustness of the image processing algorithms developed in the research were challenging. The diversity of the nature of the images and the variety and randomness of the components present in them made the task complicated. The quality of the images varied according to the lighting conditions, processing of the smear, including improper staining and the impurities in the imaging hardware.

In order to perform the experiments and for credible analysis and testing of the images, the author received special training on microscopic malaria diagnosis accredited by Institute of Biomedical Sciences (IBMS) conducted by University of Westminster in collaboration with experts from London School of Hygiene and Tropical Medicine (LHTM) and Hospital for Tropical Diseases (HTD).

A machine based learning approach was used in the research. In order to calibrate the parameters and set up the threshold values, the algorithms were manually trained using 50 images in the ADVRG database that were provided by NIMR. Initially, the morphological filtering algorithm and the ARR transform algorithm were trained together.

Automated Malaria Diagnosis Using Mobile Phones

Once a complete diagnostic tool was developed by combining these algorithms with the rest of the algorithms for object identification and estimation, a total of 268 images containing 16,673 objects were processed and manually tested.

Out of these 268 images, the successful automatic diagnoses for 251 images were verified by experts from the corresponding sources. The results obtained from the processing of the images in the ADVRG dataset were compared with the manual diagnosis of the respective images by the experts in NIMR. Similarly, the test results for images from Anna University and PIMS were verified by the respective personnel who provided the images. However, it was not possible to obtain verification of 17 images from LSHTM and HTD.

The results obtained from the experiments were used to interpret the overall performance of the system as a diagnostic tool. Chapter 6 of this thesis explains in detail, the nature, diversity and randomness of the type of images used for processing. Section 6.2.3 demonstrates the variety of the images used and the subsequent results proving the robustness of the system. Chapter 6 also provides a detailed description on the total number of images processed and the object analysis along with RBC and WBC estimation, performance analysis conducted based on the experimental results, the diagnostic probabilities obtained from these results and the Graphic User Interface (GUI) designed for the visual analysis of the automated malaria diagnosis tool developed in the research.

1.6 Thesis Outline

The rest of this chapter mention the main contributions and related publications resulted from this research.

Chapter 2 presents a literature review of the existing studies related to automated malaria diagnosis. In addition a background survey on current trends in digital pathology, blood image analysis, mathematical morphology, fluorescent microscopy and mobile phone microscopy has also been described.

Chapter 3 deals with the description of pre-processing techniques used in this research related to image analysis of peripheral thin blood slides. It discusses the problems

Automated Malaria Diagnosis Using Mobile Phones

associated with image acquisition including contrast adjustment and non-uniform illumination and describes the solutions proposed.

Chapter 4 proposes a novel method for determining the location of blood components in the peripheral blood image and discusses the results.

In Chapter 5, a novel method which is fundamentally different from the existing segmentation techniques to differentiate the WBCs and RBCs present in the image is explained.

Chapter 6 illustrates two different approaches for parasite identification and discusses the advantages and disadvantages of each method. It also proposes a novel algorithm to identify the infected cells from a given set of foreground objects consisting of RBCs and WBCs. It also describes the procedure used to identify the life stages of *plasmodium spp.* The chapter also gives a performance measurement analysis of the entire diagnostic set-up.

Chapter 7 introduce extended studies conducted during the research. Chapter 7 present an experimental study on *P.falciparum* gametocyte detection for post-treatment malaria diagnosis. It also describes a parallel investigation that has been carried out on fluorescent microscopic images to diagnose malaria.

Chapter 8 describes the investigations conducted on mobile phone microscopy and discusses the design and implementation of the algorithms onto a mobile phone.

Chapter 9 gives a summary of the thesis, reflections, achievements, contributions to the field and suggestions for future work.

Chapter 2

Malaria Diagnosis: A Quick Review

Different areas of malaria eradication programme such as parasite diagnosis, invention of new vaccinations, improvement of existing vaccinations, improvement in pre and post treatments and prevention of transmission of disease are subjected to intense research. However, in this chapter, associated literature based on malaria parasite diagnosis that are relevant to the research work described in the thesis are been presented.

This chapter gives an insight into the field of microscopy, related diagnosis and current trends in automated diagnosis and digital pathology. The diagnosis of many blood related infections use similar techniques to those used in the diagnosis of malaria and, therefore, the algorithms developed here could be adapted for the diagnosis of these diseases.

2.1 Morphological Analysis of Human Blood Components in Thin Blood Films

In peripheral blood, definitive diagnosis of malaria is done by visual detection and recognition of the parasite in a stained sample of blood through a microscope. Studies indicate that thin blood films are preferable to thick films because they are more effective in identifying the life stages of different species. Also the colour features of the components are clearly distinguishable in thin blood images. Thick films are better for confirming the presence of parasites but RBCs are destroyed when preparing the slide and hence morphological analysis is very difficult [22], [23], and [24].

Thin blood films are peripheral blood smears that are prepared on one side of a glass microscope slide with a thin layer of venous blood [25]. The slides will be stained usually using Giemsa stain or Wright's stain, processed with methanol to preserve the RBCs and examined using a microscope. If examined through a light microscope, the *Plasmodium* species, white blood cells (WBC) and platelets or artefacts will appear as saturated dark blue-purplish whereas the red blood cells (RBC) are lightly coloured.

Automated Malaria Diagnosis Using Mobile Phones

However, the visual effect of staining varies according to the lighting and imaging conditions of the microscope and can sometimes affect the diagnosis. The examinations are mainly conducted for the quantitative analysis of blood components such as RBCs and WBCs, detecting any deformations or abnormalities in the cell structures, diagnosis of infection and other blood related issues.

Figure 2.1 shows how a peripheral blood image will look like after Giemsa staining. The RBCs, also known as erythrocytes are usually lightly stained round structures and do not have nuclei. The diameter of RBCs is approximately 6-8 μ m. They have a central paler area in the centre, roughly 1/3 of the cell diameter [25]. However in certain erythrocyte abnormalities like sickle cell anaemia, the shape and size of the cell structure is significantly different.

The WBCs or leukocytes are the biggest cell components and their size varies according to the type. They are nucleated structures and depending on the morphology of the nucleus, WBCs are divided into eosinophil, basophil, neutrophil, lymphocyte and monocyte [26]. During staining, the nucleus closely binds to the stain reagents and gives a dark blue-purplish colour to the cell. The size of the WBCs can vary between 8 and 20 μ m in diameter. The size and staining concentration helps to easily identify the WBCs.

Platelets and artefacts are small reddish components present in the plasma of the stained blood with varying size of 2-3 μ m in diameter. During microscopic examination, the microscopists mainly use the size and staining concentration of the cells to identify and diagnose.

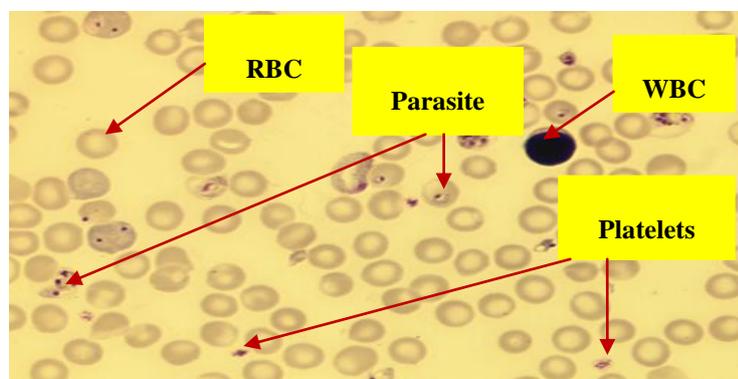
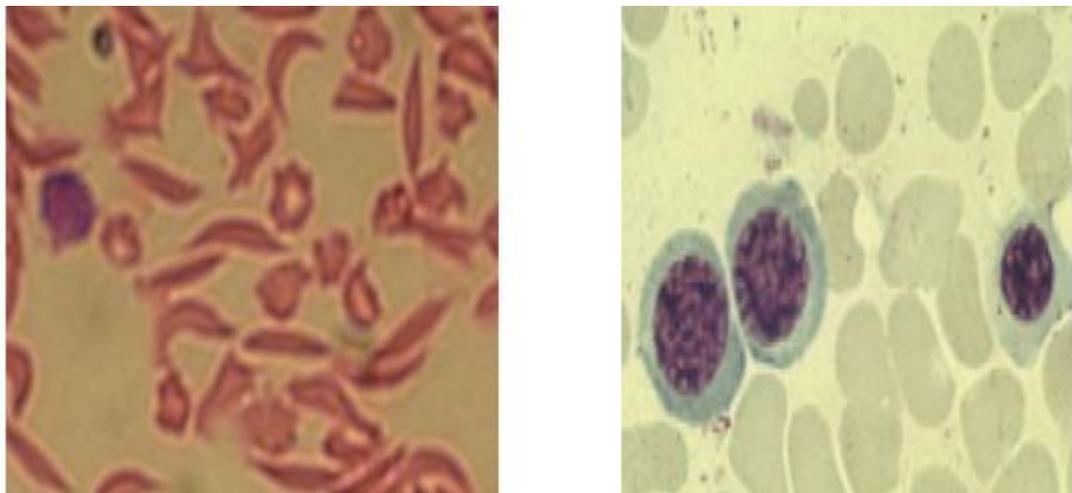


Figure 2.1: Microscopic image of a Giemsa stained thin blood smear

Automated Malaria Diagnosis Using Mobile Phones

2.2 Image Analysis of Thin Blood Slides

Automated image analysis considers several morphological and staining factors of the images for diagnosis and differentiation of the cells unlike examined by a trained person. As shown in Figure 2.1, microscopic images of thin blood films with malaria infection will often contain a mixture of large dark-blue leukocytes (WBCs), medium-sized pale-red RBCs and small purplish-red parasitic nuclei, platelets and artefacts. The cells can be easily differentiated in an image by a trained person using certain parameters such as size, shape, colour and other morphological factors (such as absence of nucleus for RBCs). However, for automated analysis the diagnosis can be spurious because these features are not always consistent. For example, Figure 2.2 describes two thin blood microscopic images in which the shapes of RBCs are highly distorted in the former (sickle cell anaemia), whereas RBCs are nucleated in the latter (Erythroblastemia). Similarly the anatomies of WBCs are different depending on their subtypes. Figure 2.3 shows examples of the main subtypes of WBCs. The dark stained area is the nucleus surrounded by very lightly stained cytoplasm within a cell membrane. In addition, some abnormal WBCs with structural deformation can be observed in some thin blood films.



(a)

(b)

Figure 2.2: Abnormalities in RBC cell structure.

(a) Sickle Cell Anaemia [26] (b) Erythroblastemia (Nucleated RBCs) [27]

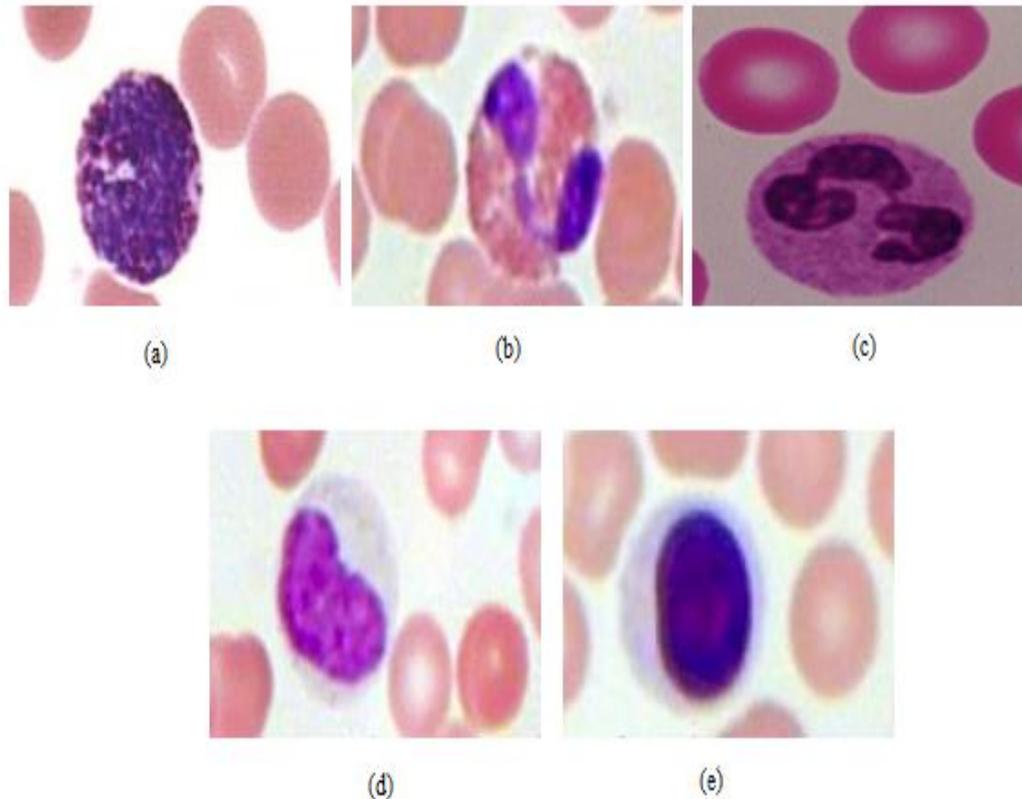


Figure 2.3: Subtypes of White Blood Cells

(a) Basophil (b) Eosinophil (c) Neutrophil (d) Monocyte (e) Lymphocyte [26]

During microscopic examination of malaria infected blood, pathologists look for small red-purplish nuclei residing in the RBC. However for automated image analysis for malaria, misdiagnosis might occur due to their structural resemblance with other components in the blood. For example, Figure 2.4 compares the size and shape of *Plasmodium spp* with other blood components such as thrombocytes, artefacts and leukocytes. The similarities in the morphology of the parasites with other components can impair the detection rate and specificity of the diagnosis. Hence proper cell structure analysis, colour information and size estimation is very important for automated image analysis.

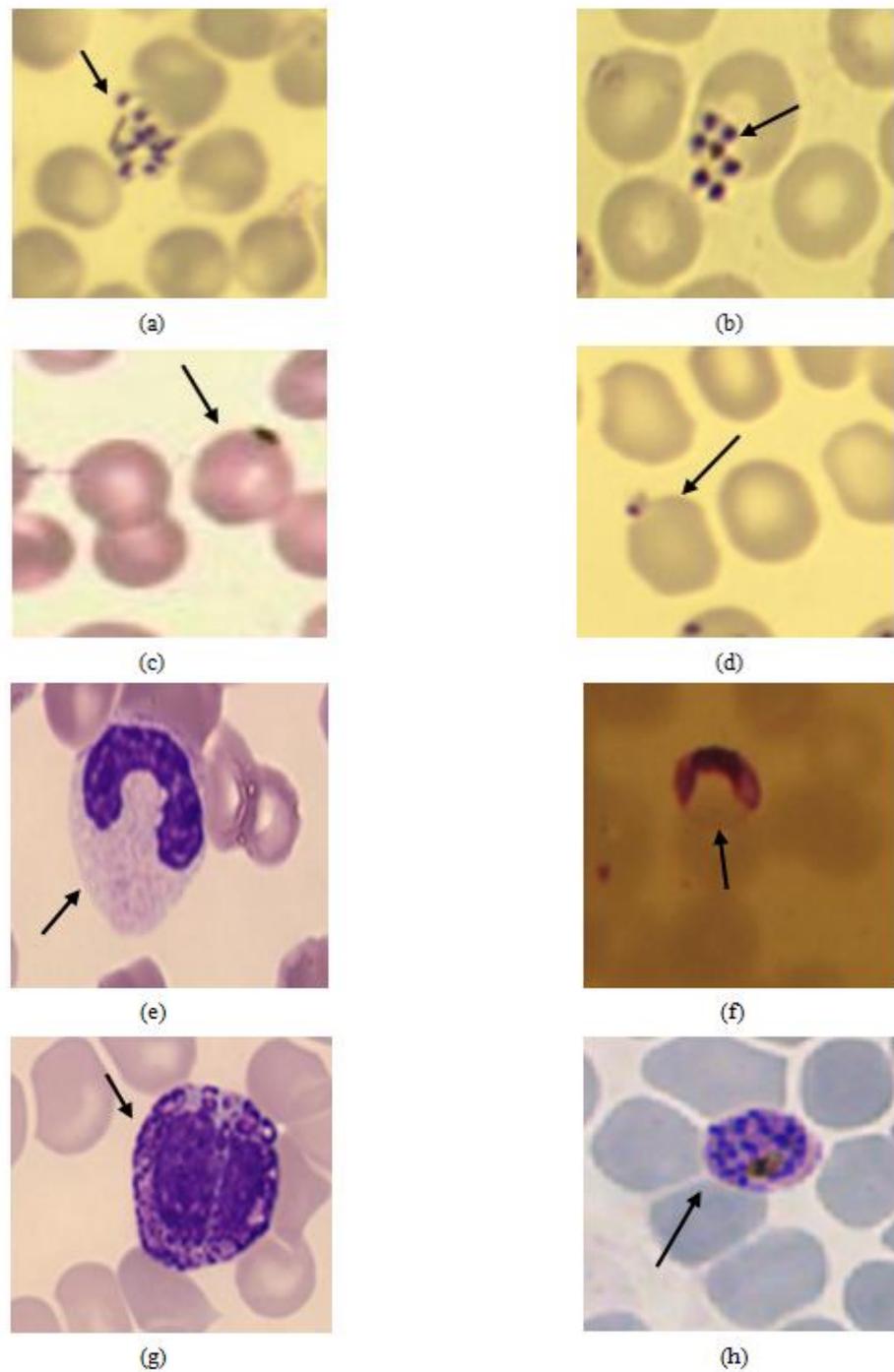


Figure 2.4: Similarities in size and shape of *Plasmodium spp* life stages with other cell components.

Fig (a) shows thrombocytes (platelet) whereas Fig (b) shows a schizont (asexual stage of merozoites with numerous nuclei) of the parasite. Fig (c) shows a platelet on top of an RBC whereas Fig (d) is the ring stage of *P.falciparum*. Fig (e) shows an Eosinophil (WBC) whereas Fig (f) is a gametocyte of *P.falciparum*. In Fig (g) the arrow points to a Basophil (WBC) whereas Fig (h) shows schizonts of *P.vivax*. Image (c) is from [26]; the rest are taken from the University of Westminster malaria image set.

Automated Malaria Diagnosis Using Mobile Phones

2.3 Performance Measurements

The performances of any diagnostic tests are determined by measuring the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). While sensitivity and specificity helps to measure the diagnostic accuracy of the tests, the PPV and NPV estimate the individual probability of the disease and vary according to the occurrence of the disease [107]. There are different interpretations of these measurements as in literature [108] and [109], in which the result is based on concluding whether a given blood sample is infected or not. However, in this research, the performances are measured based on whether an RBC is infected or not or in other words, by detecting the presence of one or more parasitic nucleus in the RBC by analysing the image of a thin blood smear. It checks the following four different parameters in order to measure the sensitivity, specificity and the predictive values.

- True Positives (TP): Detect an infected RBC (Test positive for a positive cell).
- True Negative (TN): Discard a non-infected RBC (Test negative for a negative cell).
- False Positive (FP): Detect a non-infected RBC (Test positive for a negative cell).
- False Negative (FN): Discard an infected RBC (Test negative for a positive cell).

The total number of occurrences of the above outcomes leads to the measurement of the following:

Sensitivity: Measures the proportion of infected cell detection. For our research, sensitivity tells how good the diagnostic test is to detect infected cells among the total number of infected RBCs. It can be calculated using the following formula:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.1)$$

Higher the sensitivity, greater the chance to detect the infection (or less chance to miss any infected RBCs).

Automated Malaria Diagnosis Using Mobile Phones

Specificity: It gives the probability of negative result for non-infected RBCs. It gives the proportion of number of non-infected cells discarded among a total of non-infected RBCs. Specificity can be calculated as:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2.2)$$

Higher the specificity, lesser the chances are to misdiagnose a healthy cell. If a healthy person is misdiagnosed with malaria, it can lead to serious consequences. It is equally dangerous to give a negative result to an infected person. Hence both specificity and sensitivity plays a vital role in the performance measurement of any diagnostic tests.

Positive Predictive Value (PPV): The PPV indicates the consistency of a positive result, or in other words, predicts the chances of a cell being infected if it is tested positive. It can be calculated as:

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2.3)$$

Negative Predictive Value (NPV): NPV is the opposite of PPV. It indicates the reliability of a negative result. It gives the probability of a patient being healthy when given a negative result. It predicts the chances of a cell being non- infected if it is tested negative.

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}} \quad (2.4)$$

Both PPV and NPV depend on the prevalence of the disease and the population of the positive and negative samples, or in other words, they are sensitive to the distribution of samples.

Accuracy: The overall procedure of performance measurement leads to the estimation of the accuracy of the system on a concrete set of samples. For any evaluation methods, accuracy estimation is necessary in order to improve the performance. The accuracy of the technique adapted in this research can be calculated using the above performance measurements.

Automated Malaria Diagnosis Using Mobile Phones

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (2.5)$$

Using the above formula, the accuracy can be calculated which will represent the average error occurred in a given set of samples containing unequally distributed infected and non- infected RBCs.

2.4 Current Trends in Digital Pathology

Digital pathology deals with virtual microscopy or Whole Slide Imaging (WSI) which is an automated, fast and high resolution imaging process. Glass slides with human tissues are observed and digitised to extract the information for diagnostic, research and clinical purposes. With the increasing demand for skilled microscopists and experts, coupled with higher processing time, manual microscopy contributes a fair share to the workforce pressure. Hence virtual microscopy using digital slide scanners can be a reliable diagnostic aid in the laboratory and enhance the speed of diagnosis, reduces labour and holds large amount of data for quantitative analysis [28].

Digital pathology equipments (digital slide scanners) are furnished with image acquisition hardware, image compression and processing and data storage facility. For image acquisition, the system follows conventional microscopic techniques. The prepared slides can be examined using different magnification depending on the type of tissue. For example the magnification used for blood slides are different from muscular tissue biopsy. Once the images are taken, they undergo compression for ease of storage. The images are processed using various image processing algorithms for cell segmentation, pixel classification, image registration for mapping and pattern recognition.

Digital pathology is mainly used for the WSI of human tissues for tumours and cancers [29], [30], and [31]. For these applications the WSI is performed under 40x magnifications. For fluidic or micro fluidic examination such as thin blood analysis, the magnification has to be increased to 100x and this could lead to large image data set and thereby more memory and storage requirements. Hence very few slide scanners such as [32], [33] has the WSI facility for thin blood smears with 100x magnification. Nevertheless there is an emerging trend towards automated thin blood smear image analysis and

Automated Malaria Diagnosis Using Mobile Phones

digitisation of thin blood smears are happening in the area of virtual microscopy, to which this research is hoping to contribute.

2.5 Automated Diagnosis of Malaria in Thin Blood Films

In the field of malaria diagnosis there is a great deal of research activity covering molecular diagnosis, pigment analysis, and blood smear analysis. However, this Section concentrates mainly on research conducted on peripheral blood images. The most recent work in this field is the work of Parkhi et al. [34] in which the diagnosis is performed using mathematical modelling using linear programming. The method aims to distinguish between parasites and healthy components of the blood by defining a mathematical model for the infected cells and graphically representing the result. The paper, however, does not provide enough experimental results, information on the amount of image data used and the accuracy achieved. The work also fail to address issues related to poor image quality, pre-processing requirements and problems with classification of stained and unstained components in the blood.

The work carried out by Vink et al. in [35] is exactly not a thin blood image analysis. However it is a valuable addition to computer aided malaria diagnosis where a vision-based malaria diagnosis system is developed by scanning a finger prick cartridge. The blood is stained using Acrylic Orange (AO) which is a nucleic acid binding fluorescent stain. In order to distinguish between the parasites and other blood components, a machine-based learning approach is used where trained classifiers are used to differentiate the parasites. The method is highly efficient for quantitative analysis with high specificity but has limited sensitivity. Especially in cases of low parasitemia, the method will fail to diagnose due to insufficient fluorescent staining due to lack of adequate nucleic acid. In addition the method is limited to infected cell identification without life stage recognition and the classifiers are trained only for *P.falciparum* detection and no other species identification has been addressed.

In the work of Prasad et al. in [36], malaria infected blood image is analysed and a decision support system was developed for use under the Android operating system in mobile phones for remote consultation. The method first segments the RBCs, using morphological erosion of binary image, discards any WBCs by applying a threshold, and

Automated Malaria Diagnosis Using Mobile Phones

parasites are detected using a 2D trained classifier and standard deviation. It also addresses the presence of other stained components such as platelets and artefacts and removes them using pre-classification techniques based on size and pixel connectivity. The method, however, has to send the results to remote experts who can then confirm the findings. The algorithm was first developed in Matlab where each image took approximately 120 seconds to process due to the complex neural network architecture used for classification.

An automated and unsupervised malaria diagnosis system has been developed by Purwar et al. in [37]. The method uses conventional edge detection techniques for RBC segmentation combined with Chan-Vese based boundary detection algorithms for region segmentation. The Hough transform [51] is used to determine predefined shapes of RBCs and other stained particles. A non-hierarchical probabilistic k -means clustering classifier is used to detect the parasites. The method possesses improved sensitivity but for a limited data set (approximately 40 images). It also fails to demonstrate the system accuracy on a variety of samples with different imaging conditions. The method is quite complex and involves a series of morphological operation including closing, thinning and spur removal.

The contributions of University of Westminster in the field of blood image analysis and malaria diagnosis have already been discussed in Section 1.6. Ruberto et al. [19][38], K.N Mohano Rao [20],[39],[40] and Tek et al. [42] have analysed and studied thin blood images. Ruberto proposed infected blood analysis using morphological operators and automated thresholding techniques using granulometry and regional extrema for estimation of size distribution [38]. Rao proposed an improved granulometry algorithm called area granulometry to estimate the size of objects and modified distance transform to deal with under and over segmentation issues which normally arise during differentiation of components in biomedical images [39],[40]. Tek worked mainly in the field of malaria diagnosis and developed a computerised malaria diagnosis system with image acquisition hardware. Tek proposed minimum area watershed segmentation method for segmenting the foreground objects in the blood image and addressed the issue of colour normalisation and noise reduction.

The aim of the research described in this thesis is similar to the above studies but is methodologically different from all of them. The main objective of the research has been to develop an unsupervised, less complex algorithm for malaria detection, capable of being

Automated Malaria Diagnosis Using Mobile Phones

implemented in a low-cost device such as a mobile phone or tablet. The work has the same objective as that described in [36] but is technically different because it avoids the use of trained classifiers and neural network architectures. Instead it uses less complex morphological operations, thereby reducing time and cost. For example, the processing time for each image analysis and diagnosis takes less than 40 seconds in Matlab which is just one third of the time taken in [36]. Moreover, the method provides on the spot diagnosis rather than sending the data for expert clinical analysis. The following chapters in this thesis explain the complete diagnostic process, related case studies and extensions of the work, techniques and challenges dealt within this research.

Chapter 3

Data Analysis and Morphological Processing of Peripheral Blood Images

In this chapter a detailed study and description of the image data used for the automated diagnosis of malaria and the pre-processing requirements is provided. As mentioned earlier, the aim of the research is to diagnose malaria using mobile phone microscopy. This is achieved through the following operations:

- Design and deploy suitable hardware for image acquisition;
- Develop image analysis and processing software;
- Integration of image processing software onto the mobile phone hardware platform.

This research primarily concentrates on developing reliable and fast image processing software for the blood image analysis. Even though the study did not attempt to build any hardware for image acquisition, the issues have been explored for future development and a possible design is explained in Chapter 10. The image processing software is developed using a relay of newly developed and existing morphological operations. These operations can be used not only for malaria diagnosis but also for the normal blood analysis and/or spatial analysis of images from other disciplines.

The images used for experimental purposes are acquired from the National Institute of Medical Research (NIMR), UK, Hospital for Tropical Diseases (HTD), London and Applied DSP and VLSI Research Group (ADVRG) Laboratory database of University of Westminster. A set of image were also acquired from the thin blood slides provided by the National External Quality Assessment Service (NEQAS) for Microbiology, using a mobile phone camera placed on the eyepiece of a microscope.

3.1 Image Acquisition and Processing

Examination of thin blood smears should follow specific procedures in order to gather maximum information. As shown in Figure 3.1 the blood smear is spread from thick end, where the cells are more clumped, to thin feathered end. The area from the centre

Automated Malaria Diagnosis Using Mobile Phones

approaching towards the feathered end is called the 'Zone of Morphology' (also called as Monolayer) that carries maximum information which is useful for inspection and diagnosis [41]. The microscopists first analyse the slide with a lower objective magnification (usually 20x) for monolayer detection which is then followed by examination under higher magnification. The zone of morphology which is the working area of the specimen for light microscope examination is usually considered to be 2 cm² [25]. The microscopists will scan this area, assessing each field with a diameter of 200µm at a time so that any morphological anomaly could be quantified and no repetitions of the fields occur. The scanning pattern is shown in Figure 3.1 below.

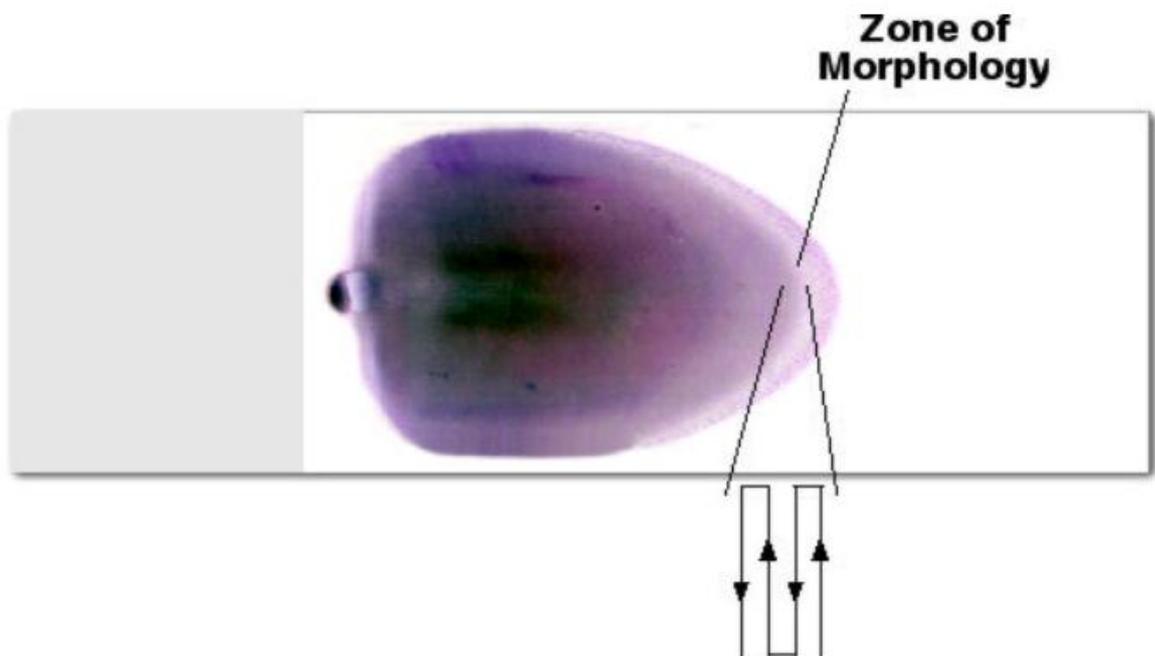


Figure 3.1: Scanning pattern of peripheral blood smear for morphological evaluation.

Reproduced from [25].

3.1.1 Digital Imaging of Thin Blood Films

Hypothetically, the digital imaging of thin blood smears is performed analogously as the above described scanning technique. The camera situated at the eyepiece of the microscope captures images while the slide will be moved in the XY direction similar to the scanning pattern in Figure 3.1. However the total number of images to scan the whole slide (WSI) depends on the magnification and field of view of the microscope. Ideally 1000x (100x Objective and 10x eyepiece) magnification is used for malaria diagnosis to calculate the

Automated Malaria Diagnosis Using Mobile Phones

parasitemia. For malaria diagnosis, WSI is not necessary since more than half of the smear does not carry the morphological information required for diagnosis.

According to [42] and [43], the total number of images captured at 100x objective magnification to cover the entire zone of morphology or monolayer is around 32,500, for a 1300x1030 pixel 2/3 inch CCD camera. If each image is considered to be each field, around 30,000 fields will be stacked for one slide analysis. This is not analogous to manual examination where only 100 fields are required for examination to give a negative decision and 25 fields (5000 RBCs in total at a rate of 200 RBC/field) to calculate parasitemia if an infected cell is found [16]. Apart from these, the selection of appropriate fields also remains an important aspect while automating the process. In [44] the authors propose a method to automatically select the working area on the blood smear using linear cascading classifiers. In this study, the proposed method for image analysis processes the images regardless of their non homogeneity, thereby avoiding any necessity of field selection procedure and a total of 4000-5000 RBCs per slides are examined to derive the sensitivity and specificity of the system.

3.1.2 Image Quality

The quality of the image is very important for automated diagnosis. Several parameters such as the size of CCD sensors, staining, auto focusing, lighting conditions, image compression level and enhancement play a vital role in achieving the desired image quality. The built-in electronics components and imaging sensors along with slide movements to capture successive fields can produce significant spatial and temporal noise. Another problem is the non-uniform illumination caused by the calibration issues of the microscope. For mobile-phone microscopy this issue can be more prominent as the specimen is illuminated by the LED sensors of the image acquisition set up attached to the phone. The limitations to provide suitable adjustable microscopic hardware such as condenser adjustments and variable iris worsen the illumination issues. Also the imperfection in the mobile phone lens and mismatch of field curvatures resulting in part of the field being focused (variation in focus), and other irregularities can lead to diffraction issues. However, with the computational and imaging advancement in mobile-phone technology, high-resolution images could be taken and saved using lossless compression in

Automated Malaria Diagnosis Using Mobile Phones

order to retain maximum information. In this work, the morphological approaches used for segmentation overcome the problem of non-uniform illumination. This is further aided by optimising the RGB weights used in the gray-scale conversion. In addition, a novel morphological filtering operation is performed which removes artefacts on the image caused by dust or other particles on the microscope optics.

3.2 Analysis of Thin Blood Images

Giemsa stained thin blood images can be divided into three regions based on their intensity or value of each pixel and on the amount of light registered by the photosensitive device. These regions are (1) highly stained parasites, WBCs and artefacts, (2) lightly stained RBCs and (3) unstained plasma. In conventional processing methods the blood images are usually divided into foreground and background regions. For parasite diagnosis, stained areas are further discriminated from foreground regions using various trained classifiers. However, in this research, the blood images are segmented and foreground objects are located based on the pixel intensity level which depends on the concentration of staining and are further classified based on morphology.

Two factors are mainly considered for processing of malaria infected blood image. Firstly, the plasma is considered as the background of the image and all the other cell/components as the foreground objects. Secondly the parasites reside in the RBCs are regarded as dark component on a lighter background. Hence the RBCs in the foreground regions are further sectioned as infected (background for dark components) and non-infected (with uniformly distributed connected pixels). The following Section describes the regionalisation process based on histogram analysis to extract the components of the thin blood image with *plasmodium* infection.

3.2.1 Pixel Extraction Using Gray-Level Histogram

In Figure 3.2, a malaria infected thin blood image in Red-Green-Blue (RGB) colour space is shown along with the histogram of the image with distribution of intensities in gray scale or monochrome colour space. The image clearly shows the three regions described above: unstained plasma, lightly stained RBCs and deeply stained parasite nuclei, platelets and artefacts. The image can be binarised in order to separate the three regions. The simplest

Automated Malaria Diagnosis Using Mobile Phones

approach to image binarisation is thresholding, which can be performed in various ways as described in [44], [45], [46]. However, since image binarisation is not a prominent procedure in this thesis and is being performed to have an initial rough estimation of the regions in the image, a most instinctive approach of thresholding using gray level histogram is adapted.

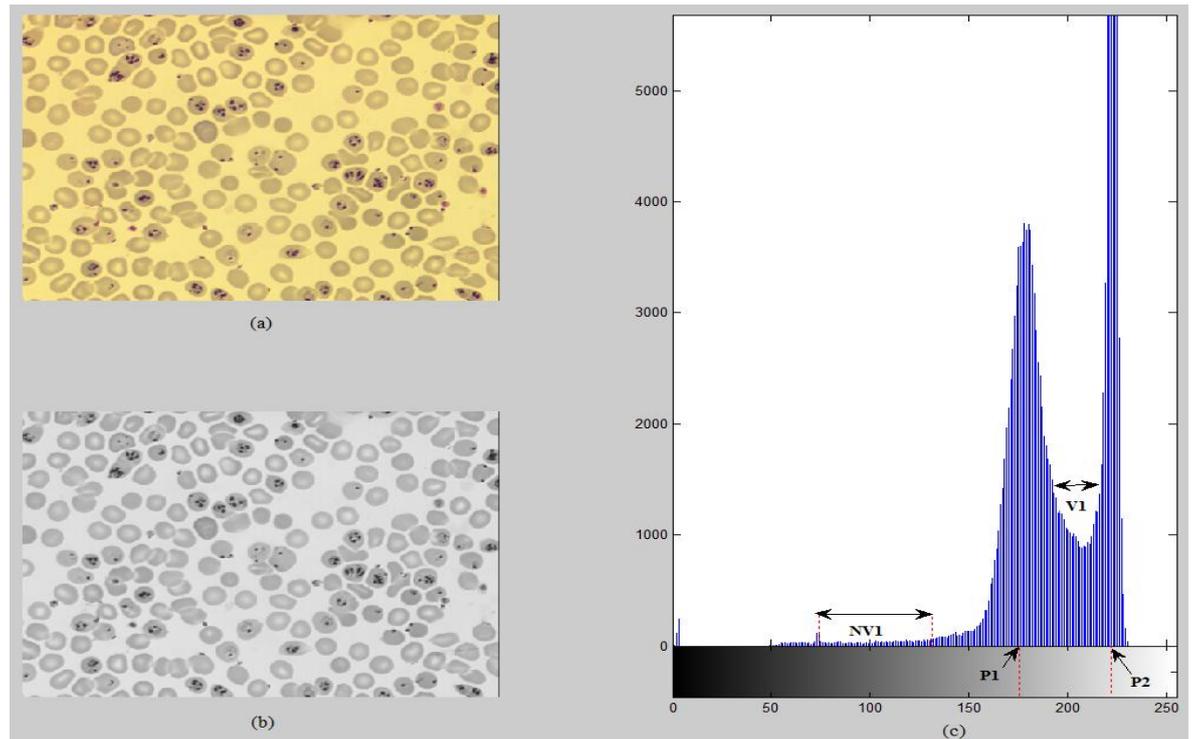


Figure 3.2 Peripheral Malaria Infected Thin Blood Image.

(a) Thin blood image in RGB colour space. (b) Gray scale image. (c) Histogram of the image. P1 and P2- Magnitude (intensity value) of the peaks representing foreground and background objects, V1- Valley between foreground and background object peaks used to set the threshold for separating the components, NV1- flat and broad valley with artefacts, platelets and noise. (Subset of foreground components)

Automated Malaria Diagnosis Using Mobile Phones

The procedure uses two threshold values, which is calculated as follows:

$$T1 = \frac{P1+P2}{2} \quad (3.1)$$

$$T2 = \frac{\min(NV1)+\max(NV1)}{2} \quad (3.2)$$

Where P1 and P2 are the intensity values of the two peaks and the min (NV1) and the max (NV1) are the pixel intensity values at the beginning and end of the valley, NV1, respectively. Using the threshold $T1$, a binary image is extracted that has been used as a background mask. A complement of the background mask can be used as a foreground mask. The threshold $T2$ is used to extract the parasite, platelet and other noises in the image. A non-zero difference of the this image and the background mask will produce a binary mask of the stained particles other than RBCs. Using the gray-level image as the marker, the binary masks will extract the background, foreground and stained components in the foreground region as shown in Figure 3.3.

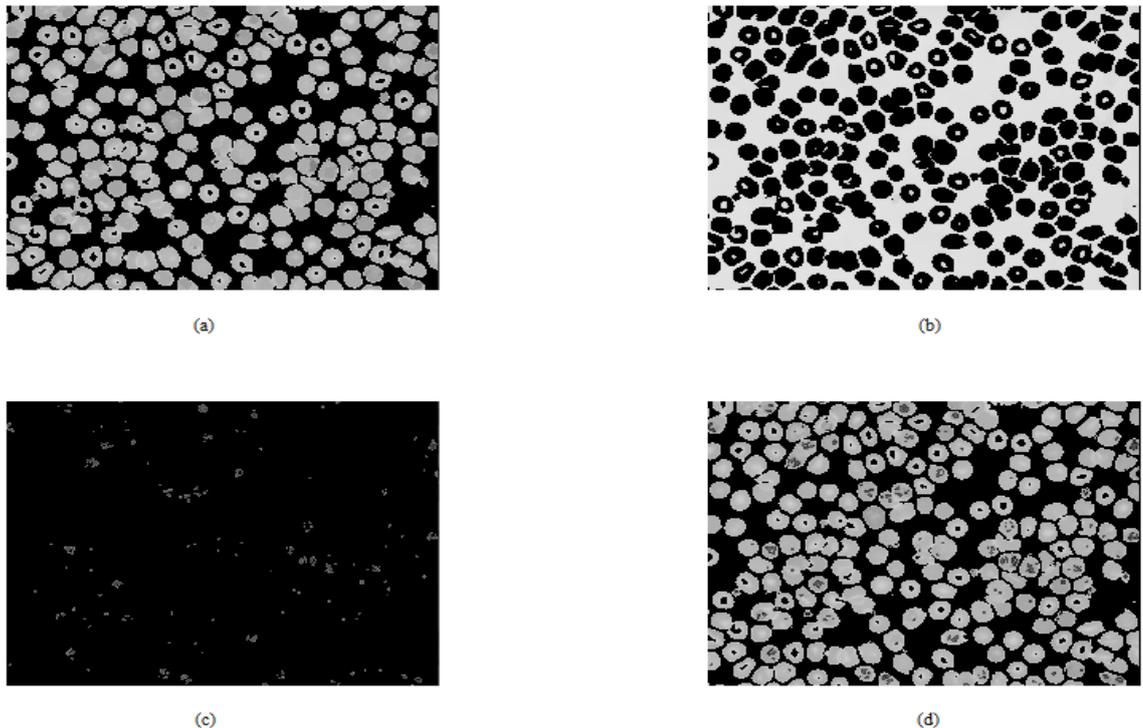


Figure 3.3 Foreground, background and stained pixel extraction.

- (a) Gray level representation of foreground objects (RBC) (b) Gray scale representation of the background (Plasma). (c) Gray level representation of deeply stained components (Parasite nuclei, artefacts and platelets). (d) Gray level image with (a) and (b) superimposed.

Automated Malaria Diagnosis Using Mobile Phones

Even though the method is a useful tool to separate the components, it is not dependable to give robust segmentation of the blood components. The peaks of the histogram do not always give reliable estimates of the threshold values because of various factors such as introduction of noise, non-uniform illumination, high level of pigmentation and difference in staining concentration. The thesis deals with these issues in the following chapters and the rest of the Sections explain the pre-processing techniques that were experimented prior to the segmentation of the image.

3.3 Improved Gray-Scale Conversion Using Contrast Adjustment

As discussed in Section 3.2, the preliminary aim of blood image analysis for malaria parasite detection is to recognize different objects present in the image prior to differentiating them as parasites and non-parasites. The foreground region of an infected blood image consists of RBCs, WBCs, parasites, platelets and any artefacts or noises induced by various other imaging factors. This research introduces a sequence of image processing techniques to differentiate these components and remove those that are not necessary for diagnosis. For reduced complexity and faster computing, the image analysis and processing is performed on gray-scale images rather than RGB colour images.

3.3.1 Gray-Scale images of Thin Blood Smears

Conversion of a colour image to gray scale is a common process in image processing which involves converting a 24 bit colour value to 8 bit gray value [47]. The gray scale of a colour image represents the luminance of the pixel value ranging from 0 to 255. Different methods such as simple averaging $((R+G+B)/3)$, maximum and minimum averaging $([\text{Max}(R, G, B) + \text{Min}(R, G, B)]/2)$ and weighted averaging (averaging the R, G, B values with chosen weights) are commonly used for gray scale conversion. For the processing of thin blood images, choosing an appropriate gray scale conversion technique is very important since any loss of information during the conversion of one colour space to other can impair the accuracy of the diagnosis. In this study the weighted average conversion method is used because it preserves more feature details than other conversion techniques. This method is based on the sensitivity of the human eye and uses optimised weights for each component.

Automated Malaria Diagnosis Using Mobile Phones

The formula used for gray scale conversion is given below:

$$\text{Gray} = W_R R + W_G G + W_B B \quad (3.3)$$

Where W_R , W_G and W_B are the optimum weights of Red (R), Green (G) and Blue (B) component values of the image respectively. The weights of all the three channels must sum to 1. This reduces the degrees of freedom from 3 to 2 in choosing appropriate weights. The standard gray scale conversion uses a weighted sum of R, G and B components with $W_R = 0.299$, $W_G = 0.587$ and $W_B = 0.114$ [51]. Previously published work suggests that maximum weight for the green channel gives better contrast for blood images since it maintains high frequency feature information and distinguishes contrasts better than other channels [19], [48]. In [49], the authors prove that using the green channel alone by keeping W_R and $W_B = 0$ gives higher contrast between structures under different backgrounds and thus is best suited for blood imaging. In the following Section of this chapter, a prior analysis of the contrast of the images and an experiment to find the optimum weights of R, G and B channels suitable for conversion of blood images to gray scale without losing the morphological information has been explained. The Section also describes a simple contrast enhancement technique using weighted gray-scale conversion for thin blood images which avoids the oversaturation of certain features in the image that usually occurs in global contrast enhancement techniques.

3.3.2 Contrast Enhanced Gray-Scale Conversion of Blood Images

In this experiment the optimum weights for R, G and B are identified for gray-scale conversion by plotting the contrast of the image as a function of the mix. Since image contrast is the difference in intensity between the max and minimum value and depends upon the domain distribution of light and dark regions in the image [50], this experiment is well suited for blood images which possess a range of intensity levels due to the difference in staining concentration.

Contrast of an image could be measured as a ratio of difference in higher and lower luminance to average luminance of the image. There are different techniques to achieve this. In [51] two contrast measurement techniques, Weber contrast measurement and Michelson contrast measurement, are described. Both deal with the spatial distribution of

Automated Malaria Diagnosis Using Mobile Phones

the dark and light features and background of the image. Weber contrast is normally applied on images having uniform background and is calculated by measuring the ratio of difference in intensities between the background and the features to the background intensity. Michelson contrast is commonly used for extreme dark and light patterns. It calculates the ratio of difference in maximum and minimum intensities to twice the average intensity. Weber contrast is best suited for images with uniform background intensities whereas Michelson contrast deals with images of high spatial frequency. Since this research deals with two dimensional (2D) blood images of size $M \times N$ with a non-uniform background and foreground features of varying intensities, both methods when applied, did not measure the actual contrast and hence are not suitable for blood image analysis. In fact Michelson contrast measurement technique provided better result than Weber technique, but since it deals with upper and lower limits of intensities, the results were not informative for images with low level infections where extreme feature differences were not present. Consequently, for this application, contrast was measured using a Root Mean Square (RMS) approach which does not depend on the spatial frequency content of the image [51].

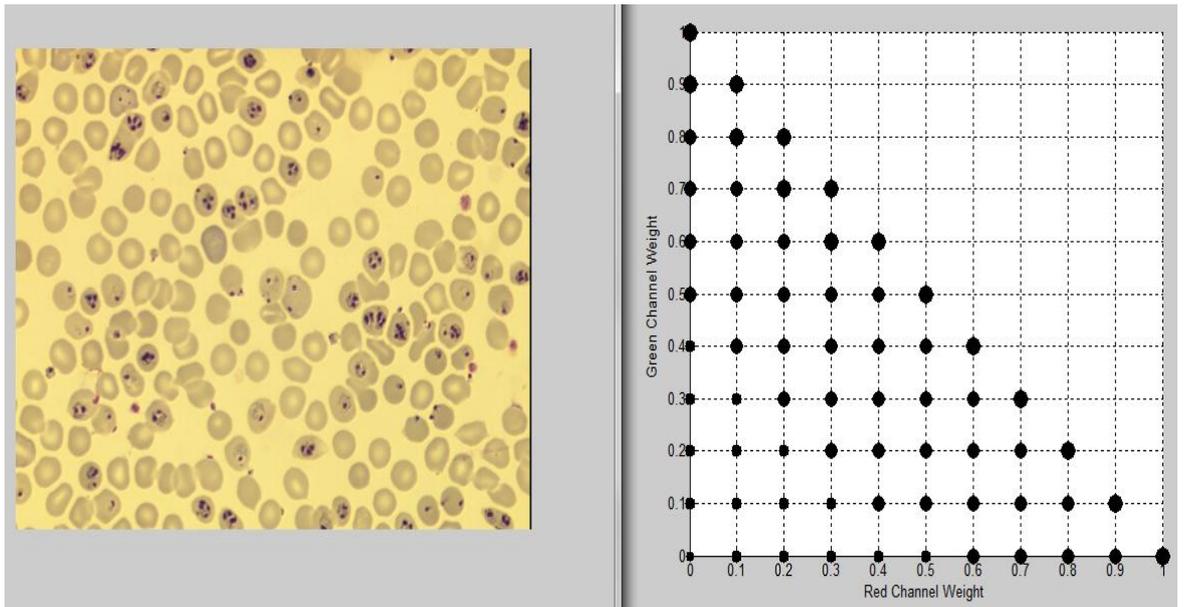
RMS contrast measures the standard deviation of pixel intensities in the image and it is formulated as follows:

$$C_{\text{RMS}} = \frac{1}{\sqrt{MN}} \sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I_{ij} - I_{\text{avg}})^2} \quad (3.4)$$

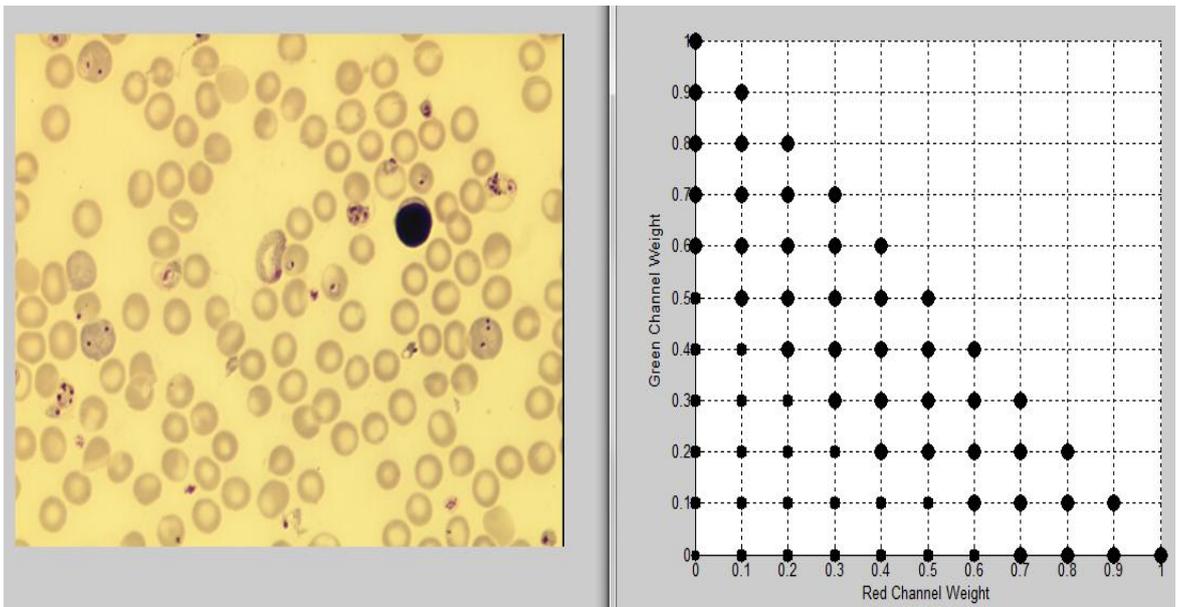
Where I_{ij} is the normalised pixel intensities at the i^{th} and j^{th} element of the image of size M by N and I_{avg} is the mean normalised pixel intensity value.

Figure 3.4 shows the plots of the contrasts of the blood image with respect to the different weights of R and G components ranging from 0 to 1. The weight of B component depends on the other two in order to get a total sum of 1. The images of varying illumination and staining concentration have been chosen to study the effect of optimum weights and the resultant plots are given below.

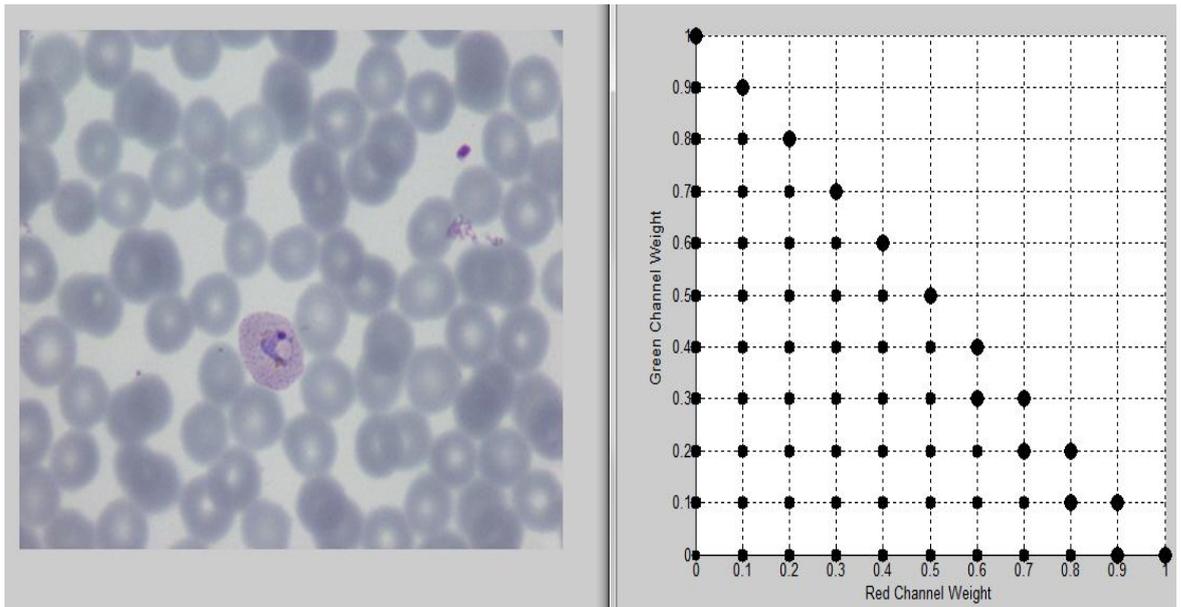
Automated Malaria Diagnosis Using Mobile Phones



(a)



(b)



(c)

Figure 3.4: Contrast plot as a function of red and green weights for gray scale conversion.

Images in (a) (b) and (c) are prepared under different illumination and staining concentration. The stained component features in the images are also different. The x-axis of the plot represent the weights of red (W_R) and y-axis represent the weight of green channel (W_G), both ranging from 0 to 1.

Figure 3.4 represents three images taken under different illumination and /or staining concentration. For example, the illumination of image in (c) is very different to that of (a) and (b). The concentration of staining in all the three images is different and so does the features of the blood components. The contrast values were plotted in 3D with Red, Green and Blue value form the X, Y and Z axis respectively but for ease of understanding, a triangular domain is shown with red and green weights. The red and green weights form the co-ordinate variables with blue along the plane (Z axis). The plot displays RMS contrast as black circles at the location specified by the x, y and z components. The area of each marker represents the contrast value, hence, larger the circle higher the contrast at those weights of R, G and B. The plot indicates that better contrast is achieved along the line where weight of blue = 0 ($W_B=0$) and the optimum weights could be chosen as a mix of red and green along the line which indicates $W_R+W_G=1$. Also to be noted is that even though the green component of the images ($W_G=1$, $W_R=0$ and $W_B=0$) provides the highest contrast for images (a) and (b), for image (c), the maximum contrast is high at both ($W_G=1$, $W_R=0$ and $W_B=0$) and ($W_G=0$, $W_R=1$ and $W_B=0$). This might be presumably because the features of the stained blood components in (c) are not as sharp as the images in (a) and (b) due to difference in the illumination spectrum. Based on the conclusions

Automated Malaria Diagnosis Using Mobile Phones

derived from the above image analysis, for particular illumination set ups, this experiment should be part of the standard procedure to calibrate the conversion weights and thereby enhancing the contrast and preserving the feature information simultaneously.

Another experiment for contrast enhanced gray scale conversion was also conducted as part of the pre-processing set up where a single image is segmented and undergoes gray-scale conversion using different sets of weights applied to different portions of the image in order to avoid over-saturation of certain areas and to increase the appearance of light-dark transitions. The analysis of Figure 3.4 revealed that, for the regions of non-infected RBCs, a higher weight for the red channel defines features better whereas for stained components a higher weight for the green channel preserves more information. Hence the image is segmented by thresholding, using the histogram approach described in Section 3.2.1, into three different regions: plasma, RBCs and deeply stained areas and suitable weights are applied which are derived by measuring the contrast at respective combinations. The contrast enhanced images along with its red, green and blue components are shown in Figures 3.5-3.7.

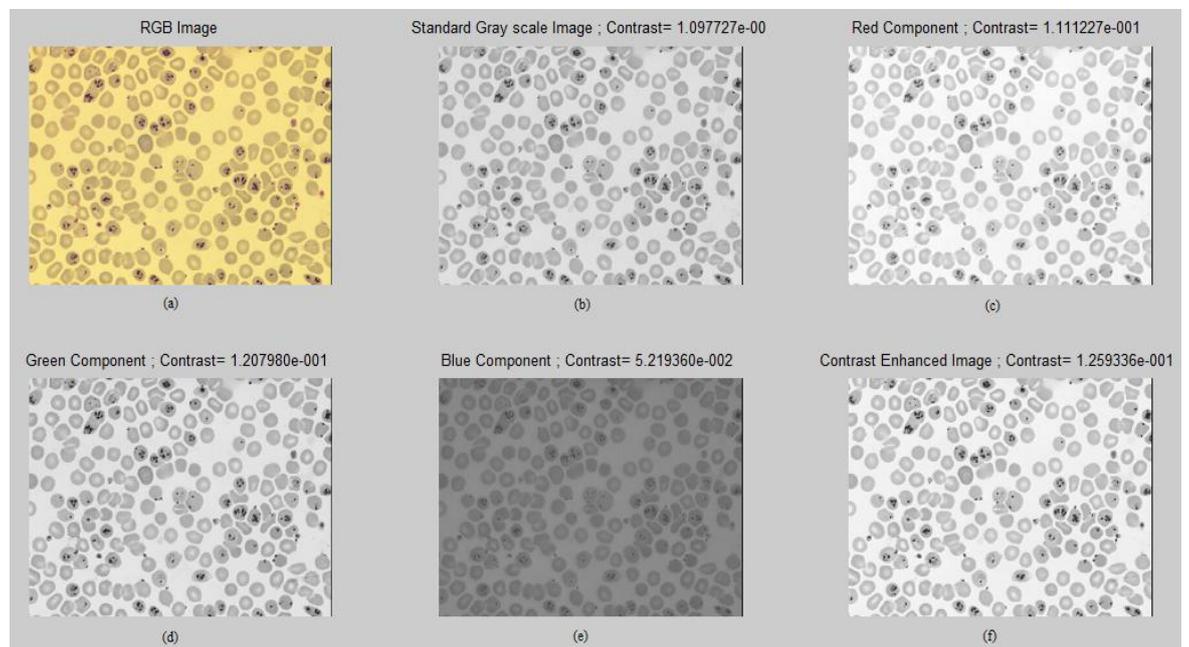


Figure 3.5: Contrast enhancement using segmented weighted averaging.

(a) Colour Image. (b) Standard gray-scale conversion using $W_R=0.2989$, $W_G=0.5870$ and $W_B=0.1140$. Image contrast = 0.1097. (c) Red component of the image $W_R=1$, $W_G=0$ and $W_B=0$. Image contrast = 0.1112. (d) Green component of the image $W_R=0$, $W_G=1$ and $W_B=0$. Image contrast = 0.1207. (e) Blue component of the image $W_R=0$, $W_G=0$ and $W_B=1$. Image contrast = 0.0521. (f) Contrast-enhanced image using segmented averaging. Image contrast = 0.1259.

Automated Malaria Diagnosis Using Mobile Phones

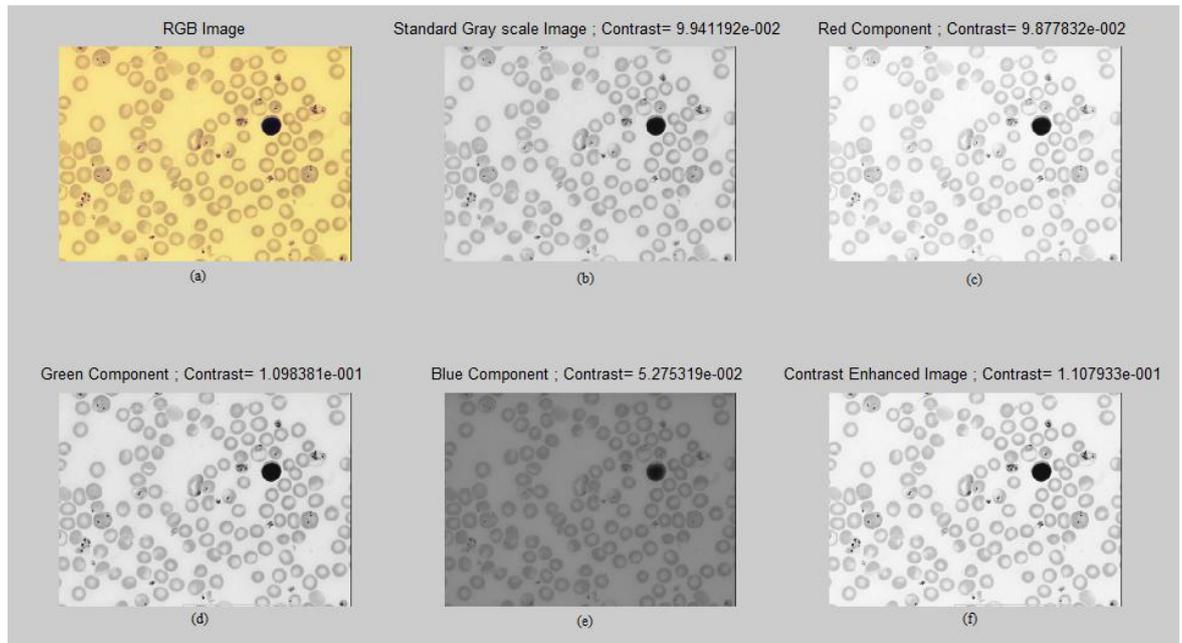


Figure 3.6: Contrast enhancement using segmented weighted averaging.

(a) Colour Image. (b) Standard gray scale conversion. Image contrast = 0.0941. (c) Red component of the image. Image contrast = 0.0987. (d) Green component of the image. Image contrast = 0.1098. (e) Blue component of the image. Image contrast = 0.0527. (f) Contrast enhanced image using segmented averaging. Image contrast = 0.1107

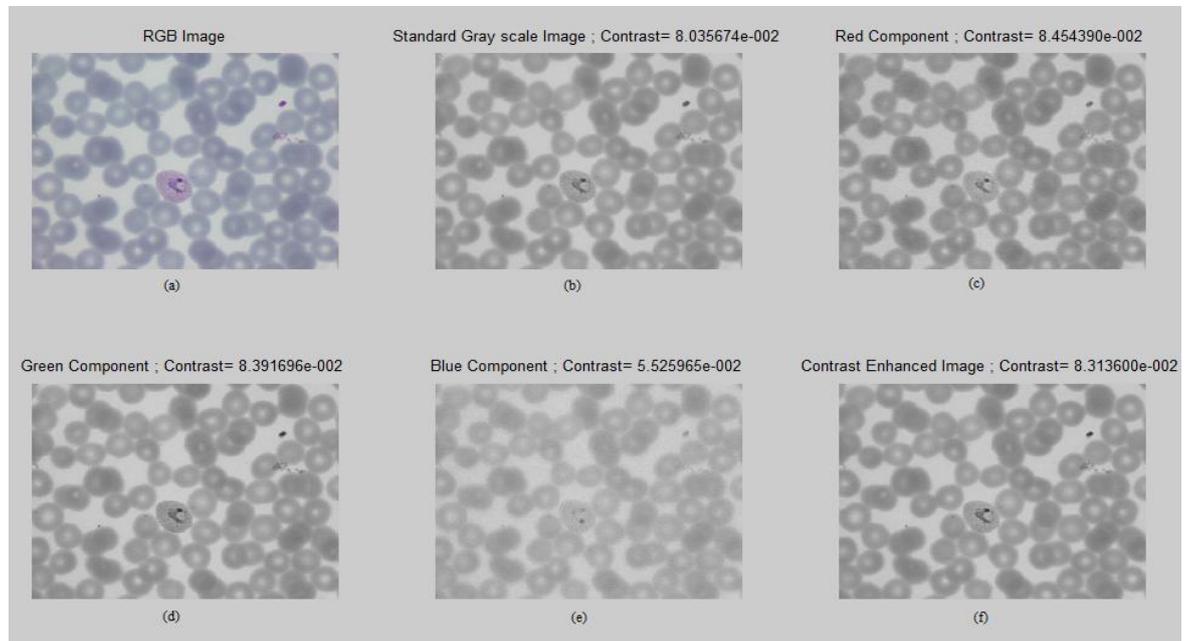


Figure 3.7: Contrast enhancement using segmented weighted averaging.

(a) Colour Image. (b) Standard gray scale conversion. Image contrast = 0.0803. (c) Red component of the image. Image contrast = 0.0845. (d) Green component of the image. Image contrast = 0.0839. (e) Blue component of the image. Image contrast = 0.0552. (f) Contrast enhanced image using segmented averaging. Image contrast = 0.0831.

Automated Malaria Diagnosis Using Mobile Phones

From the above figures, it can be concluded that segmented weighted averaging enhance the contrast, providing better feature definition than standard gray scale conversion. However, unlike Figure 3.5 and 3.6, the contrast value of the enhanced image is less than that of the red and green component in Figure 3.7 due to the difference in illumination spectrum and staining.

3.3.2.1 Conclusion

Both experiments conducted above concludes that ignoring the blue channel always enhance the image contrast. Depending on the nature of light source used, a mix of optimised weights of red and green components calibrated along the $W_B=0$ line would be an ideal choice for gray scale conversion of stained blood images.

The experiments conducted above are very useful for blood image analysis and processing, since it preserve useful colour information of the foreground objects prior to segmentation. However, in this research a novel blood cell identification technique has been developed, as explained in Chapter 4, which does not require any contrast enhancement and can work on image with poor feature quality. Hence the above experiments are not very critical at this point. Nonetheless, the analysis method described above can be used to optimise the gray scale conversion in any application involving contrast enhancement. This can be used as a calibration technique to take into account any variation in the spectral properties of the light source and staining characteristics.

3.4 Morphological Filtering

This Section explains the morphological filtering operation to remove the platelets, artefacts and noises in the Giemsa-stained blood image. As shown in the previous figures, the platelets and artefacts are highlighted by the staining. Platelets are small round or oval cell structures, 2-3 μm in diameter. Artefacts and noise are present in the image due to various factors such as chemical processing while staining and preserving the blood smear, noise from the microscope lighting or from the scanner. In order to separate the important components such as RBCs and WBCs in the image, it is necessary to distinguish and eliminate the platelets and other artefacts. Any stain/non-stain classification tool used to separate these objects tends to consider them as either the parasite or a white blood cell.

There are some fundamental criteria to distinguish them from other stained components in the image such as size and location. The size of the platelet or artefact is smaller than other stained cell components. However, the size cannot be used as the only parameter to distinguish them as the stained nuclei of the parasites are of similar size too. The location of the platelets/artefacts can be used to discriminate them from the parasitic nuclei since the parasites reside within the RBC while platelets/artefacts are usually found outside the RBCs. However there are cases in which the platelets overlap on the RBC surface and are misdiagnosed as parasites hence removal of these bodies from the image is of greater importance. The following Section describes the algorithm developed for platelet removal, in which all the blood components less than the RBC size will be removed by performing morphological filtering.

3.4.1 Morphological Closing

Morphological closing operation involves two primitive operations- erosion and dilation. A morphologically closed image is obtained by performing dilation followed by erosion using a Structuring Element (SE). Closing operation is usually performed to smooth the image and fill gaps and holes. In this research, the closing operation is performed to remove all the stained components smaller than the RBCs and to remove the centre light patch within the RBC, thereby preserving a uniform intensity distribution in the RBC regions. However, unlike conventional closing which involves the same structuring element for both dilation and erosion, two different structuring elements are used to perform these operations.

Dilation (denoted by the symbol(\oplus) is a procedure which replaces the original intensity value with the largest (i.e. lightest) value in the neighbourhood. For an image with darker objects in the lighter background as in the case of blood images, dilation reduces the size or fully eradicates any darker object smaller than the SE. Let \mathbf{I} be the gray scale image and se be the structuring element. The dilated image \mathbf{I}_d is calculated using the formula:

$$\mathbf{I}_d = \mathbf{I} \oplus se \quad (3.5)$$

Automated Malaria Diagnosis Using Mobile Phones

The dilation of the gray-scale image \mathbf{I} at any location (x,y) by the structuring element se is calculated to be the maximum value of the neighbourhood outlined by se whose origin is at (x,y) and is given as [51] :

$$\mathbf{I} \oplus se = \max_{(p,q) \in se} \{f(x-p,y-q)\} \quad (3.6)$$

Erosion and dilation are dual operations which complement and reflect with respect to connected sets [51]. Erosion (denoted by the symbol (\ominus)) expands the size of darker object in a lighter background up to the size of the structuring element as it will pick the minimum value from the neighbourhood.

For a gray-scale image \mathbf{I} with a structuring element se , the eroded image \mathbf{I}_ε is calculated using the formula:

$$\mathbf{I}_\varepsilon = \mathbf{I} \ominus se \quad (3.7)$$

And for location (x,y) ,

$$\mathbf{I} \ominus se = \min_{(p,q) \in se} \{f(x+p,y+q)\} \quad (3.8)$$

Where x and y are the location of the pixel undergoing erosion. A closing operation (dilation followed by erosion) attenuates the darker regions (stained) depending on the size of the structuring element and hence could be used to remove the artefacts, platelets and other noises but retain the RBCs and WBCs. The closing operation performed on a thin-film blood image using a disk shaped structuring element with a radius the same as that of the radius of the RBCs is shown in Figure 3.8.

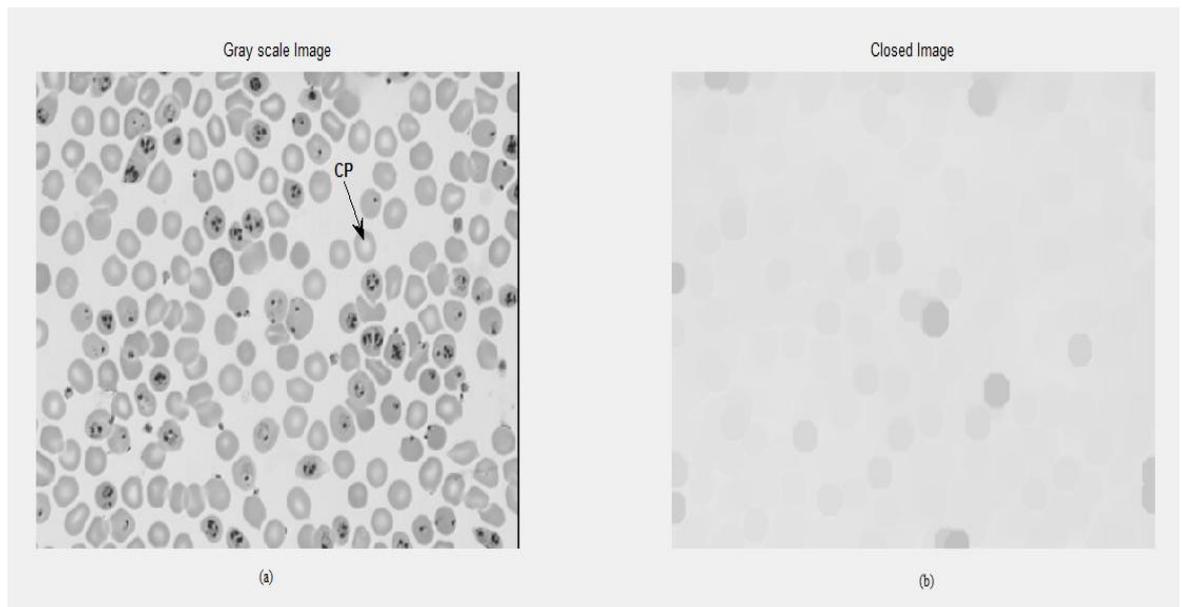


Figure 3.8: Morphological closing of a gray-scale thin blood image.

(a) Gray-scale image. CP is the light centre patch within the RBC. (b) Closed image.

As observed in Figure 3.8, the closed image (b) has lost the small particles so all the platelets, artefacts and noises have been eliminated. However, the RBCs have become much fainter and almost 80% have disappeared completely! This is due to the presence of the light patch in the centre of most RBCs (CP Figure 3.8(a)). This attenuates the actual RBC intensity during dilation. To avoid this situation and to retain a uniform pixel intensity value in the RBC region, a novel approach is used where two different structuring elements are used to perform for dilation and erosion.

3.4.2 Morphological filtering for the removal of platelets and other noises

In this approach the gray-scale image undergoes dilation using an annular ring structuring element followed by erosion of the dilated image using a disk shaped structuring element. The two structuring elements used for dilation and erosion are shown in Figure 3.9. The inner and outer diameter of the dilation ring is set to 35% and 70% of the typical RBC diameter respectively. The erosion disk has a diameter of 70% of the RBC diameter. The radius of the structuring element depends on the radius of the RBC, so that any component smaller than the RBC will be removed upon completion of these operations.

Automated Malaria Diagnosis Using Mobile Phones

As part of the initial calibration procedure, the radius of the RBCs are determined by performing a granulometry operation which estimates the particle size distribution using a series of closing operation with structuring elements of increasing size and computing the difference between successive closings. Because of the hole in the structuring element, the dilation operation does not replace the darker pixel values of the RBC rim with the lighter values in the centre. Figure 3.10 compares the effect of using a disk-shaped SE and a ring-shaped SE on the same image shown in Figure 3.9(a). The intensity of the RBCs is hugely reduced in (a) but is retained in (b). The erosion operation on the dilated image will eventually restore the size of the RBC and the resultant closed image will have all the foreground blood cells with uniform intensity distribution around RBC as shown in Figure 3.11.

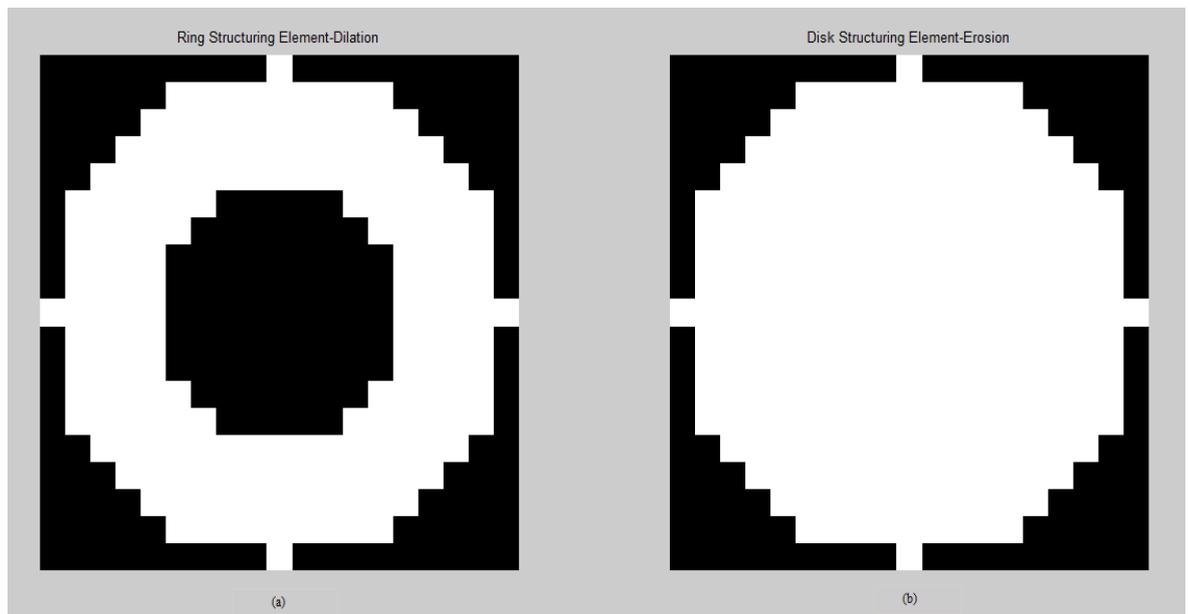


Figure 3.9: Structuring elements for morphological filtering.

(a) Ring-shaped SE for dilation (b) Disk-shaped SE for erosion

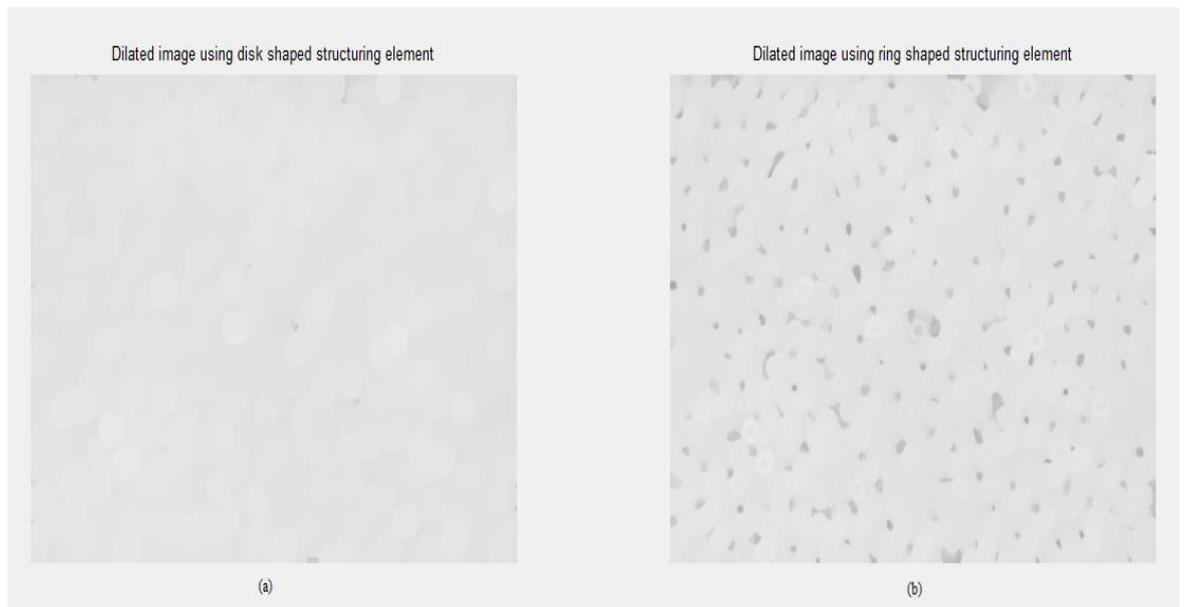


Figure 3.10: Morphological dilation

(a) Dilation using disk-shaped structuring element (b) Dilation using ring-shaped structuring element where the RBC intensity information is retained.

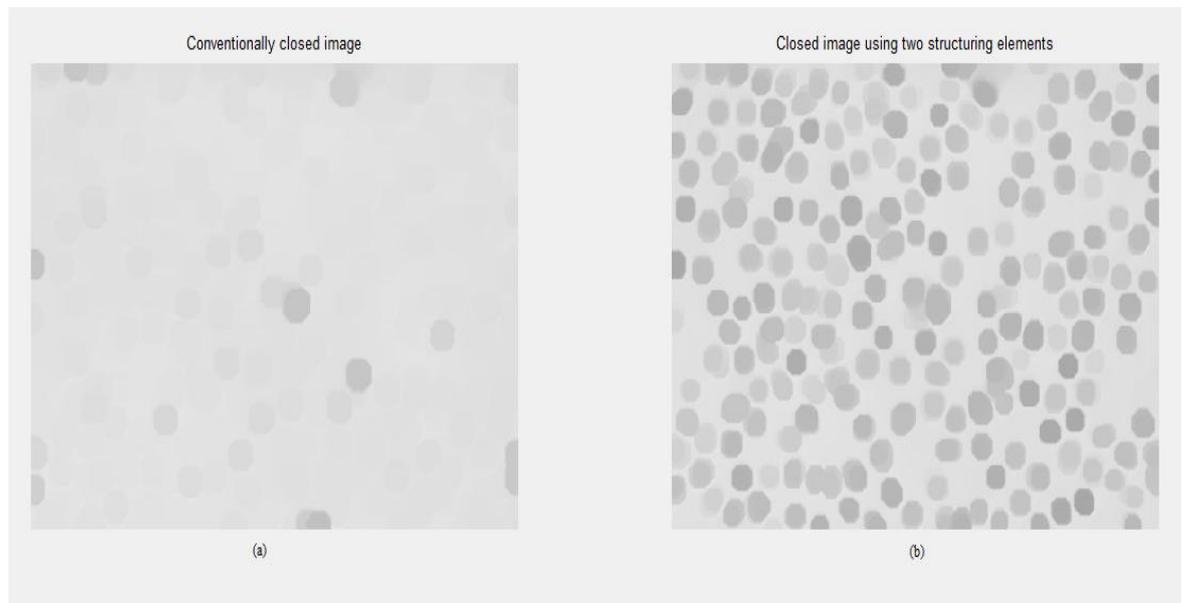


Figure 3.11: Morphologically closed images.

(a) Image closed using the same structuring element for dilation and erosion (b) Image closed by performing dilation using a ring structuring element followed by erosion using a disk-shaped structuring element.

Automated Malaria Diagnosis Using Mobile Phones

As shown in Figure 3.11 (b), the resultant closed image retains all the RBCs and WBCs present while filtering out the other components.

The method is highly reliable in tackling common issues faced during the blood image processing such as heterogeneous intensity distribution of foreground objects, light central pallor in the RBC, and artefacts and unwanted platelets created during blood preparation and imaging. Unlike conventional image processing algorithms which uses complex pre-processing procedures which involves filtering using median filter, subtraction technique or morphological closing or opening, contrast enhancement or colour normalisation procedures for better visibility and uniform illumination of foreground components and techniques such as ‘hole filling’ and connected component extraction algorithms to remove the light centre pallor within the foreground components [35],[37],[42],[53], the method reduces the complexity of the pre-processing procedure by dealing the issues simultaneously while retaining the essential features of foreground objects (RBC and WBC).

3.5 Summary

The chapter deals with thin blood image analysis and discussed the regionalisation of objects in the image through histogram analysis. A case study on gray-scale conversion was conducted and a tool to identify the optimum weights for a contrast enhanced gray-scale conversion has been proposed. The chapter also introduced a new algorithm for morphological filtering of the blood images as a pre-processing tool for segmentation and thereby eliminating any additional requirements for noise removal. Conventional morphological closing on blood images removes the unwanted components but also leads to loss of valuable information. The proposed morphological filtering preserves the necessary information of foreground components while removing the noise and artefacts. This method could be subjected to modifications to adapt to the pre-processing of other pathological images as well such as tissue analysis and cell differential analysis.

Chapter 4

Blood Component Detection and Estimation

This chapter describes a novel segmentation procedure used to identify the location of the RBCs in the given area of the thin-film image and estimate the population. Segmentation is the process of partitioning the image into different segments and it is usually application specific. The segmentation procedure used in this chapter refers to partitioning the image into foreground objects of RBCs and WBCs and the background plasma. Different morphological segmentation tools for blood cell segmentation were described in Chapter 2. In this chapter some additional relevant studies have been summarised.

Hierarchical partitioning is the main strategy for segmentation and operates deductively or inductively. In the deductive approach, the higher level object plane is segmented first and the process then proceeds to segment within the substructure components. For malaria diagnosis, this approach first segments the image into foreground cellular objects and background plasma. Further segmentation steps separate the RBCs from other blood cells and then segments the stained components within the RBCs. On the other hand, with the inductive approach, first the stained objects are identified and then further classified through segmentation of the RBC regions to determine whether they are inside a RBC (i.e. a parasite) or outside (i.e. most likely a platelet) [21].

The main issue associated with segmentation is the over or under segmentation of the cell regions [21], [110]. Under segmentation is the process in which the segmented area contains further group of objects to be segmented. Undefined cell boundaries due to cell overlapping or closely placed cell boundaries result in under segmentation. Over segmentation further segments the extracted objects in the background to subcomponents. It is usually caused by the lack of understanding of the cell parameters used for segmentation. For instance, parameters such as circular shape and size of the RBCs are commonly used for RBC segmentation. However an enlarged circular artefact or noise can affect the

Automated Malaria Diagnosis Using Mobile Phones

segmentation accuracy and can lead to over segmentation. Other factors include non-homogeneity and disfigurement of the cells as well as algorithm-specific issues.

An inductive segmentation technique based on granulometry and automatic thresholding has been explained by Ruberto et al. in [38]. In this study they applied granulometry, computed from a series of morphological opening with pre-defined structuring elements of increasing size to estimate the size of the cells and uses the information to determine the connective sets where nuclei of parasites are detected using regional extrema. However the method does not provide detailed evaluation of segmentation of WBCs and enlarged platelets. The work undertaken by Rao et al. in [39] and [40], follows the deductive approach for segmentation in which a modified granulometry technique called area granulometry is proposed. This estimates the area of the cells in order to differentiate the regions. Area granulometry considers the volume of pixels removed during consecutive morphological opening operations of increasing size. The area information obtained is used to segment the image into foreground and background regions containing cellular objects and plasma respectively. While able to exclude WBCs from the foreground, the method is prone to over segmentation as it considers all the stained components including as parasites. Also the accuracy rate of area granulometry is very much depended on cell overlapping. For images containing highly overlapped cells, the method fails to accurately measure the size of the components.

A modified segmentation technique called minimum area watershed transformation technique for automated diagnosis of malaria is described by Tek et al. in [54]. In this method, peaks obtained from the intensity histogram of the images provide a double threshold value which is then used for the segmentation process. For a less complex image, this process can be minimized by taking the image histogram of the original greyscale image which then gives two peaks and thereby the threshold for binarisation. The threshold value is usually obtained by calculating the centre point between the two peaks. This method utilizes the area information obtained from area granulometry such that, as the segmentation proceeds, a region is not recognized until it has achieved a certain area. The minimum area watershed segmentation procedure is used as the primary segmentation operator which is then followed by marker extraction using the Radon transform [51], [52] to detect the cell centres. Even though the method addresses over and under segmentation

Automated Malaria Diagnosis Using Mobile Phones

issues to a great extent, it is not applicable to highly concentrated regions of the blood film as the cell regions cannot be clearly defined and the Radon transform fail to pick up the cell centres. The area granulometry technique is also not very productive for an image with heavily clumped cells.

A recent study was carried out by Pearl et al. as described in [55] in which the cell boundaries are detected using the Hough transform [51], [52] and are extracted using fuzzy curve tracing. In the work done by Sharif et al. in [57] the WBCs are segmented by applying a morphological operator on YCbCr colour space and a marker-controlled watershed segmentation technique is used to extract the RBCs. A blood cell segmentation approach using evolutionary methods is described by Osuna et al. in [58], in which differential and artificial bee colony evolutionary methods are used to perform segmentation using histogram information and a minimum distance estimator. Similar ideas can be found in works published by Dorini et al. [59], and Yi et al. [58], which deal with blood cell segmentation for different applications. However, all of the above studies lack detailed evaluation of the performance of the segmentation strategy especially under non-uniform lighting conditions and field concentrations. Techniques that use inductive strategies for malaria diagnosis should elaborate the pre-processing techniques that deal with platelets and artefacts rather than assuming all the stained components are RBCs and WBCs. Similarly, in the deductive approach, the differentiation of WBCs from RBCs has to be addressed. In addition, cell overlapping remains a serious issue for all these techniques.

In this chapter a novel foreground partitioning technique called Annular Ring Ratio transform method will be described. The method uses a deductive approach and does not follow any of the earlier segmentation algorithms; instead it uses a foreground-object differentiating technique. A morphological approach based on the spatial distribution of the gray-level intensities is used and the cell radius acts as the vital attribute in most of the steps in the algorithm. Compared to the previous works done in the field, the proposed method is less complex and is computationally more efficient since it avoids recurring morphological opening or closing ,intense mathematical operation such as those used in Hough and Radon transform and marker extraction tool apart from the two parameterised structuring elements used for morphological operations . Apart from the morphological filtering none of the images used for processing has undergone colour normalisation and/or other enhancements.

4.1 Annular Ring Ratio Transform Method

The Annular Ring Ratio (ARR) transform method is a novel technique developed to identify the blood components in the image. The method is based on the local variation of pixel intensity values and information on the size of the cells that are obtained using granulometry. The process consists of firstly applying the ARR transform method to create a ratio-transformed image followed by a peak detection algorithm to locate the peaks in the ratio-transformed image. This provides the coordinates of the cell centroid. The ARR transform method is performed on the morphologically filtered blood image described in Chapter 3. This provides an image devoid of parasites, platelets and other stained components smaller than the size of the RBC and has only RBC and WBC regions with uniformly distributed intensity values.

4.1.1 Definition

The Annular Ring Ratio (ARR) Transform calculates, for each pixel, the ratio of two regions. The first region is an annular ring, as shown in Figure 4.1(a), centred on the pixel in question. The outer diameter of this ring is made approximately 25% larger than that of an RBC, (but smaller than a WBC) and the inner diameter is approximately 90% of that of a RBC. The second region is a disc shaped region of diameter equal to the inner perimeter of the first region as shown in Figure 4.1(b).

Let I_o be the average intensity of the image within the outer annular ring region and I_i is the average inner disc intensity within the inner disc region. Then the Annular Ring Ratio, ARR , for the pixel in question, is defined as:

$$ARR = \max \left[\left(\frac{I_o}{I_i} - 1 \right), 0 \right] \quad (4.1)$$

This is computed for each pixel in the image. Since we are interested in objects in which the inner region is darker than the outer, areas where this is not so are clamped at zero. This has the effect of reducing irrelevant background clutter. Dark circular objects against a lighter background are transformed into light circular objects whose intensity increases towards the centre of the object against a black background. The brightest point occurs at

Automated Malaria Diagnosis Using Mobile Phones

the centre of the object and this locates the centroid of each RBC or WBC if any. This is for positive images. For negative images of light objects against a dark background, the ratio I_o/I_i is inverted in Equation 4.1. Figure 4.2 illustrates the appearance of a transformed image. Note that, since the transform works with a ratio of intensities in the vicinity of the cell, it is independent, both of the average level of illumination but also any variation of illumination or film thickness across the whole image.

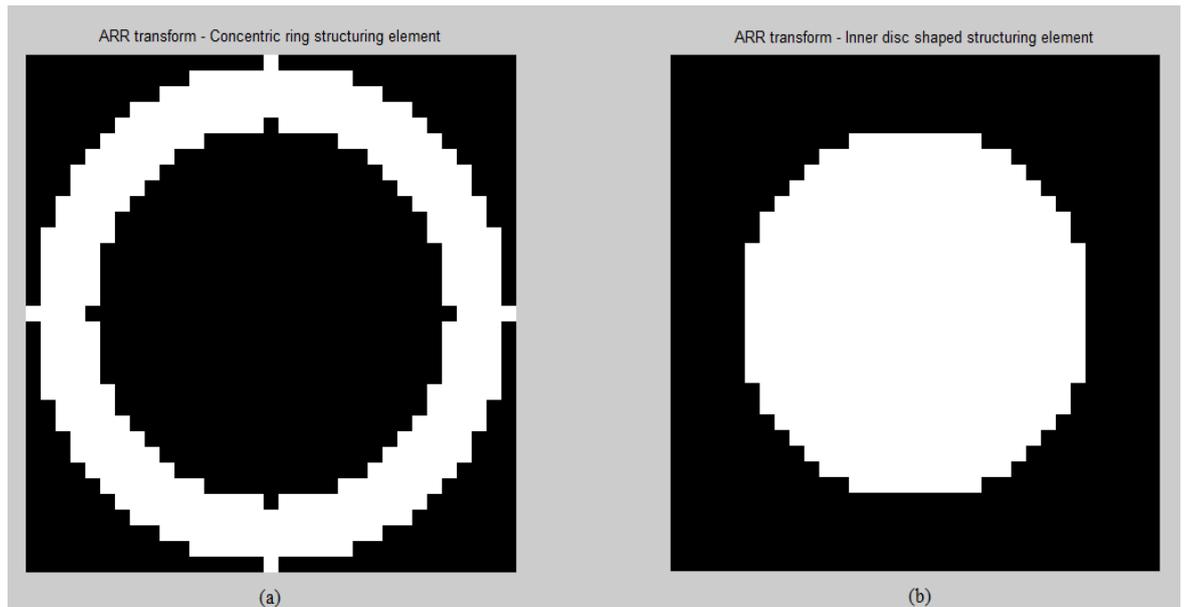


Figure 4.1: Structuring elements for the ARR method

(a) Annular ring outer structuring element; (b) Disc shaped inner structuring element.

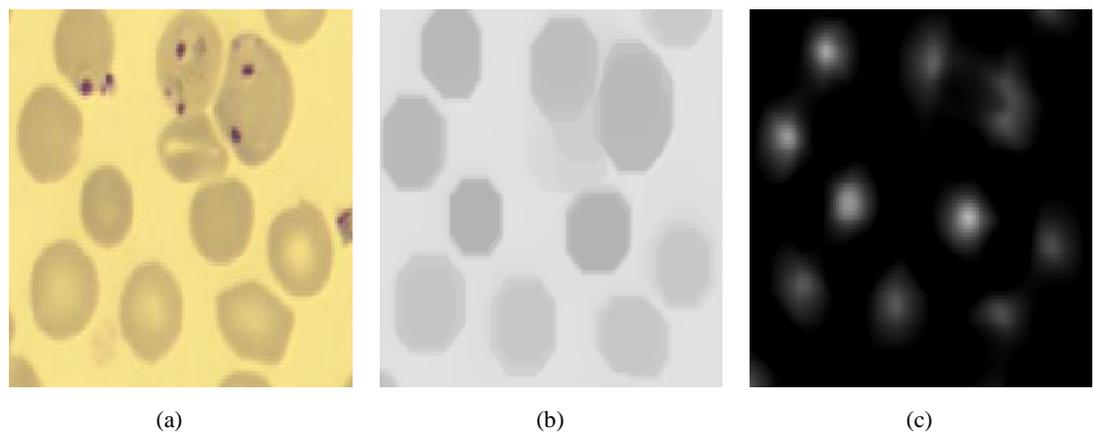


Figure 4.2: Illustration of ARR method with a positive image with bright objects on a lighter background.

(a) Original Image (b) Image processed by morphological dilation and erosion using two separate structuring elements. (c) Ratio-flattened closed image. The ratio of pixel intensities within structuring elements in 4.1 (a) and 4.1 (b) are taken.

4.1.2 The Algorithm

The pseudo-code for the ARR transform method is given in Figure 4.3. The neighbourhood of each pixel to be processed is defined by the Euclidian distance measurement based on the diameter of the structuring elements so that '*souter*' and '*sinner*' returns the sum of intensities of all the pixels in the neighbourhood of the outer and inner circular areas of the structuring elements respectively. Two counters to count the pixels mapped with the corresponding structuring elements are iterated during each pass. The boxed area represents the computation of the intensity ratio. The algorithm creates a transformed image with blob like structures which generates bright peaks at the centre of each cell as shown in Figure 4.2(c). The following Section illustrates a peak detection algorithm to detect the centre of the cells and there by locating the coordinates of the centre pixel.

Automated Malaria Diagnosis Using Mobile Phones

Annular Ring Ratio Transform

Ringratio creates the ratio of the outer ring average intensity to inner disc intensity

I = grayscale image to be processed

outrad = radius of the annular ring structuring element(in pixels)

inrad = radius of the inner disc structuring element(in pixels)

p = pixel intensity at respective coordinates

couter = counter to count the pixels within the region of annular ring.

cinner = counter to count the pixels within the inner disc.

[ymax, xmax] = size(*I*);

define vertical search range – *y1:y2*

define horizontal search range – *x1:x2*

Check Euclidian distance from centre to check whether the region is mapped under inner disc or annular ring

While *n1=y1:y2* and *m1=x1:x2*

$rad2 = (n1 - n)^2 + (m1 - m)^2;$

if $rad2 \leq (outrad)^2$

if $rad2 > (inrad)^2$

souter = *souter* + *p*;

couter = *couter* + 1;

else

sinner = *sinner* + *p*;

cinner = *cinner* + 1;

end;

end;

Ringratio(p) = ***souter* * *cinner* / (*sinner* * *couter*);**

Figure 4.3 Pseudo Code of ARR Transform method.

4.2 Peak Detection Algorithm

The peak detection algorithm was developed in order to detect and locate the bright centre points within the ratio-transformed image, which are assumed to be the centroid of the cell. The output of the peak detection algorithm is not an image but a list of coordinates, each one identifying the location of a blood cell. This makes this approach fundamentally different to other segmentation algorithms.

4.2.1 The Algorithm

To locate the bright peaks in the ratio transformed image, the algorithm uses a structuring element as shown in Figure 4.4. The diameter of the structuring element is approximately 75% of the diameter of an RBC. This will in effect process all the pixels within the diameter except the centre pixel and calculates the local maxima. If the maximum intensity within the neighbourhood is smaller than that of the centre pixel, it is considered to be the centroid of the cell and the co-ordinates are registered. Thus the locations of all the peaks are found by searching for the local maxima in the ARR transformed image. A small threshold is applied to the peaks to remove background clutter.

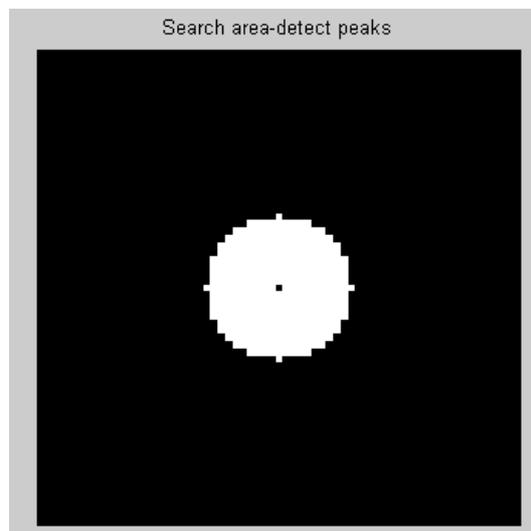


Figure 4.4: Structuring element which defines the search area to detect the peaks.

Automated Malaria Diagnosis Using Mobile Phones

The pseudo-code for the peak detection algorithm is given in Figure 4.5. As explained above, the centre pixel value will not be considered while estimating the local maxima and the condition for peak is only satisfied if the centre pixel intensity value is greater than the maximum of its neighbourhood.

Peak Detection Algorithm

Ringpeak detects the centre point of each cell in the ratio-transformed Image.

Input = ratio-transformed image.

sarad = radius (in pixels) of the search area

peaks = A list of (x,y) coordinates of the cell centres.

I_thresh= threshold value of the ratio of intensities.

P= pixel intensity value at the respective coordinates.

Sd= structuring element (search disc)

[ymax, xmax]=size(I);

create the search disc sd

*for n=1:2*sarad+1*

*for m=1:2*sarad+1*

rad2=(n-sr-1).^2+(m-sr-1).^2;

if rad2<=(sarad)2

sd(n,m)=1;

define vertical search range -y1:y2

define horizontal search range -x1:x2

if (I(n,m)>I_thresh) && (max(max(I(y1:y2,x1:x2),sd))<p)

peaks(np,1)=m;peaks(np,2)=n;m and n are the coordinates of each centre pixel.

Figure 4.5 Pseudocode of Peak Detection Algorithm.

4.2.2 Experimental Results

Both ARR transform method and peak-detection algorithm are applied to a variety of blood slide images acquired using different image acquisition and laboratory set-ups. These source images are taken at a range of resolutions, with variations in staining concentration and illumination. Also the images are randomly picked up irrespective of the field density and concentration of blood components.

The image dataset contain set of images supplied by: National Institute of Medical Research (NIMR), UK, University of Westminster (UoW) database in Applied DSP and VLSI Research Group (ADVRG) laboratory, the Hospital for Tropical Diseases (HTD) and UoW microscope imaging set up from thin blood smears supplied by National External Quality Assessment Service (NEQAS). The image set contained a total of 5000 foreground blood components which includes infected RBCs, non-infected RBCs and WBCs. These cells were manually evaluated prior to processing and were compared with the automated results. An average of 95.7% of cells was identified in total.

The following Figures 4.6-4.9 demonstrates the performance of the algorithm and its robustness in dealing with images of varying size, format, illumination, field concentration and colour properties. Figures 4.6-4.7 represents the different stages of processing of the algorithm on images in TIFF format with a resolution of 512 x 381 pixels.

As explained earlier, the RBC diameter is the main attribute used to set the sizes of the structuring elements for morphological filtering, ARR transform method and peak detection algorithms. A granulometry technique was used to calibrate the process by determining the average radius of the RBCs in pixels prior to working with images having different resolutions. For images in Figure 4.6 and 4.7 and RBC radius of 13 pixels were used. However the image in Figure 4.8 was processed with an RBC radius of 40 pixels. The image in Figure 4.9 was processed with an RBC radius of 27 pixels.

The image in Figure 4.6(c) is the filtered image as described in Chapter 3. Figure 4.6(d) is the ratio-transformed image obtained by applying the ARR transform method to the filtered image in (c). The brighter points indicate the cell centres. These bright centres are located by the peak detection algorithm which lists the coordinates of the centre pixels.

Automated Malaria Diagnosis Using Mobile Phones

Figure 4.6(e) illustrates the final result with red dots marking the centroid on the original image using the location information obtained from the peak detection procedure.

Figure 4.7 is an image with a deeply stained WBC which will be detected along with the RBCs. Unlike the image in Figure 4.6, the image shown in Figure 4.7 has a smaller field density and a reduced concentration of foreground components. Figure 4.8 is a high-resolution JPEG image of size 1200x1600 pixels. The image is taken under poor lighting conditions so the contrast is poor and the illumination non-uniform. Furthermore, in certain regions the cells are highly overlapped, as shown by the text arrows- OV1 and OV2 in Figure 4.8(a), making parasitemia estimation difficult. In fact, the image would be rejected by manual microscopist for RBC estimation due to its poor quality. Nevertheless, the ratio transform successfully detected 91% of the total foreground components in the image.

Figure 4.9 is another example of varying imaging and staining conditions. The image is in JPEG format with a size of 1030x1300 pixels. The colour properties and staining concentration of the image is different from that of Figures 4.6, 4.7 and 4.8. The cells are not clearly defined and the image has many artefacts present. The morphological filtering successfully removes the noise but the information on the cell regions is lost. The ARR transform method detects 87% of the cells in the image.

Note that, apart from morphological filtering, none of the images evaluated underwent any pre-processing techniques such as illumination correction, colour normalisation or contrast enhancement.

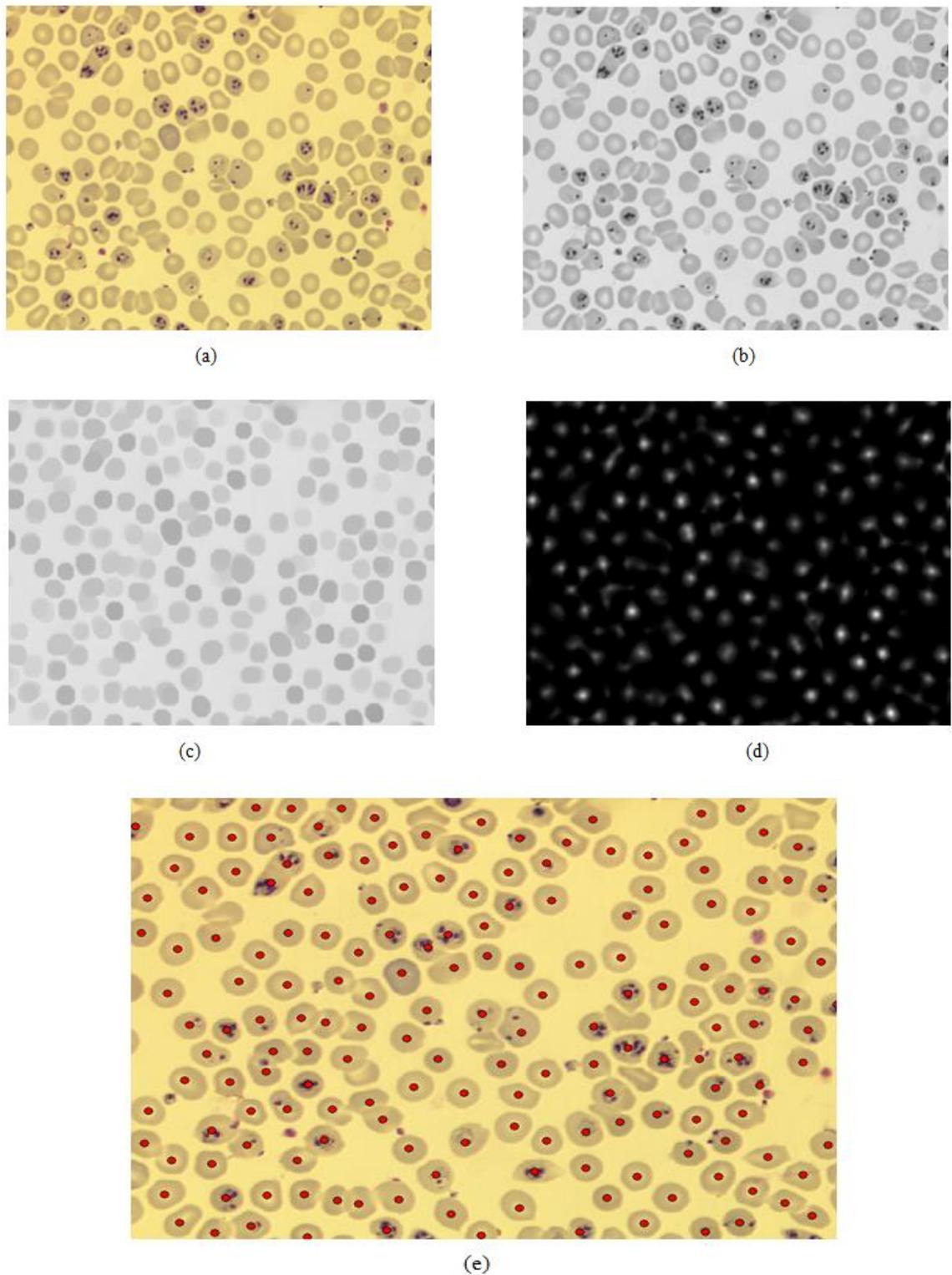


Figure 4.6: Demonstration of the ARR transform method on a TIFF image of size 512 x 381 pixels.

(a) Original image; (b) gray-scale image; (c) resultant image after morphological filtering to remove platelets, artefacts and noise; (d) ratio transformed image; (e) resultant image after performing peak detection technique on (d). The detected cells are marked with red dots.

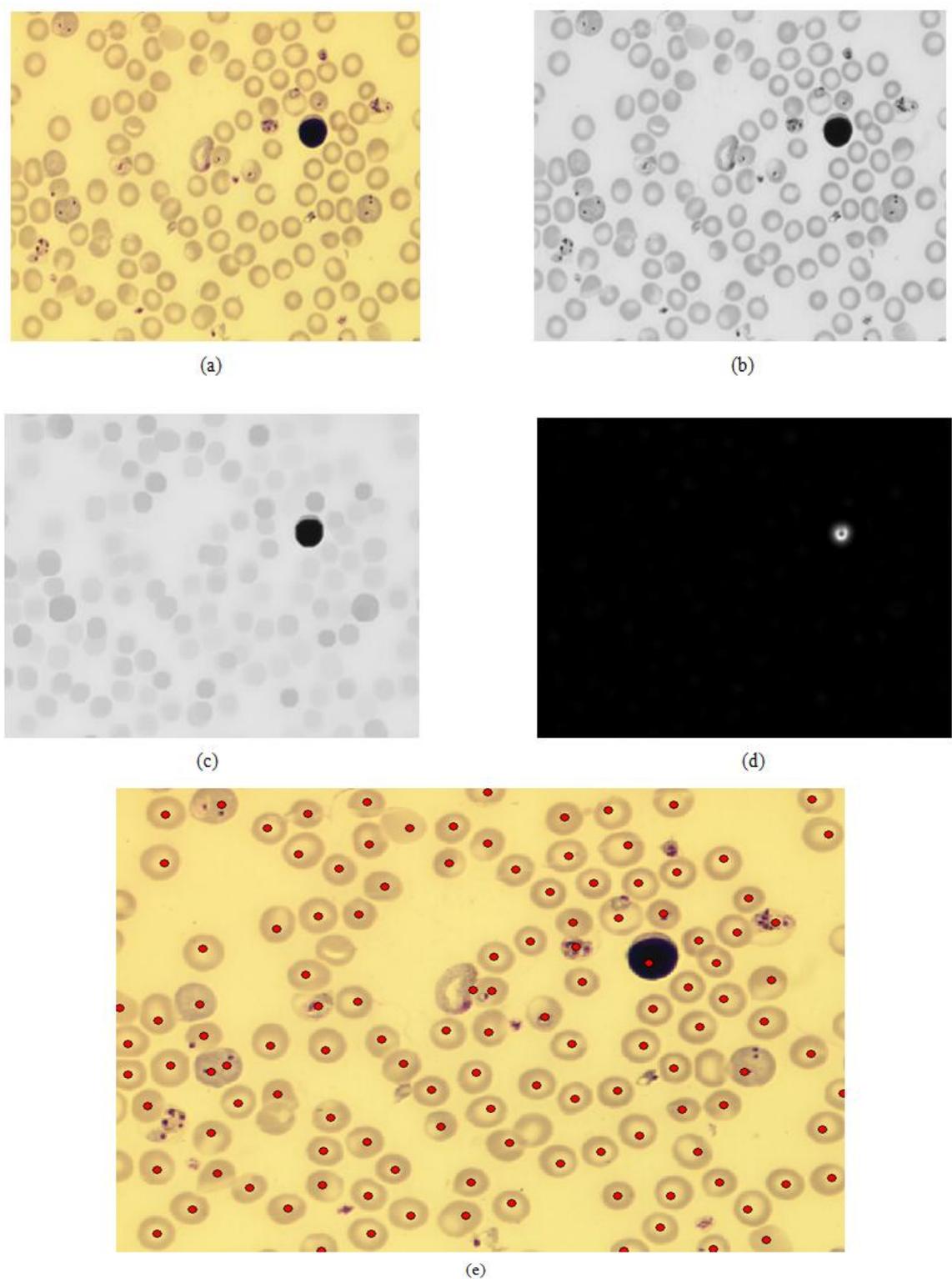


Figure 4.7: Demonstration of the ARR transform method on a TIFF image of size 512 x 381 pixels with a different field concentration.

(a) Original Image; (b) gray-scale image; (c) resultant image after morphological filtering to remove platelets, artefacts and other noise; (d) ratio-transformed image; (e) resultant image after performing the peak detection technique on (d). The detected cells are marked using red dots.

Automated Malaria Diagnosis Using Mobile Phones

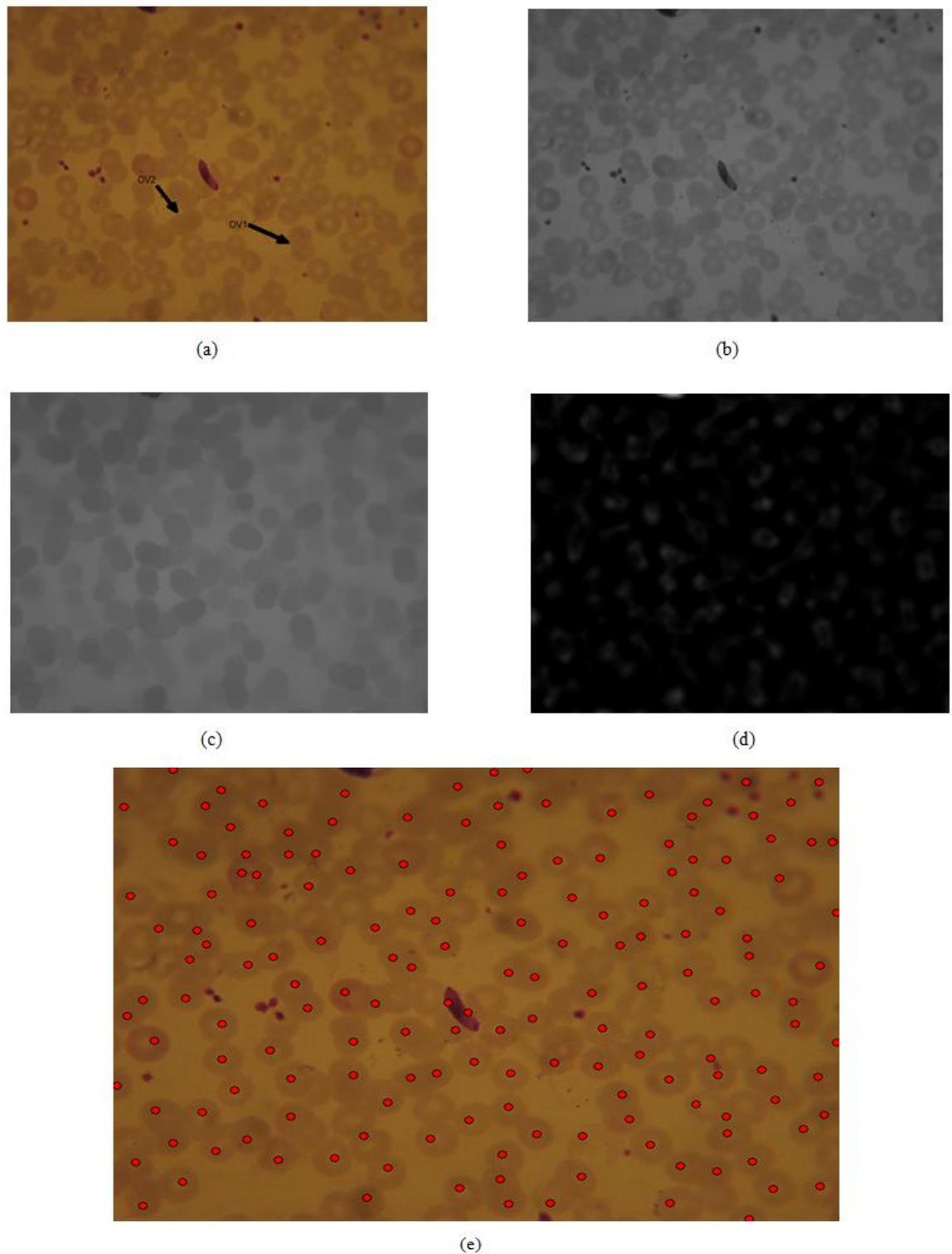


Figure 4.8: Demonstration of the ARR transform method on a JPEG image of size 1200 x 1600 pixels with non-uniform illumination.

(a) Original image; (b) gray-scale image; (c) resultant image after morphological filtering to remove platelets, artefacts and other noise; (d) ratio transformed image; (e) resultant image after performing the peak detection technique on (d). The detected cells are marked using red dots.

Automated Malaria Diagnosis Using Mobile Phones

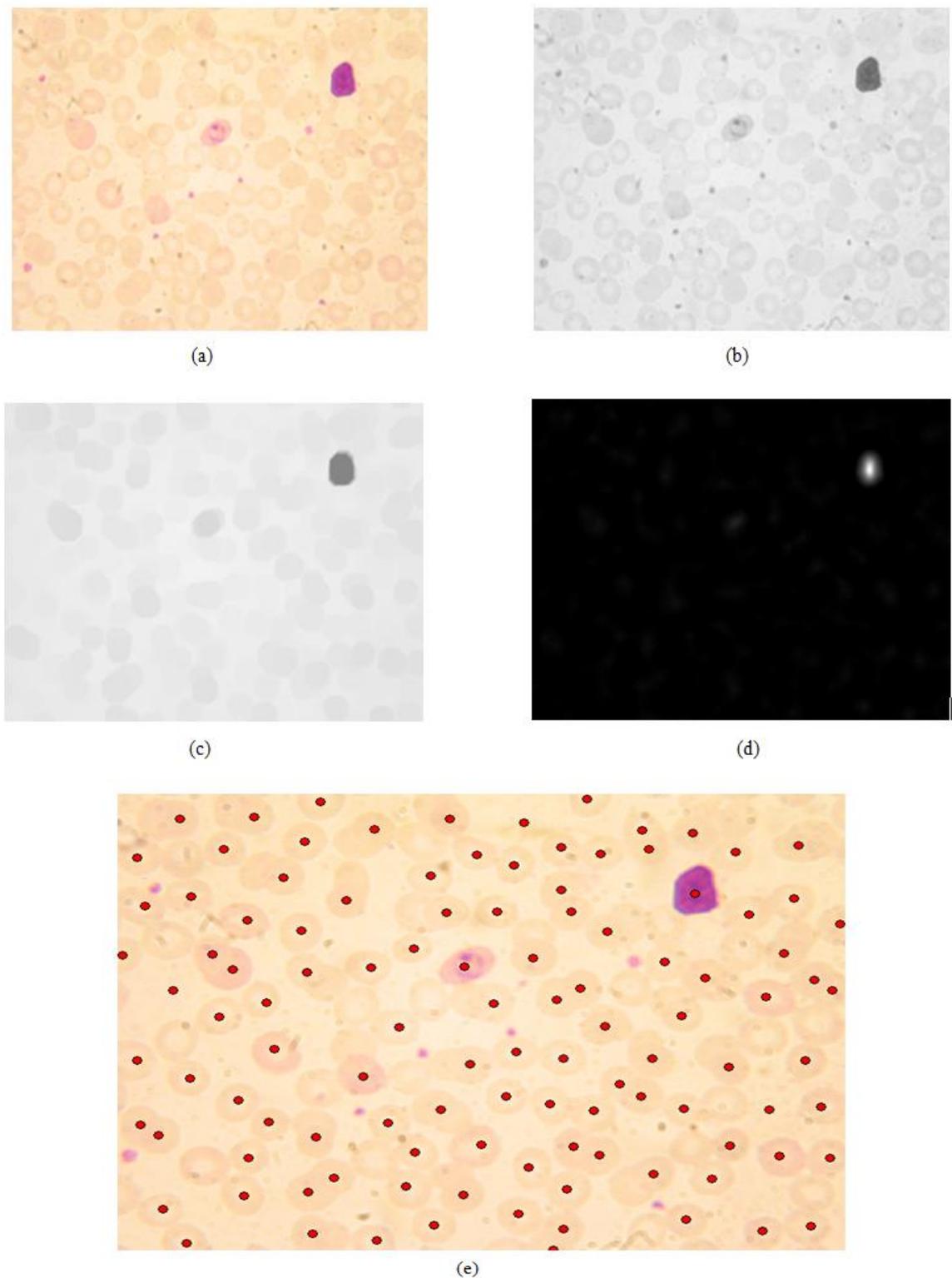


Figure 4.9: Demonstration of the ARR transform method on a JPEG image of size 1030 x 1300 pixels with a different staining concentration.

(a) Original image; (b) gray-scale image; (c) resultant image after morphological filtering to remove platelets, artefacts and other noise; (d) ratio transformed image; (e) resultant image after performing the peak detection technique on (d). The detected cells are marked using red dots.

4.3 Comparison with Segmentation Methods

The ARR transform method along with the peak detection algorithm is a simple alternative to earlier segmentation algorithms reported in the literature [38], [39], [54], [55], [57], [58] and [59]. It is fundamentally different to all of them. For example, a conventional segmentation tool, using edge detectors and connected pixel component extractors, segments the images in Figure 4.6(a) and 4.8(a) as demonstrated in Figures 4.10(a) and (b) respectively. These figures are classic examples of under-segmentation. The solution used so far to rectify under-segmentation problems is to use watershed segmentation which divides images to unique regions based on their regional minima [117], [118]. However, for an infected blood image, watershed segmentation is not very effective due to the presence of heterogeneous grey-level regions which leads over-segmentation of the images. Hence, additional markers are required and marker-controlled watershed segmentation proved to be more effective for blood component segmentation as described in [54] and [57]. The results of images segmented using the marker-controlled segmentation technique using a gradient operator is shown in Figure 4.10(c) and (d). If compared with the same images processed using the ARR transform method as shown in Figure 4.6(e) and 4.8(e), it can be seen that the marker-controlled watershed segmentation algorithm failed to identify all the cells. In these cases, the failures are due to the presence of light centre pallor within the RBCs in Figure 4.6(a) and non-uniform illumination in Figure 4.8(a). Watershed segmentation also has difficulty dealing with poorly stained images.

The possible solutions to these problems are further extension of the architecture using illumination correction algorithms, hole filling algorithm to remove the light centre pallor as in [37] and introducing effective marker extraction tools such as Radon transform as in [54] or Hough transform as in [55]. This further increases the architectural complexity and thereby the computational cost of the segmentation algorithm. On the other hand, the ARR transform method successfully overcomes the problems of poor illumination and staining and requires minimal architecture for object estimation.

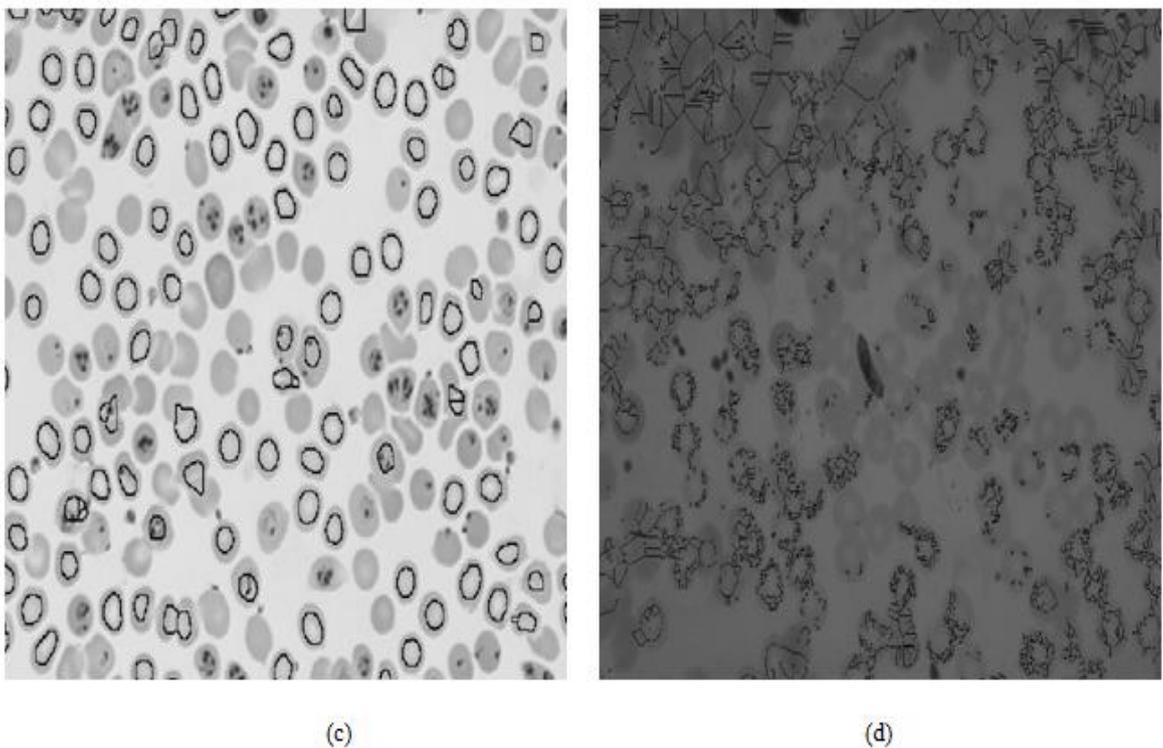
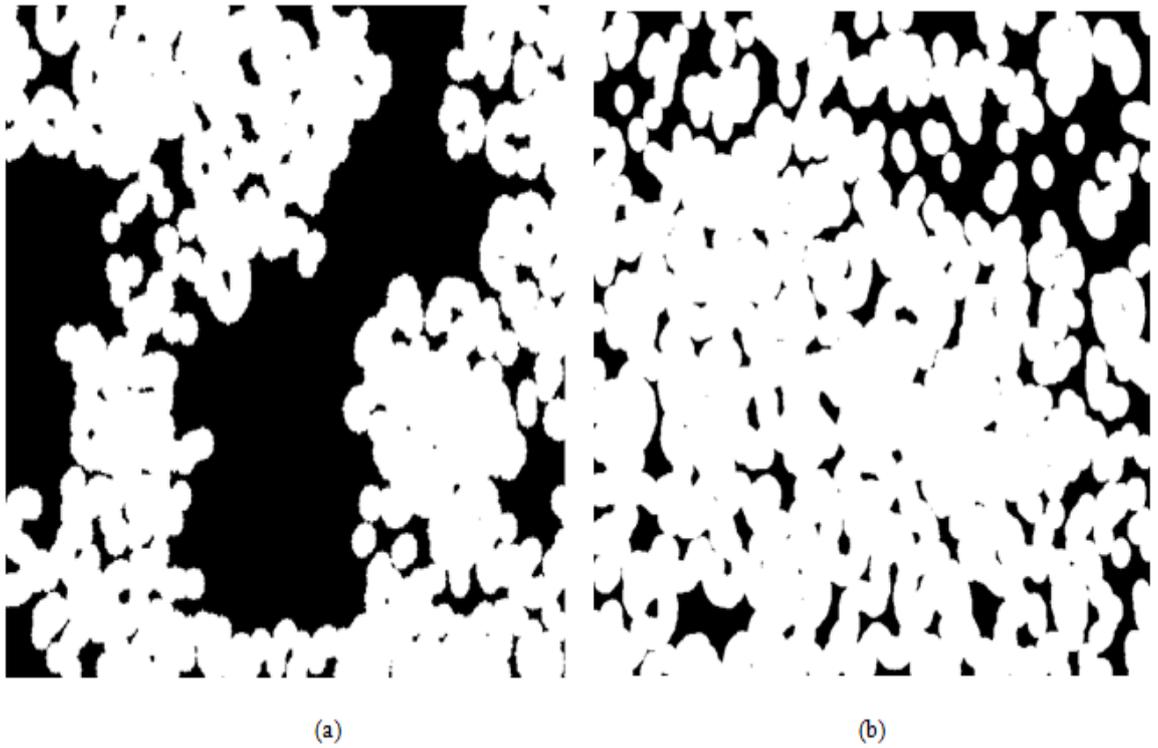


Figure 4.10: Comparison of ARR transform method with other segmentation algorithms.

(a)-(b) Segmentation using edge operators and connected pixel extractors. (c)-(d) Segmentation using Marker controlled watershed algorithm.

Automated Malaria Diagnosis Using Mobile Phones

Based on the above experiments, the following table highlights the major advantages of the ARR transform method compared to the watershed segmentation algorithm.

Table 4.1 : Comparison of ARR transform method with Marker controlled watershed segmentation method.

Features	Marker Controlled Watershed Segmentation	ARR transform method
Pre-processing algorithms used	<ul style="list-style-type: none"> • Colour normalisation algorithm • Subtraction/filtering algorithm for noise removal • Non-uniform illumination correction algorithm • Image smoothing (optional) 	<ul style="list-style-type: none"> • Morphological filtering algorithm for noise removal.
Structuring elements required	Yes	Yes
Edge operators required	Yes	No
Other morphological operations involved	<ul style="list-style-type: none"> • Regional minima and maxima calculations • Morphological image reconstruction • Recurring morphological closing and opening operations • Marker extraction • Image masking and superimposing. 	<ul style="list-style-type: none"> • Localised intensity ratio calculation • Local maxima calculation for peak detection
Computational time (Cpu time) in Matlab (512x381 tiff image)	47 seconds	24 seconds

4.4 Summary

This chapter described a novel image processing technique to locate the foreground components in a blood image. The method uses the ARR transformation algorithm to calculate the ratio of outer to inner intensities based on an annular ring structuring element. The cells centres are further located using a peak detection algorithm which searches the local maxima in the neighbourhood of specified structuring element.

The method has several advantages compared to the existing segmentation procedures that were discussed earlier in the chapter. Depending on the size of the structuring element, the object to be segmented can be chosen easily. The method is less complex as it does not involve any pre-processing techniques other than the morphological erosion and dilation to remove the platelets and artefacts. Unlike conventional segmentation procedures, the method provides the location information of the foreground components (RBC and WBC) and gives an estimation of the cell population. The method is independent of the average level of illumination and its variations and performs well with images of varying illumination and thickness.

The major issues with cell segmentation are colour inconstancy and non-uniform illumination. Both play a central role in extracting the objects from the background. In [42], Tek et al. propose solution to these issues. Due to various properties of microscopic imaging set-up, the colour, intensity and size of the cells can vary considerably. However, since the ARR transform method measures the ratio of intensities, the problem of non-uniform illumination or film thickness is eliminated.

As the monochrome image is a function of the colour image with a distribution of over and under saturated coloured pixels, the ratio transform of the gray level image calibrate the colour and tonal imbalances and hence avoid the necessity of any colour normalisation or calibration techniques. The only calibration process absolutely necessary for different imaging set-ups is to establish the size of the RBCs in the image. All other parameters are scaled from this measurement. The other calibration procedure that can be done is to determine the optimum weights in the colour to gray-scale conversion. However, our experience to date is that this is not very critical.

Automated Malaria Diagnosis Using Mobile Phones

Another problem associated with segmentation is under and over segmentation of the cells depending on the field concentration. The ratio transform does not face any issue with partially overlapped regions with cell centres apart by $\frac{3}{4}$ of the RBC radius as demonstrated in Figure 4.6. However for closely packed cells with cell centres almost overlapping, the method considers them as a single cell. A possible solution initially considered was to define the cell regions (using the distance measurement from cell centres given by the peak-detection algorithm) by evaluating pixel connectivity based on the size of the RBC or WBC and categorically assign the new centroid of the cells. However, this could debase the overall diagnosis competency of the automated system when it comes to parasitemia estimation. Hence, a better option would be to discard the overlapped areas and select only those areas with evenly distributed cells in a given field.

Since parasitemia in thin blood films is the ratio of total number of infected cells to total number of RBCs, a total of 1000 RBCs from 4-5 fields consisting of 200-250 cells are used for estimation. For manual diagnosis the pathologist roughly estimates the number of RBCs within a field and uses this to compute the parasitemia. However, for both manual and automated diagnosis, overlapped cells present a problem for RBC number estimation. It is difficult to count the number of RBCs in an overlapped area and usually parasites are less visible. By selecting the appropriate zone of morphology, the ARR transform method can estimate the cell centres and give the exact number of clearly differentiated RBCs examined for the presence of parasites. It can select as many fields as necessary until it counts a total of 1000 RBCs and can give an accurate parasitemia measurement. This is an important advantage of the ARR transform method over existing methods including manual diagnosis. The future work on this area is to develop an automated set up to select the appropriate zone of morphology and a user friendly radius selection tool which initiates the calibration of the entire procedure.

Chapter 5

White Blood Cell and Red Blood Cell Differentiation

The differentiation of WBCs and RBCs is a common procedure in peripheral blood analysis. The process, in its own or as a prerequisite to other clinical procedures, helps to estimate the WBC count, RBC count and to assess other analytical and clinical procedures which involves- diagnosing anaemia, checking the blood immune system, as well as any abnormalities of these blood cells. This chapter explains the procedures adapted to identify and differentiate the WBCs and RBCs in the image which were already located by the ARR transform method explained in Chapter 4. The main aim of undertaking these studies in this research are to identify and locate the RBCs in a peripheral thin blood image and thereby to detect the presence of any *plasmodium* malaria parasite within them. In addition, the total number of erythrocytes needs to be calculated in order to calculate the parasitemia of the blood which is a measure of infection. In order to classify the RBCs using an automated system, the other components in the image have to be identified and discarded. This chapter discusses various object extraction techniques analysed and adapted to discriminate the WBCs and RBCs. and explains a novel algorithm for blood images to differentiate and estimate them.

5.1 Definition in terms of Cell Differentiation

The ARR transform method followed by the peak detection algorithm detects and marks the RBCs and WBCs as shown in Figure 4.7(e). It is important to recall that the output of this process is a list of coordinates while the other segmentation techniques produce a binary image from which the location of the cells has to be extracted. Two most obvious attributes to differentiate the RBCs from the WBCs are the cell radius and staining concentration. Various image segmentation techniques were analysed which can utilise these attributes to discriminate the cells.

Two basic approaches of segmentation are edge-based segmentation which deals with discontinuities and region-based segmentation based on similarities [51], [52]. Since

the boundaries of each stained components in the blood image are sufficiently different from each other, the local discontinuities could be used as an important attribute in differentiating them. However, with blood cell images, the foreground objects are textured and the intensity distributions are not uniform within each component. Moreover, the parasitic nuclei within an infected RBC cause higher levels of intensity variation than those occurring at the RBC boundaries. These spurious changes in intensity values make it difficult to find a unique boundary for each cell using edge information. Consequently, edge-based segmentation, while useful for images with uniform intensity, can only be used with blood images when combined with other segmentation techniques as in [45] or for pre-processing as in [61].

The region-based segmentation techniques in which the image is partitioned into regions using a set of predefined criteria were also considered and various parameters were investigated as conditions for creating sub regions of leukocytes and erythrocytes in the image. In the analysis of blood images, size and intensity are the two factors most easily identified by the human eye. In order to utilise these parameters for machine vision, a region-based segmentation approach using thresholding is the easiest solution as demonstrated in Figure 5.1. This displays the effect of thresholding using information obtained from the gray-level histogram. The idea was based on the work published in [44].

5.1.1 Segmentation based on thresholding

The intensity histogram shown in Figure 5.1(b) has two well defined peaks which groups the foreground and background pixels into dominant modes. In order to extract them, a suitable threshold has to be selected as explained in Section 3.2.1. Using the threshold value $T1$ obtained from the histogram as shown in Figure 5.1(b), the segmentation of WBC can be performed using the following equation:

$$P(x,y)=\begin{cases} \text{WBC}, & I(x,y) < T1 \\ \text{Background}, & I(x,y) \geq T1 \end{cases} \quad (5.1)$$

Automated Malaria Diagnosis Using Mobile Phones

Similarly, to extract all the foreground components from the background, following equation can be used:

$$P(x,y)=\begin{cases} \text{RBC+WBC}, & I(x,y) < T2 \\ \text{Background}, & I(x,y) \geq T2 \end{cases} \quad (5.2)$$

Where $P(x, y)$ is the segmented pixel and $I(x, y)$ is the gray-scale intensity value of the image at any point (x, y) . Subtracting the image in Figure 5.1(c) from 5.1(d) should extract all the RBCs in the image.

However, this approach is not an effective differentiation tool especially for blood images with WBCs having area mostly occupied by its cytoplasm. The intensity distribution of the cytoplasm of the WBC falls within the RBC gray-scale value and hence will leave a binary ring around the location of the WBC as pointed out by the red arrow shown in Figure 5.1(e).

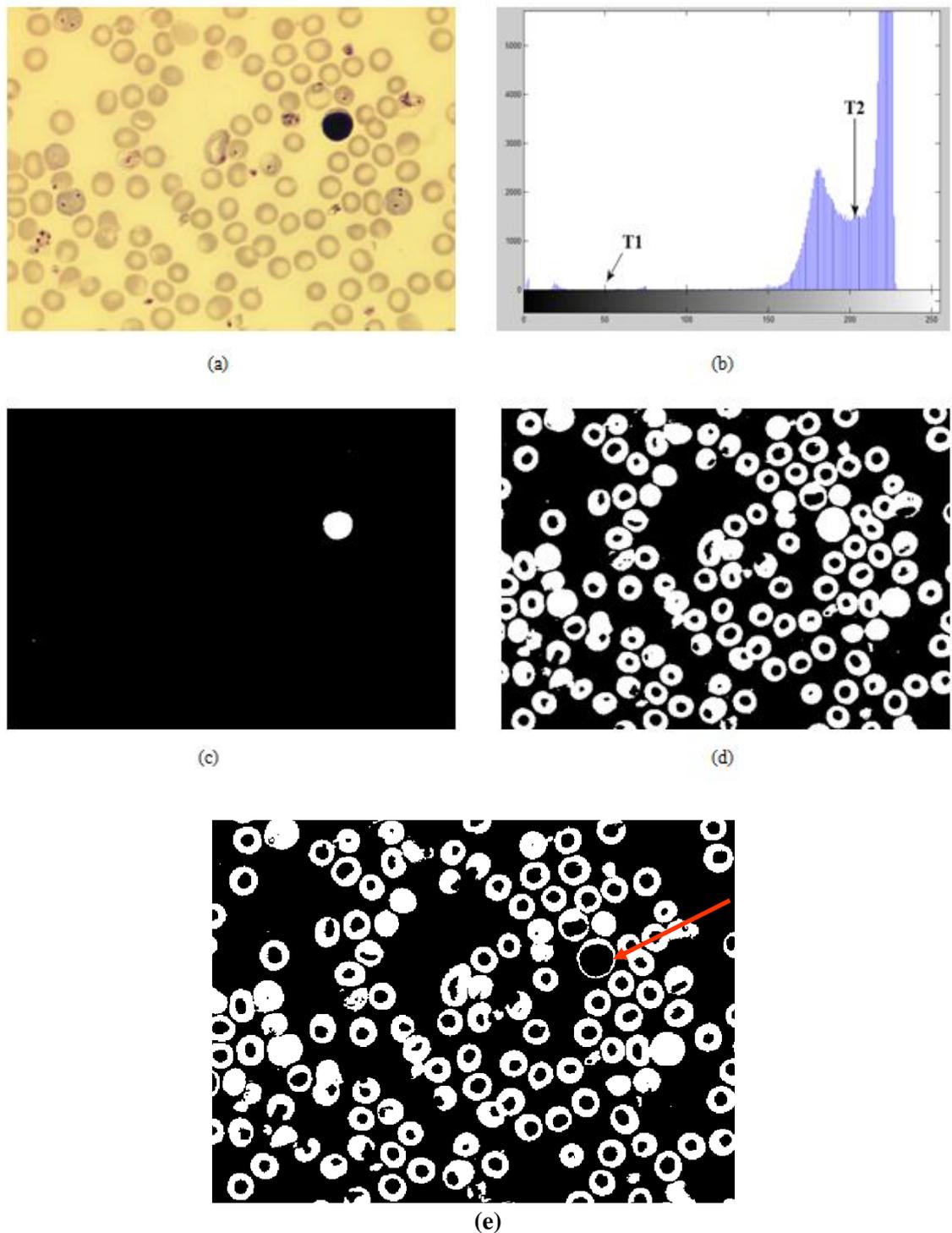


Figure 5.1: Application of global thresholding for blood cell extraction using intensity histogram.

(a) Original Image. (b) Intensity histogram of gray-scale image. (c) Extracting WBC using global threshold T1. (d) Extracting RBCs and WBC using global threshold T2. (e) Image with RBCs extracted after subtracting (d) from (c). The image clearly preserves some information on WBC since the threshold T1 only picks up the deeply stained nucleus of the WBC (Red arrow pointing at the WBC). The ring shaped feature around the location of the WBC is the cytoplasm of the WBC which has approximately the same staining concentration as the RBCs.

Automated Malaria Diagnosis Using Mobile Phones

A combination of the above said thresholding approaches were also attempted which segments the images into different regions based on the following equation:

$$P(x,y)=\begin{cases} \text{WBC,} & \text{if } I(x,y) \leq T1 \\ \text{RBC,} & \text{if } T1 < I(x,y) \leq T2 \\ \text{Background,} & \text{if } I(x,y) > T2 \end{cases} \quad (5.3)$$

Where $P(x, y)$ is the segmented image and $I(x, y)$ is the gray-scale intensity value of the image at any point (x, y) . Even though thresholding is the easiest way of segmentation, the threshold selection process has a high degree of uncertainty for poorly processed images and images attenuated by noises produced from platelets, artefacts and stained nuclei of the parasite. The threshold values $T1$ and $T2$ are selected from the valleys of the intensity histogram and hence clearly depend on the depth of the valleys and the separation of the peaks. For blood images, the selection of the threshold based on valleys are greatly affected by the non uniformity of the illumination source, noise in the image and the presence of variety of stained components.

Figure 5.2 and 5.3 shows examples of images from the UoW image data set and the corresponding histograms which demonstrate that the thresholding operations were unable to cope with the poor processing and variations in illumination and contrast. Figure 5.2(d) and 5.3(d) show the images with the foreground objects detected and located using Annular Ring Ratio method. Despite the poorly processed cells and poor contrast and illumination, the ARR transform method detected 90% of the cells.

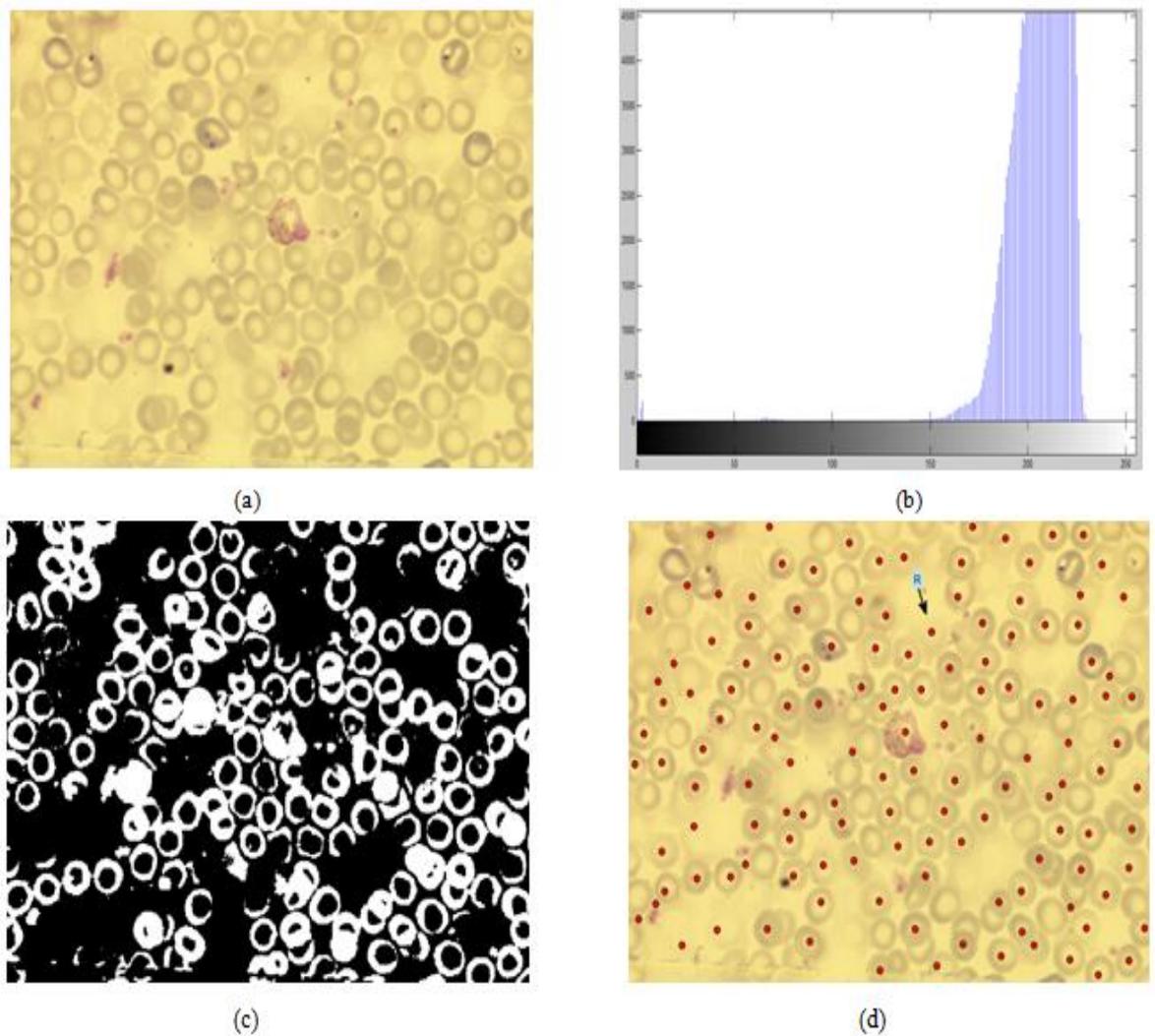


Figure 5.2: Effect of poor processing and illumination on histogram.

(a) Original Image. (b) Histogram of the image. It is difficult to find a proper threshold for this image from the histogram due to lack of definite peaks. (c) Binary image after thresholding. (d) Output image with the RBCs detected using Annular Ring Ratio transform. Despite the poorly processed cells and poor contrast and illumination, the method detected 136 cells out of 152 including the deteriorated RBC marked by the arrow.

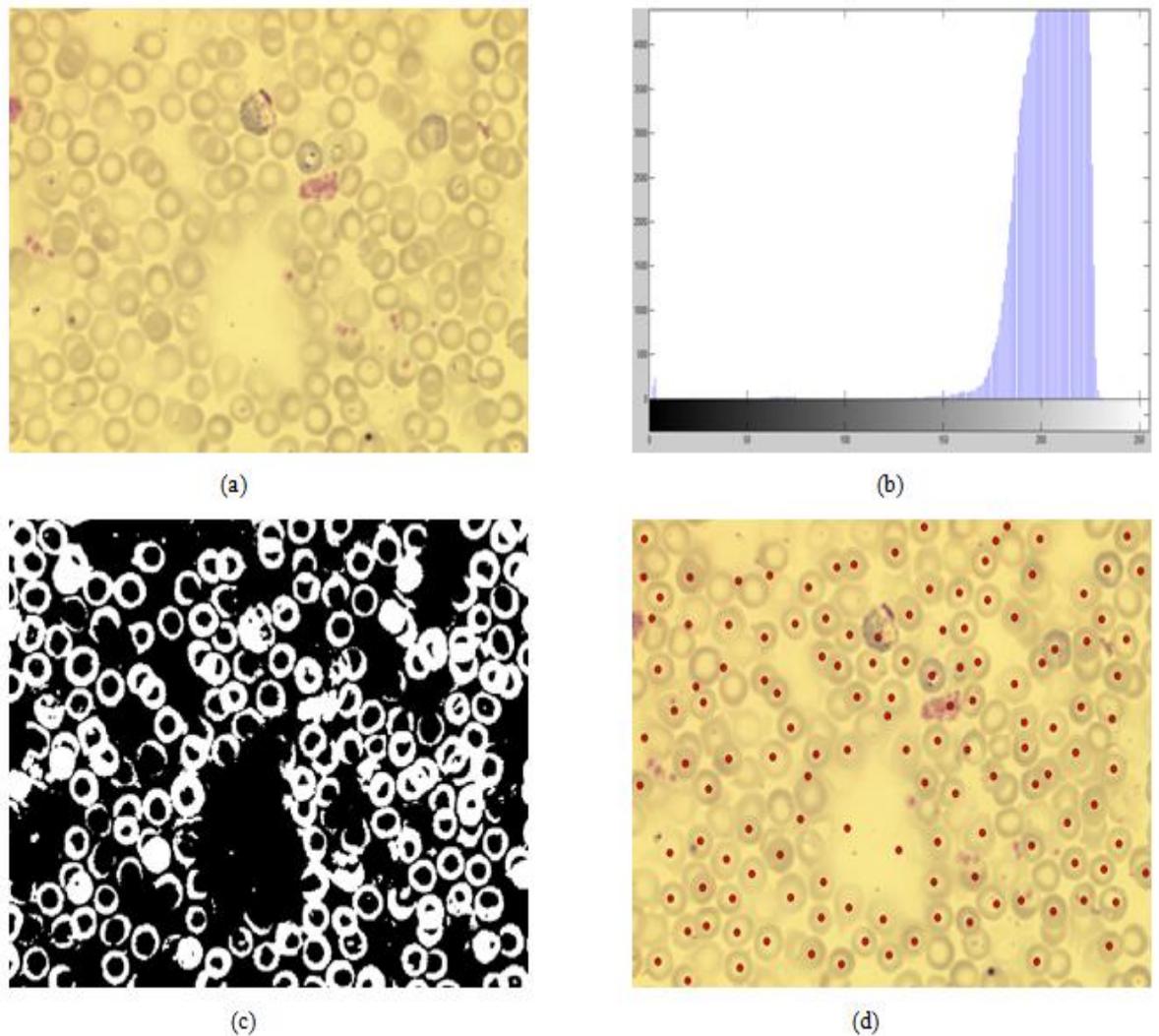


Figure 5.3: Effect of poor processing and illumination on histogram.

(a) Original Image. (b) Histogram of the image. It is difficult to find a proper threshold for this image from the histogram due to lack of definite peaks. (c) Binary image after thresholding. (d) Output image with the RBCs detected using Annular Ring Ratio transform.

In [54], Tek et.al. describes a morphological segmentation technique called minimum area watershed segmentation which is a modified threshold-based watershed segmentation procedure. It interprets the image in the dimensions of spatial coordinates and intensity. The process consists of pre-processing the image, construction of binary watershed lines and marker extraction based on predefined criteria. The method segments foreground and background regions based on the thin blood film thresholding method developed by Rao [20]. It utilises the cell area information to perform consecutive area opening to estimate the non-zero difference which roughly separates the foreground regions. The threshold values of foreground and background objects were then found from the histograms

Automated Malaria Diagnosis Using Mobile Phones

calculated from the corresponding regions' pixel values. The method provides a stable segmentation result but has some disadvantages such as over segmentation which can be corrected using external and internal markers. However, in addition to the pre-processing steps involved, marker extraction can be a complex procedure depending of the predefined attribute used for marker selection. The predefined criteria can be intensity values, connectivity, location, texture content, relative distance and morphology [51]. Hence considering the hardware complexity and computational speed, much simpler but reliable algorithms were considered.

The following Sections describe two methods that attempt to differentiate RBCs and WBCs in an image in which the foreground components, equal to or greater than the size of the RBCs, have already been located by the ARR transform method.

5.2 WBC and RBC differentiation algorithm

The algorithm to discriminate RBC, WBC and background regions based on Equation (5.3) uses two threshold values that are selected from the valleys between the peaks of the intensity histogram of the gray-level image. However, for blood images, the selection of the thresholds dependent upon the location of the valleys is greatly affected by the non-uniformity of the illumination source, noise in the image and the variability of the degree of staining particularly with heavily infected cells. For a non-infected blood image, since the intensity variation between the stained RBCs and WBCs are considerable, the second condition ($T1 < I(x, y) \leq T2$) of Equation (5.3) is fulfilled. Depending on the level of noise present in the image, the technique can be effective in extracting the foreground components from the background but only for lightly infected images. However, the intensity distribution of the platelets and artefacts due to staining can also fall within the condition $T1 < I(x, y) \leq T2$ leading to the false identification of these as WBCs.

In order to avoid these false identifications, the algorithm was applied in conjunction with the ARR transform method such that only those pixels located as the centre of a cell by the ARR transform method are processed. This avoids false prediction of stained pixels (outside any cell) as being RBCs or WBCs.

Hence the algorithm is modified to

$$P_{(X,Y)} = \begin{cases} \text{WBC,} & \text{if } I_{(X,Y)} \leq T1 \\ \text{RBC,} & \text{if } T1 < I_{(X,Y)} \leq T2 \\ \text{Background,} & \text{if } I_{(X,Y)} > T2 \end{cases} \quad (5.4)$$

Where (X,Y) are the coordinates located by the ARR transform method as the centre of the foreground components and $I_{(X,Y)}$ is the pixel intensity value at (X,Y) .

For infected blood or non-uniform illumination however, the method shows varying accuracy based on the level of infection. The sensitivity was considerably high for low-level infections but for higher parasitemia, since the pixels of parasite nuclei possess intensities which come within the threshold $T1$, the process leads to false identification of heavily infected cells as WBCs.

5.2.1 Differentiation based on local Mean and Variance

Since the thresholding operations were proved not to be solely depended upon for the discrimination of the foreground objects into RBC and WBC, a different set of attributes such as the variance of regions were considered. Since variance is not a sufficient predicate to segment a region, an algorithm with local mean and variance as the decision predicates was developed. Similar algorithms already exist in the field for image segmentation in conjunction with other attributes and techniques. In [61], the authors describe an improved fingerprint segmentation using mean and variance to avoid the high background noise interference. In [62] a similar technique with global mean, local mean, variance and coherence were developed which adapts a rule-based system to regionalise the image. Both methods calculated the predicates on every pixel in the image. The algorithm developed in this study is faster than those described in [61] and [62] as it considers only those pixels located by the ARR transform method (the centre pixels of RBC and WBCs) and avoids the rest. This has an additional advantage of avoiding the false identification of platelets and artefacts present in the image.

Automated Malaria Diagnosis Using Mobile Phones

The local mean and variance of the located pixels are calculated as follows:

$$\mu(X, Y) = \frac{\sum_{(x,y) \in \mathbf{R}(X,Y)} I(x, y)}{n_{\mathbf{R}}} \quad (5.5)$$

$$\sigma^2(X, Y) = \frac{\sum_{(x,y) \in \mathbf{R}(X,Y)} (I(x, y) - \mu(X, Y))^2}{n_{\mathbf{R}} - 1} \quad (5.6)$$

where X, Y are the listed coordinates of the RBC and WBCs located by ARR transform method, $\mathbf{R}(X, Y)$ is the circular region around the centre and $n_{\mathbf{R}}$ is the total number pixels in the region \mathbf{R} , $I(x, y)$ is the pixel gray-scale intensity value at (x, y) . $\mu_{(X, Y)}$ and $\sigma^2_{(X, Y)}$ are the respective mean and variance of the pixel intensities within the region $\mathbf{R}(X, Y)$.

The average intensity value of the pixel neighbourhood within a WBC is much lower than that within both normal and infected RBCs. Similarly the variance of the pixel intensity of RBC is comparatively higher than the WBC because of the light centre patch as well as infection. Hence RBC and WBC regions can be differentiated using these two parameters by setting appropriate thresholds. The threshold values were determined through manual experimentation and evaluation on a set of random images from the UoW database. This approach in fact also speed up the computation of the algorithm especially for images of large size, as only selected (by the ARR transform method) pixel neighbourhoods were processed. The results obtained using this method is given below in Figure 5.4.

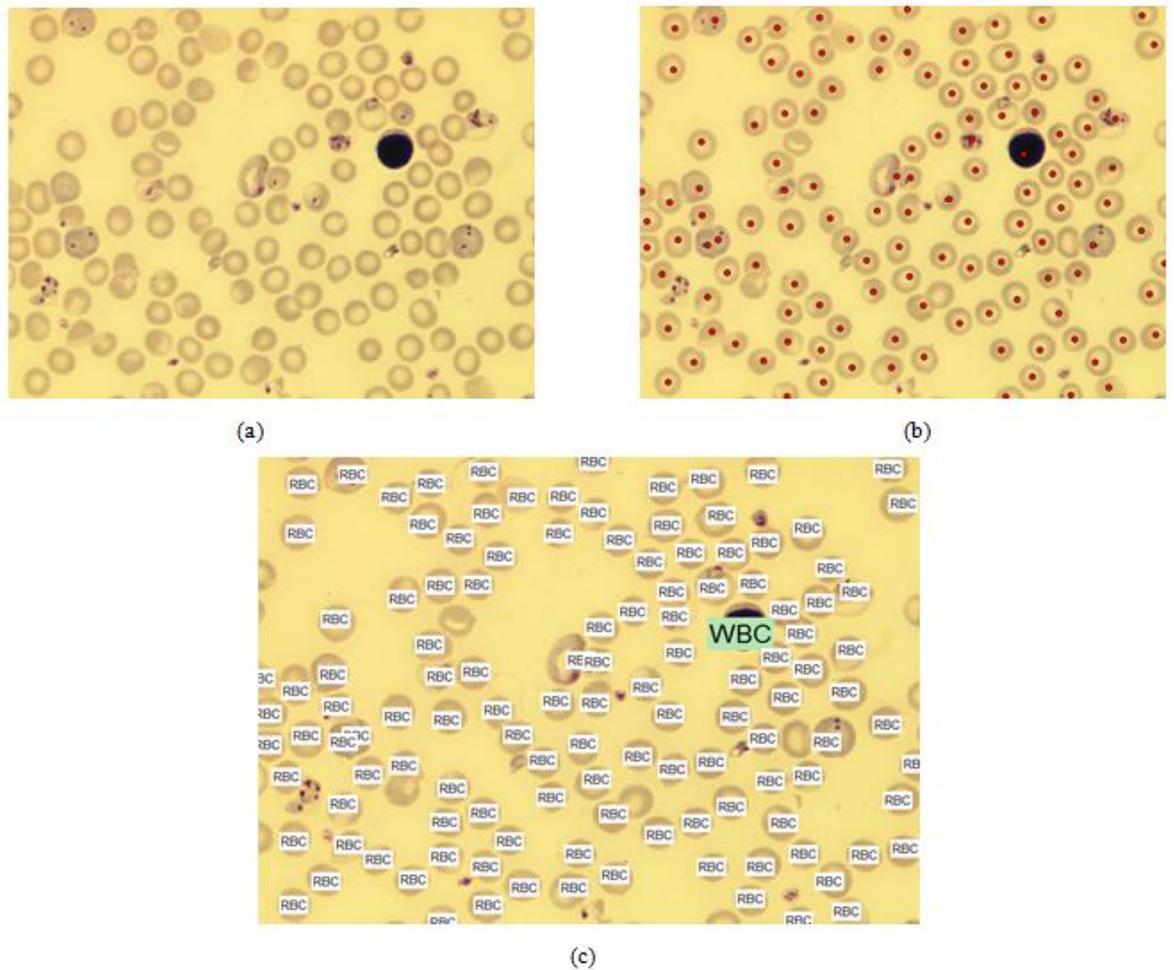


Figure 5.4: The RBC and WBC differentiation algorithm using localised mean and variance.

(a) Original Image; (b) Image processed using the ARR transform method which locates the foreground objects greater than or equal to the size of the RBC. Both RBCs and WBC are located. (c) The resultant image in which RBCs and WBCs are identified by calculating the local mean and variance of all the located pixel neighbourhoods marked in (b).

Even though this method is quite efficient in distinguishing between healthy RBCs and WBCs in an image, it is less reliable with infected blood images. For the early stages of infection (ring trophozoites and trophozoites) the attributes (mean and variance) are useful to discriminate between RBCs and WBCs. But on a heavily infected RBC (with schizonts and gametocytes) the mean and the variance matches with that of the WBC and hence are falsely identified as WBCs. In addition, since the mean and variance are calculated on the pixel neighbourhood with the size of RBCs, the presence of any large artefacts having the size of RBC will also lead to incorrect identification. The following figures 5.5 and 5.6 show two examples of the demerits of this method.

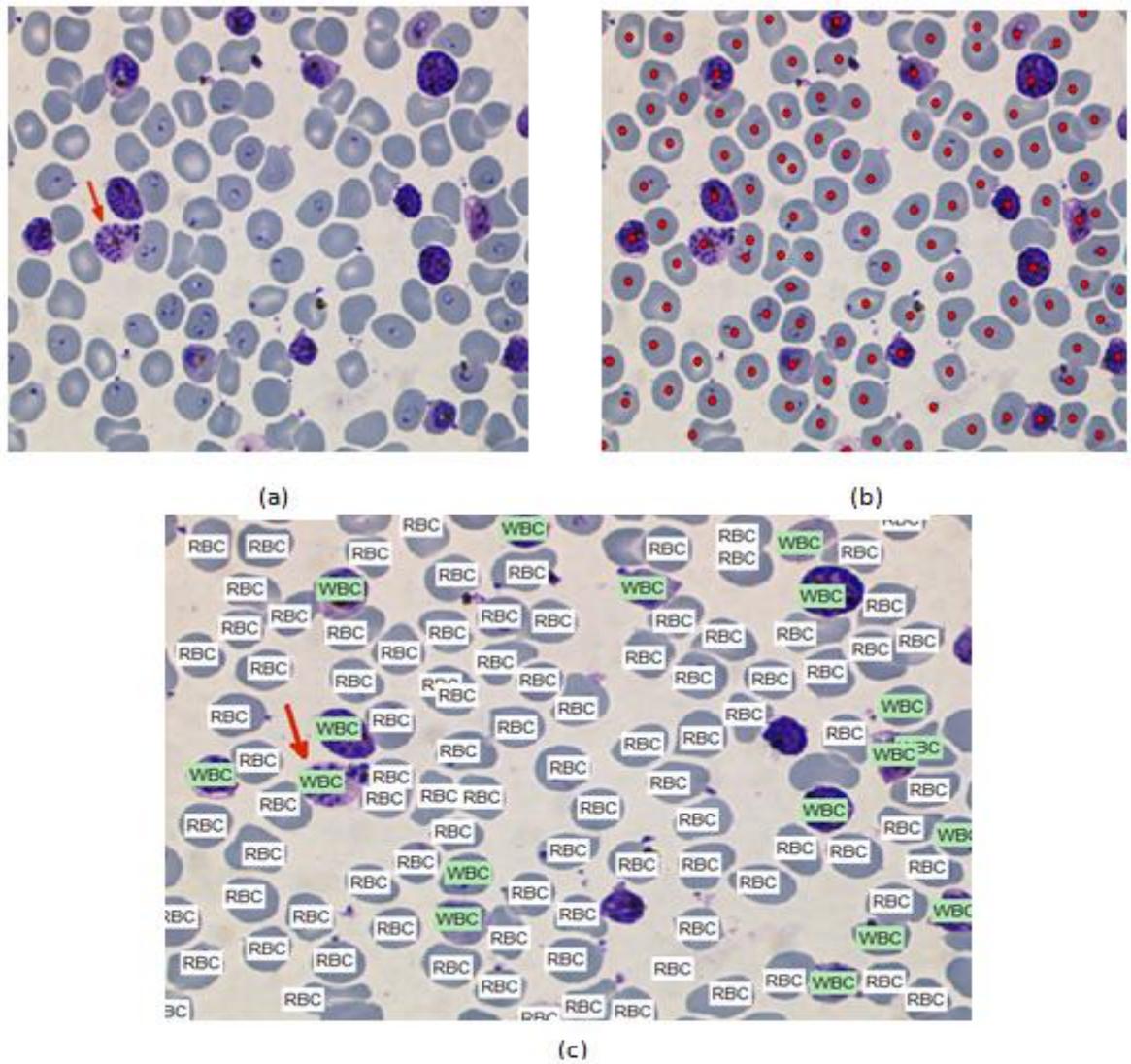


Figure 5.5: Disadvantages of Mean and Variance method.

(a) The input image. The red arrow points to a heavily infected RBC. (b) Image processed by the ratio-transform method which detected all the foreground objects equal to/greater than the size of RBC. (c) Resultant image after applying Mean and Variance method. The infected RBC pointed by the red arrow has been falsely identified as a WBC since its mean and variance falls within the range of mean and variance of a WBC.

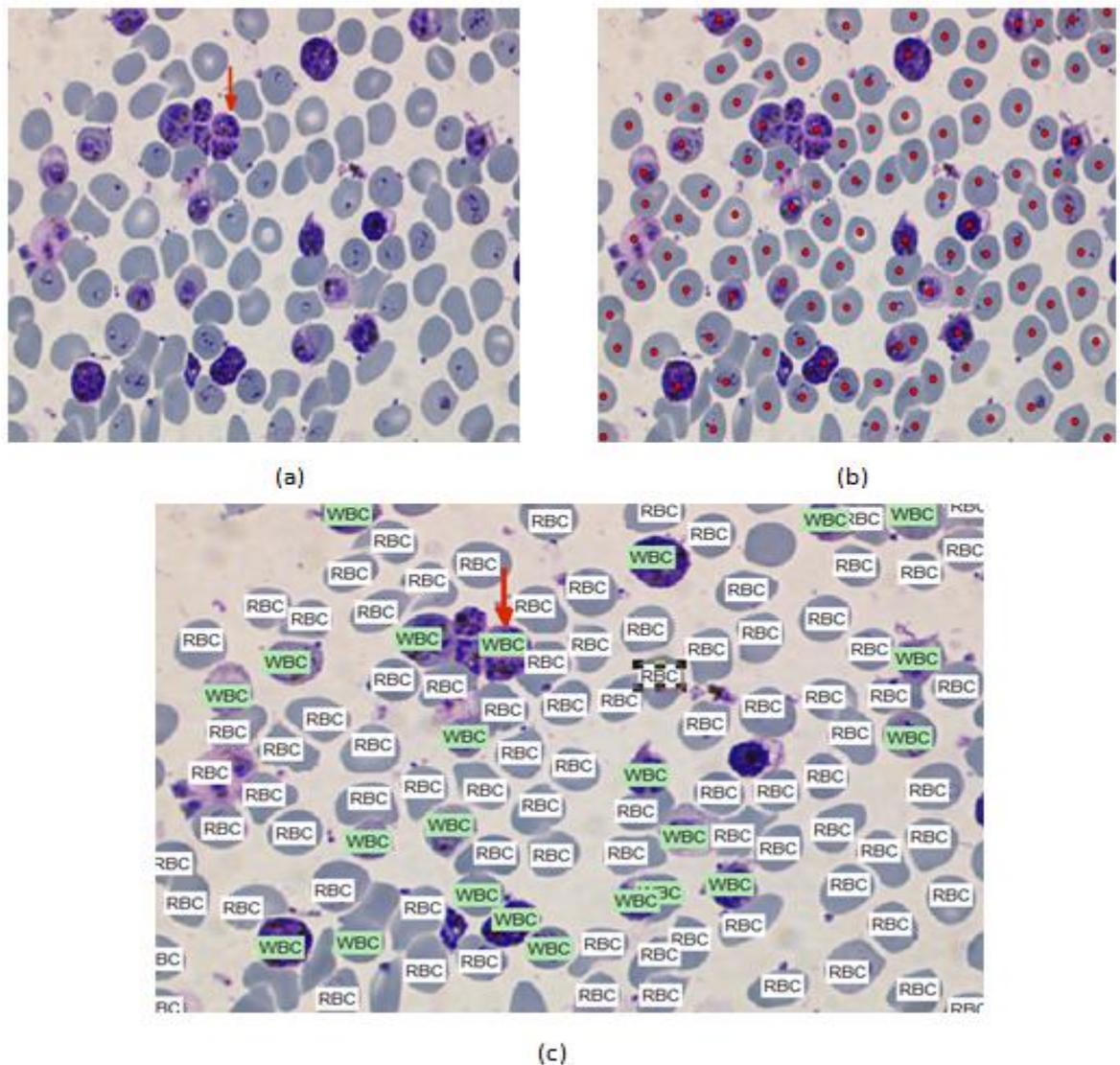


Figure 5.6: Disadvantages of Mean and Variance method.

(a) The input image. The red arrow points to a deconstructed cluster of infected RBCs with schizont life stage. (b) Image processed by ratio-transform method which detected all the foreground objects equal to/greater than the size of RBC. (c) Resultant image after applying Mean and Variance method. The infected RBC pointed by the red arrow has been falsely identified as a WBC because its mean and variance falls within the range of the mean and variance of the WBC. Other potential infected cells in the image were also wrongly identified as WBCs.

5.2.2 Modified Algorithm based on Mean and Cell Area

In order to effectively discriminate infected RBCs from WBCs, a modified algorithm with mean gray-level intensity and size of the cells has been developed which addresses the issues related to the previous methods. The algorithm uses the information obtained from the morphological filtering (dilation using ring shaped structuring element followed by erosion using disk shaped structuring element) explained in Chapter 3. The filtered image is a morphologically closed image, $\mathbf{I_E}$, (The term ‘closed’ is used, even though the image is dilated and eroded using different structuring elements.) as shown in Figure 5.10(b). The peak detection algorithm applied on a ratio transformed $\mathbf{I_E}$ will provide the location of the centroid of each cell. Since the WBC radius is greater than normal RBCs, the closing operation (with a structuring element of radius 70% of the RBC) will not remove the WBCs, as shown in Figure 5.10(b). The histogram (Figure 5.10(c)) of the image in Figure 5.10(b) clearly shows two peaks representing the WBCs and so can be used to distinguish the WBCs from the other cells. The modified algorithm is developed such that, instead of the gray-level image, the closed image $\mathbf{I_E}$ is processed. The mean intensity of the region connected to each pixel of the closed image at the location provided by the ARR transform method is calculated. The located pixels with a connected pixel neighbourhood having an area greater than that of a RBC and an average intensity less than a selected threshold will be diagnosed as a WBC and those above the threshold is considered as a RBC. The flow chart representing the algorithm is given in Figure 5.7.

The flow chart shows a single iteration of the algorithm. The total number of iterations depends on the total number of peaks detected by the peak detection algorithm of the ARR transform method. The parameters $A_{(X,Y)}$ is the area containing the connected pixels surrounding the location (X,Y) and $A_{(RBC)}$ is the area of RBC calculated from the granulometry method.

Automated Malaria Diagnosis Using Mobile Phones

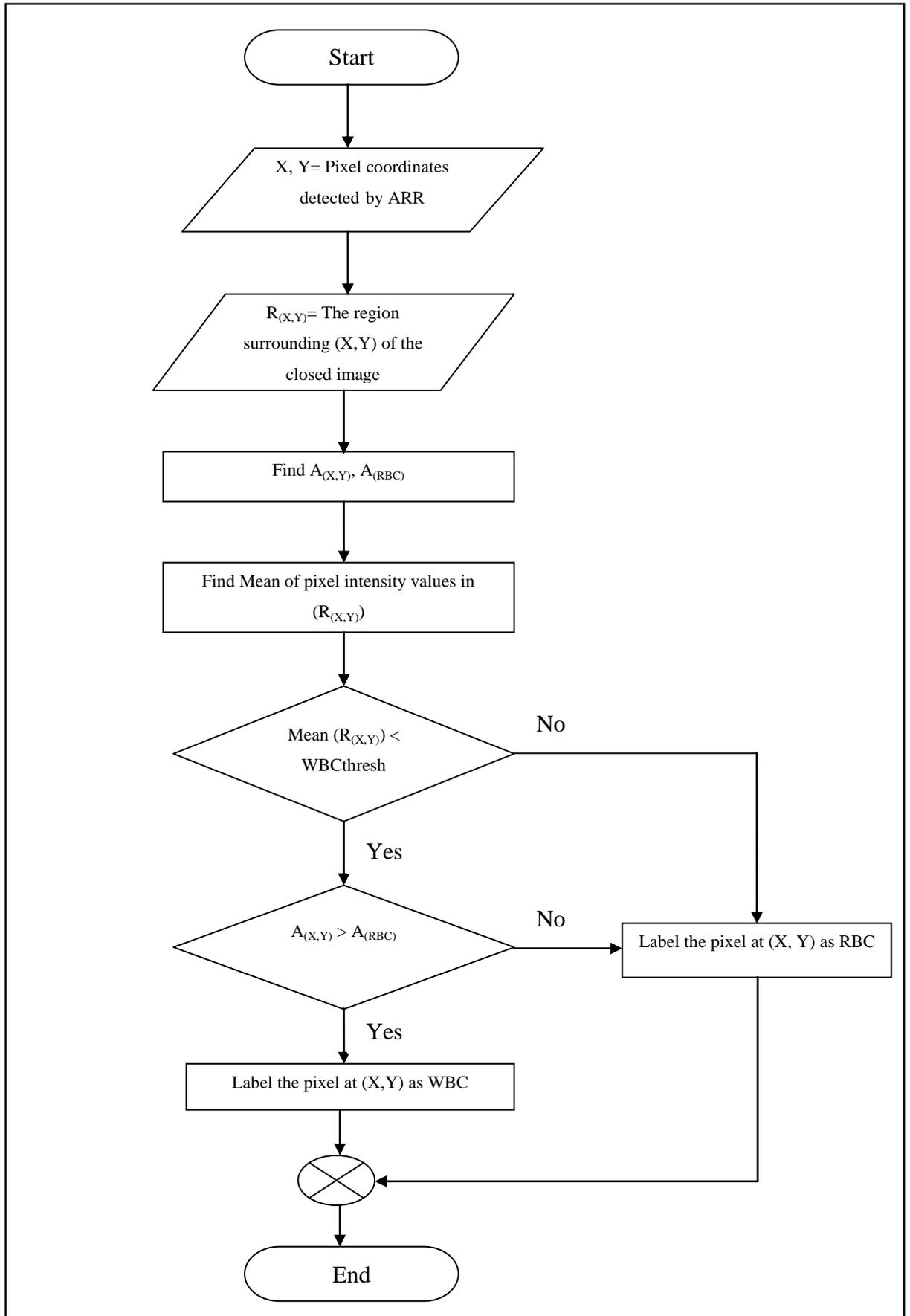


Figure 5.7: RBC and WBC differentiation algorithm Flowchart

5.3 Experimental Results

The result obtained after applying the algorithm on images of various density, staining concentration and illumination are given in Figures 5.8-5.12. Figures 5.8 and 5.11 are taken under same imaging conditions but the concentration of cells is different. Figure 5.11 is a poorly smeared slide with overlapping and destroyed cells. Figure 5.9, 5.10 and 5.12 are taken under same staining and imaging conditions. However, Figures 5.9 and 5.10 does not contain any WBCs. Figure 5.12 has three different types of WBCs present in them.

In Figure 5.8 and 5.11, the histogram shows two distinct peaks, one below 50 and one above 150. From the histogram information, it is easy to distinguish the cells by setting the WBC threshold between 50 and 150. However, in Figures 5.9 and 5.10, the staining concentration on infected cells is very high. The histogram information shows the intensity values spread out from 80 to 250. For certain types of images such as the one shown in Figure 5.12, the histogram is less informative than the others. The intensity level of RBC and WBC will overlap with each other. Hence the additional attribute of the size of the cells is a very useful predicate to reliably discriminate the cells. After a set of manual experimentation, the WBC threshold has been set to 150.

5.3.1 RBC Estimation

The information after differentiation can be used to estimate the total number of RBCs in the image. The total number of peaks detected by the peak detection algorithm applied to the ARR transformed image will give the total number of foreground components comprising the WBCs and RBCs. The cell differentiation algorithm explained above locates and label the WBCs and RBCs in the image. The difference in the total number of peaks and the number of located WBCs will give the total number of RBCs in the image.

$$\text{Number of RBCs} = \text{Total number of peaks} - \text{Number of WBCs}$$

This information is used to estimate the parasitemia of the blood which is the ratio of infected RBCs to the total number of RBCs.

Automated Malaria Diagnosis Using Mobile Phones

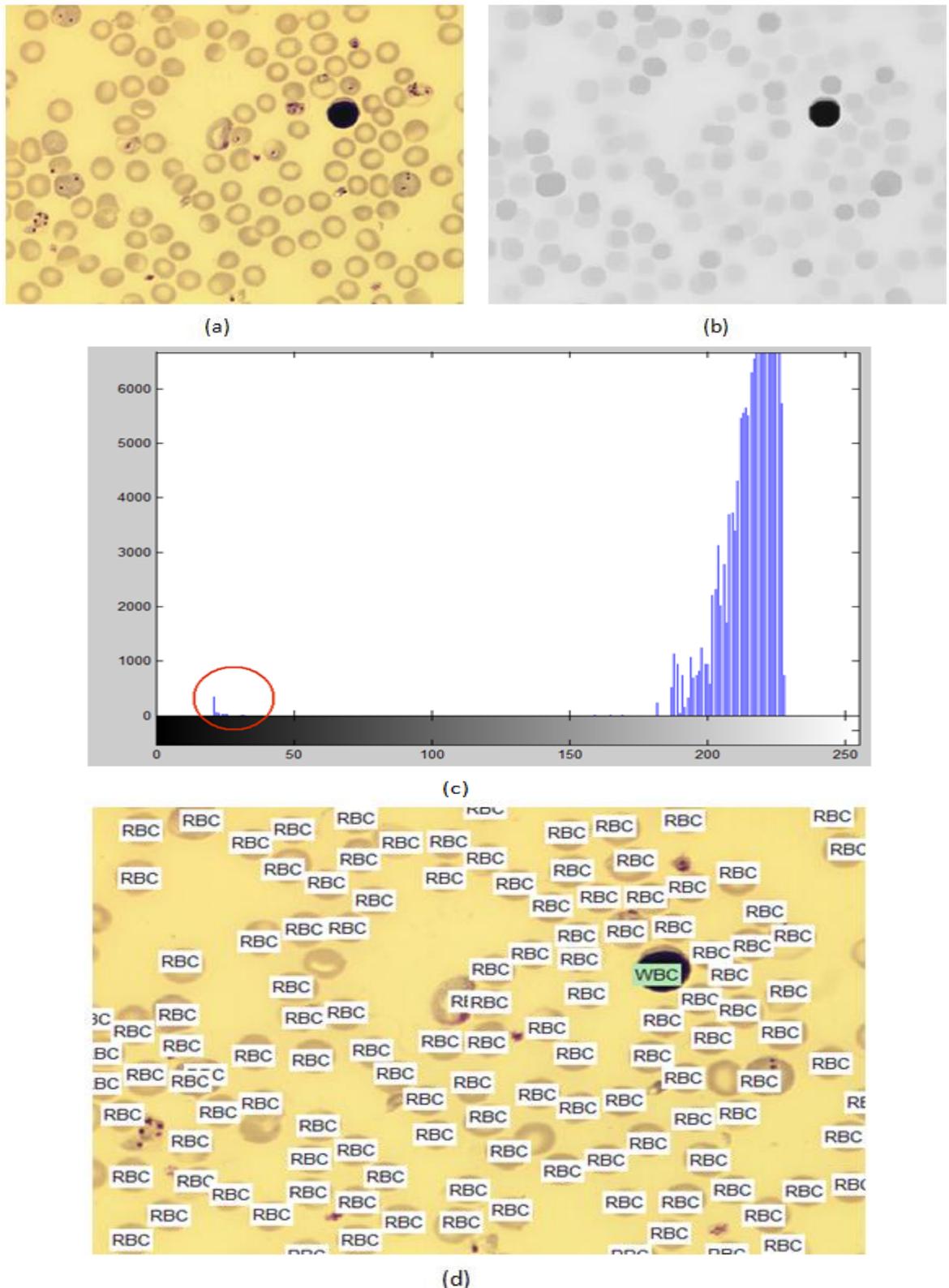


Figure 5.8: WBC and RBC differentiation algorithm.

(a) RGB Image. (b) Morphologically closed gray-scale image. (c) Histogram of the image in (b). The histogram clearly has two peaks which distinguish the WBC from RBC + background intensities. The circled area shows the WBC intensity range. (d) The resultant image with WBC and RBC differentiated.

Automated Malaria Diagnosis Using Mobile Phones

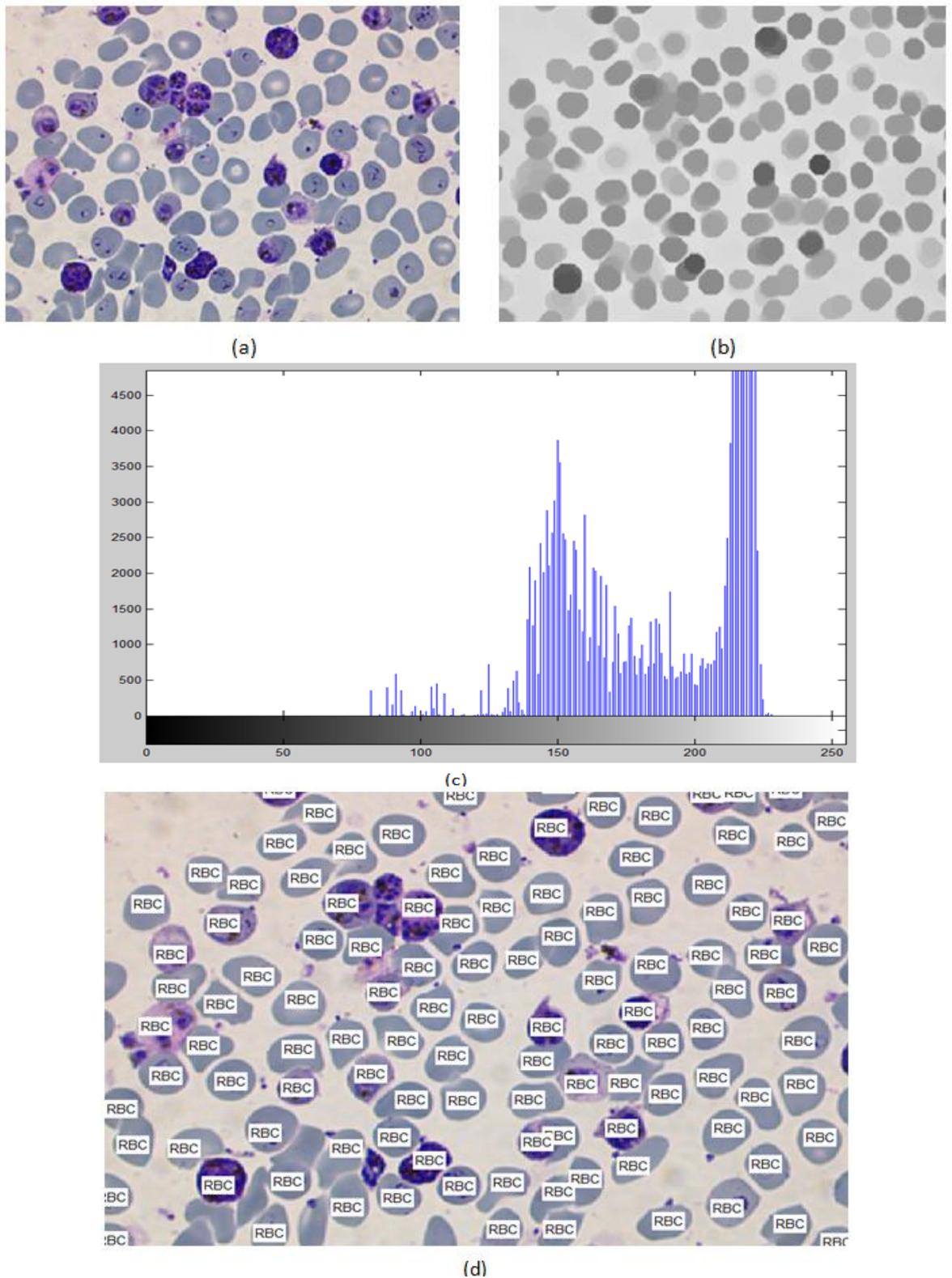


Figure 5.9 WBC and RBC differentiation algorithm.

(a) RGB Image. (b) Morphologically closed gray-scale image. (c) Histogram of the image in (b). Since there are no WBCs in the image, the intensity values are distributed between 80 and 250 which represent the RBCs and the background. (d) The resultant image with WBC and RBC differentiated.

Automated Malaria Diagnosis Using Mobile Phones

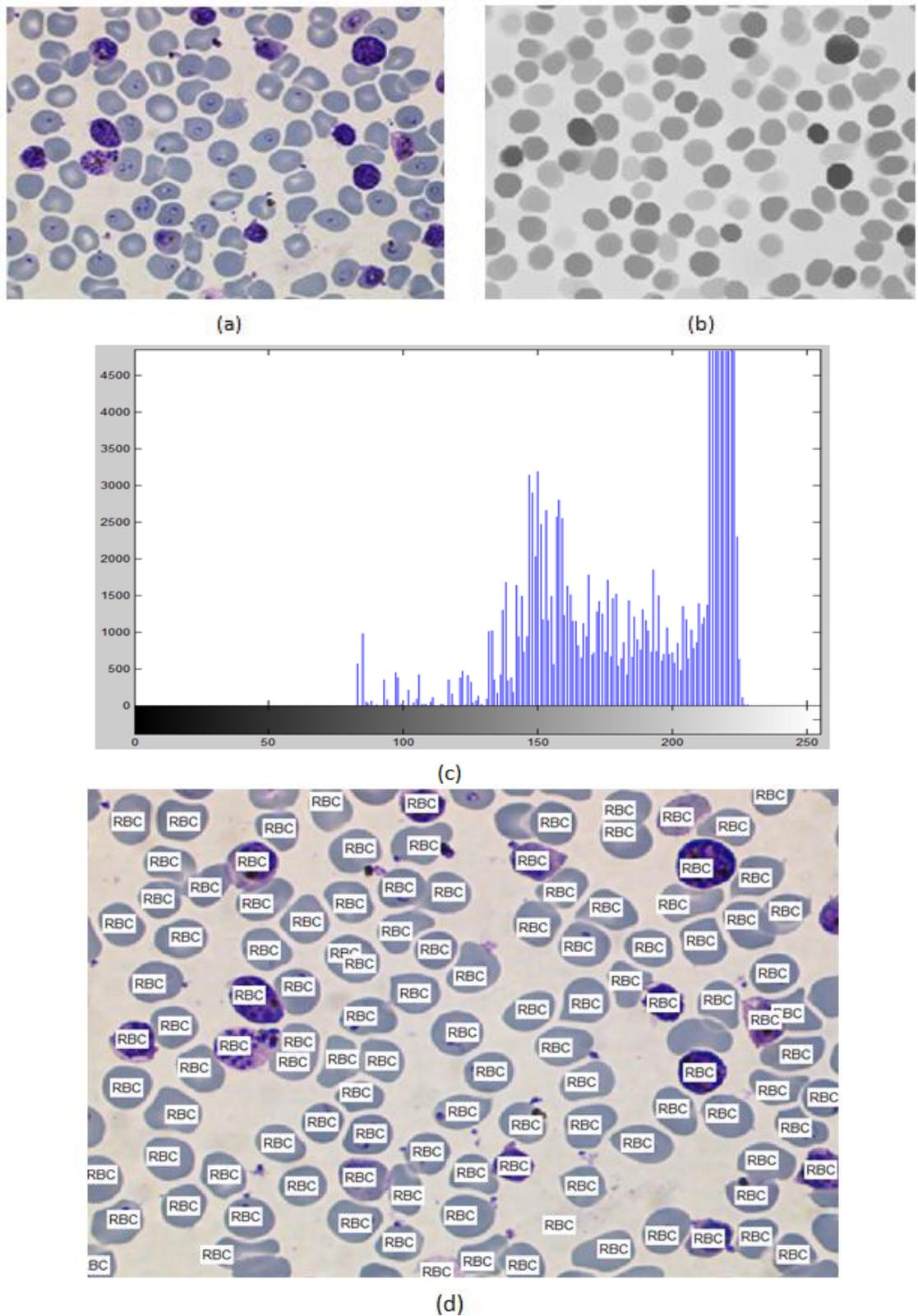


Figure 5.10: WBC and RBC differentiation algorithm.

(a) RGB Image. (b) Morphologically closed gray-scale image. (c) Histogram of the image in (b). In the absence of WBCs the intensity levels in the histogram are distributed well above 50. (d) The resultant image RBCs differentiated.

Automated Malaria Diagnosis Using Mobile Phones

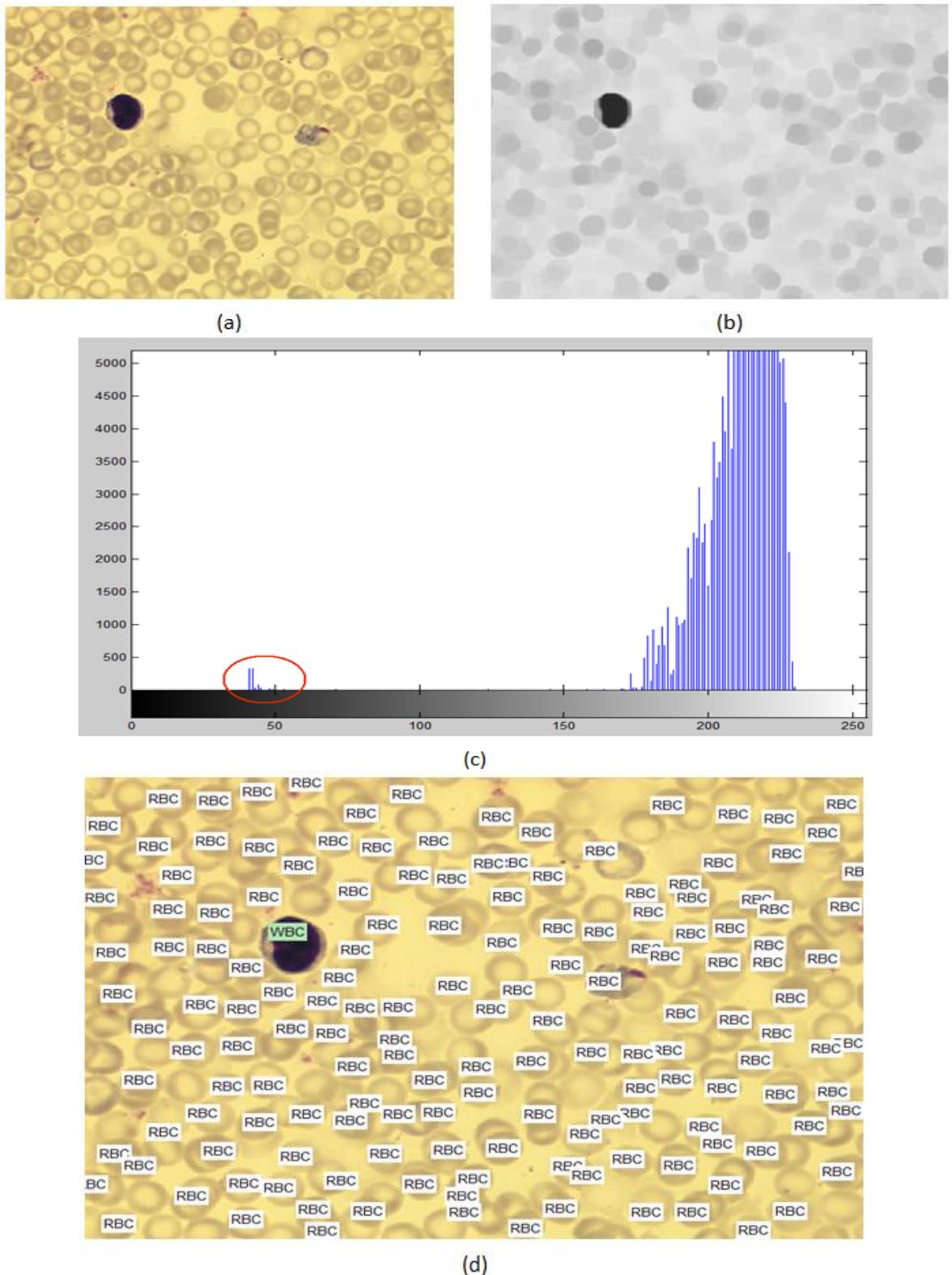


Figure 5.11: WBC and RBC differentiation algorithm.

(a) RGB Image. (b) Morphologically closed gray-scale image. (c) Histogram of the image in (b). In the presence of a WBC, the histogram contains two peaks. The WBC threshold falls within the circled area. (d) The resultant image with WBC and RBC differentiated.

Automated Malaria Diagnosis Using Mobile Phones

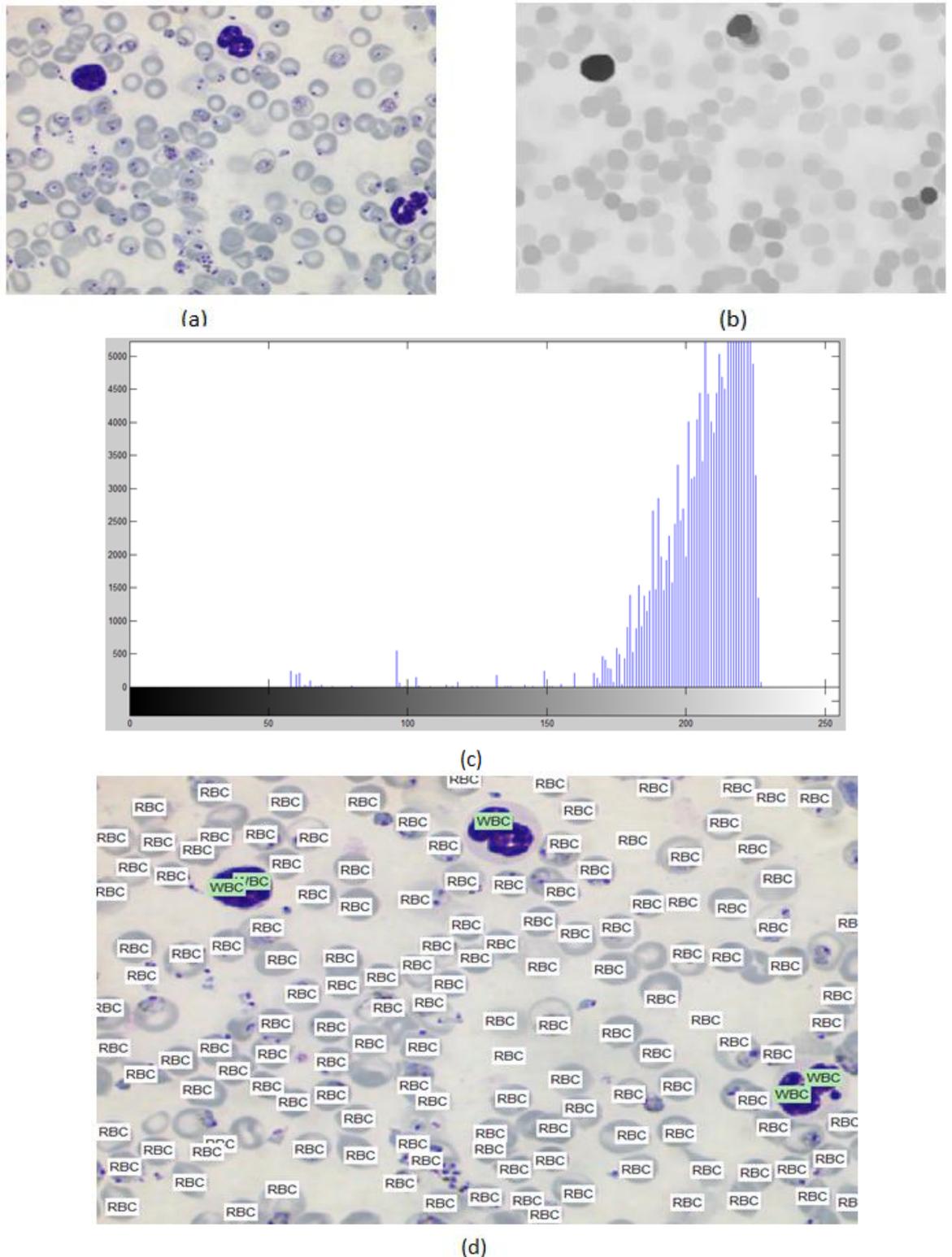


Figure 5.12: WBC and RBC differentiation algorithm.

(a) RGB Image with three types of WBCs. (b) Morphologically closed gray-scale image. (c) Histogram of the image in (b). The histogram does not give distinctive peaks hence setting the threshold is tricky. However, the WBCs are distinguished using the size and mean intensity of the pixel neighbourhood. (d) The resultant image with WBC and RBC differentiated.

5.4 Discussion and Summary

In this chapter a novel RBC and WBC differentiation algorithm has been presented. Unlike existing segmentation techniques, the algorithm aims to discriminate and estimate the population of the RBCs which will facilitate the parasitemia calculation. The study explored different possibilities of separating the RBCs and WBCs which are already identified as foreground components by ARR transform method. The diversity and unpredictable nature of the images were the major challenge faced during the research to deal with the problem. However, an effective RBC (and WBC) identification tool has been presented which will use the mean intensity and size of the cells as the main predicates for differentiation. The novelty of the method lies in the fact that instead of processing the gray-scale image, the mean intensity of the closed (morphologically dilated and closed using different structuring elements) image is used. Instead of calculating the local mean of all the pixels, only those pixels located as the centroid of the foreground components are processed which is then followed by measuring the area of those pixels that are assumed as WBCs.

The method is a cell-differentiation procedure rather than a segmentation technique. Hence, any performance comparison with conventional segmentation procedures is not valid. However, in comparison with existing segmentation methods such as explained in [54], [61] and [62], which involve identification of image components, the method is simpler and computationally effective. It does not require any pre-processing operations for colour normalisation and/or noise filtering and avoids the necessity for binary watershed lines, marker selection and extraction (c.f. Chapter 4). Even though the calculations involve estimating the mean intensity, the computation is limited to selected pixels rather than the entire population. Also the area estimation is performed only on those pixels that are suspected to be WBC.

The RBC estimation (counting) is another advantage of the method which is an important criterion for parasitemia calculation and is one of the main characteristic features in malaria diagnosis. Apart from malaria diagnosis, the method can also be used as a standalone WBC or RBC detection/estimation tool for various blood imaging applications including the processing of human and animal blood. The method could also be used for

Automated Malaria Diagnosis Using Mobile Phones

platelet detection with required modification on the cell radius parameter used in the ARR transform method.

The method is highly sensitive to RBC radius which can vary under different imaging conditions. A radius calculation algorithm using granulometry is used prior to the processing of the image. However, under a consistent image acquisition set up and initial calibration, the system will automatically perform the operations and is insensitive to variations in staining concentration and non-uniform illumination.

Chapter 6

Malaria Parasite Identification and Life-stage Recognition

The gold standard for malaria parasite identification is the microscopic examination of a thin/thick blood smear and detecting the stained particles in the blood film. The blood smears are stained using the ‘Romanovsky’ stain technique which is a mixture of polychrome and eosin [63]. In thin blood smears, Giemsa stain is used where as in thick films, Field’s staining technique is preferred, and both belong to the Romanovsky staining technique. For automated malaria diagnosis, the parasite identification is performed by extracting the stained objects in the image and hence stained pixel extractions is one of the most important aspect of automated diagnosis.

6.1 Stained pixel recognition

In a malaria-infected blood image, there are different categories of stained regions. The largest stained object is the cytoplasm of the WBC, possessing a large area of connected stained pixels. The smallest object will be the artefacts, platelets and parasitic nuclei. The RBC region falls within the WBC and platelets, and is the most common stained region. These regions can be easily distinguished using the parameters such as the size and staining concentration. However, the perception of the latter depends on the lighting and imaging conditions. Using a standard lighting set-up, the WBCs and artefacts appear as dark blue-purple components, parasites as purplish pink and platelets as saturated pink colour. The RBCs however, will be very lightly stained using the Giemsa-staining in thin blood films and hence any stained pixel extraction tool will not consider them as a stained component.

The malaria parasites entering the blood stream from the liver hide in the erythrocytes (RBCs) to avoid being detected and destroyed by the leukocytes (WBCs). The parasite once inside the RBC matures and divides asexually. Depending on the level of maturity, their appearance within the erythrocytes varies. Figure 6.1 shows the different stages of malaria parasite infection. The nucleus of the parasite undergoes staining when processed with Giemsa stain in a thin blood smear and appears as a purplish-pink colour.

Automated Malaria Diagnosis Using Mobile Phones

Figures 6.2-6.5 shows images taken under different imaging, processing and staining conditions. Both parasites and WBCs are deeply stained with purplish blue colour as shown in Figure 6.2, but the saturation level of staining in the WBCs is higher than that of parasitic nuclei. In spite of this distinctive feature, the identification of the parasites based on staining concentration can often leads to false detection. Since WBCs are considerably larger than parasites, size can be an important attribute in differentiating them. But, as shown in Figure 6.3, there are cases where the stained artefacts present in the image hold similar characteristics to the stained nuclei of WBCs. A detailed view of stained particles and related characteristics has been given in Chapter 2.

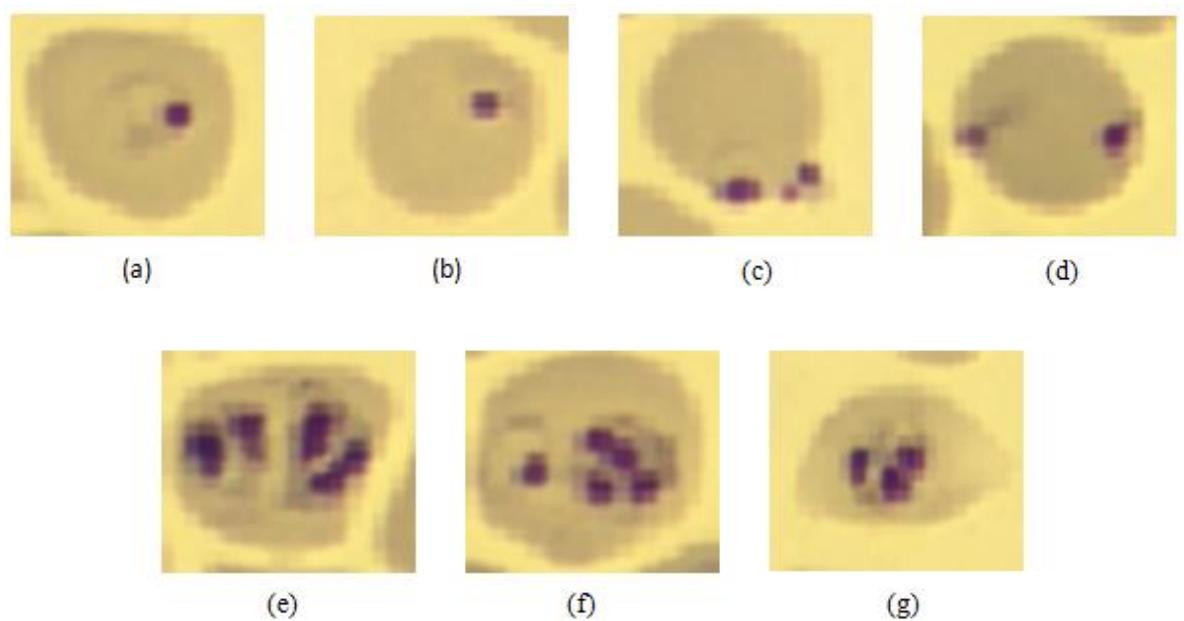


Figure 6.1: Variable Appearances of parasitic nuclei in the RBCs

(a) Ring Trophozoite with a ring of cytoplasm surrounding the nucleus. (b) Trophozoite with invisible cytoplasm. (c) Accolè form-Trophozoite parasitizes the marginal portion of the RBC (d) Trophozoites. (e), (f), (g) Various appearances of the Schizont life stage. All these infected cells were present in a single peripheral thin blood image.

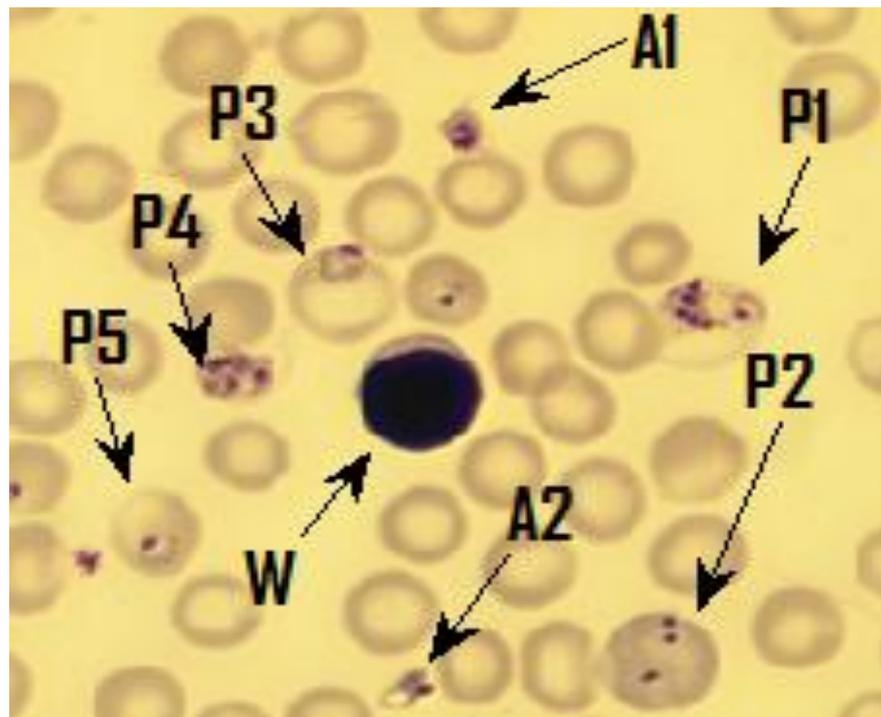


Figure 6.2: Stained particles in a blood image. W-White Blood Cells, P-Parasites, A-Platelets and Artefacts.

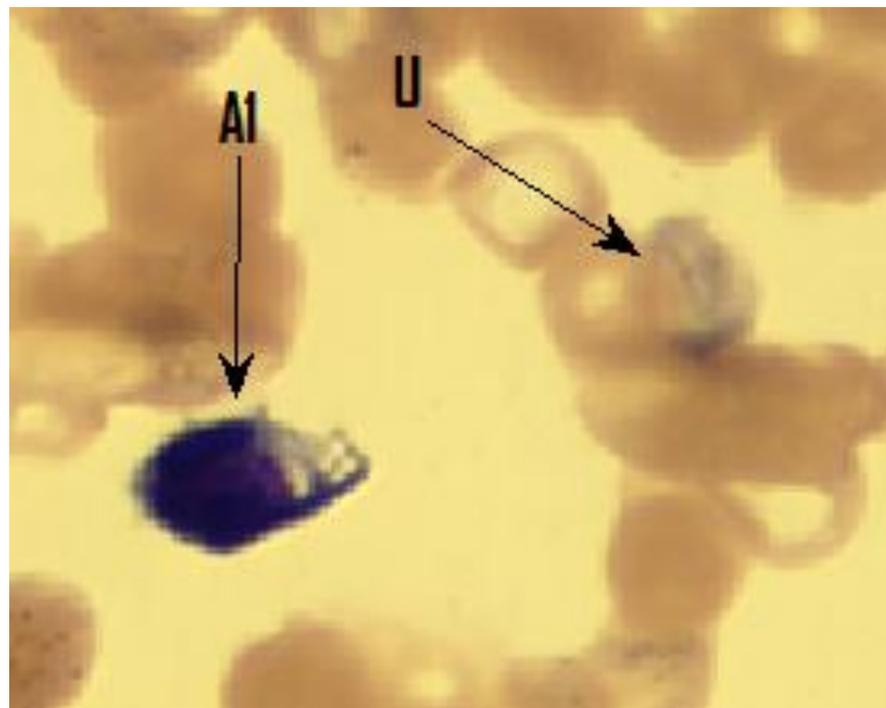


Figure 6.3: Example of images with different staining and illumination set up.

The artefact in this image has similar staining and size characteristic to the WBC shown in Figure 6.2. A-Platelets and Artefacts, U-Unknown.

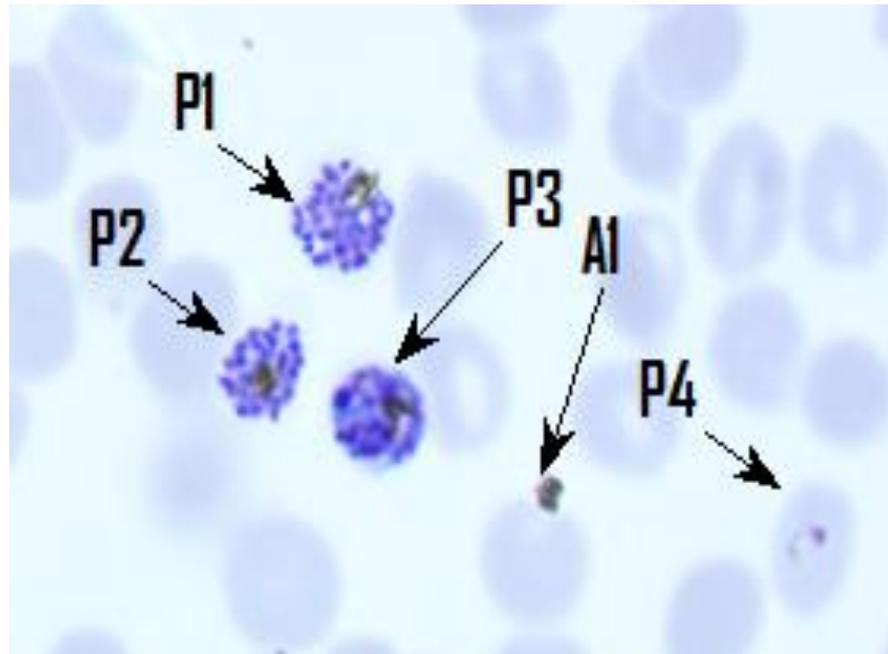


Figure 6.4: Example of images with different staining and illumination.

This image has a different illumination and staining concentration to the image in Figure 6.2. P-Parasites, A-Platelets and Artefact.

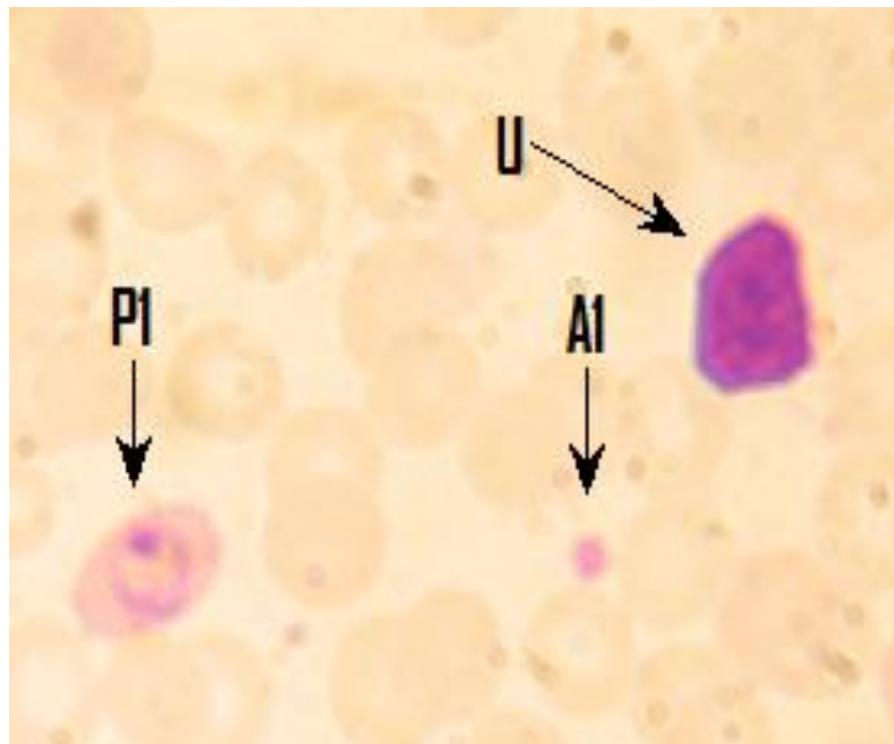


Figure 6.5: Example of images with different staining and illumination set up.

P-Parasites, A-Platelets and artefacts and other noises, U-Unknown.

Automated Malaria Diagnosis Using Mobile Phones

As explained in Chapter 2, there are various approaches for stained pixel extraction. In [19], Ruberto et al. explains an inductive strategy for detecting the stained pixels first, followed by its residing regions. In [20], a deductive approach is explained in which the regions were initially segmented followed by extraction of the stained pixels. Both approaches used histogram based thresholding for pixel extraction. In [42], an inductive strategy is adopted but, instead of histogram-based thresholding, a Bayesian pixel classifier was employed. Similar decision-based classification approaches were widely used thereafter for parasite extraction. In [64], Sheikhsosseini et al. explains an automated diagnosis of malaria which involves stained pixel extraction using intensity and colour followed by segmentation of the objects using a stained circle matching algorithm. The pre-processing step involves non-linear diffusion filtering followed by a series of decision rules to diagnose the presence of parasites. Similar literature using decision-based classifiers was analysed in Chapter 2 [34]-[37].

In this Chapter a novel approach to parasite detection using the modified ARR method is described. The method classifies the parasitic nuclei within the RBC without the help of any rule-based classifier. The method also performs life-stage recognition of the malaria parasite. The work is published in [68, 69] with narrower contexts. The chapter also discuss the results of the overall procedures and a performance measurement has been conducted which will be explained in Section 6.4. It concludes with a discussion based on performance analysis with possible improvements.

6.2 Identification of Infected Cells

The infected cells are the RBCs in which the parasite resides. In order to calculate the parasitemia (level of infection), the detection of infected cells is necessary. It is safer to locate the RBCs before searching for the stained parasitic nuclei. Otherwise any stained pixel classification tool will give a false positive identification of artefacts, platelets and WBCs present in the image that are outside the RBCs. The detection of infected cells is a challenging step purely because of the inconsistency in its visual characteristics in the image. Depending on the level of infection, the appearance and spatial geometry of the infected RBCs vary. In addition, poor processing of the smear and imaging conditions also make the classification difficult. Hence, rather than developing a generic stained pixel

Automated Malaria Diagnosis Using Mobile Phones

extraction tool, this work has focussed on developing a semi-adaptive tool (explained in Chapters 4 and 5) that firstly locates the RBCs and then differentiates the infected cells.

6.2.1 Identification of infected cells based on size and morphology

Morphologically, an infected cell has one or more stained nuclei and depending upon the maturity of the parasites, an RBC will contain 1-8 nuclei. Each nucleus is a connected group of pixels in an RBC region, with saturation based on staining. Smaller size, circular shape and saturated colour are the distinctive features of these nuclei. A classification method based on these features was initially developed which utilizes a modified ARR method to detect the stained pixel groups within the RBCs. The RBCs are already located by the ARR transform method. The radius of each cell and the nucleus depends on the size of the pixel (in nm). For example, for an image of 512x381 pixels, each pixel is approximately 0.25 μ m, the RBC radius is 12-13 pixels and radius of each nucleus is approximately 2-3 pixels.

The ARR transform method calculate the ratio of intensities of a circular object using a concentric ring structuring element. In order to detect the RBCs, a concentric ring structuring element with outer radius 25% larger than RBC radius and inner radius 90% of the RBC radius was used which will then locate all the RBCs (and WBCs). Similarly, the ARR method can be modified to detect the circular parasite nuclei by using a concentric structuring element with the above specification applied on radius of the nucleus. Hence the modified ARR transform method will have a concentric ring structuring element with outer radius 25% larger than, and inner radius 90% of the parasitic nuclei radius. Thus in the lighter background of RBC, the peak detection algorithm picks up the centroid of the nuclei/chromatin dots of the parasite from the ratio transformed image. Since the centroid location of the RBC is already known, the ring ratio will be performed in the neighbourhood of the RBC centre. This will eliminate any risk of false detection of platelets /artefacts that are outside the RBC radius.

The peak detection algorithm is performed to identify the centre pixel of each nucleus and according to the number of nuclei present within each RBC, life stages can be recognised. The structuring element used for the ARR transform method as well as peak detection algorithm will be approximately equal to the size of the nucleus (2-3 pixels).

Automated Malaria Diagnosis Using Mobile Phones

Figure 6.6 shows the result obtained by using the ARR transform method using a structuring element with smaller radius based on the radius of the parasite nuclei. The locations of the centroid of the nuclei picked up by the peak detection algorithm on the ratio transformed image are marked with red crosses. The background is the RBC that has already been detected (as explained in Chapter 5).

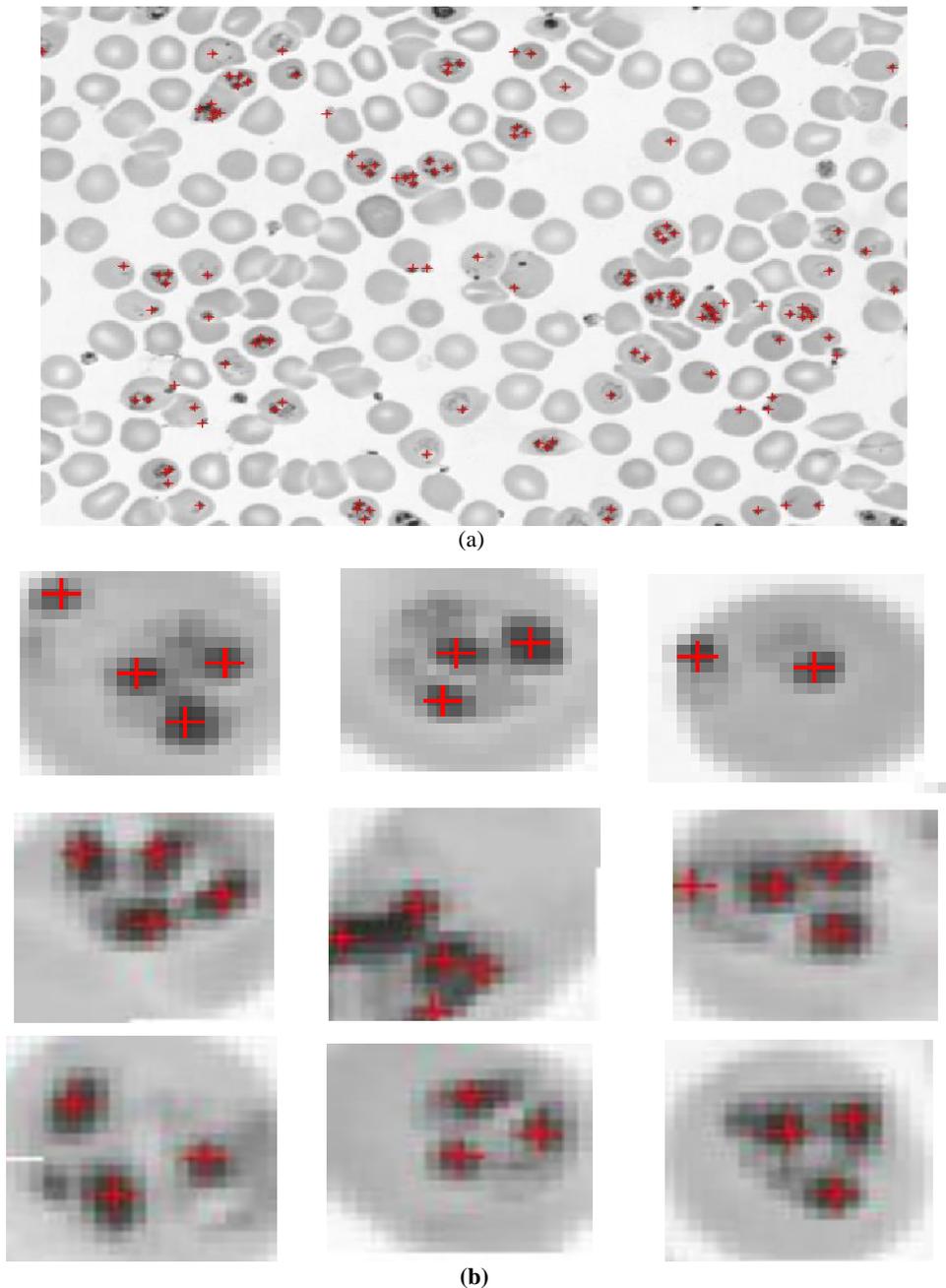


Figure 6.6: Infected cell identification using modified ARR method.

(a) The image processed to detect the infected cell. All the RBCs with the red crossing have been diagnosed as infected cells. (b) Zoomed sections of the image in (a) in which the parasite nuclei are detected within the RBC and are marked using red crossing.

Automated Malaria Diagnosis Using Mobile Phones

Hence to conclude, the algorithm performs the ARR transform method twice, firstly to detect the foreground components, RBCs and WBCs and secondly to detect the stained nuclei within the RBCs. The latter is performed using a concentric ring structuring element based on the size of the nuclei and the former using a structuring element based on the size of the RBCs.

The method was tried on a set of 15 images having around 1500 RBCs. Even though the sensitivity to detect the infected nuclei was more than 90%, the specificity of the algorithm was very poor. This is primarily due to the fact that the ARR method is very sensitive to sharp variations in the gray scale intensities which often happen when the cells are highly overlapped leaving sharp shadows which will be picked by the ARR method. In addition, the area in which the modified ARR transform method is applied is also crucial. In order to search for the stained nuclei, the modified ARR transform method is applied in the neighbourhood of all the located centre pixels. The neighbourhood has an area equal to the area of a typical RBC. However, in any image, there will be a few RBCs with deformities that occur due to several factors such as parasite invasion (Spherocytes) , destruction while processing of the smear, and genetic or acquired disorders (Sickle Cell Anaemia, Thalassemia) [65]. The edges of these abnormal cells with the centre-pixel neighbourhood less than the normal RBCs area will be picked up by the ARR transform method and will falsely identify as set of nuclei. This will highly impair the accuracy of the overall diagnostic set up. Hence in order to improve the accuracy, a more reliable and robust method based on the morphology and colour features was developed which is applied along with the modified ARR transform method to detect the infected cells. The method performs visual perception analysis on Hue-Saturation-Value (HSV) colour space and the details have been explained in the following Section.

6.2.2 Infected-Cell Identification using the Hue-ARR Algorithm.

The Hue-Saturation-Value (HSV) colour space analysis has been used in various image segmentation procedures due to its distinctive feature perception quality. Unlike RGB images, HSV images are closer to the human perception of objects and describe the colour sensation more precisely [51]. The hue and saturation components can be used as dominant attributes to extract pixel features as they are more consistent and robust in perceiving the

Automated Malaria Diagnosis Using Mobile Phones

colour features due to staining than the RGB interpretation [66]. They are used in several classification algorithms such as k-means clustering to group the pixels [67].

Like RGB images, the HSV colour images are also multichannel images with vector valued pixels. For images with high spatial correlation like the thin blood images, the hue component of the image can be used for processing, which will produce scalar valued pixels as in gray-level (monochromatic) images [52]. Nevertheless, unlike gray level images, hue component of the HSV image describes the pure colour of the image components and is well suited for describing colours in terms of practical human interpretation [51]. Even though monochromatic images are measurable, easily interpretable, and speed up the image processing operations, the RGB to gray-scale conversion often leads to the loss of colour-carrying information and the colour features are not very well differentiated. Since the most distinctive attribute of the parasitic nuclei is its stained colour features, a processing algorithm based on the hue component of the HSV colour- space was considered as it retains maximum colour description.

6.2.2.1 Definition

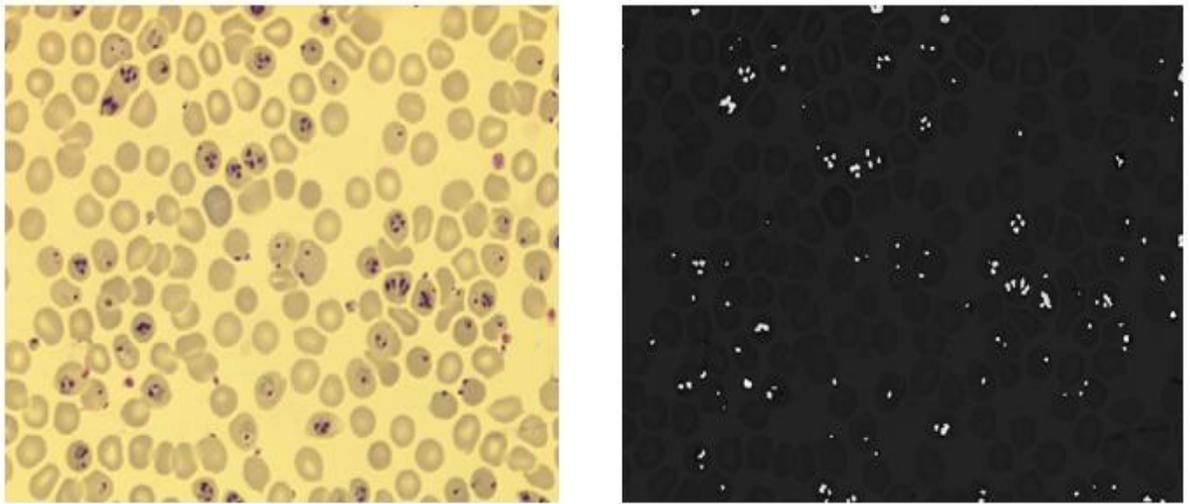
The hue component from a given RGB image can be obtained by performing the following transformation, pixel by pixel:

$$Hue = \begin{cases} \theta, B \leq G \\ 360 - \theta, B > G \end{cases} \quad (6.1)$$

Where,

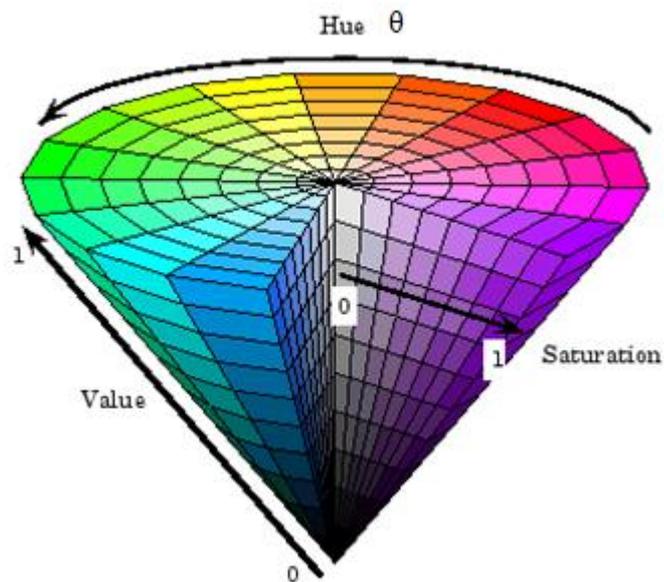
$$\theta = \cos^{-1} \left\{ \frac{(R-G) + (R-B)}{2(\sqrt{(R-G)^2 + (R-B)(G-B)})} \right\} \quad (6.2)$$

and R, G and B are the Red, Green and Blue component of the original RGB colour space respectively [51]. Figure 6.7 (c) represents the graphical representation hue and the resultant hue image obtained when applied on an RGB image is shown in Figure 6.7(b).



(a)

(b)



(c)

Figure 6.7: Hue component representation.

(a) RGB Image (b) Corresponding Hue component (c) Symbolic representation of Hue.

The hue component will highlight the parasite and other deeply stained components in the image as shown in Figure 6.7(b) and can be used to detect the stained nuclei in the image. In [19] the stained pixels are detected by the histogram thresholds on the top-hats of the hue and saturation channels of the image. A similar approach was initially analysed in

Automated Malaria Diagnosis Using Mobile Phones

which the hue component of the image is extracted at first, followed by image binarisation using thresholding. By detecting the regional maxima, the stained nuclei can be detected. However, the disadvantage of this method is false extraction of the artefacts and platelets. Even when the image does not contain any stained pixels, the regional maxima and thresholding will extract other pixels as stained pixels. Hence the approach was modified and a combination of the modified ARR method and hue component thresholding was applied.

6.2.2.2 Algorithm

The algorithm works as follows. Using the modified ARR method described in the previous Section 6.2.1, the RBC region and the corresponding suspected parasitic nuclei will be identified. The suspected regions where the nuclei have been detected will then be compared with the region surrounding the RBC locations in the binarised Hue image to give a final verdict on whether the region contains an infected cell or not. The hue component of the HSV image is at first converted to a binary image using thresholding. Unlike gray to binary conversion, the hue to binary conversion avoids the information loss to a great extent. Compared to luminance or grey level values ranging from 0 to 255, the hue value of a pixel falls within 0 to 1. Hence for a blood image with stained pixels of varying intensity range, the thresholding of the hue component to binary is more feasible because it interprets the true colours of the pixel and the filtering of stained parasitic nuclei from other stained components is easier. Once the binary image is obtained, a localized maxima search locates the white pixels. Sequentially, the modified ARR algorithm (as explained in Section 6.2.1) is applied to the located RBC regions and those regions with suspected nuclei are marked. If the location of the identified white pixels (of the binarised Hue image) falls within the marked RBC region (by the modified ARR method), the cell is identified as infected. The ARR transform method performed (as explained in Chapter 4) has already provided the centre of each RBC. Hence the search can be defined in terms of RBC radius surrounding the already located centroid. The Pseudo code for the algorithm is explained in Figure 6.8 (a). Figure 6.8 (b) shows the resultant image obtained after identification of the infected cells. The red circle marks the infected cells identified.

Automated Malaria Diagnosis Using Mobile Phones

HUE-ARR algorithm to detect the infected cells

Pks = Coordinates of the centre pixel of the infected RBC (suspected by the ARR modified transform)

$[x,y]$ = size of the image.

HueBin= Binarised Hue Image

\forall pixels coordinates x, y {

L =Length (Pks);

For $k=0:L-1$

If $(x,y) \in R(k)$, (where $R(k)$ is the region surrounding the k th coordinates in Pks).

If $\text{Max} [\text{HueBin}_{(R(x,y))}] = 1$, (where $R(x,y)$ is the region surrounding the pixel coordinates at (x,y) of the binarised hue image).

$Pks(k)$ = infected cell coordinate;

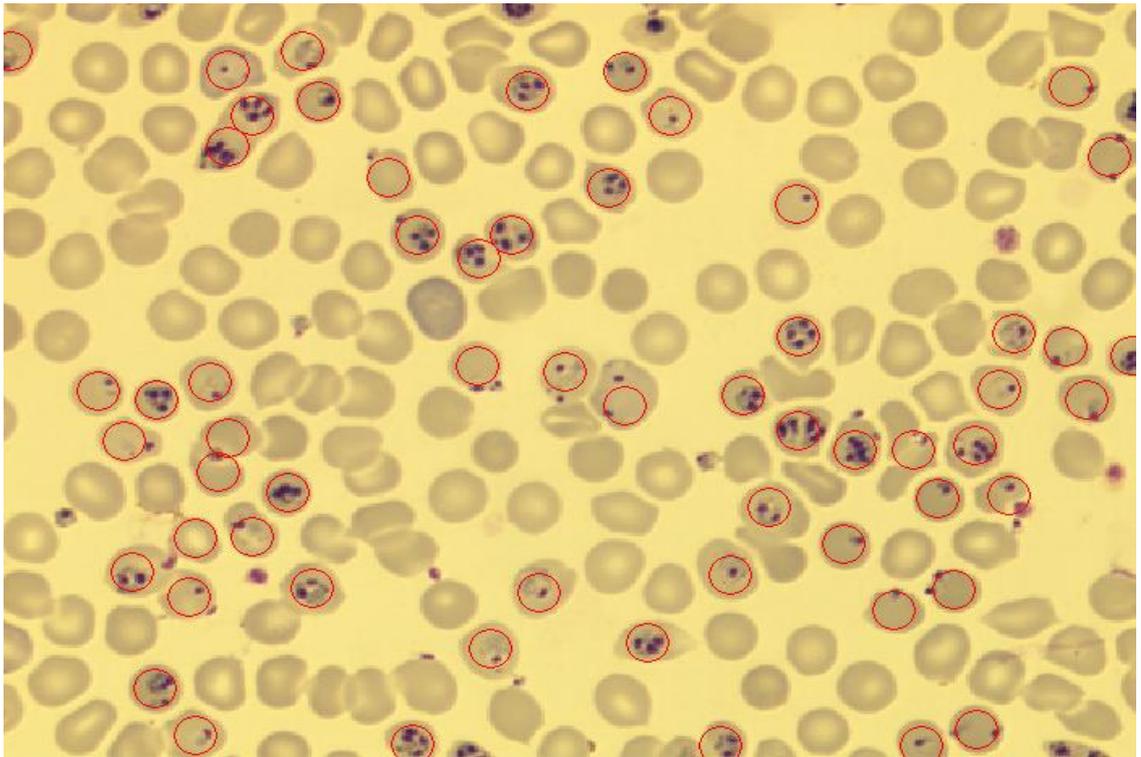
else

$Pks(k)$ = non-infected cell coordinate;

$k=k+1$;

end;

(a)



(b)

Figure 6.8: Hue-ARR algorithm. (a) Pseudo code of the algorithm. (b) Resultant image in which the infected cells are circled.

6.2.3 Experimental Results

The Figures 6.9-6.19 shows selected results obtained using the Hue-ARR algorithm. The blood slides are processed under different laboratory set ups and the images are developed under different staining and imaging conditions. Hence the images shown in this Section are carefully chosen irrespective of the diagnostic results, in order to give the reader an impression on the versatility of the inputs that a diagnostic tool has to expect. All these images were processed in Matlab [111] environment. Depending on the size of the images, the time taken for diagnosis varies. For an image of size 512x381 pixels, the time taken by the processor for a complete diagnosis is approximately 28 seconds.

Figures 6.9, 6.10 and 6.11 are obtained from the same slide, however, the density and distribution of the cells in each of them varies. In Figure 6.11, the cells are very closely packed with overlapping RBCs.

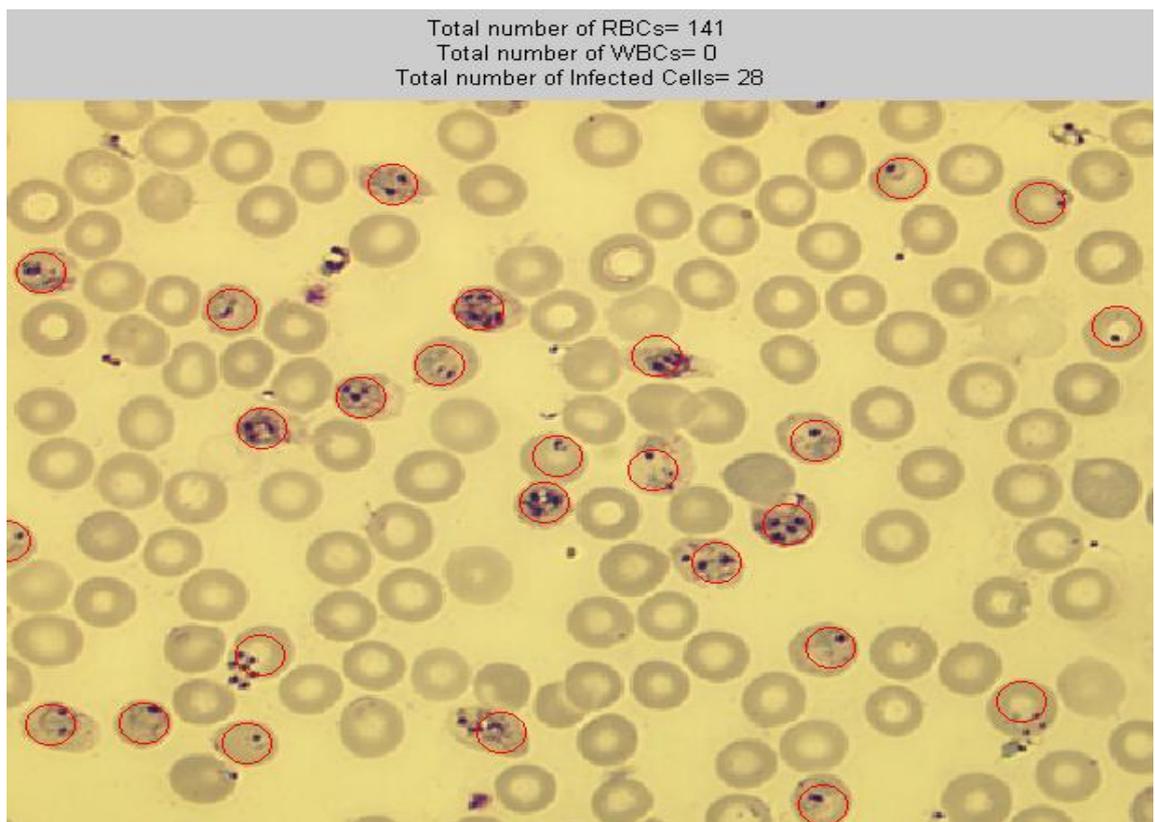


Figure 6.9: Processed image in which the cells that have been identified as infected are circled.

Automated Malaria Diagnosis Using Mobile Phones

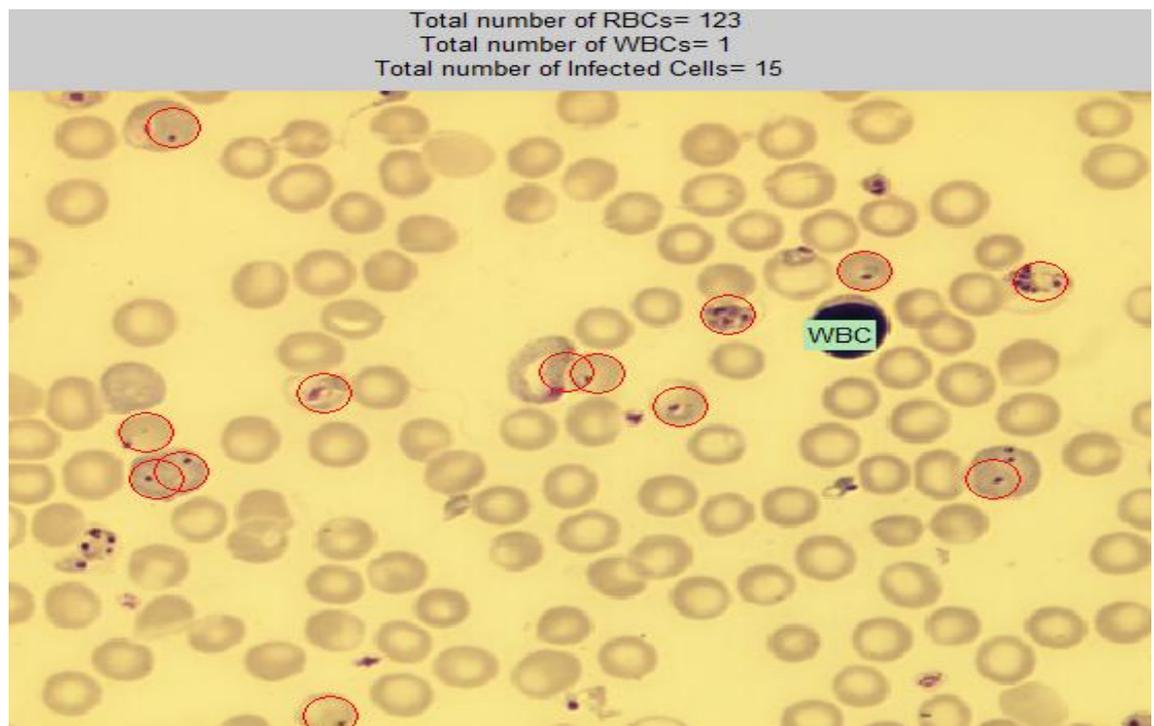


Figure 6.10: Processed image in which the cells that have been identified as infected are circled and WBC differentiated.

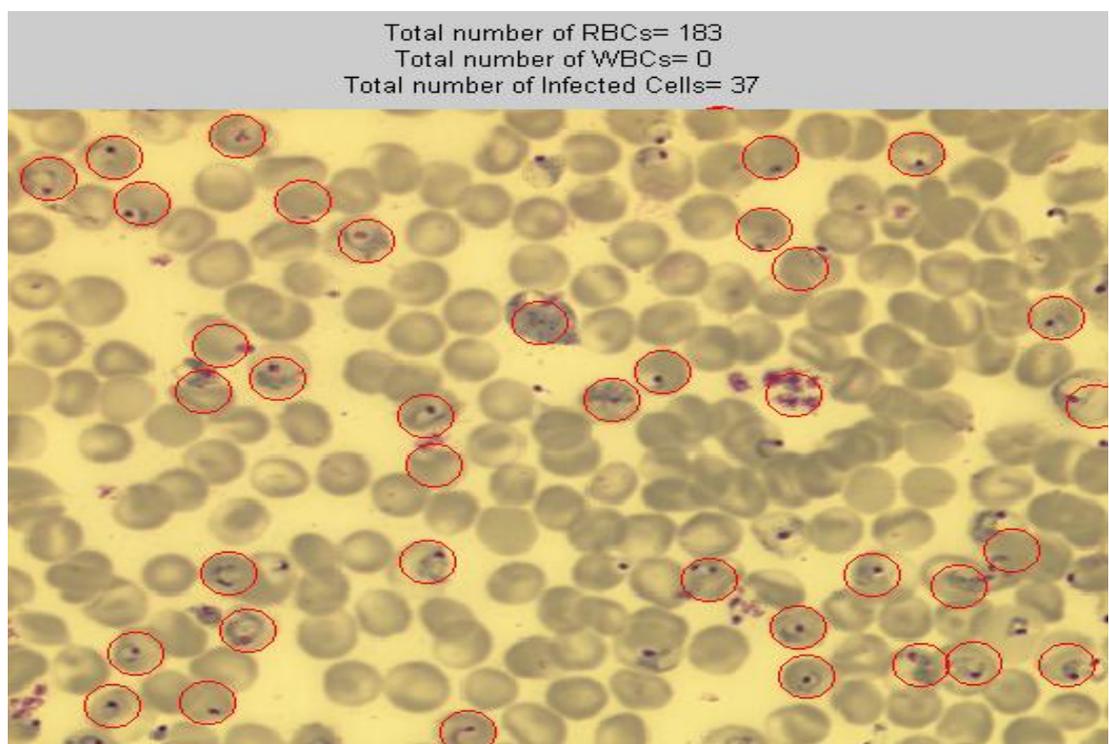


Figure 6.11: Processed image in which the cells that have been identified as infected are circled.

Due to the cells being closely packed, some of the infected nuclei are missed. However, more than 90% of the infected cells are found.

Automated Malaria Diagnosis Using Mobile Phones

Figures 6.12, 6.13 and 6.14 are acquired under different illumination set-ups and hence the image contrast and concentration of the staining is different from the others. Due to their higher resolution, the radius in pixels is larger for the processing of Figure 6.12 than in the other images. However, there are two false positives at the WBCs. Figure 6.13 is a negative image, with no infected cells. However some of the cells in the image have stained particle within an RBC which was suspected as being an artefact. The Hue-ARR algorithm diagnosed all the cells in the image as negative and the diagnosis was confirmed by the experts at the Pondicherry Institute of Medical Sciences, India who provided the image.

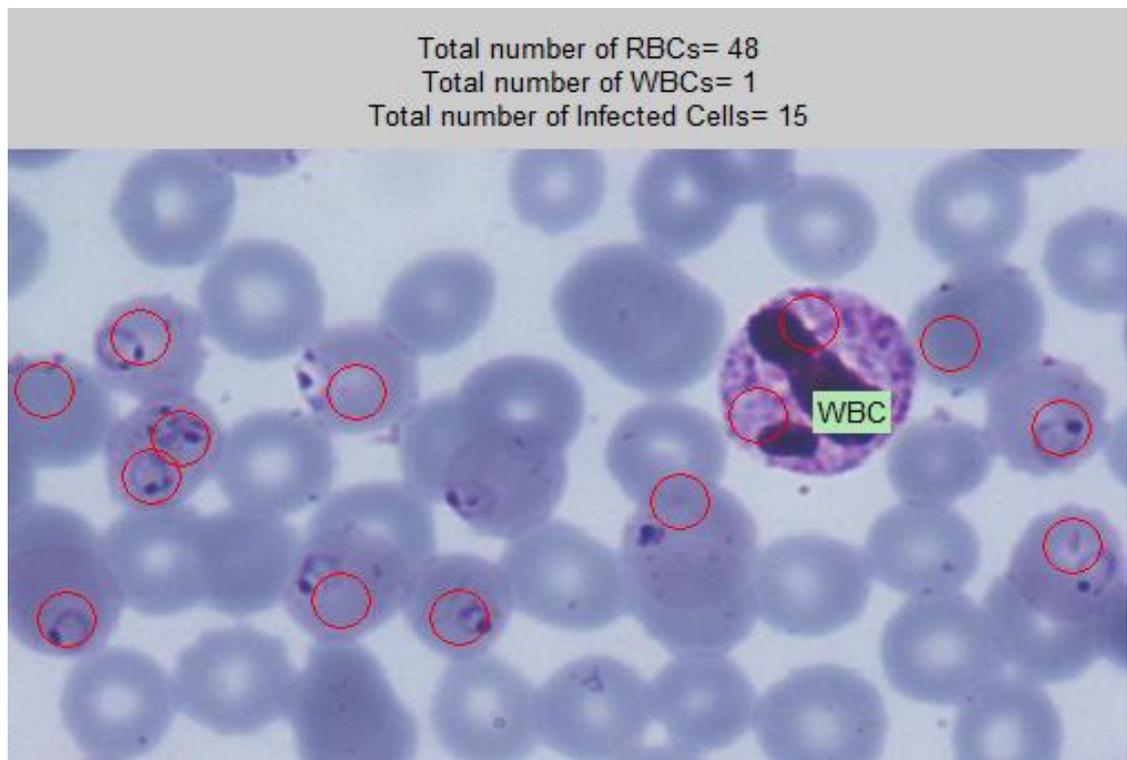


Figure 6.12: Processed image in which the cells that have been identified as infected are circled and WBC differentiated.

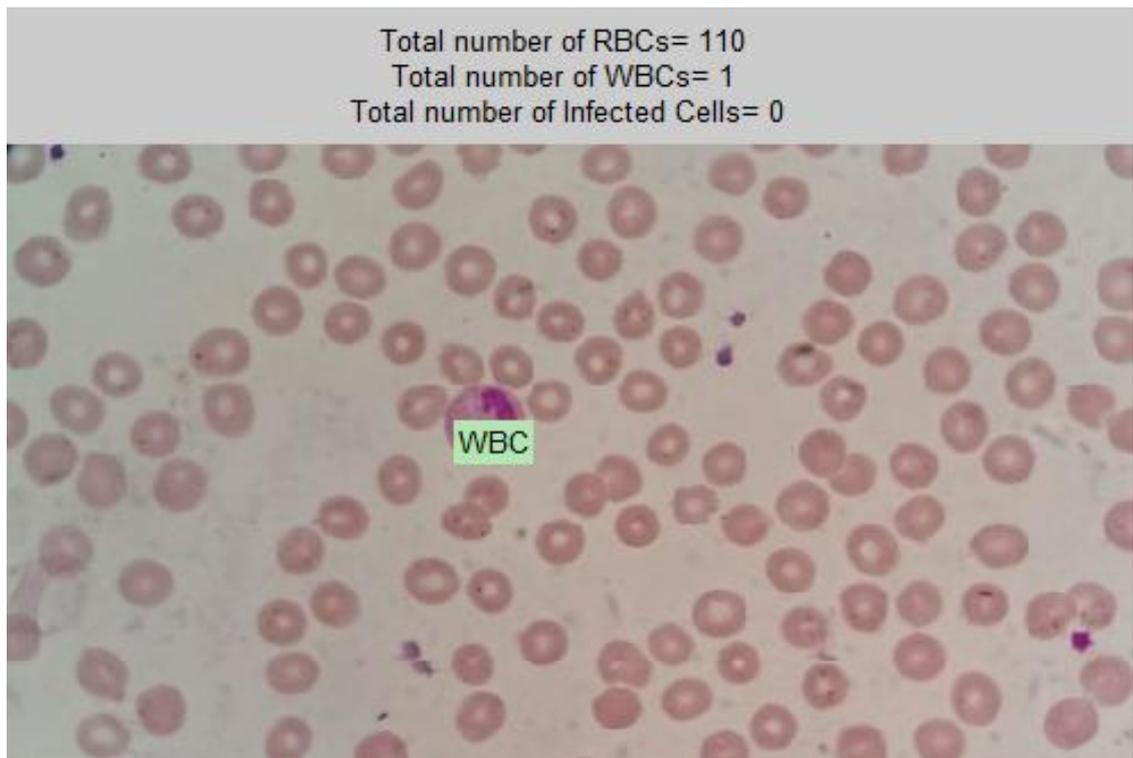


Figure 6.13: Non-infected image with WBC differentiated.

The darker dots on the RBCs are confirmed as artefacts.

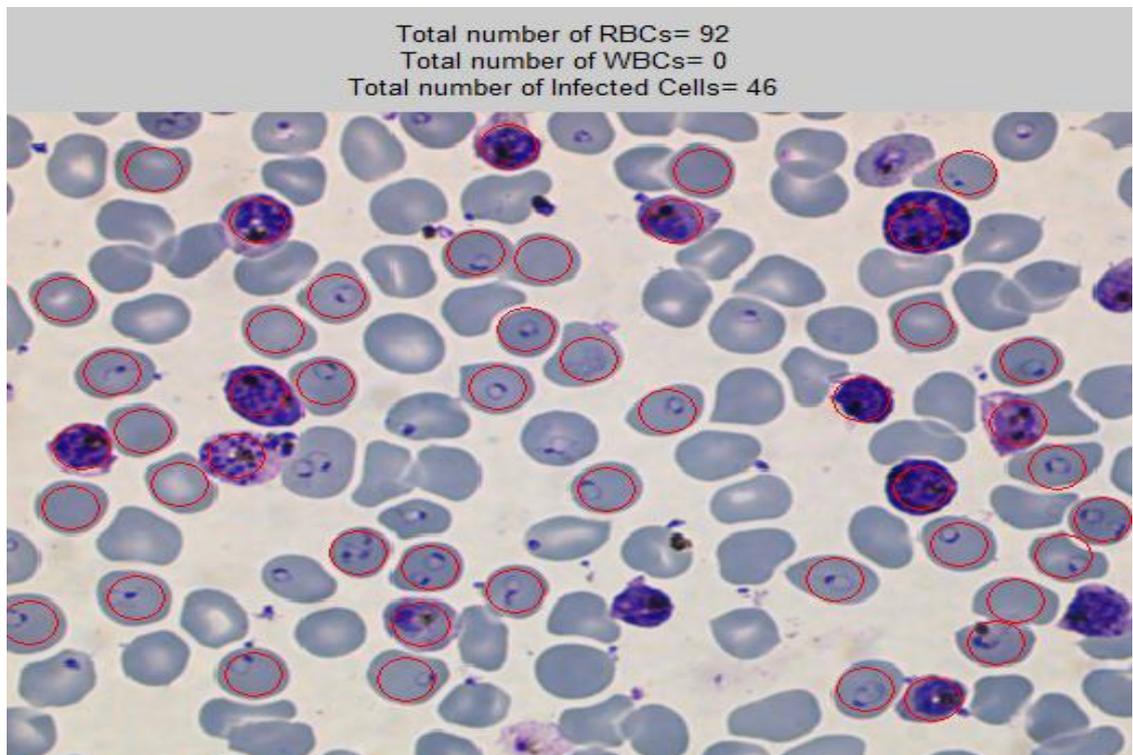


Figure 6.14: Processed image in which the cells that have been identified as infected are circled.

The image has enlarged infected RBCs that can sometimes be misdiagnosed as WBCs.

Automated Malaria Diagnosis Using Mobile Phones

Figures 6.10, 6.12, 6.13, 6.15 and 6.16 contain WBCs of different sizes and spatial orientation and are identified using the WBC differentiating algorithm described in Chapter 5.

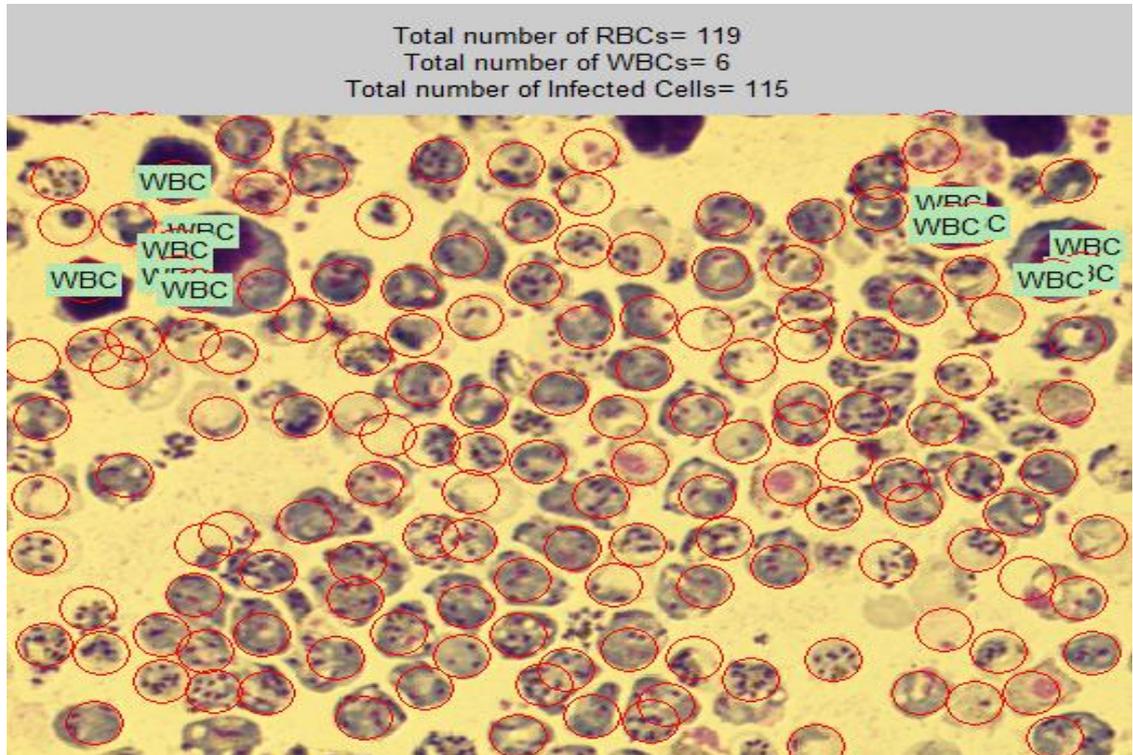


Figure 6.15: Processed image in which the cells that have been identified as infected are circled and WBCs differentiated.

The image is of a heavily infected blood with around 97% of the cells are infected.

Automated Malaria Diagnosis Using Mobile Phones

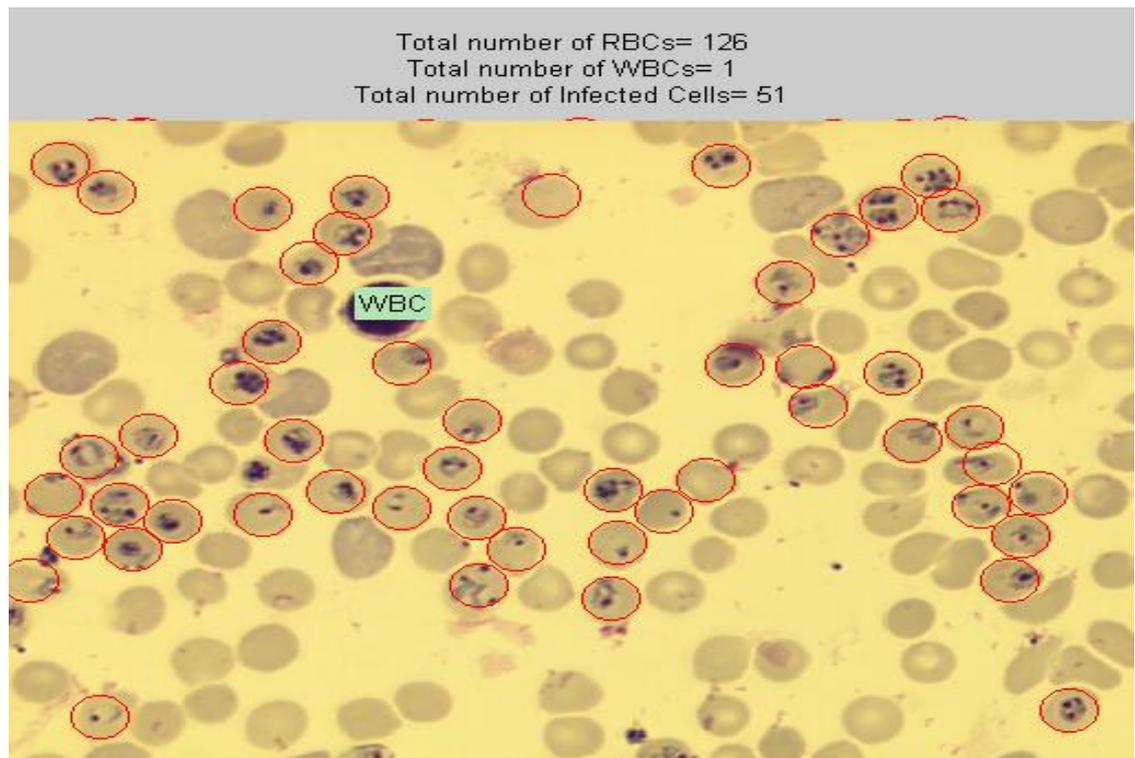


Figure 6.16: Processed image in which the cells that have been identified as infected are circled and WBCs differentiated.

The image has the monolayer of the smear with unevenly distributed cells.

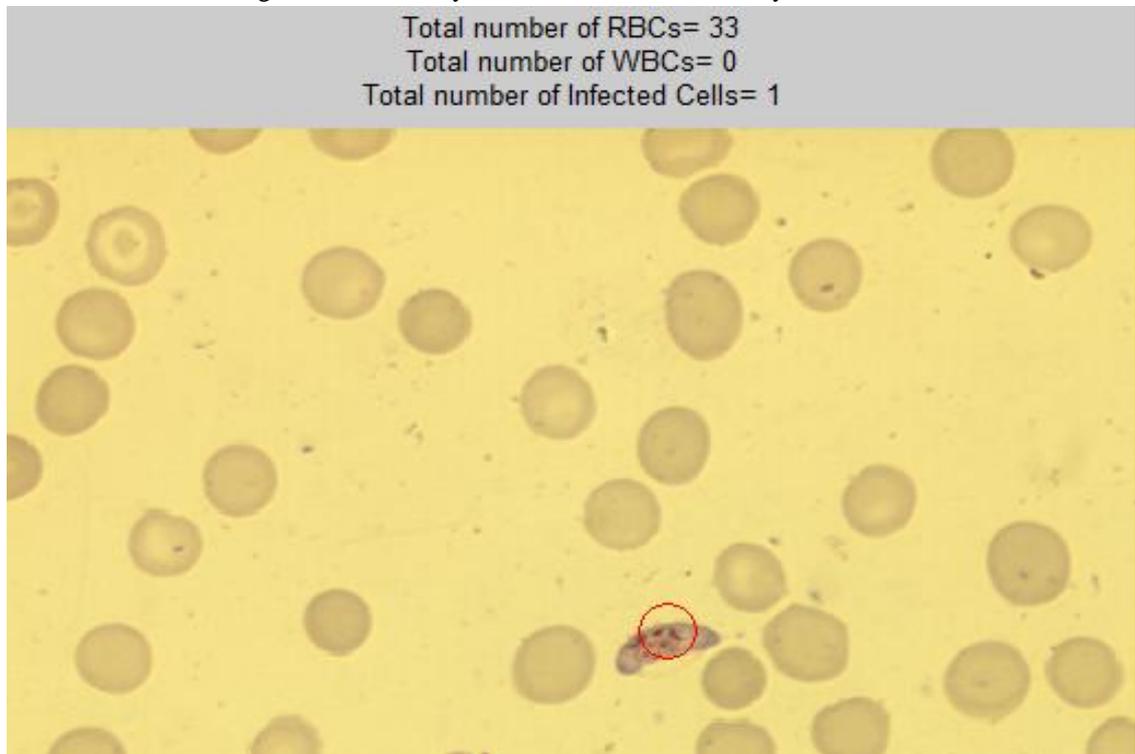


Figure 6.17: Processed image in which the cell that has been identified as infected is circled.

The image is a mostly deserted monolayer with very few cells and a gametocyte of *P.falciparum*

Automated Malaria Diagnosis Using Mobile Phones

The monolayer of Figure 6.17 has very few cells. However, since it contains a parasite, the processing of the image is crucial. Due to the uneven and noisy background, the ARR transform method is prone to falsely identify the deserted areas as RBCs.

Figures 6.15, 6.16, 6.18, and 6.19 are images of very poorly processed slides. In Figure 6.18, the cells are heavily overlapped with indistinguishable boundaries. There is an artefact similar to a WBC and a distinct parasite, which is a gametocyte of *P.falciparum*. The parasite identification, both manual and automated, is comparatively easy for this image because of the discrete parasite. However, in general, it is very difficult to obtain an accurate diagnosis for images with a large number of overlapped or degraded cells and is not an ideal choice for parasitemia (ratio of infected RBCs to total number of RBCs) estimation.

Figure 6.15 is a heavily infected blood smear with around 97% of RBCs infected by the parasite. The image contains parasites of varying level of maturity. Figure 6.16 is also a poorly processed image with unevenly distributed cells and background noise. Some of the artefacts present in the images in Figures 6.15, 6.16 and 6.19 are due to the nuclei being released to the plasma due to the rupturing of the RBCs. For such images, the algorithm mostly detects only those nuclei that are present within the RBC region. In Figure 6.19 most of the RBCs are destroyed mainly due to bad slide preparation and/or prolonged storage of blood before smear preparation.

Automated Malaria Diagnosis Using Mobile Phones

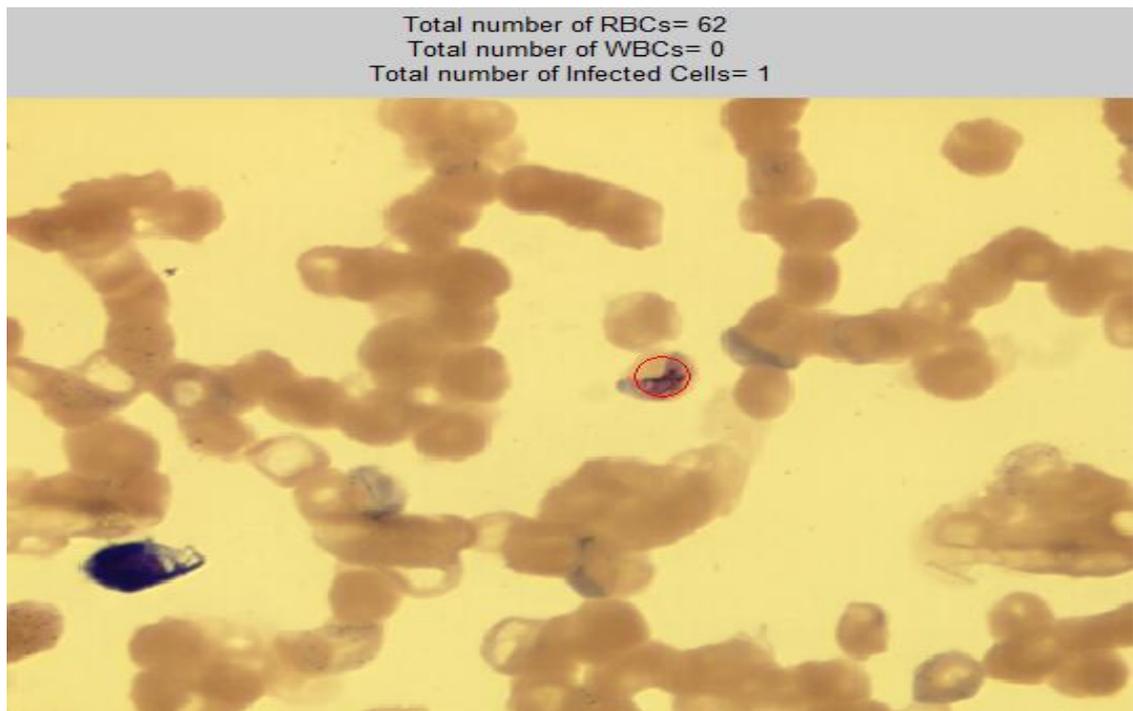


Figure 6.18: Processed image in which the cell that has been identified as infected is circled. The image has highly overlapped cells with indistinguishable boundaries and hence the identification of RBCs is very difficult.

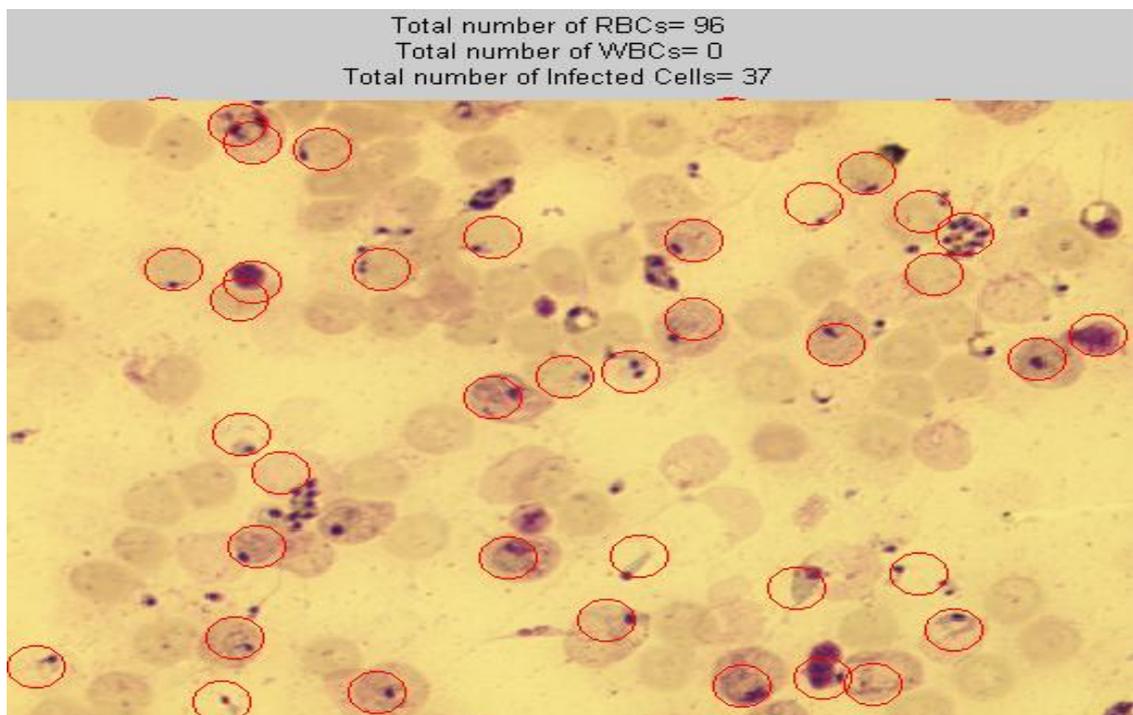


Figure 6.19: Processed image in which the cells that have been identified as infected are circled.

The blood slide is in a very poor condition and most of the RBCs have been destroyed. The background is full of noise due to artefacts, pigmentation and nuclei of parasite from ruptured infected cells. This makes the diagnosis very difficult.

Automated Malaria Diagnosis Using Mobile Phones

Along with the diagnosis, the algorithm estimates the total number of RBCs, WBCs and infected cells as shown in the following examples. A detailed report on the performance measurements of the algorithm is provided in Section 6.3.

6.3 Performance Analysis

Testing and validating the performance of a diagnostic tool is extremely important in global health care. For an automated system, parametric, semi-parametric and non parametric approaches are used to analyse the performance. For malaria diagnosis, the accuracy (validity) of the system is dependent upon reliable identification of the RBCs and classification of them into parasitized and non-parasitized cells in order to determine whether a blood sample is infected and the degree of parasitemia. As explained in Section 2.3, the accuracy of any diagnostic test is determined by calculating the True Positive ($TP = \Pr\{\text{Positive result and a positive sample}\}$), True Negative ($TN = \Pr\{\text{Negative Result and a negative sample}\}$), False Positive ($FP = \Pr\{\text{Positive result and a negative sample}\}$) and False Negative ($FN = \Pr\{\text{Negative result and a positive sample}\}$) [70]. The following Figure 6.20 demonstrates these parameters.

The overall performance of the diagnostic tool is determined by measuring the Sensitivity (The probability of the diagnosis to be positive for a positive sample $= \frac{TP}{TP+FN}$), Specificity (The probability of the diagnosis to be negative for a negative sample $= \frac{TN}{TN+FP}$), Positive Predictive Value (PPV- The probability of the person to be infected for a positive diagnosis $= \frac{TP}{TP+FP}$) and Negative Predictive Value (NPV- Probability of the person to be non infected for a negative diagnosis $= \frac{TN}{TN+FN}$) along with the accuracy ($= \frac{TP+TN}{TP+FP+TN+FN}$) [71]. Both PPV and NPV depend on the prevalence of the disease and the population of the positive and negative samples, which means, they are sensitive to the distribution of samples. All the above factors leads to the overall performance of the system and indicates the accuracy which represents the average error occurred in a given set of samples containing unequally distributed infected and non- infected cells.

There are different interpretations of the performance measurements as mentioned in Chapter 2, but all relies on the diagnosis which concludes whether a given blood sample is infected or not. In [72], Oliviera et.al. proposed an abstract for automated malaria

Automated Malaria Diagnosis Using Mobile Phones

diagnosis combined with artificial intelligence which achieved 74% sensitivity. In his thesis Tek has described a non-parametric performance measurement using a KNN classifier [42]. Tek worked on a set of images with a total of 669 parasite classes and 3431 non-parasite classes. In this research, the performances are measured based on testing whether an RBC is infected or not, or in other words, by detecting the presence of one or more parasitic nucleus in the RBC by analysing the images of different fields of thin blood smears.

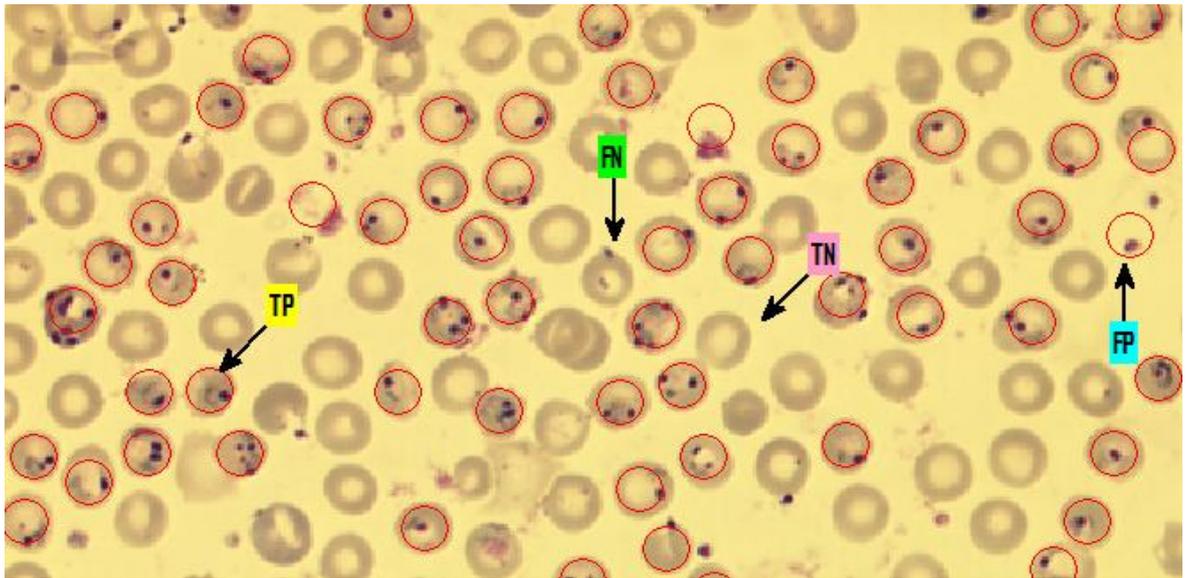


Figure 6.20: Parameters used for performance analysis.

TP- True Positive: Positively identifying an infected cell. TN- True Negative: Confirms a non-infected cell as negative. FP- False Positive: Confirming a non-infected cell as infected. FN-False Negative: Negative result for an infected cell.

In this study, parametric analyses of images from various sources were conducted. The image data set consisted of images supplied by the National Institute of Medical Research (NIMR), UK, Hospital for Tropical Diseases (HTD), London, Pondicherry Institute of Medical Science (PIMS), India, Centre for Medical Electronics, Anna University, Chennai, India and image dataset of Applied DSP and VLSI Research Group (ADVRG) Laboratory at the University of Westminster. A set of images were also acquired from the thin blood slides which belongs to the National External Quality Assessment Service (NEQAS) for Microbiology using a mobile phone camera placed on the eyepiece of a microscope. These slides were provided by the specialists of London

Automated Malaria Diagnosis Using Mobile Phones

School of Hygiene and Tropical Medicine (LSHTM) and HTD. The images were acquired using Canon A360, HTC-1, Samsung Galaxy S2, S3 and Apple Iphone-4.

The application developed in this research for malaria diagnosis progressed mainly through three stages: RBC and WBC estimation and differentiation, infected cell identification and life stage recognition. Hence instead of evaluating the performance of one diagnostic system, three separate analyses were conducted and the sensitivity, specificity, PPV, NPV and accuracy of RBC detection, WBC detection and parasitic cell identification were measured.

Table 6.1: Performance measurement results

Type of Component	Sensitivity	Specificity	PPV	NPV	Accuracy
	$\frac{TP}{TP + FN}$	$\frac{TN}{TN + FP}$	$\frac{TP}{TP + FP}$	$\frac{TN}{TN + FN}$	$\frac{TP + TN}{TP + FP + TN + FN}$
Parasites	93.4%	96.4%	89.3%	97.9%	95.7%
RBC	98%	61%	95%	78%	94%
WBC	85%	99%	90%	99%	99%

The images were manually analysed on a cell-by-cell basis at first and the results of this analysis were used as a measure of the true diagnosis. Given that the images were acquired using a variety of equipment with different resolutions, the system had to be calibrated according to the RBC radius for each image depending on its size and resolution. However, this was not a problem for images acquired from the same acquisition set-up once initial calibration was completed. A granulometry algorithm was developed to get the information of the RBC radius. The pattern spectrum of the algorithm will provide the necessary information on the size of the components present in the image.

A total of 268 images containing 16,673 objects in all were processed. The images were used in their original condition and no colour normalisation or image enhancement was performed prior to processing. There were 16 images containing P.ovale, 2 images of

Automated Malaria Diagnosis Using Mobile Phones

P. malariae and 38 images of *P. vivax*. One image contained a mixed infection of *P. vivax* and *P. falciparum*. The rest of the images were of species *P. falciparum*. A total of 15,342 RBCs, 52 WBCs, 1279 artefacts/platelets and 4036 parasitized cells were manually identified in the images.

The application successfully diagnosed 3766 parasitized cells and 12,195 cells were identified as non-infected. The non-infected class included RBCs, WBCs and platelets. The method failed to identify 270 infected cells and a total of 442 non-infected cells were falsely identified as positive. A total of 15021 RBCs were successfully estimated and 321 RBCs were missed. A total of 44 WBCs were positively identified while 8 WBCs were missed. For malaria diagnosis, the method possesses 93.4% sensitivity, 96.4% specificity and an overall accuracy of 96%. The analytical results obtained for infected cells, RBCs and WBCs are provided in Table 6.1.

The test results of the images obtained from the image dataset of ADVRG were verified by from the look-up table already available. Similarly the results obtained by processing the images provided by Anna University and PIMS were already presented and verified. The experimental results on images provided by LSHTM and HTD have not yet been verified. However, the author has received specialised training from the experts of LSHTM and HTD, and has previously worked on these slides to perform diagnosis and parasitemia estimation during the course ‘Updates in Blood and Tissue Parasitology’ conducted by University of Westminster in 2010 and the performance was verified by the experts.

6.3.1 Parasitemia estimation

Parasitemia is the extent of infection and can be calculated with respect to WBCs or RBCs. In thin blood film examination the parasitemia is estimated after positively identifying the parasite. Parasitemia is an important aspect during the course of the treatment of the disease. Consecutive parasitemia estimation is often performed after the initiation of the therapy [15]. Once the parasites are detected, few more fields will be evaluated until 1000 RBCs are estimated. For example, for an infected blood slide containing 200 RBCs per field, an estimation of infected cells in 5 fields will be enough to measure the parasitemia. In this case the parasitemia is estimated as

Automated Malaria Diagnosis Using Mobile Phones

$$\text{Parasitemia} \cong \frac{\text{Number of Infected RBCs found}}{1000} \quad (6.3)$$

Depending on the number of RBCs per field, the total number of fields to be evaluated varies. In this research, the Hue-ARR algorithm estimates the total number of infected cells and RBCs as demonstrated in Section 6.3. Hence for a single blood smear with infected blood, depending on the number of RBCs per image, the parasitemia can be easily calculated using the method. However for poorly processed and/or illuminated images such as in Figure 6.15 and 6.17, the RBC estimation will be complicated leading to the miscalculation of parasitemia.

6.3.2 Diagnostic Probabilities

Most of the existing literature in the field describes the performance measurement experimented on a set of objects in an image data set rather than per-smear analysis. However, it is necessary to distinguish between the probabilities associated with the identification of a single cell as being parasitized, the probabilities associated with the result of analysing a number of cells from a single patient and the probabilities associated testing a number of patients.

6.3.2.1 Single Measurement Statistics

Because there is the possibility of a WBC or an artefact being misidentified as an infected RBC and vice versa, it is necessary to consider all the cells found, not just the RBCs. For the analysis of a single cell, let R be the event that the cell is a RBC, let I be the event that the cell is parasitized and let D be the event that it is diagnosed as being parasitized (whether correctly or not). Since only RBCs, can be parasitized, I implies R . However, this is not true of D . Then the probability of the cell being a RBC is p_R and the probability of the cell being parasitized is p_I . Note that this is the probability taken over all the objects found including WBCs and artefacts as well as the RBCs. Consequently, p_I is not the same as the parasitemia, $p_{I/R}$, which is defined as the proportion of RBCs that are

Automated Malaria Diagnosis Using Mobile Phones

parasitized. Since $p_{\bar{R}} = 0$ (a non-RBC cannot be parasitized), the relationship between the parasitemia and p_I is simply,

$$P_I = P_{IR} = P_{I|R}P_R \quad (6.4)$$

The accuracy of an individual measurement is characterised by two parameters: $p_{D|I}$, the probability of a positive diagnosis given that the cell is parasitized, i.e. the sensitivity of a single measurement, and $p_{\bar{D}|\bar{I}}$, the probability of a negative diagnosis given that the cell is not parasitized, i.e. the specificity of a single measurement. Note that these are assumed to be a property of the measurement technique but are also a property of the quality of the image and the distribution of the number of parasites in each cell. The latter in turn is dependent on the life stage and species of parasite.

The joint probabilities of the four outcomes for the measurement of a single cell expressed in matrix form are:

$$\begin{aligned} \mathbf{P}_{DI} &= \begin{pmatrix} p_{\bar{D}|\bar{I}} & p_{\bar{D}|I} \\ p_{D|\bar{I}} & p_{D|I} \end{pmatrix} = \begin{pmatrix} p_{\bar{D}|\bar{I}} & p_{\bar{D}|I} \\ p_{D|\bar{I}} & p_{D|I} \end{pmatrix} \cdot \begin{pmatrix} p_{\bar{I}} & 0 \\ 0 & p_I \end{pmatrix} = \begin{pmatrix} p_{\bar{D}|\bar{I}} & 1 - p_{D|I} \\ 1 - p_{\bar{D}|\bar{I}} & p_{D|I} \end{pmatrix} \cdot \begin{pmatrix} 1 - p_I & 0 \\ 0 & p_I \end{pmatrix} \\ &= \begin{pmatrix} p_{\bar{D}|\bar{I}}(1 - p_I) & (1 - p_{D|I})p_I \\ (1 - p_{\bar{D}|\bar{I}})(1 - p_I) & p_{D|I}p_I \end{pmatrix} \end{aligned} \quad (6.5)$$

The diagnostic probabilities are, therefore:

$$\begin{aligned} \mathbf{P}_D &= \begin{pmatrix} p_{\bar{D}} \\ p_D \end{pmatrix} = \begin{pmatrix} p_{\bar{D}|\bar{I}} & p_{\bar{D}|I} \\ p_{D|\bar{I}} & p_{D|I} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} p_{\bar{D}|\bar{I}} & p_{\bar{D}|I} \\ p_{D|\bar{I}} & p_{D|I} \end{pmatrix} \cdot \begin{pmatrix} p_{\bar{I}} \\ p_I \end{pmatrix} = \begin{pmatrix} p_{\bar{D}|\bar{I}} & 1 - p_{D|I} \\ 1 - p_{\bar{D}|\bar{I}} & p_{D|I} \end{pmatrix} \cdot \begin{pmatrix} 1 - p_I \\ p_I \end{pmatrix} \\ &= \begin{pmatrix} p_{\bar{D}|\bar{I}}(1 - p_I) + (1 - p_{D|I})p_I \\ (1 - p_{\bar{D}|\bar{I}})(1 - p_I) + p_{D|I}p_I \end{pmatrix} \end{aligned} \quad (6.6)$$

Using Baye's law [113], the Positive Predicted Value, PPV, for a single measurement is given by:

Automated Malaria Diagnosis Using Mobile Phones

$$p_{I|D} = \frac{p_{DI}}{p_D} = \frac{p_{D|\bar{I}}p_{\bar{I}}}{p_{D|\bar{I}}p_{\bar{I}} + p_{D|I}p_I} = \frac{p_{D|I}p_I}{(1 - p_{\bar{D}|\bar{I}})(1 - p_I) + p_{D|I}p_I} \quad (6.7)$$

Similarly, the Negative Predicted Value, NPV, for a single measurement is given by:

$$p_{\bar{I}|\bar{D}} = \frac{p_{\bar{D}|\bar{I}}}{p_{\bar{D}}} = \frac{p_{\bar{D}|\bar{I}}p_{\bar{I}}}{p_{\bar{D}|\bar{I}}p_{\bar{I}} + p_{\bar{D}|I}p_I} = \frac{p_{\bar{D}|\bar{I}}(1 - p_I)}{p_{\bar{D}|\bar{I}}(1 - p_I) + (1 - p_{D|I})p_I} \quad (6.8)$$

The accuracy of the measurement, p_A , is defined as probability of a true measurement, which is given by:

$$p_A = p_{\bar{D}|\bar{I}} + p_{D|I} = p_{\bar{D}|\bar{I}}(1 - p_I) + p_{D|I}p_I \quad (6.9)$$

6.3.2.2 Experimental Results

Analysing 16673 objects and comparing the automatic recognition system against the analysis of a trained microscopist provided the following experimental results in Table 6.2:

Table 6.2: Experimental results to extract the statistical parameters.

True Negative	12195	$p_{\bar{D} \bar{I}} = 0.731$
False Negative	270	$p_{\bar{D} I} = 0.016$
False Positive	442	$p_{D \bar{I}} = 0.027$
True Positive	3766	$p_{D I} = 0.226$
Total (inc. WBCs)	16673	1
RBCs	15342	$p_R = 0.920$

From the table following statistical parameters for a single measurement can be extracted:

- Probability of infection, $p_I = p_{\bar{D}|\bar{I}} + p_{D|I} = 0.242$
- Probability of detection, $p_D = p_{\bar{D}|\bar{I}} + p_{D|I} = 0.253$
- Sensitivity, $p_{D|I} = p_{D|I} / p_I = 0.934$

Automated Malaria Diagnosis Using Mobile Phones

- Specificity, $p_{\bar{D}\bar{I}} = p_{\bar{D}\bar{I}} / (1 - p_I) = 0.964$
- Positive Predicted Value, PPV, $p_{I|D} = p_{DI} / p_D = 0.893$
- Negative Predicted Value, NPV, $p_{\bar{I}|\bar{D}} = p_{\bar{D}\bar{I}} / (1 - p_D) = 0.979$
- Accuracy, $p_A = p_{\bar{D}\bar{I}} + p_{DI} = 0.957$
- Parasitemia, $p_{I|R} = p_I / p_R = 0.263$

Note that these have been obtained from many different slides, many of which were of poor quality so this is an underestimate of the accuracy of the analysis process. Also the parasitemia is the average parasitemia over all the slides as they had different levels of infestation.

The sensitivity and specificity are properties of the measurement process so the PPV and NPV can be estimated over a range of parasitemia using Equations (6.3), (6.6) and (6.7) assuming that the proportion of RBCs remains constant. The following Figure 6.21 plots these values, as a percentage, for $0 < p_{I|R} < 1$.

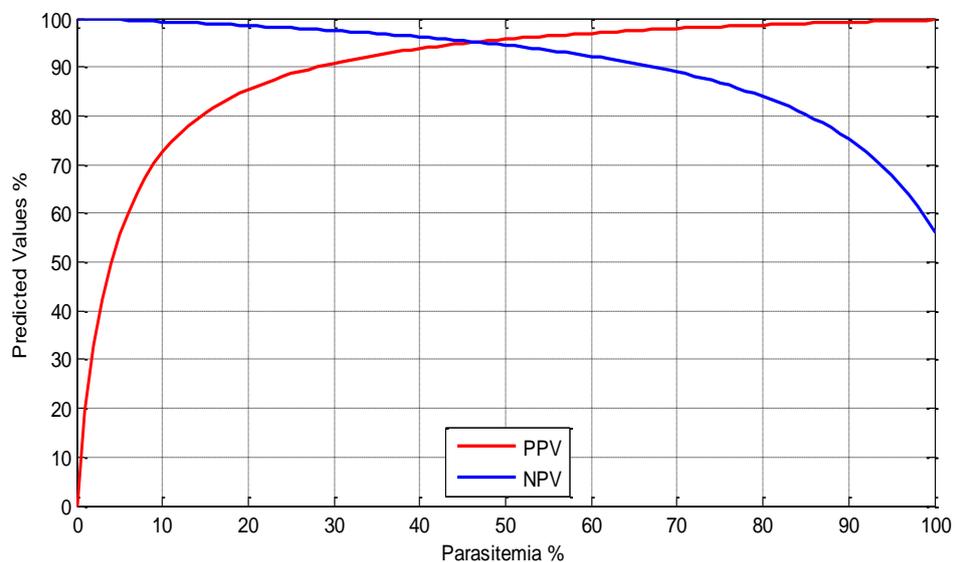


Figure 6.21:Plot to estimate PPV and NPV

Not surprisingly, this shows that, the lower the parasitemia, the more likely a positive diagnosis is false. For these parameters, if the parasitemia is less than 4%, it is more likely

Automated Malaria Diagnosis Using Mobile Phones

that a positive identification of a parasitized cell is false. Conversely, the higher the parasitemia, the more likely a negative identification of a healthy cell is false.

In addition, the accuracy can be estimated over a range of parasitemia using Equations (6.3) and (6.8) assuming that the proportion of RBCs remains constant. The following Figure 6.22 plots this, as a percentage, for $0 < p_{IR} < 1$:

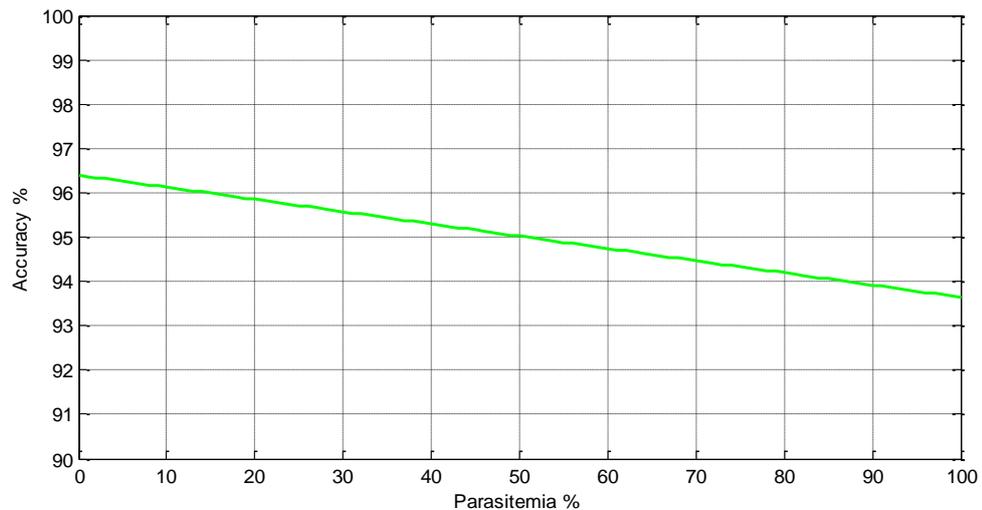


Figure 6.22: Plot to estimate the accuracy over a range of parasitemia.

It can be seen that the accuracy linearly interpolates between the specificity (for $p_{IR} = 0$) and the sensitivity (for $p_{IR} = 1$). Since these are similar probabilities, the accuracy does not change significantly over a wide range of parasitemia.

However, during image analysis where the performance is analysed among a number of cells, the four outcomes identified above can be represented by the random vector:

$$N = (N_{\bar{D}\bar{I}}, N_{D\bar{I}}, N_{\bar{D}I}, N_{DI}) \text{ where } N_{\bar{D}\bar{I}} + N_{D\bar{I}} + N_{\bar{D}I} + N_{DI} = m \text{ and}$$

$N_{\bar{D}\bar{I}}$ is the number of clean RBCs and non-RBCs diagnosed correctly as parasite-free,

$N_{D\bar{I}}$ is the number of clean RBCs and non-RBCs diagnosed wrongly as parasitized,

$N_{\bar{D}I}$ is the number of parasitized RBCs diagnosed wrongly as parasite-free,

N_{DI} is the number of parasitized RBCs diagnosed correctly as parasitized, where m be the number of cells analysed.

Automated Malaria Diagnosis Using Mobile Phones

Given that the analysis of each cell can be treated as independent, the Probability Moment Function (PMF) for this vector is:

$$p_N(n_{\bar{D}I}, n_{D\bar{I}}, n_{\bar{D}I}, n_{DI}) = \frac{m!}{n_{\bar{D}I}! n_{D\bar{I}}! n_{\bar{D}I}! n_{DI}!} p_{\bar{D}I}^{n_{\bar{D}I}} p_{D\bar{I}}^{n_{D\bar{I}}} p_{\bar{D}I}^{n_{\bar{D}I}} p_{DI}^{n_{DI}} \quad (6.10)$$

$$, \quad n_{\bar{D}I} + n_{D\bar{I}} + n_{\bar{D}I} + n_{DI} = m$$

The marginal PMF on the number of parasitized cells in the sample population is the binomial distribution:

$$p_{N_I}(n_{\bar{I}}, n_I) = \frac{m!}{n_{\bar{I}}! n_I!} p_{\bar{I}}^{n_{\bar{I}}} p_I^{n_I} \quad (6.11)$$

where $n_{\bar{I}} + n_I = m$, $n_{\bar{I}} = n_{\bar{D}I} + n_{D\bar{I}}$ and $n_I = n_{\bar{D}I} + n_{DI}$. The probability of a false negative, despite the analysis of each cell being 100% correct, is the probability of $n_I = 0$, and is given by:

$$p_{FNI} = p_{N_I}(m, 0) = p_{\bar{I}}^m = (1 - p_I)^m \quad (6.12)$$

The marginal PMF on the number of cells diagnosed as parasitized in the sample population is the binomial distribution:

$$p_{N_D}(n_{\bar{D}}, n_D) = \frac{m!}{n_{\bar{D}}! n_D!} p_{\bar{D}}^{n_{\bar{D}}} p_D^{n_D} \quad (6.13)$$

,

$$n_{\bar{D}} + n_D = m$$

where p_D and $p_{\bar{D}}$ are given in Equation (6.5).

It is important to note that in these statistics, p_I , is a given parameter. The situation of dealing with a population of patients with different levels of parasitemia (where p_I becomes a randomly distributed variable) is dealt with later.

In terms of the variables of this distribution, the conditions leading to a diagnosis for the patient represented by the analysed blood smear are:

Automated Malaria Diagnosis Using Mobile Phones

For a correctly diagnosed absence of malaria: $p_I = 0$ and $n_{D\bar{I}} = 0$

For an incorrectly diagnosed case of malaria: $p_I = 0$ and $n_{D\bar{I}} > 0$

For an incorrectly diagnosed absence of malaria: $p_I > 0$ and $n_{D\bar{I}} = n_{DI} = 0$

For a correctly diagnosed case of malaria: $p_I > 0$ and $n_{D\bar{I}} + n_{DI} > 0$

6.3.2.3 Probabilities of diagnostic performance for a negative (healthy) smear

An important special case is that of a healthy smear where $p_I = 0$ and so $n_{D\bar{I}} = n_{DI} = 0$; consequently the PMF collapses to the binomial distribution:

$$P_{N\bar{I}}(n_{D\bar{I}}, n_{DI}) = \frac{m!}{n_{D\bar{I}}! n_{DI}!} P_{D\bar{I}}^{n_{D\bar{I}}} P_{DI}^{n_{DI}} \quad (6.14)$$

$$n_{D\bar{I}} + n_{DI} = m$$

For a correctly diagnosed absence of malaria (true negative), $p_I = 0$ and $n_{DI} = 0$. The probability of this is the specificity of the overall diagnosis and is given by:

$$p_{TN} = P_{N\bar{I}}(m, 0) = P_{D\bar{I}}^m \quad (6.15)$$

,since $p_{\bar{I}} = 1$. Thus the specificity of the overall diagnosis is a function of the specificity of one measurement and falls with the number of cells analysed.

For an incorrectly diagnosed case of malaria (false positive), $p_I = 0$ and $n_{D\bar{I}} > 0$. The probability of this is:

$$p_{FP} = 1 - P_{D\bar{I}}^m \quad (6.16)$$

For the experimentally derived value of specificity, $p_{D\bar{I}} = 0.964$, the following Figure 6.23 shows the probabilities of a true negative and a false positive diagnosis as a function of the number of cells analysed, m .

Automated Malaria Diagnosis Using Mobile Phones

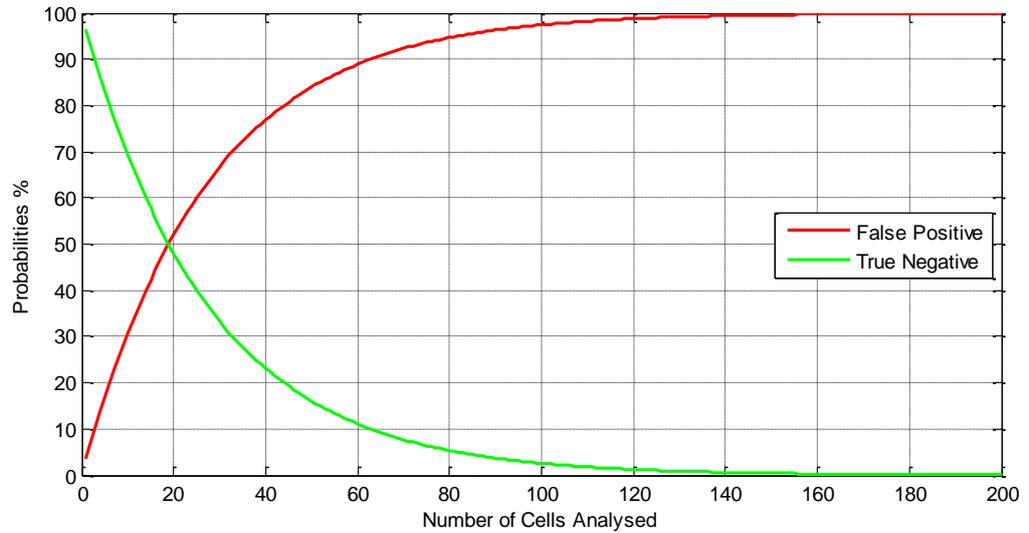


Figure 6.23: Plot of True negative and False positives against total number of cells analysed

This shows that even for a relatively small number of cells analysed, the probability of, at least one cell being falsely detected as parasitized becomes overwhelming. Consequently, in practice, a threshold for N_D has to be set, below which the patient would be assumed to be malaria-free.

Let this threshold be d_{th} , expressed as a fraction of the total number of cells examined, m . Thus, the number of positive identifications of parasitized cells required before an overall positive diagnosis is returned is:

$$n_{th} = \lfloor d_{th} m \rfloor \quad (6.17)$$

The probability of a true negative (i.e. the specificity of the overall diagnosis) is modified to:

$$p_{TN} = \sum_{j=0}^{n_{th}} p_{N|\bar{I}}(m-j, j) = \sum_{j=0}^{n_{th}} \frac{m!}{(m-j)! j!} p_{\bar{D}|\bar{I}}^{m-j} (1 - p_{\bar{D}|\bar{I}})^j \quad (6.18)$$

And

Automated Malaria Diagnosis Using Mobile Phones

$$p_{FP} = 1 - p_{TN} \tag{6.19}$$

For the experimentally derived value of specificity for a single measurement, $p_{\overline{D|I}} = 0.964$, using Equations (6.16), (6.17) and (6.18), the following Figure 6.24 shows the probabilities of a true negative and a false positive diagnosis as a function of the threshold, d_{th} , expressed as a percentage of the total number of cells examined, m , for $m = 200$ and 1000 .

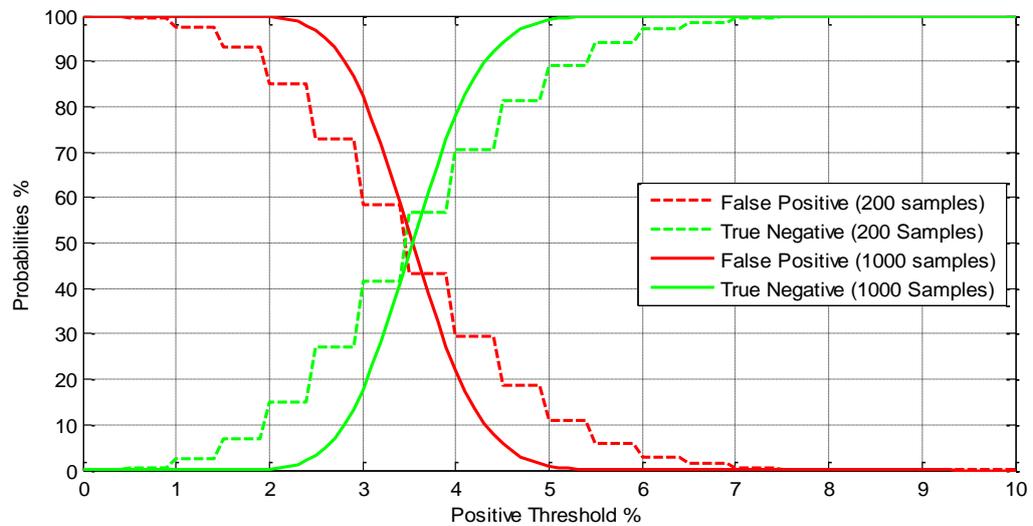


Figure 6.24: Plot of True Negative and False Positive Probabilities based on the threshold.

It can be seen that a positive threshold of 5% reduced the probability of a false positive to less than 1% when 1000 cells are analysed. Interestingly, when only 200 cells are examined, this probability at 10% is greater. The step changes for the $m = 200$ case is due to the threshold being quantized to an integer number of cells.

The interrelationship between the number of false positives and the sample size is shown in the following figure 6.25 for thresholds of 3%, 3.6% and 5%.

Automated Malaria Diagnosis Using Mobile Phones

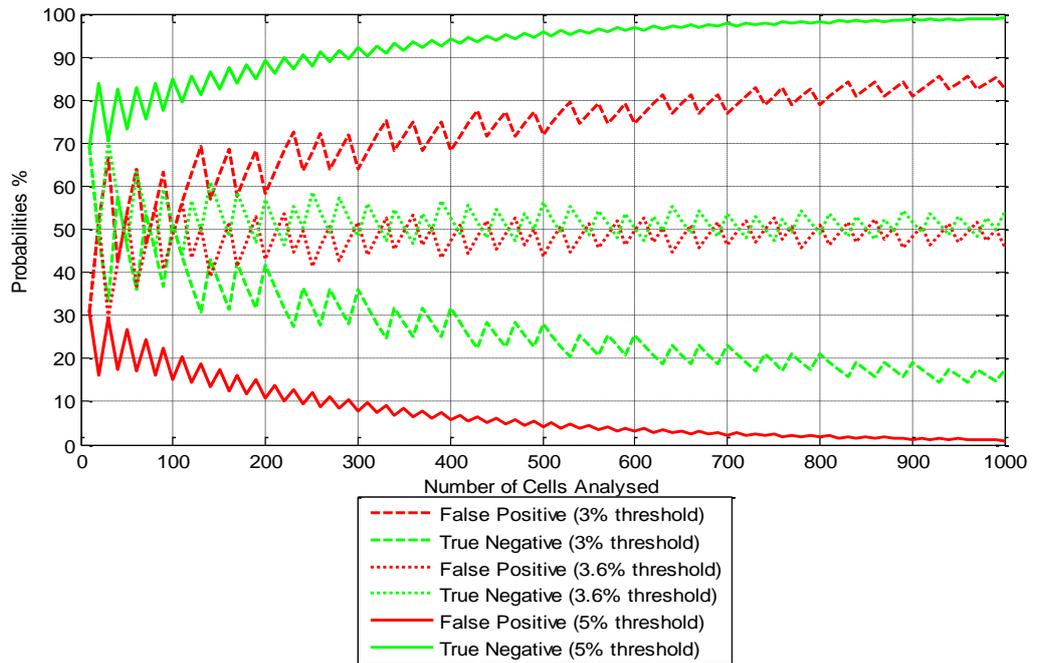


Figure 6.25: Probabilities of False positive and true negative against the number of samples.

The probabilities are plotted for 3%, 3.6% and 5% threshold

The saw tooth nature of the probabilities is due to the quantization of n_{th} . Not that the three threshold cases produce very different results which correspond to d_{th} being less-than, equal-to or greater-than $p_{D|\bar{I}}$. This shows the sensitivity of the number of false positives to the specificity of a single measurement:

$$p_{FP} \Big|_{m \rightarrow \infty} = \begin{cases} 1, & d_{th} < (1 - p_{D|\bar{I}}) \\ 0.5, & d_{th} = (1 - p_{D|\bar{I}}) \\ 0, & d_{th} > (1 - p_{D|\bar{I}}) \end{cases} \quad (6.20)$$

It is important, therefore that this threshold is set safely above the probability of a false positive of a single measurement (i.e. $1 - \text{specificity}$). However, it does provide a method of finding the specificity of a single measurement by running a series of measurements with large m over a large number of non-infected smears and then experimentally varying the threshold, d_{th} , to obtain 50% positive diagnosis. This value of d_{th} is equal to $p_{D|\bar{I}}$. This is equivalent to finding the median of the set of measured p_I . Note that this method does not require a trained microscopist to obtain a reference result.

6.3.2.4 Probabilities of diagnostic performance for a positive (Malaria infected) smear

In the case of a parasitized smear, $p_I > 0$, so the full PMF must be used.

For an incorrectly diagnosed absence of malaria (false negative), $p_I > 0$ and $n_{D\bar{I}} = n_{DI} = 0$ assuming that no threshold is set to avoid false positive diagnoses. The probability of this is given by:

$$p_{FN} = \sum_{j=0}^m p_N(m-j, 0, j, 0) = \frac{m!}{(m-j)!j!} p_{\bar{D}\bar{I}}^{m-j} p_{D\bar{I}}^j \quad (6.21)$$

More easily, this can be obtained from the marginal distribution, p_{N_D} , setting $n_D = 0$ in Equation (6.12):

$$p_{FN} = p_{N_D}(m, 0) = p_D^m = \left(p_{\bar{D}\bar{I}}(1-p_I) + (1-p_{D\bar{I}})p_I \right)^m \quad (6.22)$$

Note that this includes the possibility of the sample not containing any parasitized cells. For a correctly diagnosed case of malaria (true positive), $p_I > 0$ and $n_D = n_{D\bar{I}} + n_{DI} > 0$, again assuming that the threshold has been set to zero. The probability of this is:

$$p_{TP} = 1 - p_{FN} = 1 - p_D^m = 1 - \left(p_{\bar{D}\bar{I}}(1-p_I) + (1-p_{D\bar{I}})p_I \right)^m \quad (6.23)$$

This is the sensitivity of the overall diagnosis. Note that this is more complex to define as it is a function of both the sensitivity and specificity of one measurement, $p_{D\bar{I}}$ and $p_{\bar{D}\bar{I}}$, the number of samples, m , and the parasitemia. For the experimentally derived values of sensitivity, $p_{D\bar{I}} = 0.934$, and specificity, $p_{\bar{D}\bar{I}} = 0.964$, the following Figure 6.26 shows the probabilities of a true positive and a false negative diagnosis as a function of the number of cells analysed, m , for parasitemia of 5% and 25% using Equations (6.3), (6.21) and (6.22) assuming that the proportion of RBCs remains constant. For comparison, the probability of a false negative is shown assuming 100% sensitivity and specificity using Equation (6.11).

Automated Malaria Diagnosis Using Mobile Phones

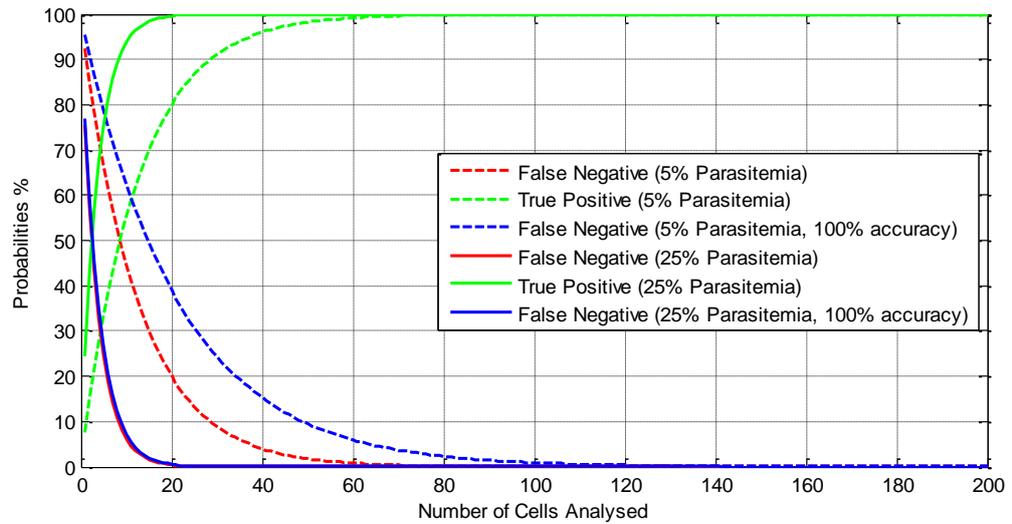


Figure 6.26: Probability plot of True Positive and False Negative against the samples at 5% and 25% parasitemia

A plot of the same probabilities against the parasitemia for $m = 20$ and 200 is shown in the following Figure 6.27:

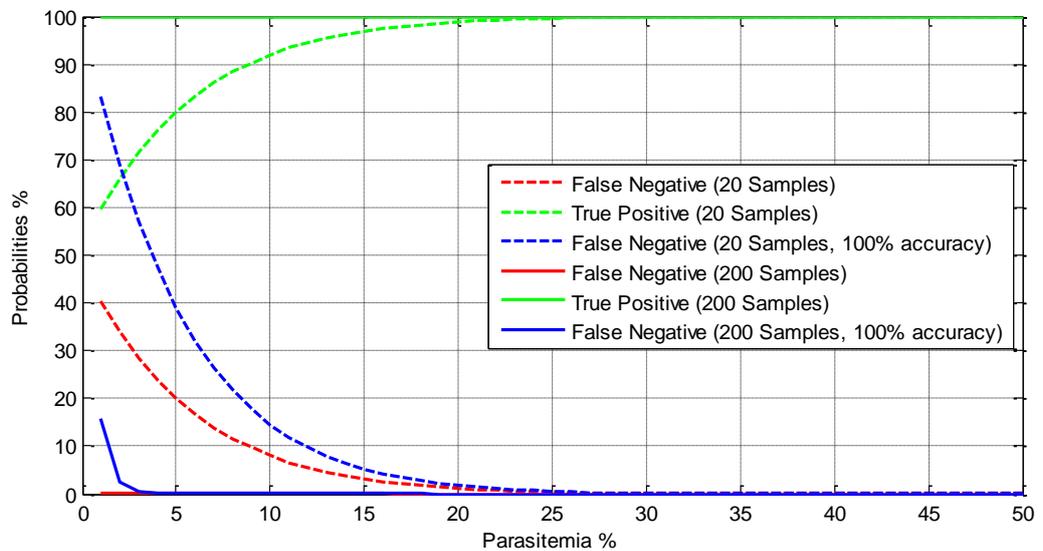


Figure 6.27: Probability plot of True Positive and False negative for 20 and 200 samples

These figures imply very good performance as far as overall sensitivity is concerned even with a small number of samples and low parasitemia. However, this is due to the threshold for detecting malaria being just 1 parasitized cell. Indeed, the non-ideal measurements appear to out-perform the 100% accurate measurements at low levels of parasitemia. This is due to the presence of false positive identifications in the non-ideal case!

Automated Malaria Diagnosis Using Mobile Phones

It has already been shown that, in order to have in any way a sensible specificity of the overall diagnosis, a threshold has to be set with $d_{th} > (1 - p_{\bar{D}\bar{I}})$. With this threshold applied, an overall positive diagnosis requires the number of cells identified as parasitized to be greater than n_{th} , where n_{th} is given by Equation (6.16). Thus, using Equation (6.12), Equation (6.21) becomes:

$$p_{FN} = \sum_{j=0}^{n_{th}} p_{N_D}(m-j, j) = \sum_{j=0}^{n_{th}} \frac{m!}{(m-j)!j!} p_{\bar{D}}^{m-j} p_D^j \quad (6.24)$$

where p_D and $p_{\bar{D}}$ are given in Equation (6.5). The probability of a true positive diagnosis is simply:

$$p_{TP} = 1 - p_{FN} \quad (6.25)$$

For the ideal 100% accuracy case, p_D becomes p_I and so Equation (6.11):

$$p_{FNI} = \sum_{j=0}^{n_{th}} p_{N_I}(m-j, j) = \sum_{j=0}^{n_{th}} \frac{m!}{(m-j)!j!} p_{\bar{I}}^{m-j} p_I^j \quad (6.26)$$

For the experimentally derived values of sensitivity, $p_{DI} = 0.934$, and specificity, $p_{\bar{D}\bar{I}} = 0.964$, the following Figure 6.28 shows the probabilities of a true positive and a false negative diagnosis as a function of the number of cells analysed, m , for parasitemia of 3% and 5% using Equations (6.3), (6.23) and (6.24) with the threshold, d_{th} , set at 5%. For comparison, the probability of a false negative is shown assuming 100% sensitivity and specificity using Equation (6.25) with the same threshold.

Automated Malaria Diagnosis Using Mobile Phones

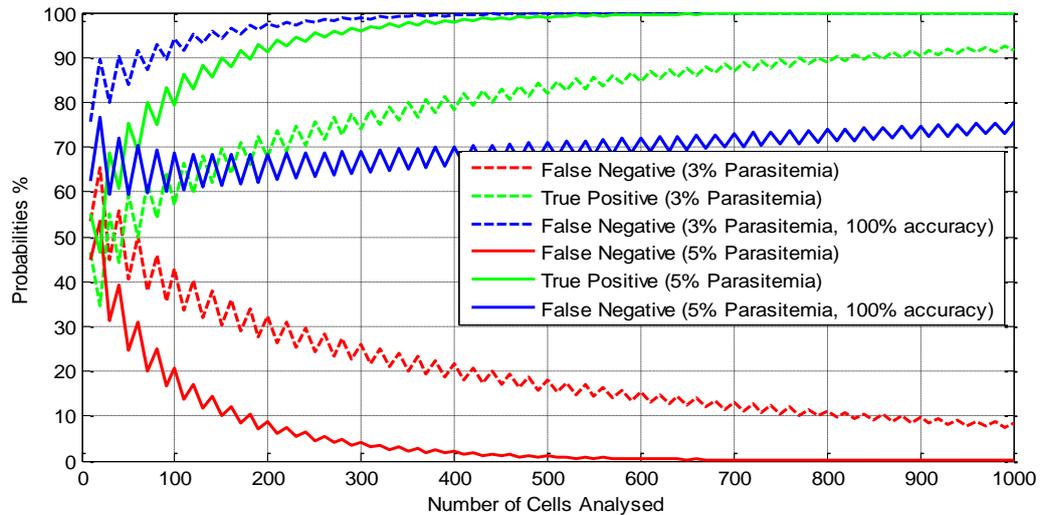


Figure 6.28: Probability plot of True Positive and False Negative against the samples with 3% and 5% parasitemia with 5% threshold.

Not that the parasitemia is low in both cases to demonstrate the overall sensitivity near the lower limit of parasitemia. The failure of the 100% accurate ‘ideal’ case is due to the threshold being set too high for this performance. (The optimum threshold in this case would be $d_{th} = 0$ which is shown in the no-threshold plot earlier.

A plot of the same probabilities with the same threshold against the parasitemia for $m = 200$ and 1000 is shown in the following Figure 6.29:

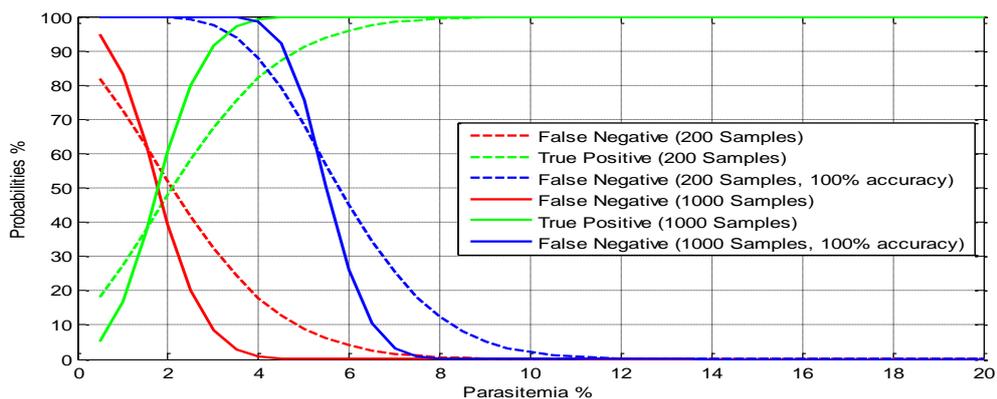


Figure 6.29: Probability plot of True Positive and False Negative against parasitemia for 200 and 1000 samples.

This reveals the danger of setting the threshold level, d_{th} , too high. This analysis shows that for the experimentally derived performance figures of the algorithm, using a

Automated Malaria Diagnosis Using Mobile Phones

threshold level, d_{th} , set at 5% and analysing 1000 cells, parasitemia down to 4% can be detected with both false negative and false positive probabilities of less than 1%.

The impact of the choice of threshold level on the probabilities of false positive versus false negative diagnoses is shown in the Figure 6.30 below with the number of cells examined, $m = 1000$ and using the experimentally derived values of sensitivity, $p_{DI} = 0.934$, and specificity, $p_{\bar{D}\bar{I}} = 0.964$ with the 100% accuracy case also plotted for comparison. The false positive probabilities are plotted for 0% parasitemia while the false negative probabilities are given for parasitemia levels of 3% and 5%.

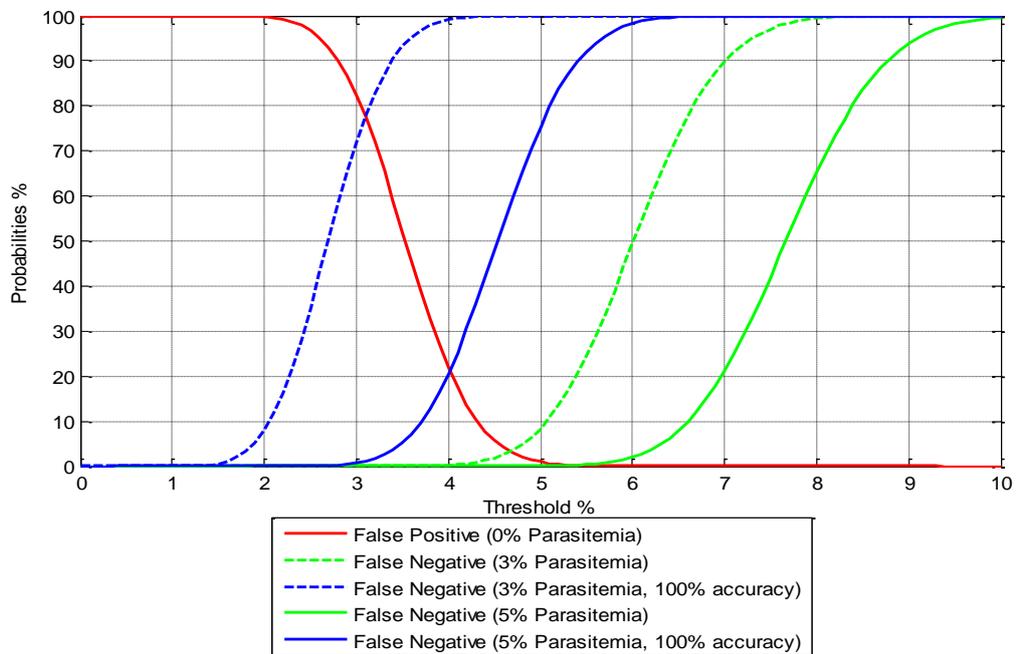


Figure 6.30: Plot to demonstrate the impact of threshold values on the probabilities.

This confirms that the algorithm is able to detect the presence of malaria with 99% confidence for parasitemia levels down to less than 5% and the use of 5% as the threshold value for the positive diagnosis. Given that the experimental values were obtained using a variety of often poorly prepared slides from different subjects using different optical acquisition systems, a single acquisition system with a consistent set of slides prepared under the similar conditions should result in better values for the measurement specificity and sensitivity and hence in a still better performance at low levels of parasitemia.

6.4 Life stage Recognition of Plasmodium Species

For the plasmodium species, life stages are determined by the number of nuclei present in the cell or in other words maturity of the parasite within the RBC. Different life stages present in the blood sample represent the degree of infection. As explained in Chapter 1, the four life stages through which the genus plasmodium transforms in a blood cell are: Ring trophozoite, Trophozoite, Schizont and Gametocyte (Male and Female) [63]. The ring trophozoite, trophozoite and schizonts are asexual stages of the parasite where as gametocytes are the sexual stage. Each of the stages exhibit distinctive morphological features [73]. Figure 6.31 shows an example of an image with all the asexual life stages of *P.falciparum* in a thin blood image. The images of other species are provided in Chapter 1. In conventional image analysis, feature extraction and classification based on neural networks are usually performed to identify the life stages [21]. In this research the recognition is carried out without artificial intelligence. Since the ultimate aim of the research is to develop a less complex architecture that can be deployed on to mobile phones, a simpler morphological approach is used which is carried out sequentially with the parasite detection.

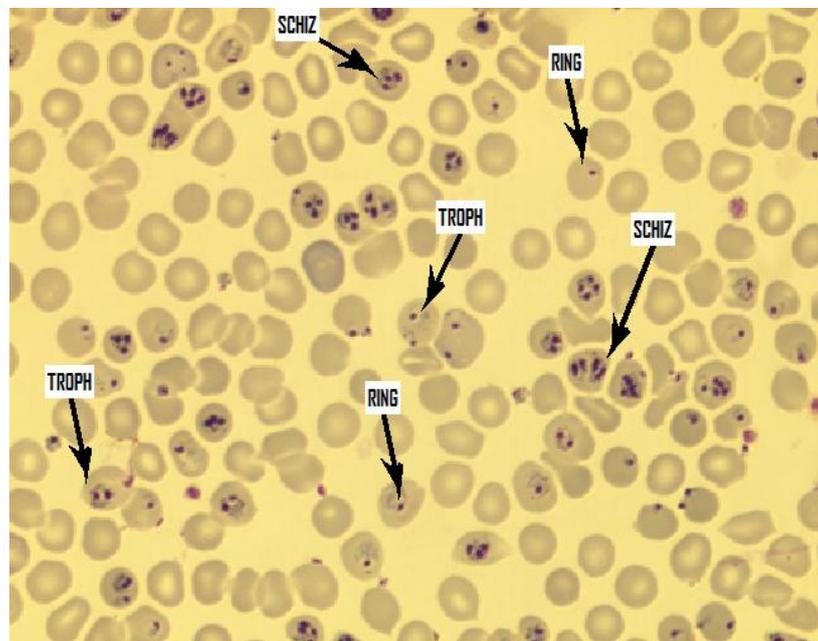


Figure 6.31: The asexual life stages in a thin blood film.

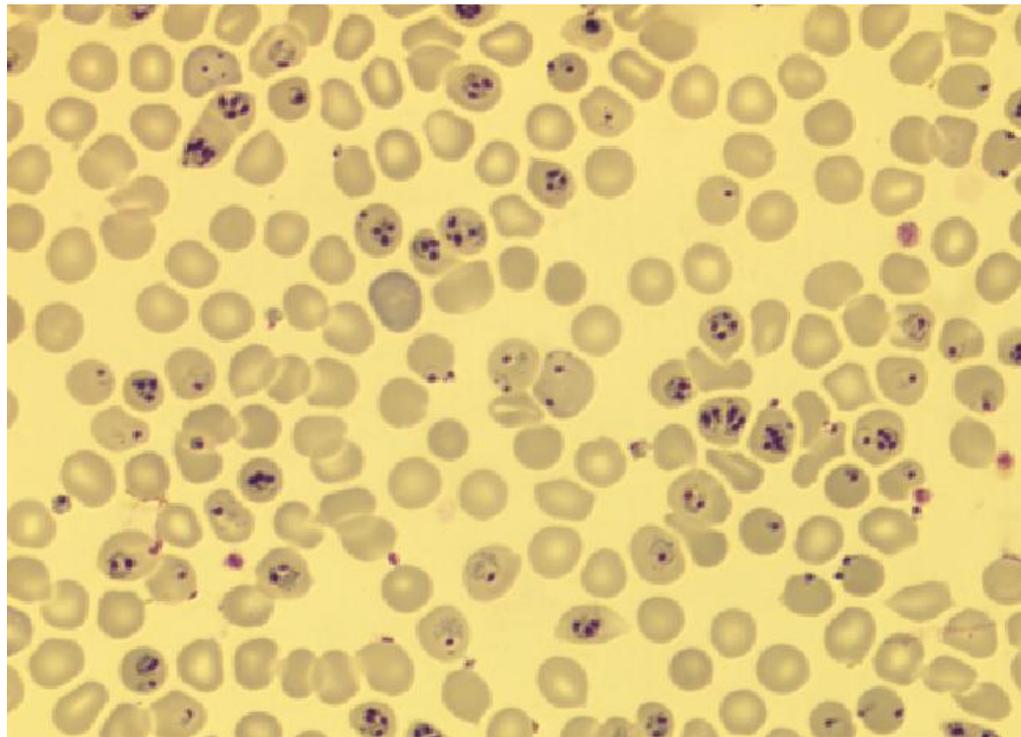
RING: Ring trophozoite with one nucleus surrounded by halo (ring) shaped cytoplasm, TROPH: Trophozoite with 2-3 nuclei and SCHIZ: Schizont with 4-8 nuclei.

6.4.1 Algorithm

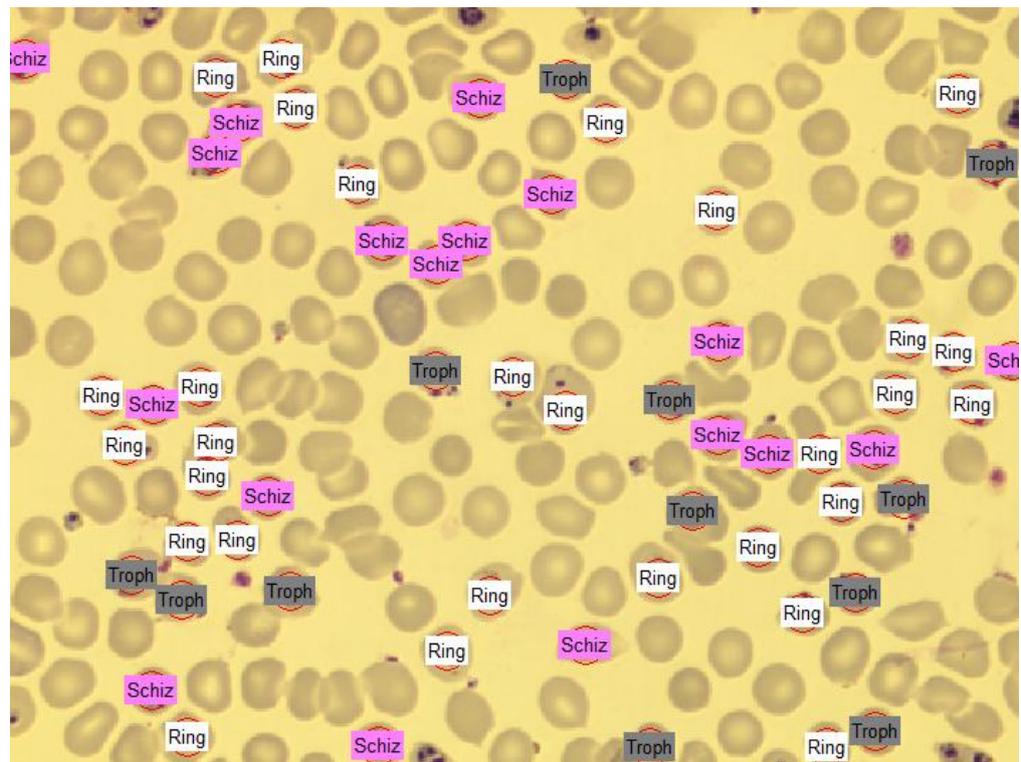
Based on the maturity of the parasite, each parasitized RBC will contain 1-9 nuclei. The Hue-ARR algorithm, described in Section 6.2, also counts the number of nuclei in each RBC using the modified ARR method (Please also refer Figure 6.6). This estimate is used to distinguish the life stages. For example, the ring trophozoites contains one nucleus in the RBC, a matured trophozoite has 2-3 nuclei and 4-8 for schizonts (In *P.vivax*, however, there would be 8-12 nuclei for a very mature schizont). The sexual life stages called gametocytes of the parasite are often present in the later stages of infection. The algorithm classifies cells having more than 8-10 nuclei as gametocyte for the species *P.vivax*, *P.ovale* and *P.malariae*. However, for the species *P.falciparum*, the gametocytes possess a distinct shape and geometry. The diagnosis of this life stage is extremely important: firstly because they are responsible for the transmission of the disease and secondly, *P.falciparum* is the deadliest of all the plasmodium species, and is responsible for 95% of the deaths caused by malaria. Hence a detailed analysis of *P.falciparum* gametocyte detection was conducted in addition to the rest of the research study and is described in Chapter 7.

6.3.3 Experimental Results

The following Figures 6.32-6.34 demonstrate the results obtained after processing the algorithm. A Graphics User Interface (GUI) was also developed in order to perform a parametric testing of the system which includes RBC identification, WBC identification, and parasite detection along with life stage recognition and parasitemia estimation. The results are explained in Section 6.4.3.



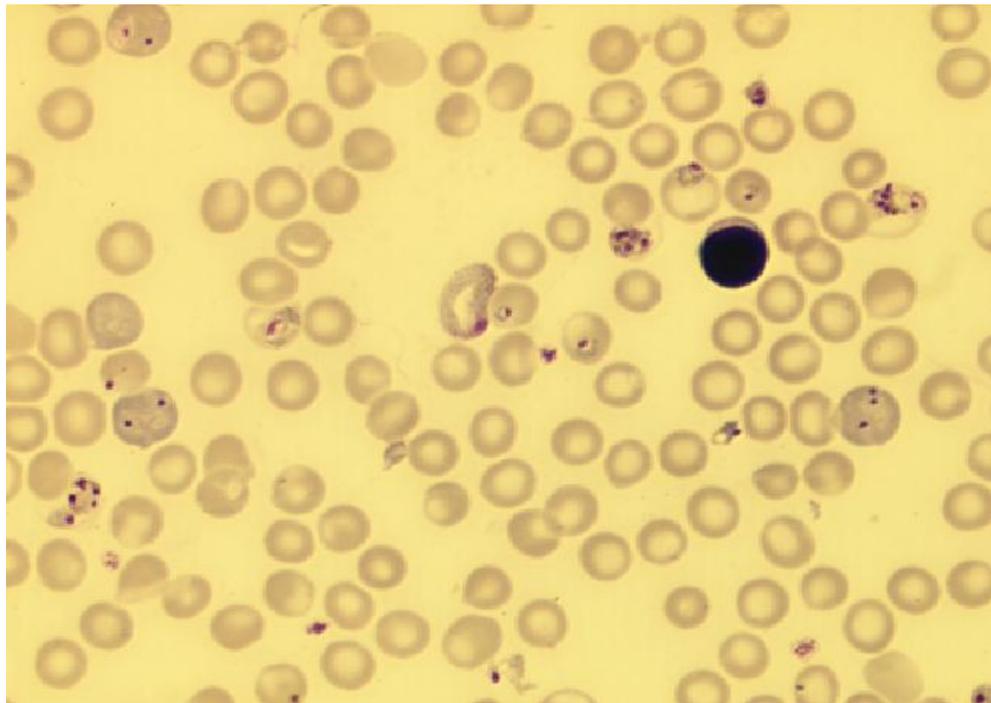
(a)



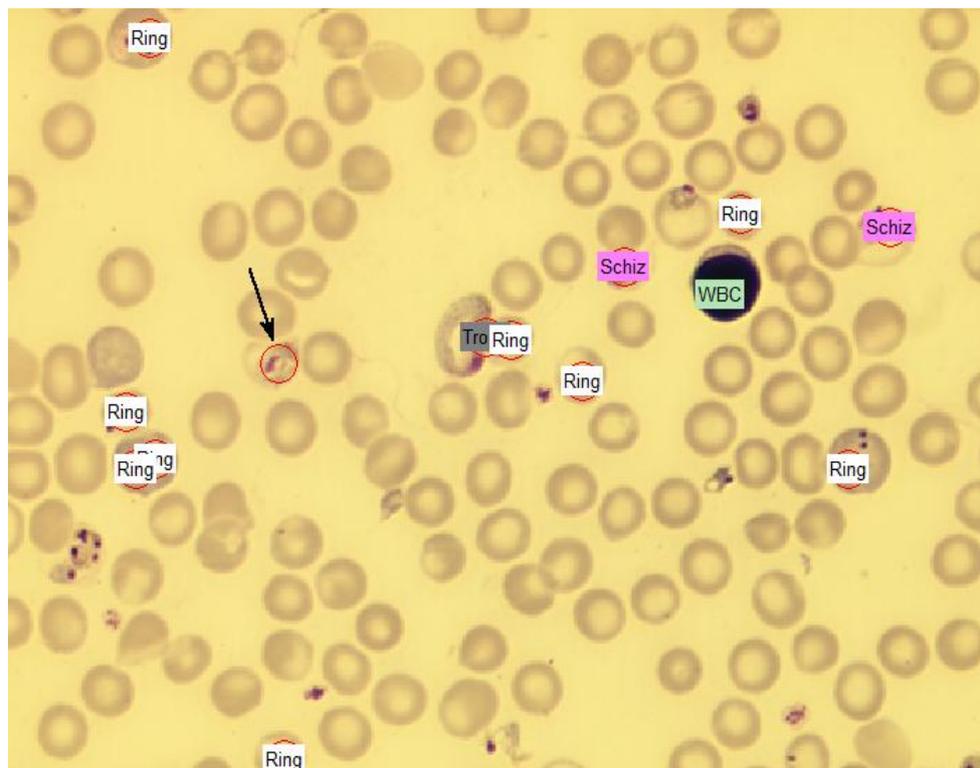
(b)

Figure 6.32: Life-recognition algorithm performed on the image in (a).

(a) Original Image (b) Resultant image with the life stages recognised. Ring- Ring Trophozoite, Troph- Trophozoite, Schiz- Schizont.



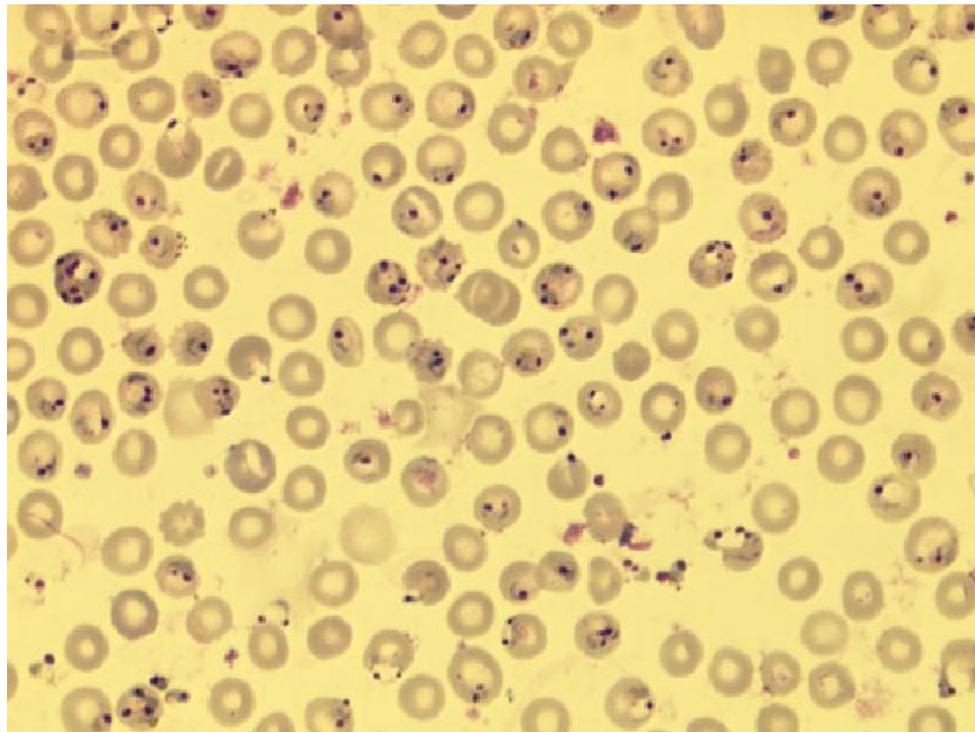
(a)



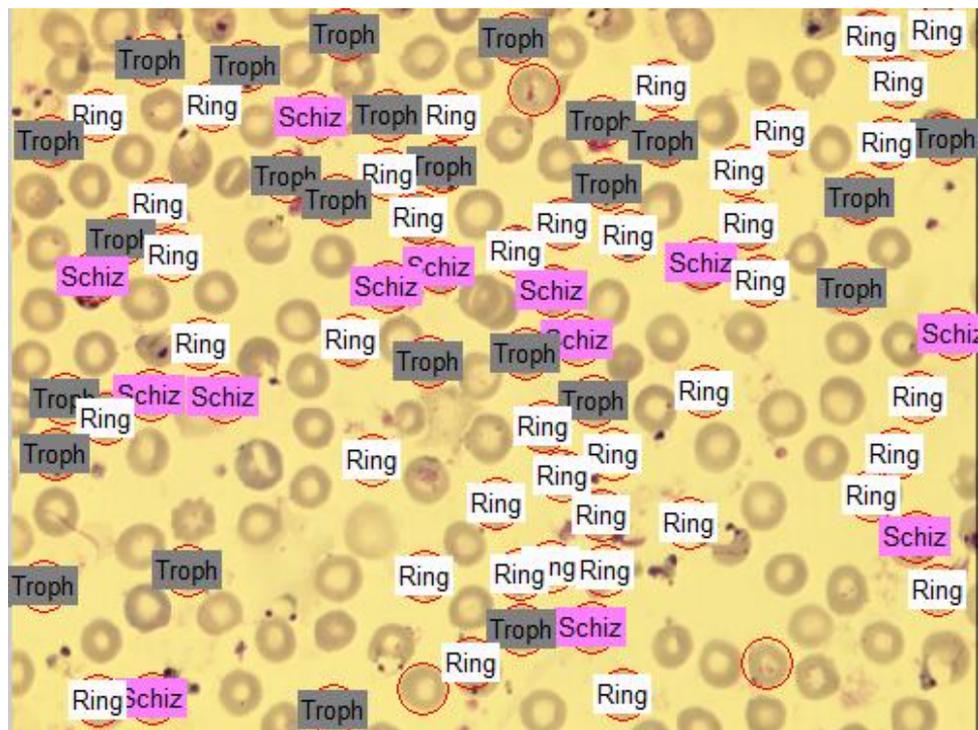
(b)

Figure 6.33: Life-recognition algorithm performed on the image in (a).

(a) Original Image (b) Resultant image with the life stages recognised. Ring- Ring Trophozoite, Troph- Trophozoite, Schiz- Schizont. The arrowed cell has been identified as an infected cell however the algorithm did not recognise its life stage.



(a)



(b)

Figure 6.34: Life-recognition algorithm performed on the image in (a).

(a) Original Image (b) Resultant image with the life stages recognised. Ring- Ring Trophozoite, Troph- Trophozoite, Schiz- Schizont.

Automated Malaria Diagnosis Using Mobile Phones

6.3.4 Algorithm Demonstration

Based on the Hue-ARR algorithm a GUI was developed to analyse and evaluate the concept of automated malaria diagnosis. The user has the options to visually analyse the image and to detect the RBCs, WBCs, infected cells, life stages and to estimate the total number of cells and eventually distinguishes the life stages using the method described in Section 6.4.1 . From the estimate of RBCs and infected cells, the parasitemia, is calculated.

In the GUI, the user has the option to load the image from the dataset and it is displayed in the top-left window. The top-right window displays the result of the processing in which all the parasitized cells will be labelled. The bottom window axis of the GUI displays various results including the intermediate results of different morphological operations which include morphologically dilated image, closed image and the ARR transformed image that are involved in the process. There are two edit windows through which the user can input different radii and threshold values for structuring element and peak detection algorithm respectively. A performance analysis can therefore be undertaken by choosing different radii for the RBCs and also different threshold values for the peak detection algorithm. The following Figures 6.35- 6.42 demonstrate the experimental results obtained.

The top two windows of the GUI represent the image to be processed (top-left window) and the processed image with infected cells marked (top-right window) respectively. The intermediate steps are shown in the bottom window. The bottom window of the GUI in Figure 6.35 shows the marked RBC while that of Figure 6.36 shows the marked WBC present. Figure 6.37 shows the infected cells marked in red circles and Figure 6.38, 6.39 and 6.40 shows examples of the three asexual life stages of the parasite. Figures 6.41 and 6.42 illustrate the intermediate results obtained from the morphological processing of the image. All the results shown in the bottom window could be viewed by selecting the appropriate radial buttons provided on either side of the window. The numerical data obtained upon processing the image, which estimates the total number of RBCs, WBCs and infected cells are provided at the output columns at the left side of the screen. The parasitemia of the corresponding processed image is also provided at the left column.

Automated Malaria Diagnosis Using Mobile Phones

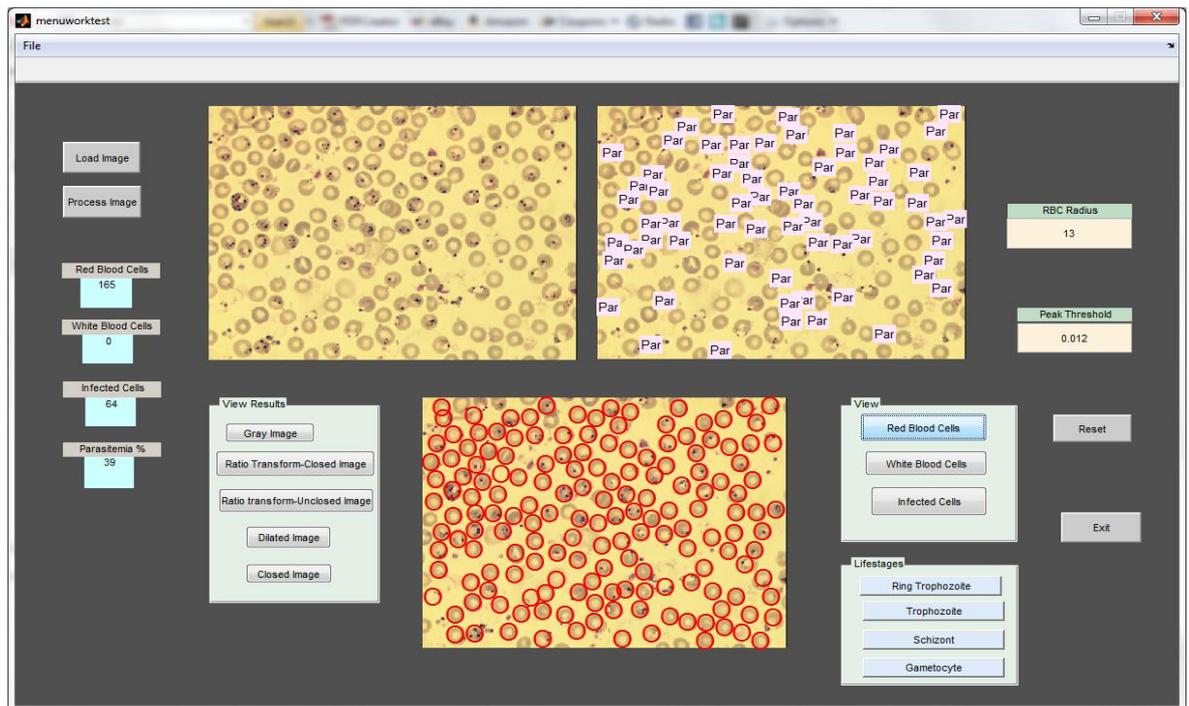


Figure 6.35: GUI Demonstration of Automated Malaria Diagnostic Set-up.

The detected RBCs are marked by red circles in the bottom window. The results can be viewed by selecting the buttons on either side of the bottom window. The numerical data provides estimates of the number of RBCs, WBCs and parasitized cells and, of the parasitemia of the input image are displayed on the columns provided at the left hand side.

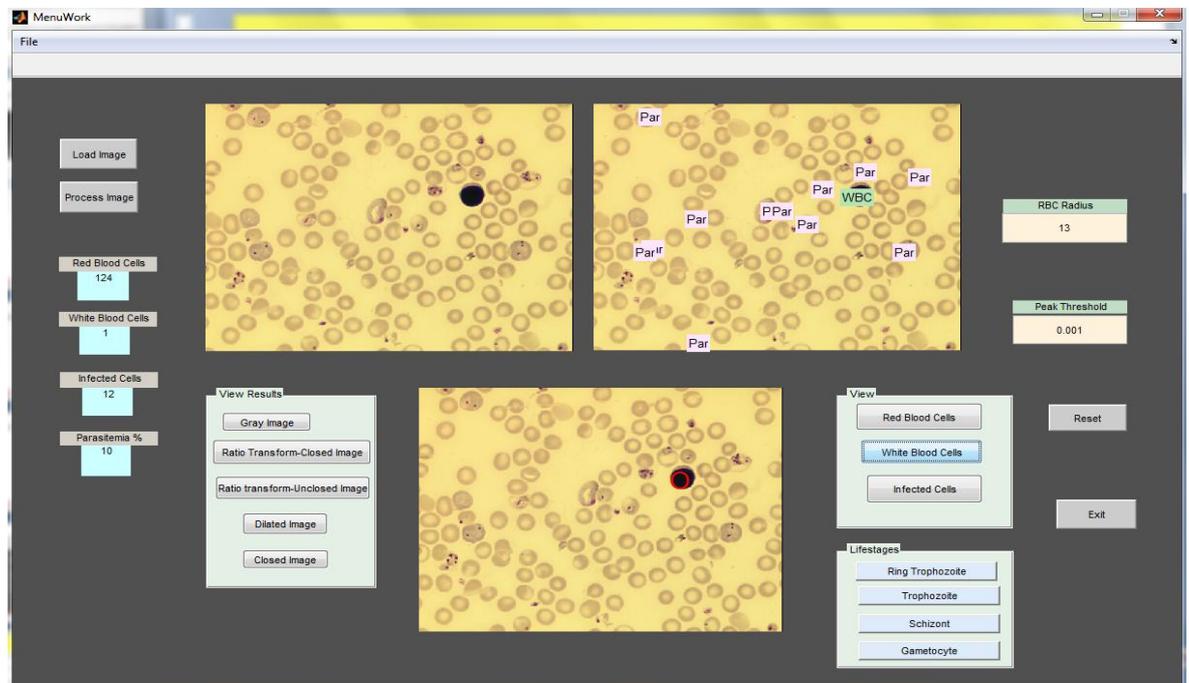


Figure 6.36: GUI Demonstration of Automated Malaria Diagnostic Set-up.

Infected blood image with WBCs present. The leukocyte is marked by the red circle in the bottom window.

Automated Malaria Diagnosis Using Mobile Phones

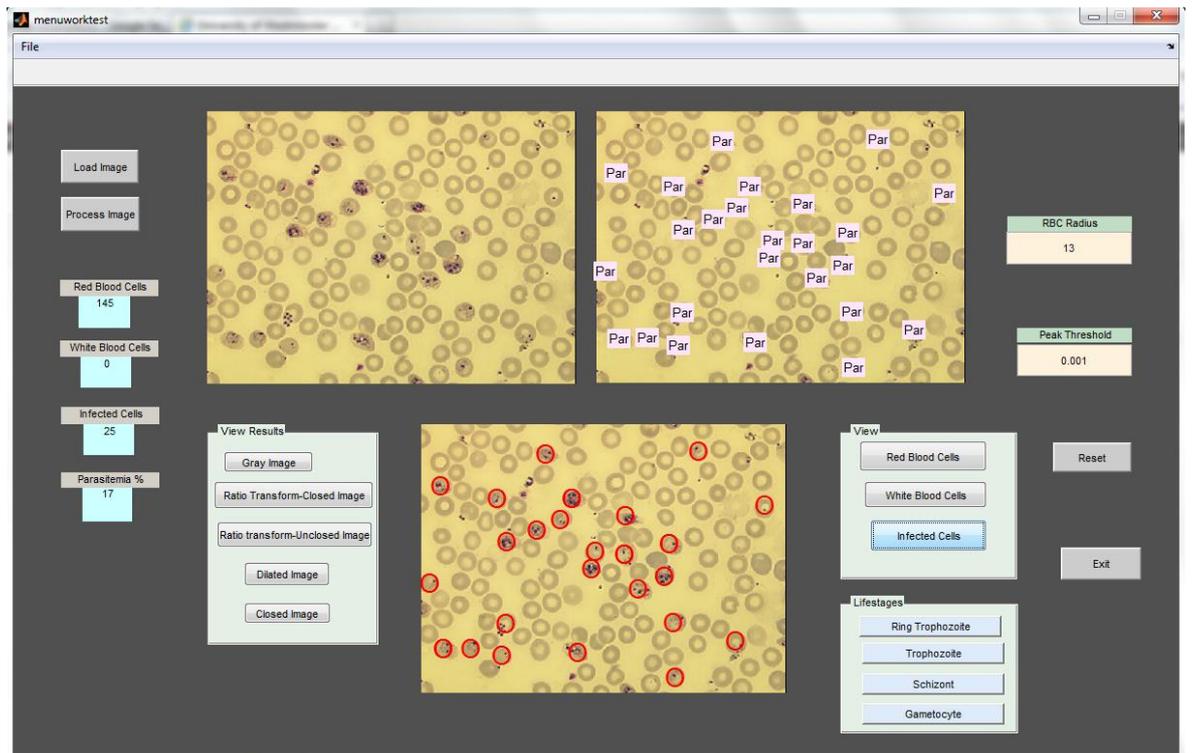


Figure 6.37: GUI Demonstration of Automated Malaria Diagnostic Set-up. The infected cells are marked by red circles and are displayed in the bottom window.

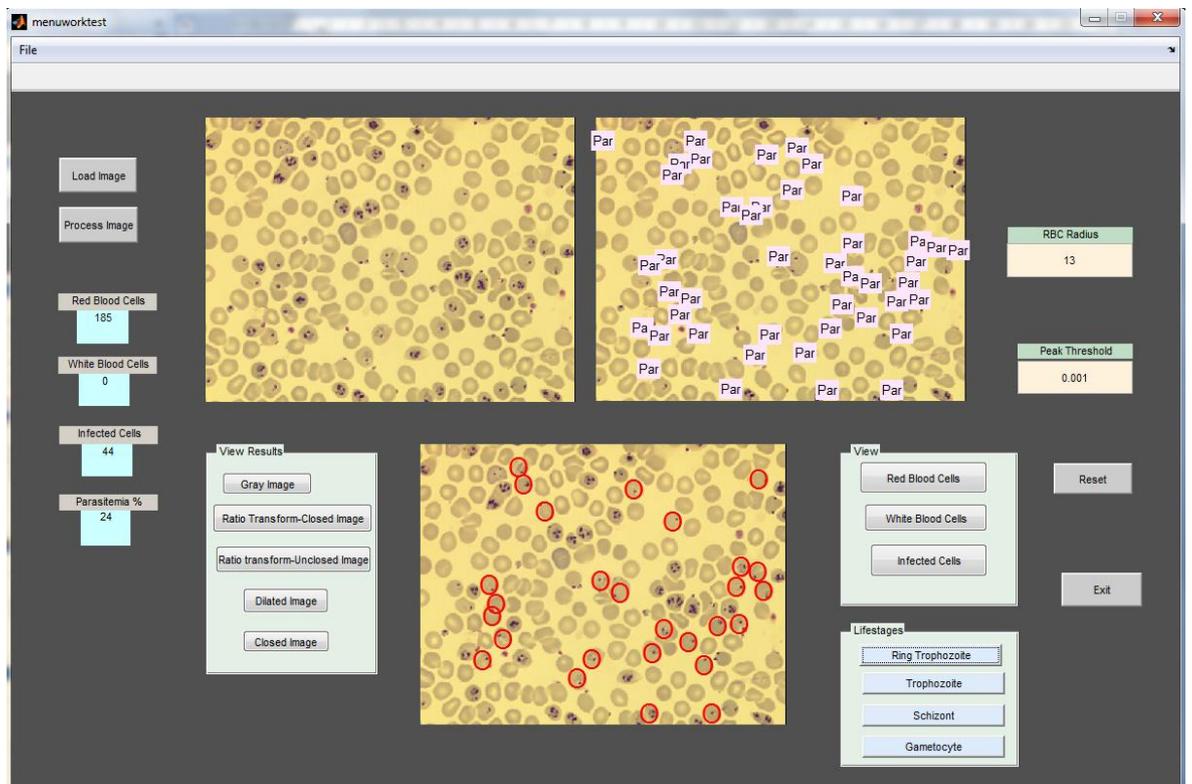


Figure 6.38: GUI Demonstration of Automated Malaria Diagnostic Set-up. The bottom window displays the Ring trophozoite life stage marked by red circles.

Automated Malaria Diagnosis Using Mobile Phones

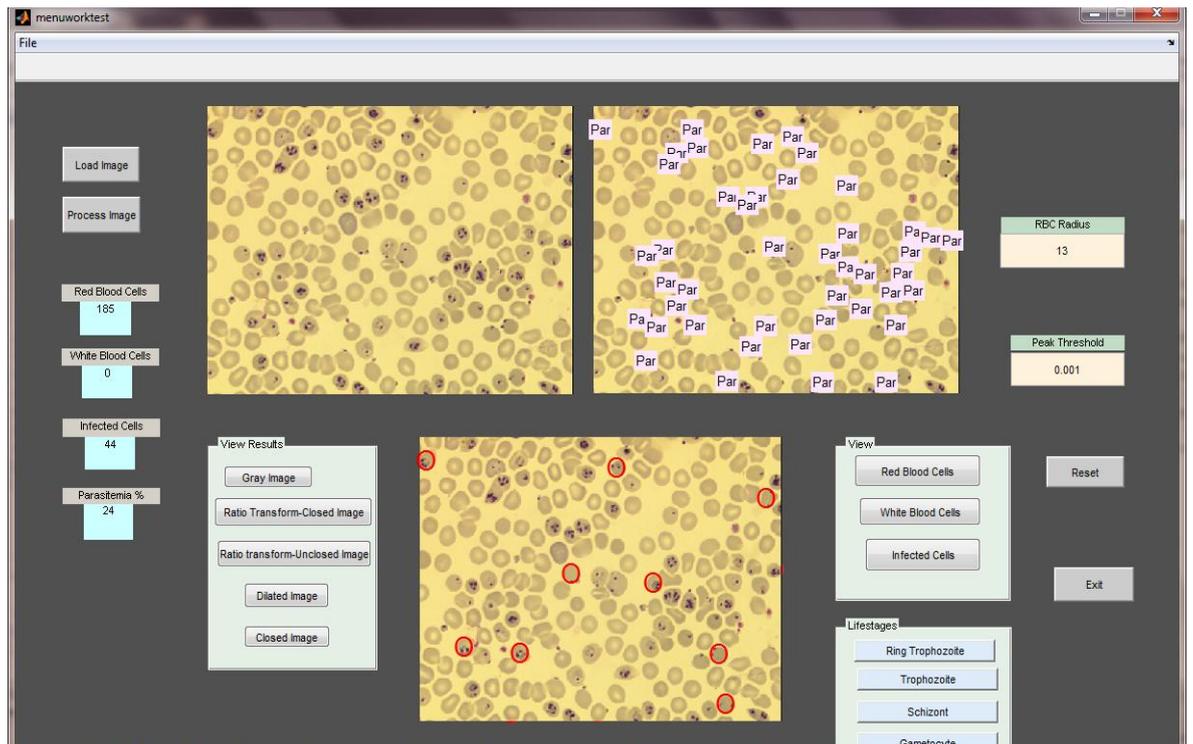


Figure 6.39: GUI Demonstration of Automated Malaria Diagnostic Set-up. The Trophozoite life stage is marked in red circles and is displayed in the bottom window.

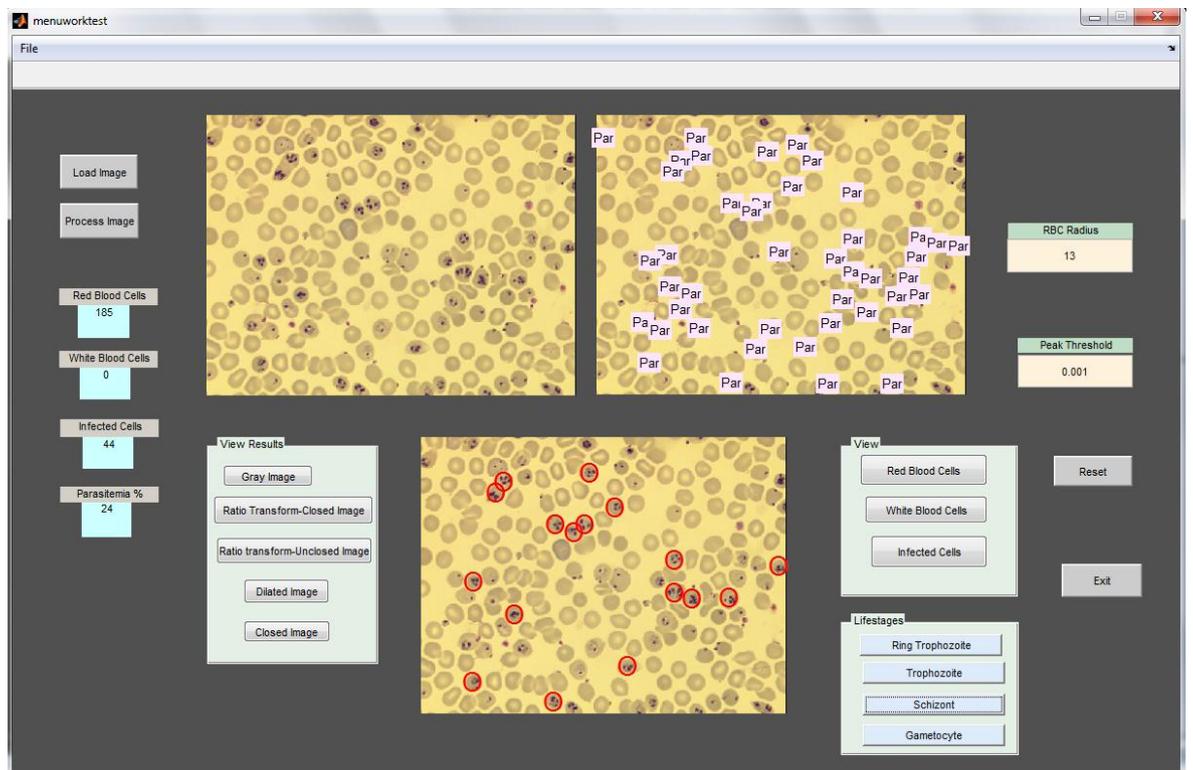


Figure 6.40: GUI Demonstration of Automated Malaria Diagnostic Set-up. The bottom window displays the Schizont life stage marked by red circles.

Automated Malaria Diagnosis Using Mobile Phones

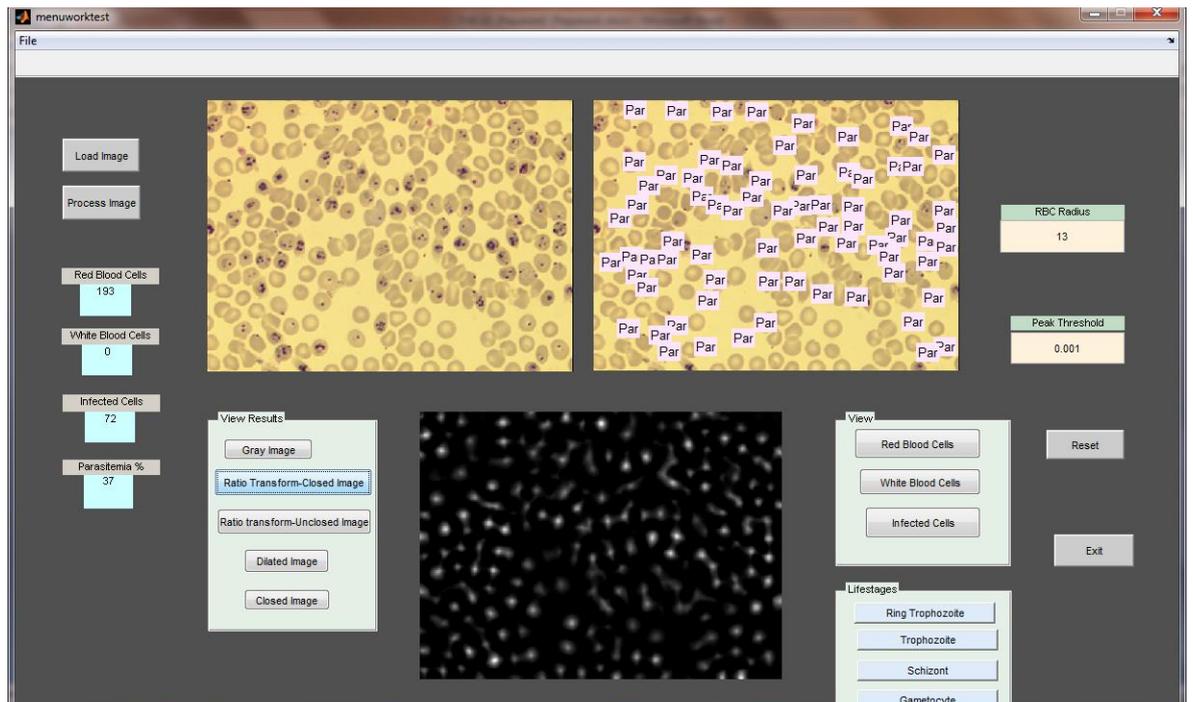


Figure 6.41: GUI Demonstration of Automated Malaria Diagnostic Set-up.

The ratio transformed image of the input is shown in the bottom window.

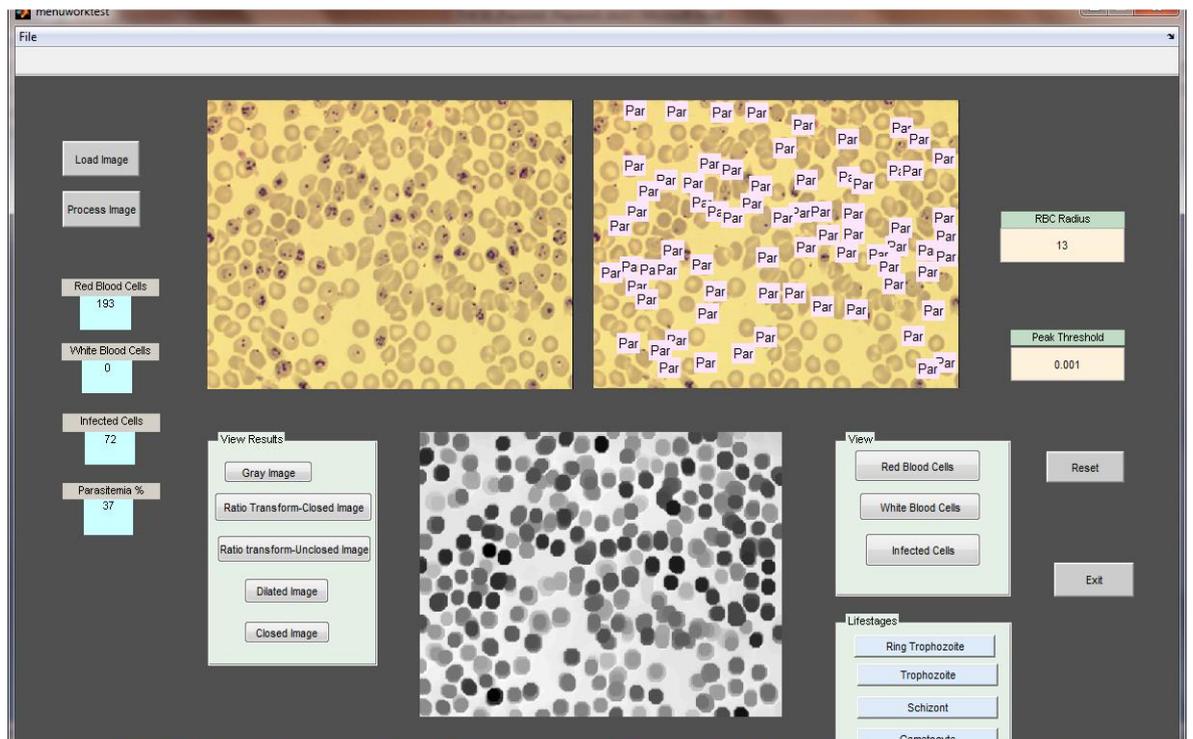


Figure 6.42: GUI Demonstration of Automated Malaria Diagnostic Set-up.

The intermediate result obtained by performing the morphological closing is shown in the bottom

Automated Malaria Diagnosis Using Mobile Phones

6.5 Discussions and Summary

In this chapter the most important function of automated malaria diagnosis- parasite detection, were addressed and an algorithm to detect the infected cells proposed. Within a limited time and limited number of image data, a performance analysis was conducted, not only for the parasite identification but also for RBC and WBC identification. The raw images were processed without introducing any enhancement techniques to correct the non-uniform illumination and staining. Hypothetically and practically it is not possible to compare the performance measurement of the automated diagnosis with manual microscopy since the efficiency of the microscopic examination depends on the experience of the microscopist. Hence, assuming that the parameters of performance measurements for manual microscopy are 100%, the maximum number of cells available was processed. For any diagnostic tool, an accuracy of 96% implies a performance comparable or even better than average microscopists. Considering the sensitivity and specificity, the system achieved approximately 94% sensitivity and 97% specificity. It is to be noted that these statistics are based on per cell measurement and does not pertain to the complete process per slide image. This in fact is promising considering the fact that more than 16,000 cells was examined to come to this conclusion. This will result in a sensitivity threshold of approximately 800 parasites/ μ l [0.6 for 10000 RBCs= $(1/0.6)*500$] compared to the average sensitivity threshold of manual microscopy by an expert microscopists, which is arguably 500 parasites/ μ l [74].

The specificity however needs improvement. Even though 97% specificity is acceptable compared to existing standards, the 3% yields a positive result for 300 non-infected objects which leads to the probability of false diagnosis of at least one healthy image in a negative specimen. However, based on the diverse nature of the sample images used for performance analysis, it can be expected that with a stable image acquisition set up in conjunction with the proposed diagnostic tool, the sensitivity and specificity could be greatly improved without any trade-offs. Similar conditions will also improve the performance of RBC and WBC detection and can be used as an additional advantage of the proposed diagnostic tool.

Automated Malaria Diagnosis Using Mobile Phones

The life stage identification was explained in Section 6.4 of the chapter. Although parasite detection is the primary aim of the research, the life stage identification was joint and contextual. Similarly, the parasitemia estimation, which is vital for any pathological analysis of infected parasitic disease, is an extension of the parasite detection. The GUI demonstrated at the end of the chapter aids the analysis of the process and provides some insight to the design of the phone-based system. It illustrates the working of the system and leave scope for further improvement and testing. A journal based on the work explained in this chapter is in preparation to be submitted to the Malaria Journal (BioMed Central).

Species recognition is another challenge for automated malaria diagnosis, which has not been addressed so far in this research. Different species of the genus plasmodium possess different morphological features and distinctive characteristics imposed on the host RBC such as enlargement of the cells. Analysis of this characteristics and feature and a possible work on species identification is pending for the future.

Chapter 7

Extensions

Automated malaria diagnosis using image processing is a vast area of study with many subsections. As demonstrated in the previous chapters, the research has been successful to a great extent in developing an effective malaria diagnostic tool. During the course of the study, few parallel investigations which were not exactly the principal aim of this research but have made its own impact and contributions to the image processing field as well as the global healthcare were carried out. They are;

- (a) *P.falciparum* gametocyte detection for post-treatment malaria diagnosis;
- (b) Automated malaria parasite identification using fluorescent image analysis.

These are extensions to the original objectives of the research and are extended applications of the Annular Ring Ratio (ARR) transform explained in Chapter 4.

7.1 P.Falciparum Gametocyte Detection

In Chapter 6, the life stage recognition algorithm for malaria diagnosis was explained and the results were demonstrated. The ring trophozoite, trophozoite and schizonts life stages are the asexual stage in which the parasite resides within the RBCs. The sexual form of the malaria parasite is the gametocyte which will be present in the blood at a later stage of the infection. The gametocytes appear between 7 and 15 days from the onset of the disease and stay for 6-7 days [75]. They aid in the transmission of the disease by entering into the gut of the mosquito during a blood meal.

Around 90% of the images used for testing the automated system contained the species *P.falciparum*. Of the four species of genus plasmodium, *P.falciparum* is the most severe and deadly. According to the studies conducted by PATH Malaria Vaccine Initiative (MVI), almost all the deaths caused by malaria are due to *P.falciparum* [78]. Unlike other species, *P.falciparum* species cause acute anaemia by destroying the RBCs. It also leads to cerebral malaria which can cause neurological complications that are fatal.

Automated Malaria Diagnosis Using Mobile Phones

The asexual forms of the parasite are symptomatic and are accountable for clinical disease and death whereas sexual forms are asymptomatic and are responsible for transmission. The antimalarial drugs for *P.falciparum* mainly targets on asexual forms due to their hazardous nature. The gametocytes however are highly resistant to these antimalarial drugs and so can reside in the blood even after the treatment [76], [77]. Since the transmission of malaria depends on these sexual forms, a fast diagnostic method is essential for control and eradication efforts.

Taking in to account the dangerous nature of the species, a parallel investigation was carried out during the development of automated malaria diagnosis, to identify the potential gametocytes of *P.falciparum* in the blood. The unique structure of the gametocyte helps to distinguish them from the other life stages. Although, the automated set up developed in this research successfully detect the gametocytes, species identification is not yet possible. An automated diagnostic tool especially designed for *P.falciparum* gametocyte was developed which mainly aims at post-treatment malaria diagnosis of this species.

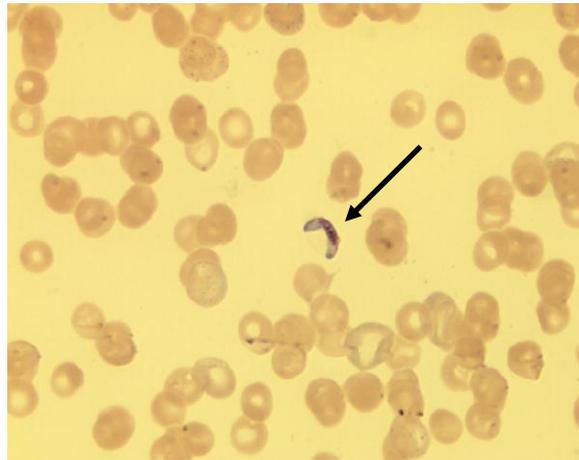
Post-treatment malaria diagnosis is usually carried out to re-evaluate the presence of the parasite in blood subsequent to the treatment of the disease. The diagnosis is performed after 4 weeks of the initiation of the antimalarial drugs. It is highly unlikely that the asexual forms will be found during this post-treatment diagnosis as they are less resistant to the drugs; however, the presence of gametocytes can be detected when further administration of gametocytocidal drugs is required [76].

Image segmentation for normal malaria diagnosis separates the potential parasites and erythrocytes (RBCs) from the image background. This aids in estimating the parasitemia. However, post malarial diagnosis does not require any RBC identification and there will not be any early life stages (asexual forms) present in the blood. Hence a segmentation algorithm which focuses on identifying only the gametocytes by distinguishing them from the WBCs has been designed. Instead of the usual segmentation procedures, it utilises a modified Annular Ring Ratio (ARR) transform method along with the component characteristics such as colour, geometry and morphology of the gametocytes. In addition, the method locates and labels the White blood cells (WBCs) present in the blood film.

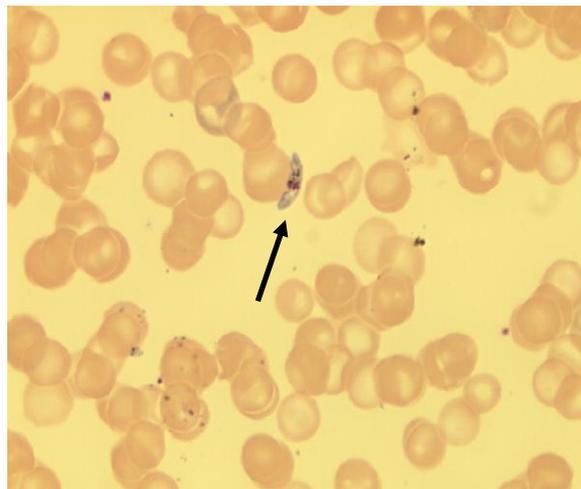
Automated Malaria Diagnosis Using Mobile Phones

7.1.1 *P.falciparum* gametocyte morphology

The *P.falciparum* gametocytes differ from the other life stages due to their crescent or sausage shape. Unlike other life stages, the RBC cell membrane in which the parasite reside may not be well defined. The orientation of the parasitic form in the image also varies. The chromatin in the cell will appear as a single mass or diffused unlike asexual trophozoic forms in which nucleus is clearly visible. Figure 7.1 shows the distinctive shape and orientation possessed by this form.



(a)



(b)

Figure 7.1: *P.falciparum* gametocyte morphology.

The parasite shows distinctive crescent shape (a) or sausage shape (b).

7.1.2 Proposed Gametocyte Detection Algorithm

Unlike the detection of earlier stages of the parasite, where it is necessary to perform a morphological dilation and erosion prior to the ARR transform method, a modified ARR transform is performed directly on to the grey-scale image for gametocyte detection. As shown in Figure 7.1, the image does not contain any other life stages so the detection of all the foreground objects is not necessary. A ratio of intensity values of every pixel in the concentric ring structuring element and inner disk is calculated. Since we are interested in objects in which the inner region is darker than the outer, the rest of cell areas are clamped to zero and thereby reducing irrelevant background clutter due to non-stained components. The brightest point as a result of the intensity ratio occurs at the centre of the object and thus results in locating only the stained components in the image instead of all the cells. In order to avoid any platelet or artefact detection, the peak detection algorithm is set to a higher threshold. The image, thus processed, will locate the gametocyte and WBCs.

The WBCs and gametocytes present in an image follow certain characteristic features as evident in Figure 7.2. The orientation and spatial geometry is different for different images. The WBCs are the largest stained component in a blood image. The asexual form of the species *falciparum* usually does not enlarge the RBCs. However, the gametocytes stretch the membrane of the RBC as they grow.

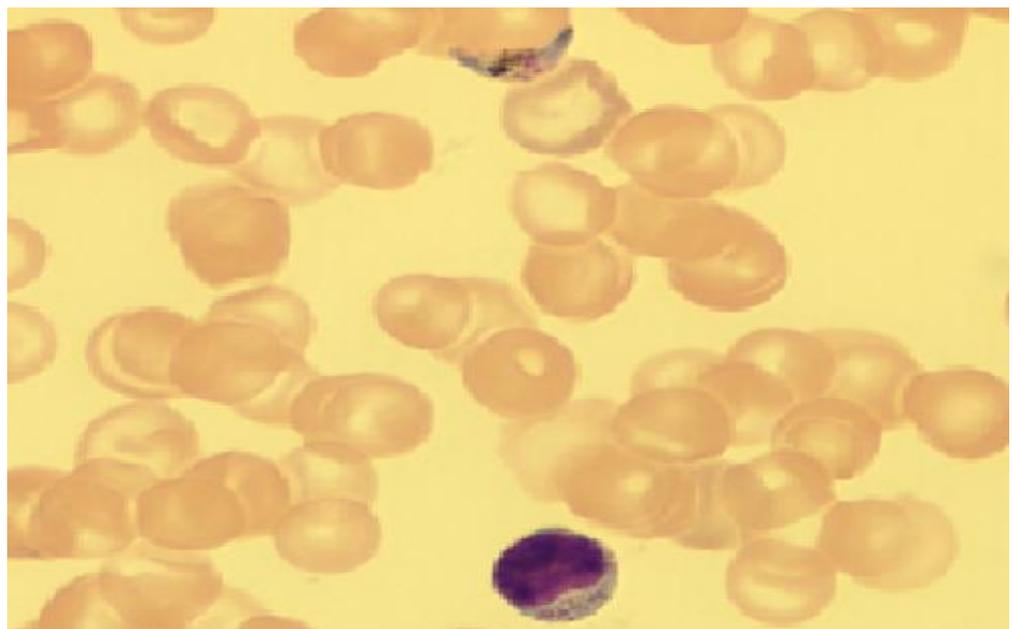


Figure 7.2: WBC and *P.falciparum* gametocyte in an image

Automated Malaria Diagnosis Using Mobile Phones

The modified ARR transform method will locate the centroid of the parasite and WBCs. In order to differentiate both components, a number of factors such as the staining concentration, size and geometry of the cells and intensity distribution were considered. The WBCs are large, deeply stained whereas the intensity distribution of the gametocytes is different. Hence once the centroid is detected and the location is obtained using ARR method, the algorithm proceeds through the following steps:

- The neighbourhood of the listed coordinates within the diameter of the object is sectored into 4 quadrants.
- The localized average intensity and standard deviation of these quadrants at each location are calculated.
- The regions with less mean intensity and less standard deviation are marked as White Blood Cells.
- The regions of higher standard deviation and average intensity are identified as gametocytes.

Let X and Y be the coordinates of the centroid located by the ARR transform method. The region surrounding the coordinate will be considered as:

$$f(X, Y) = \begin{cases} WBC, L_{wbc} < \mu(A(X, Y)) + \sigma(A(X, Y)) \\ Gametocyte, L_{gamet} > \mu(A(X, Y)) + \sigma(A(X, Y)) \end{cases} \quad (7.1)$$

Where $f(X, Y)$ is the object under processing at (X, Y) , $A(X, Y)$ is the area surrounding each coordinates and $\mu(A(X, Y))$ and $\sigma(A(X, Y))$ are the mean and standard deviation of the located neighbourhood respectively. L_{wbc} and L_{gamet} are the threshold values manually evaluated by experimenting on different images. Overall 200 images were manually evaluated in order to set up the threshold values. Also histogram evaluations of these images were carried out to finalize the values. The neighbourhood of the locations to calculate the mean and standard deviation were set according to the size of the WBC and RBCs which carries the gametocytes. The size information was obtained after performing area granulometry on these images. Figure 7.3 demonstrates the application by displaying the original image and the ratio transformed image.

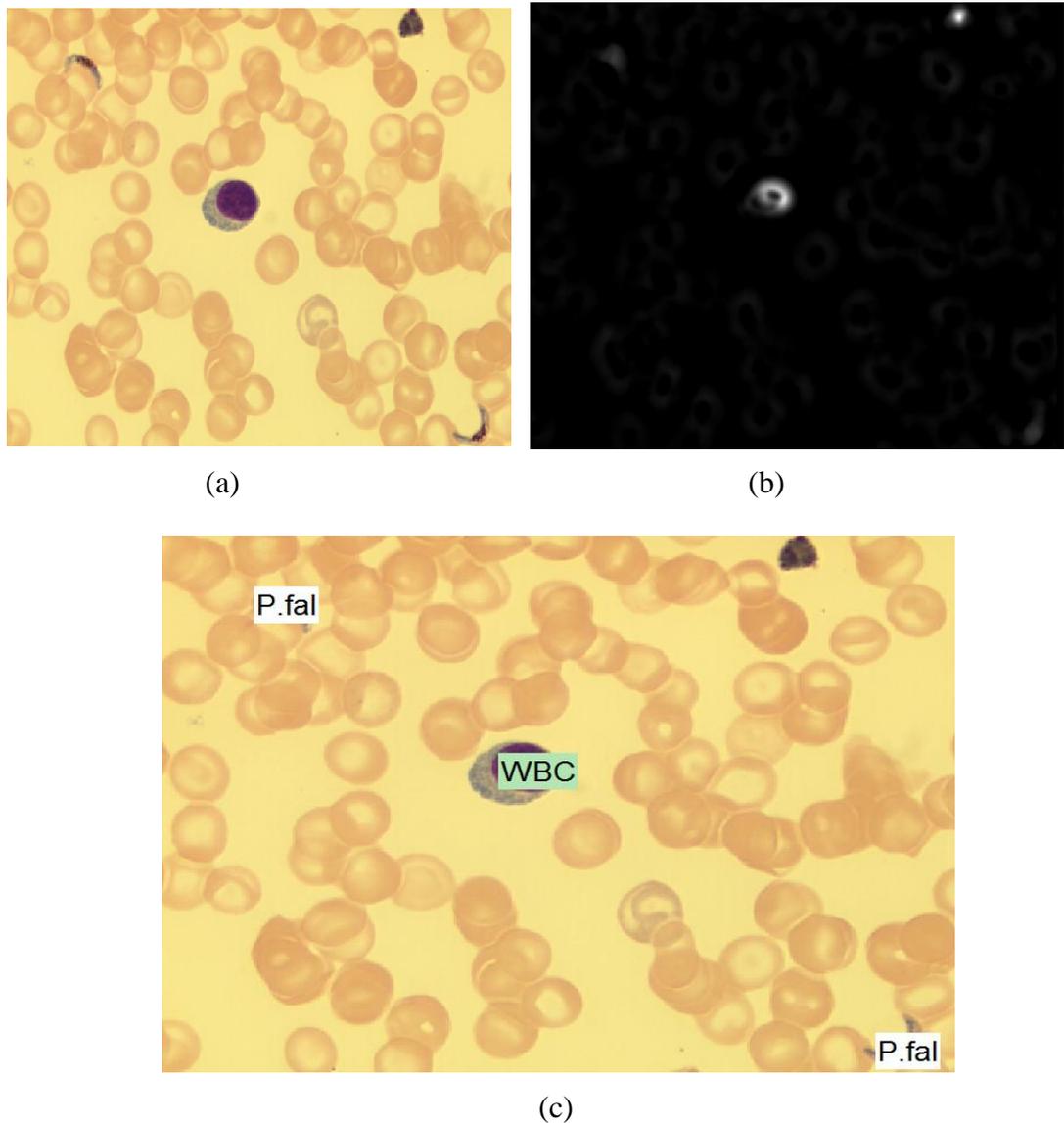


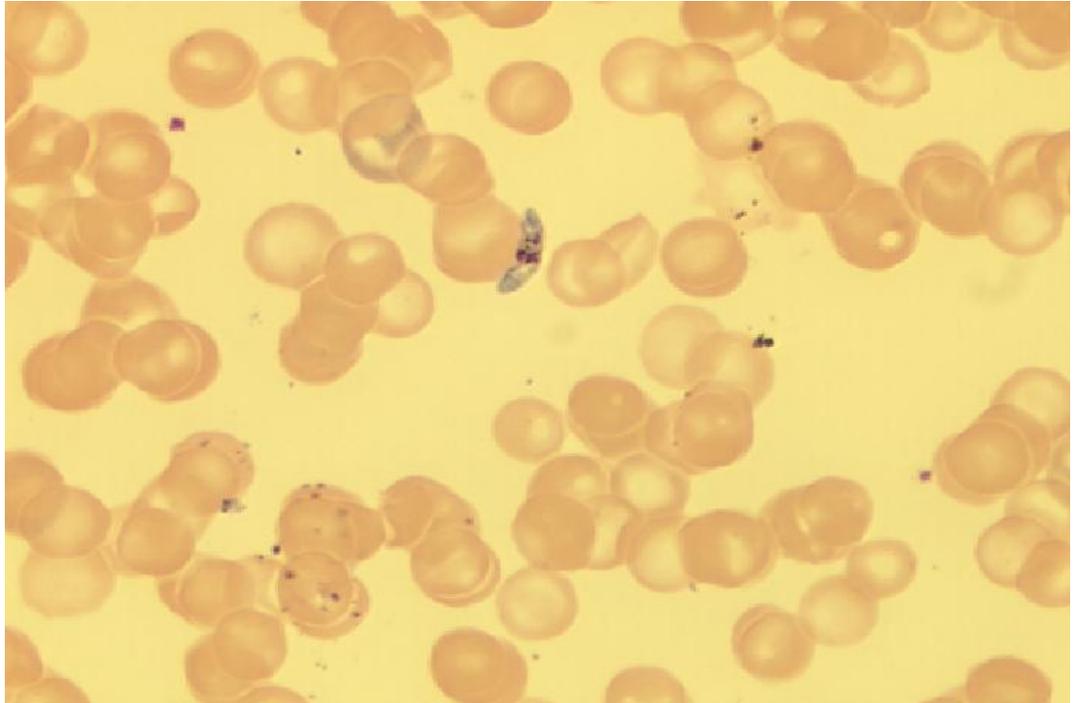
Figure 7.3: WBC and *P.falciparum* gametocyte differentiation algorithm.

(a) Original Image. (b) Ratio-transformed image which has peaks on the gametocytes, WBC and the artefact. The differentiating algorithm using the mean and standard deviation of neighbourhood intensities of the components eliminates the artefact in the top right corner and marks the gametocytes and WBCs as shown in (c)

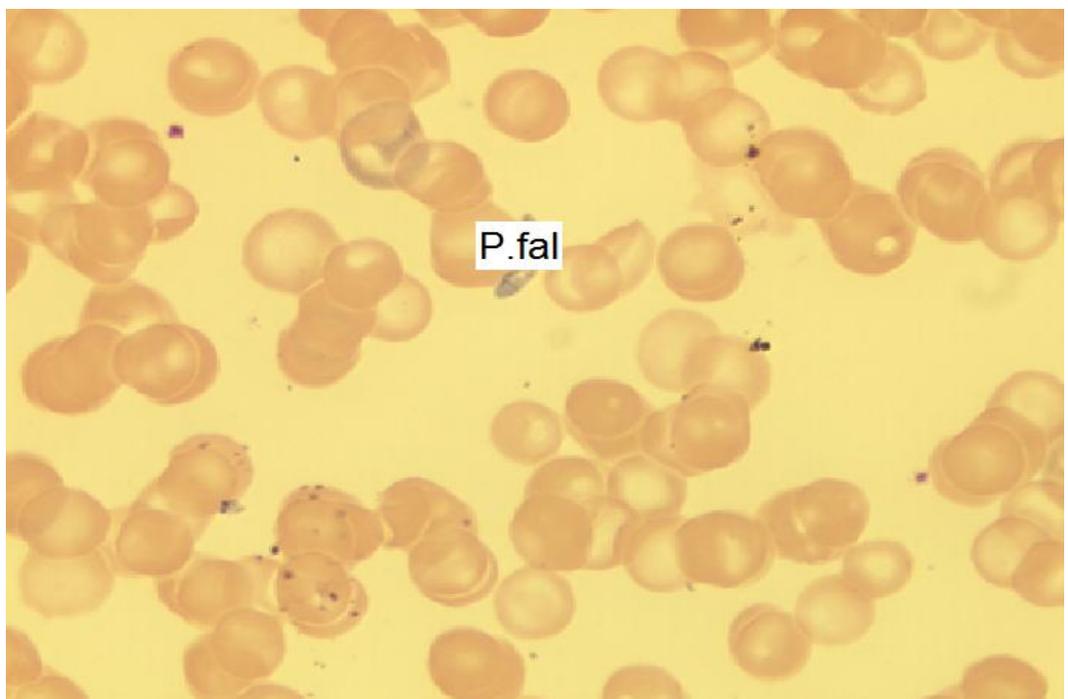
Figure 7.3 shows the results obtained after processing the image in Figure 7.3(a). Figures 7.4-7.11 shows the results obtained after applying the algorithm on images with different staining and illumination conditions. It also demonstrates the robustness of the system to the orientation of the gametocytes. The slides used were poorly processed with highly overlapping cells. The orientation of the gametocytes differs in each slide. Despite the poor slide quality, the system positively differentiates the gametocytes and WBCs and marks them automatically on the original image as the algorithm proceeds. Figure 7.12 is a

Automated Malaria Diagnosis Using Mobile Phones

manually prepared image for experimental purpose in which WBCs has been artificially pasted on to the original image. The image has all the three major types of WBCs as well as the parasite.



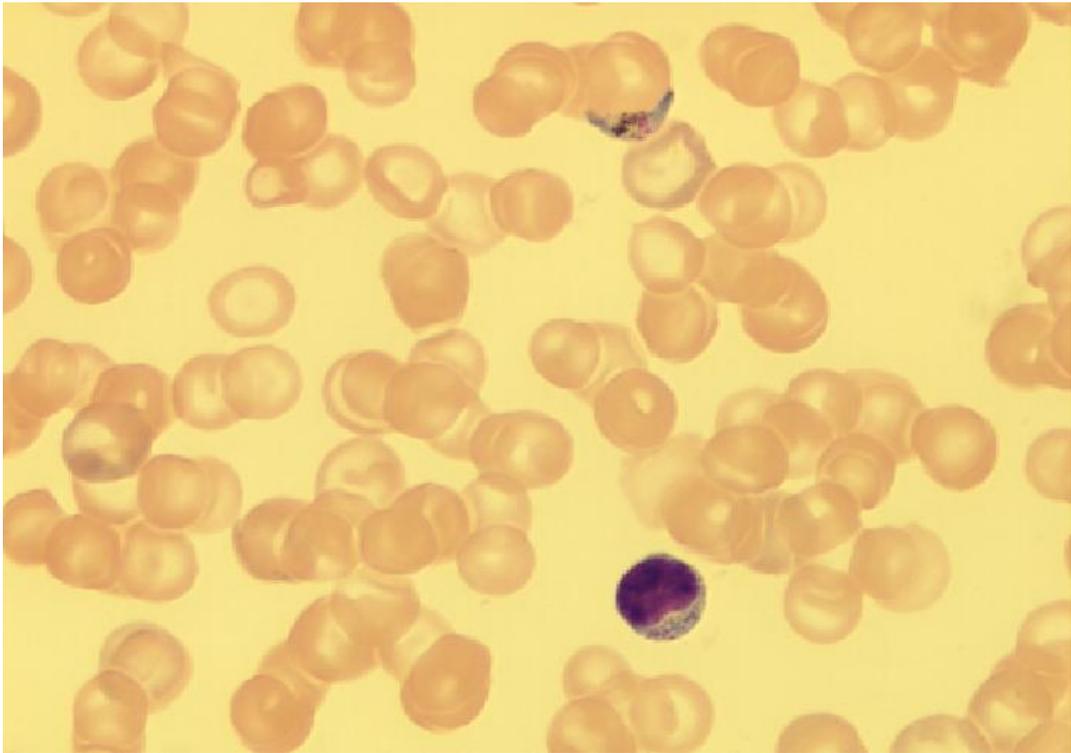
(a)



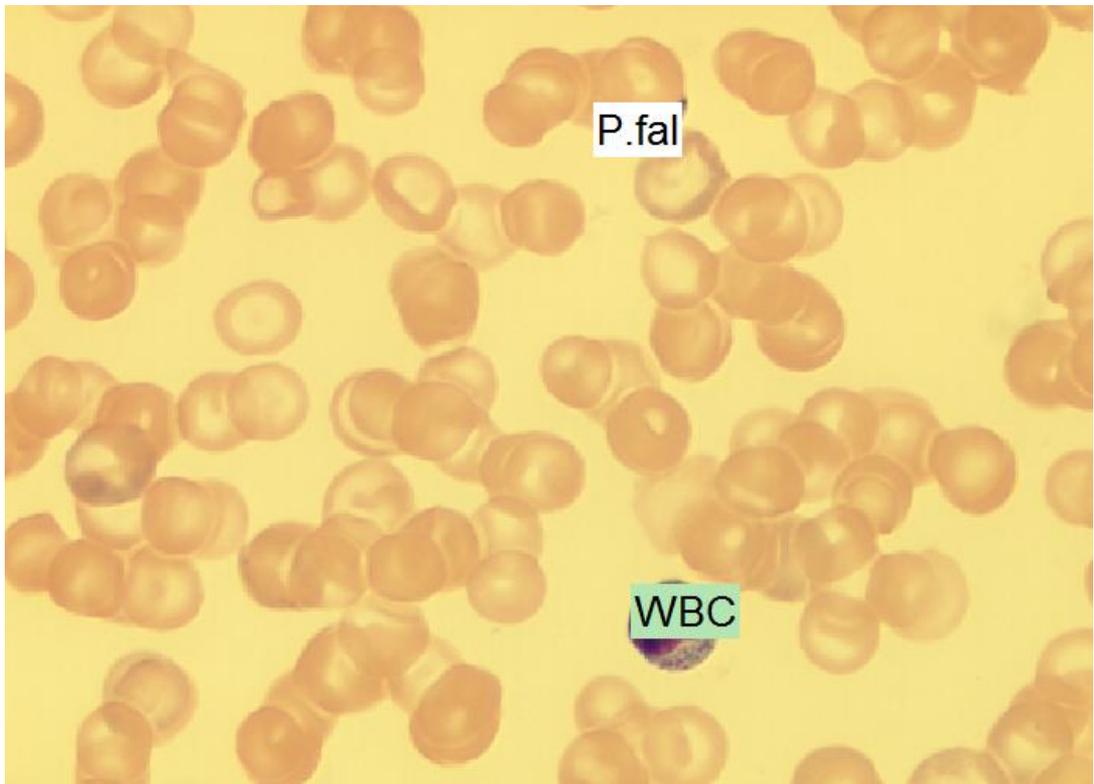
(b)

Figure 7.4: *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image



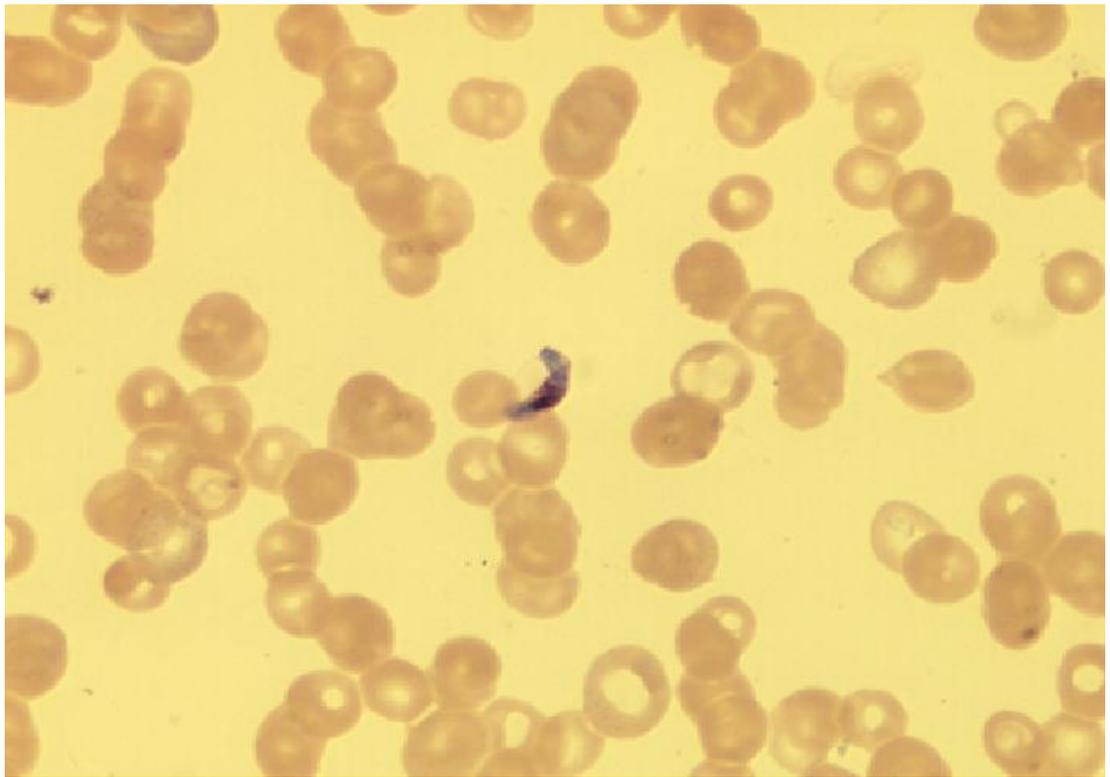
(a)



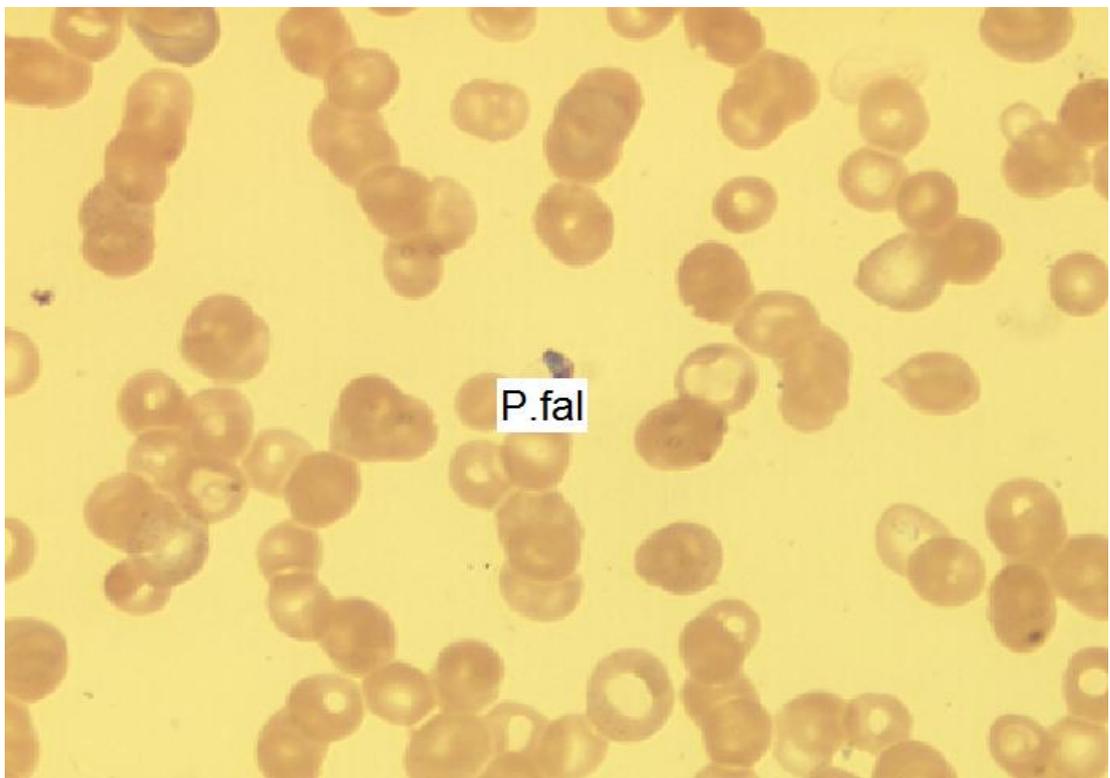
(b)

Figure 7.5: WBC and *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image



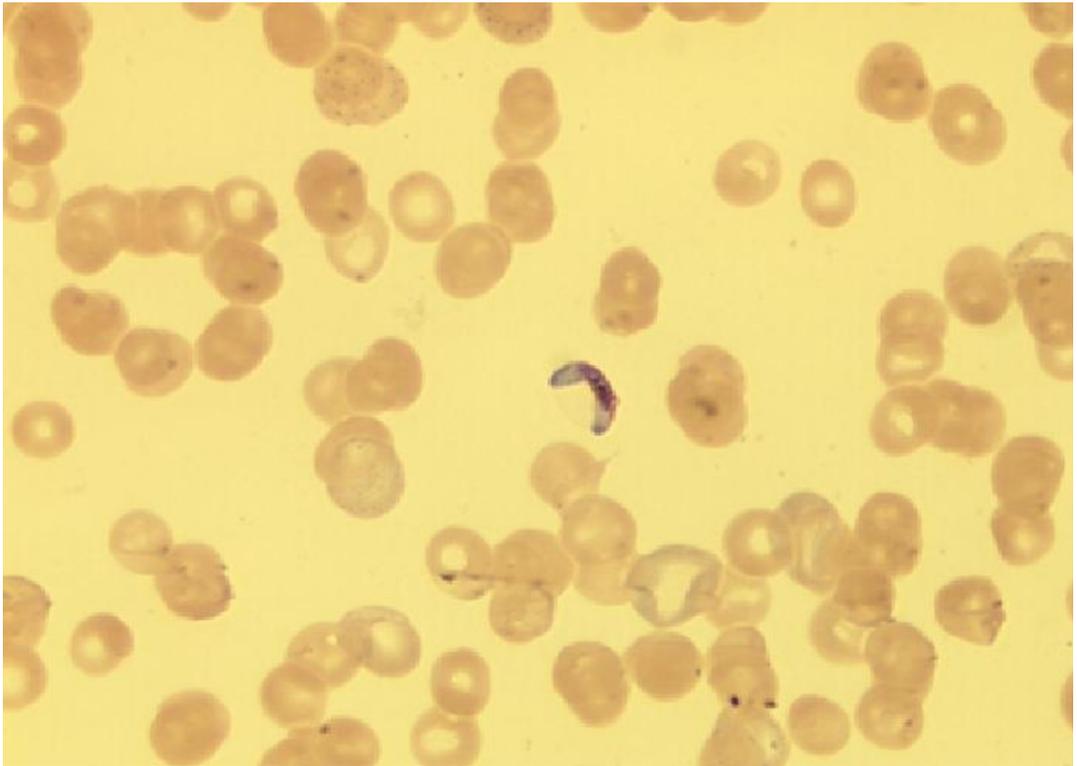
(a)



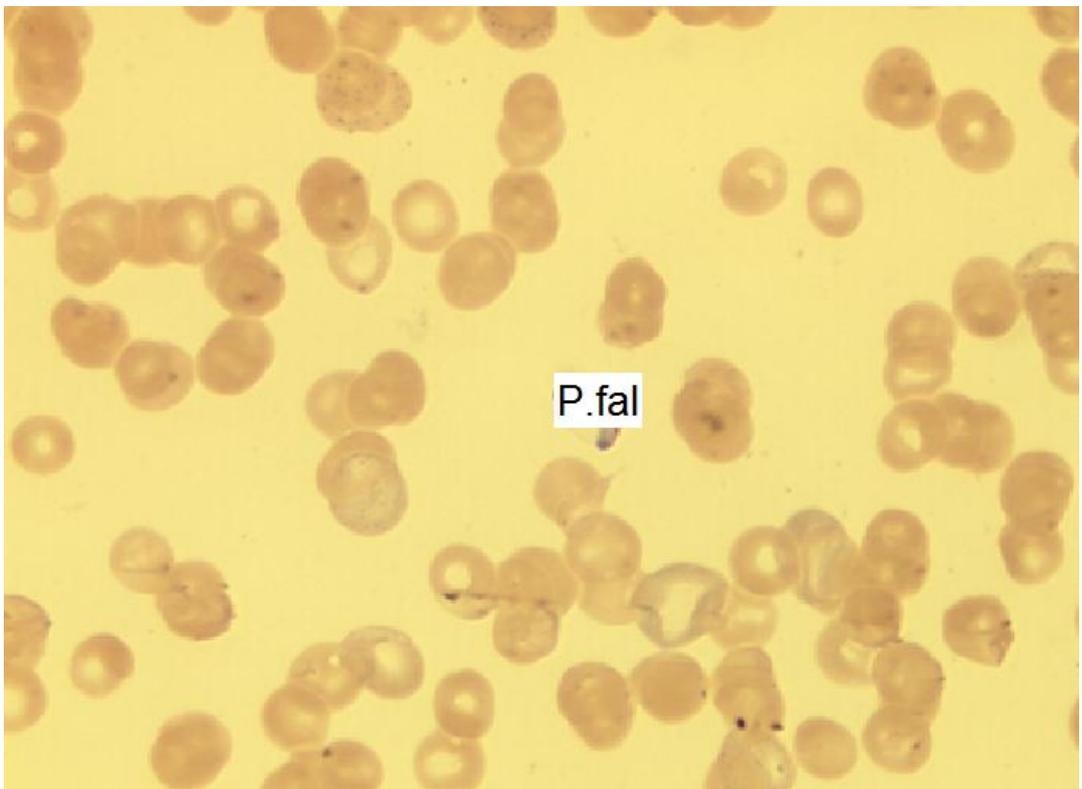
(b)

Figure 7.6: *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image



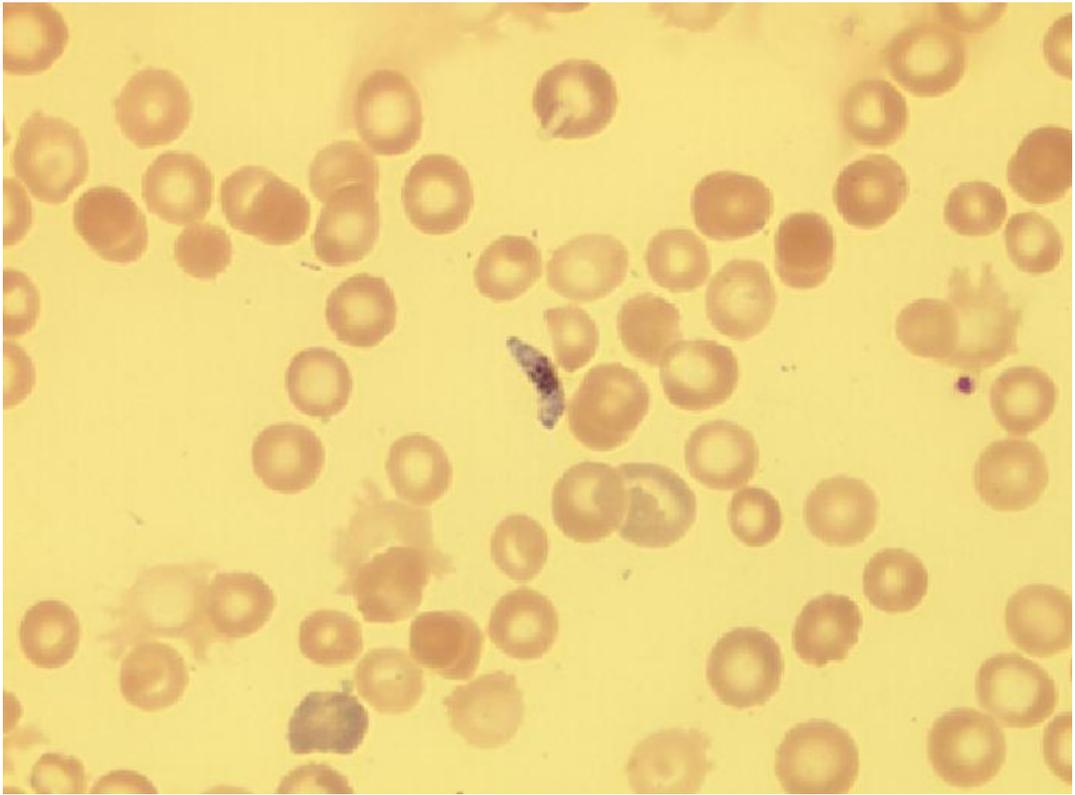
(a)



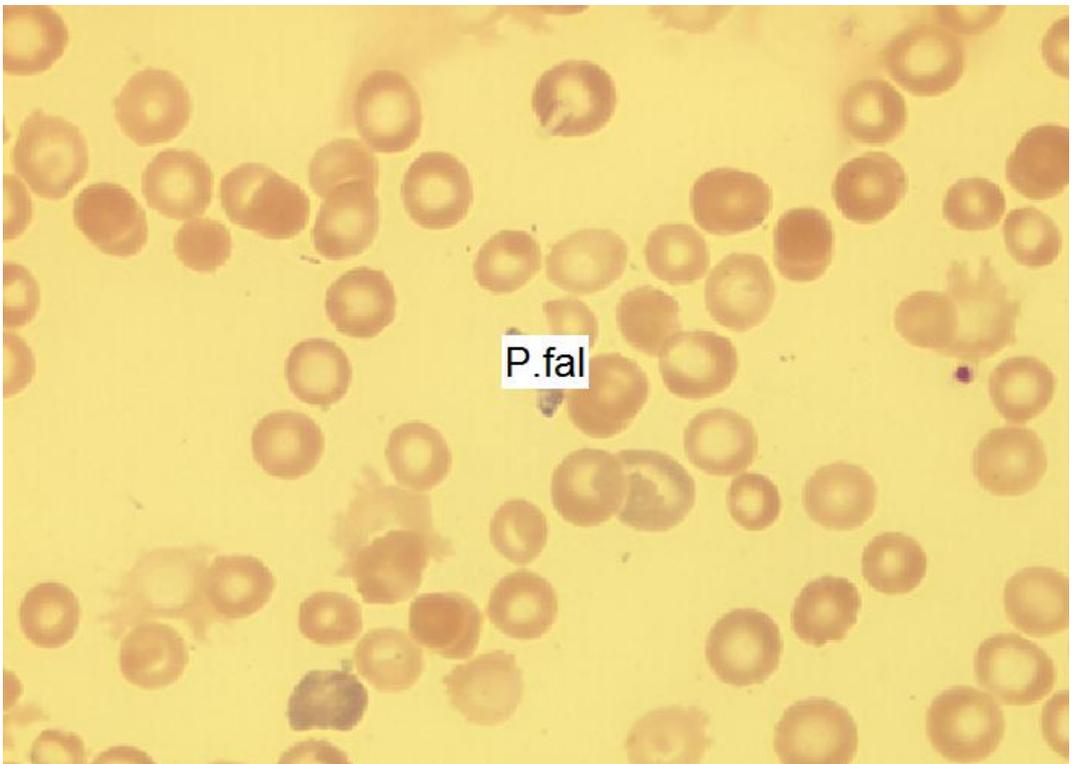
(b)

Figure 7.7: *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image



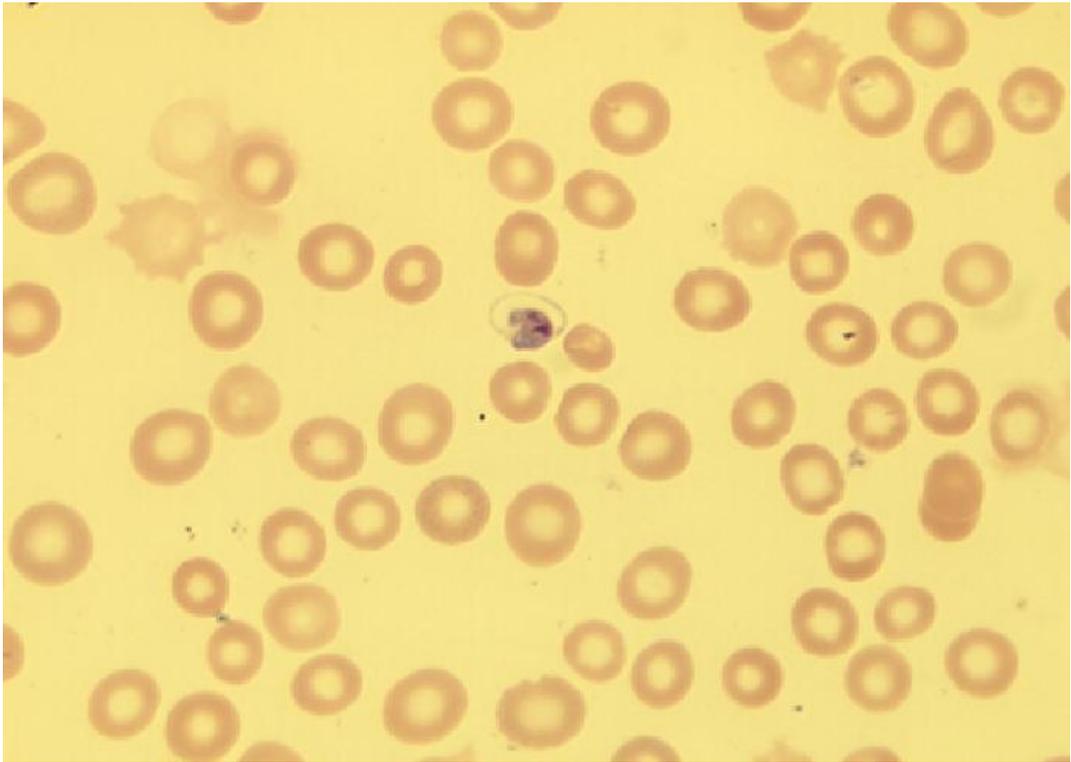
(a)



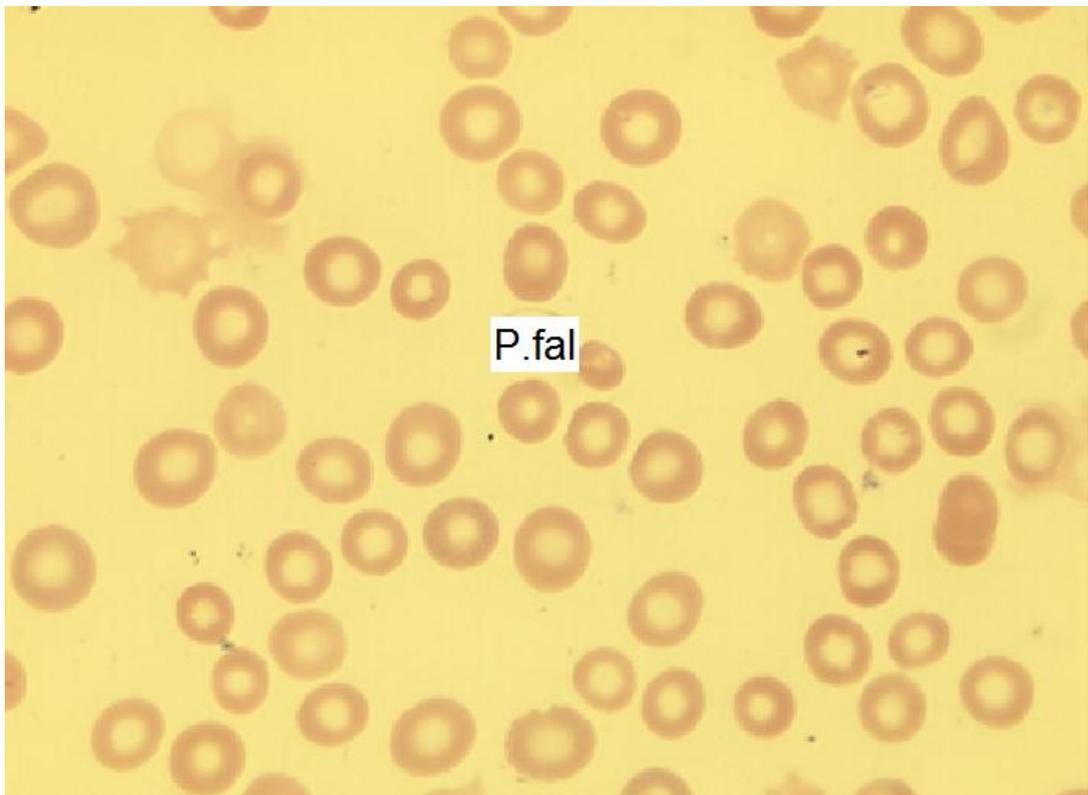
(b)

Figure 7.8: *P. falciparum* gametocyte differentiation algorithm.

(a) Original Image (b) Processed Image



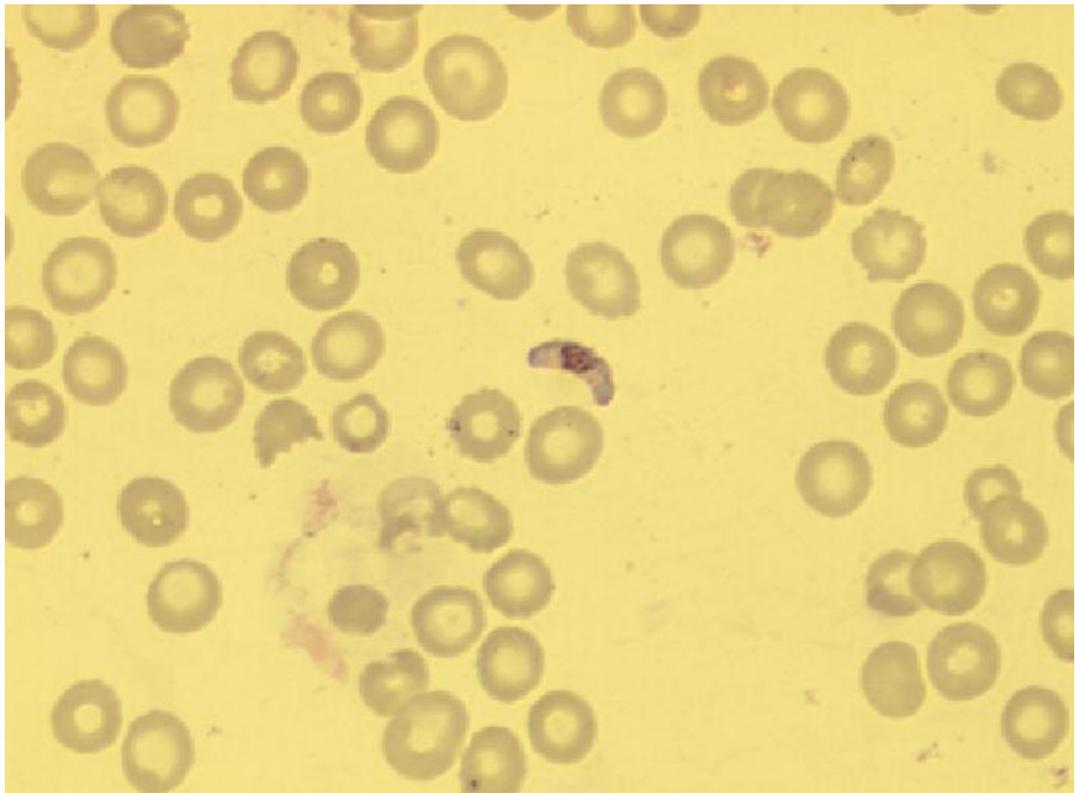
(a)



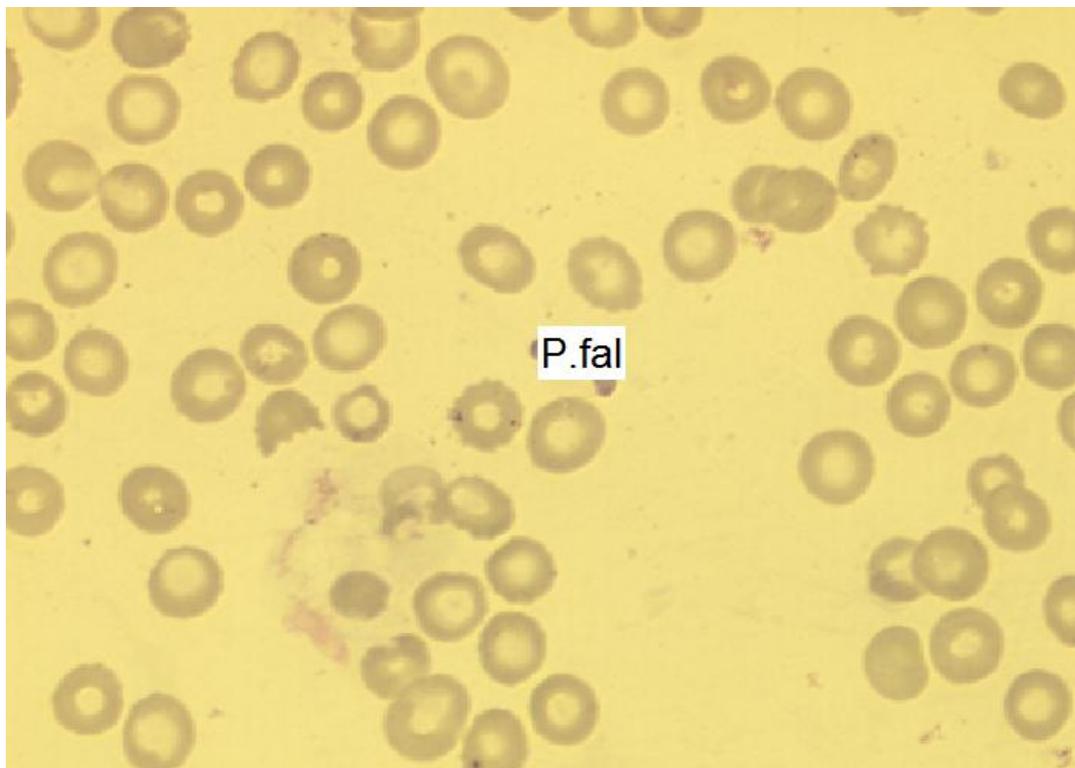
(b)

Figure 7.9: *P. falciparum* gametocyte differentiation algorithm.

(a) Original image (b) Processed image



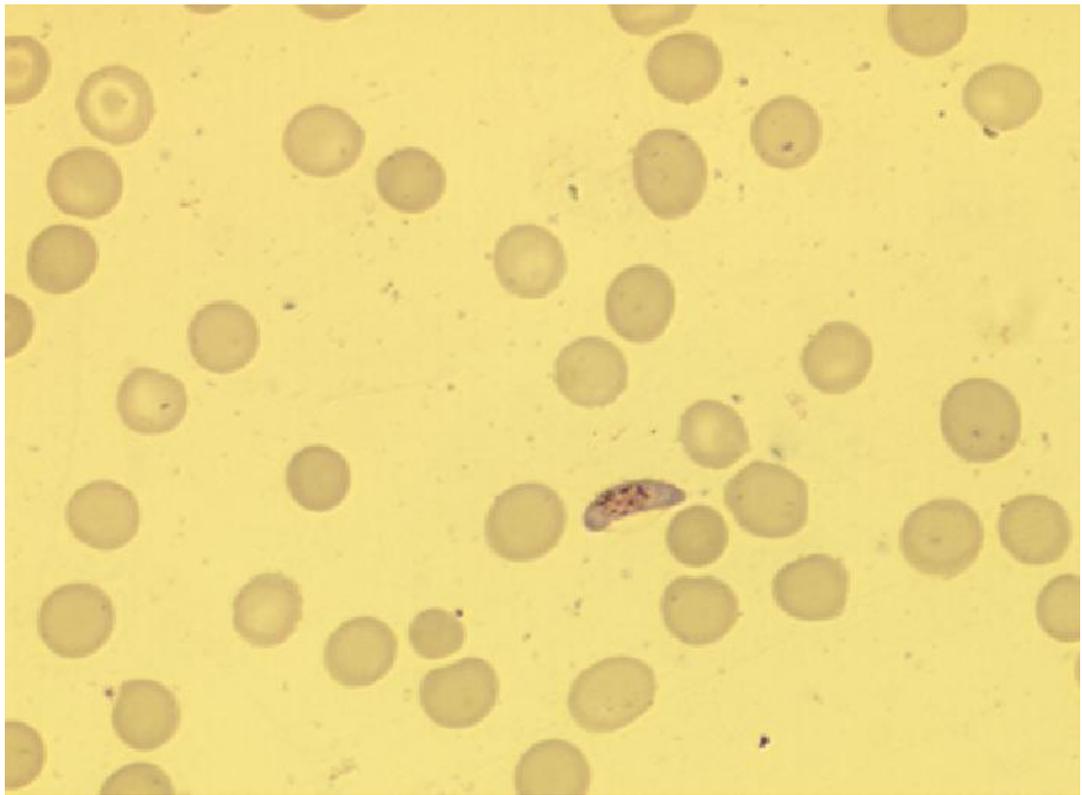
(a)



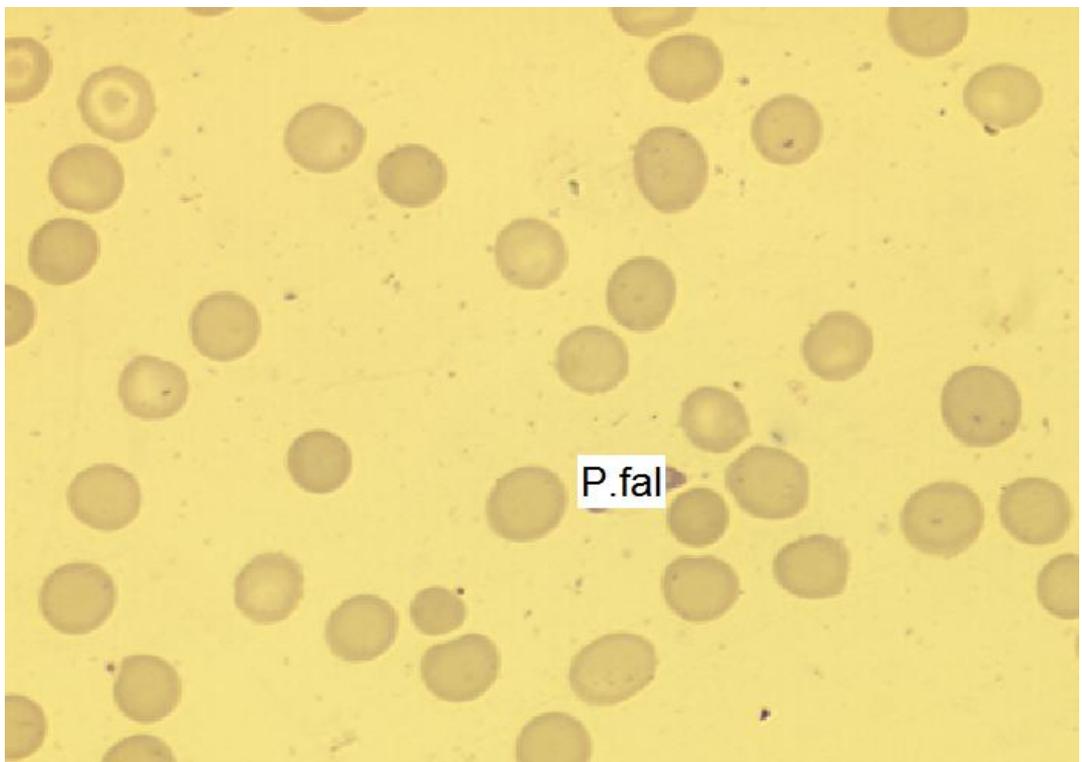
(b)

Figure 7.10: *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image



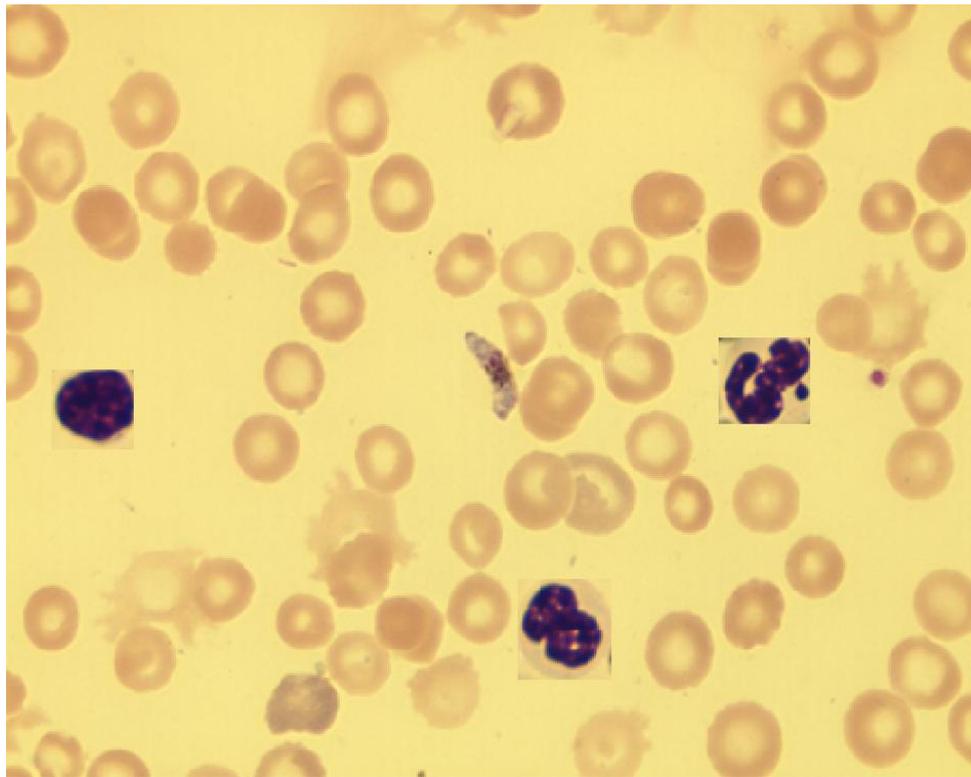
(a)



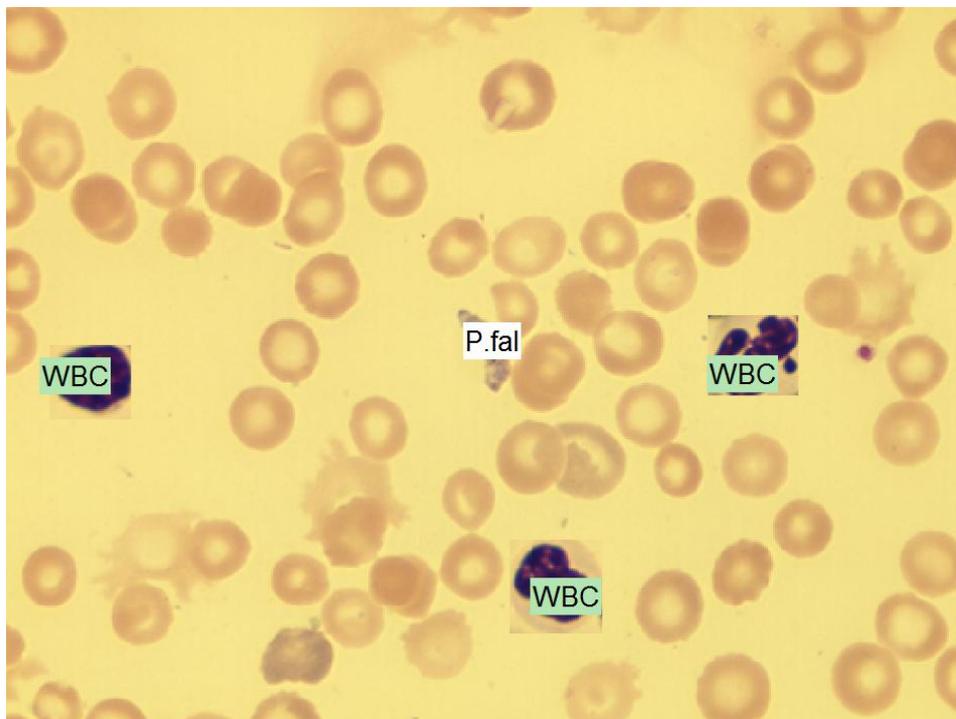
(b)

Figure 7.11: *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image



(a)



(b)

Figure 7.12: WBC and *P. falciparum* gametocyte differentiation algorithm.

(a) Original image. (b) Processed image.

7.1.3 Summary and discussion

An application specific segmentation technique to locate the *P.falciparum* gametocytes for post-treatment malaria diagnosis was explained. The method does not undergo any pre-processing procedures such as non-uniform illumination correction or noise filtering and does not involve any morphological filtering compared with the procedure devised for detecting the earlier stages of the parasite as explained in Chapters 5 and 6. Hence the complexity has been reduced to a great extent. Moreover the method is faster than other segmentation techniques which involve recurring morphological operations as well as marker extraction and rule-based classification tools. The output of the peak detection algorithm is not an image but a list of coordinates, each one identifying the location of the gametocyte and WBCs if present. This makes this approach fundamentally different to other segmentation algorithms.

The method also detects the location of the WBCs in the image and so can be used as an efficient WBC detection tool for various blood cell analyses. The processed image in Figure 7.12 demonstrates the robustness of the system in detecting the different types of WBCs. This work has been published in [79] with narrower context. The future study will look into a derivative of this method to be applicable to other species.

The following Section describes another successful application of ARR method that can be used to diagnose malaria in fluorescent images.

Automated Malaria Diagnosis Using Mobile Phones

7.2 Automated Malaria Diagnosis in Fluorescent Images

Even though the microscopic analysis of Giemsa stained blood films are the simple and faster means of malaria diagnosis, pathologists and scientists are looking for better means of evaluation. For example, fluorescent staining provides much better contrast and brightness to the parasitic cells compared to Giemsa staining. Also the fluoroscopic dyes are cheaper, visual analysis is enhanced and hence the diagnosis is faster. Fluorescent imaging is now widely used for microscopic analysis as well as rapid diagnostic tests [80], [81], [82]. In this Section, an extension of the research conducted on automated malaria diagnosis on fluorescent images as part of the research has been described.

Fluorescent staining in malaria infected blood enhances the visual perception of the parasitized cell. Apart from microscopic analysis, fluoroscopy is also used for other diagnostic procedures such as Rapid Diagnostic Tests (RDTs) and flow cytometric analysis [83], [84], [86]. The fluorescent dyes used are nucleic acid specific and binds the nucleic acid of the plasmodium species. In [81], Rebecca et.al. demonstrates the impact of different nucleic acid specific dyes on parasitic RBCs. They described a dual staining technique in which 22 nucleic-acid specific fluorescent dyes were tested on malaria-infected blood following an initial Giemsa staining. The advantage of dual staining is the double screening of the slides for species identification. The results of both bright-field and fluorescent microscopy can be compared to get more accurate results. The slides stained by only fluorescent dyes are not effective for species identification. The prior staining of Giemsa enables the comparative study and performance analysis of both bright field and fluorescent microscopy on the same slide. The experiment concluded that slides (dual) stained by SYBR Green 1 enhances the appearance of the parasites when the smear is excited by blue light.

An example of a fluorescent image stained with SYBR Green 1 is given in Figure 7.13. The highlighted components are the nuclei of the parasites. Based on the size and highlighted components, it can be concluded that the image contains a schizont life stage and few trophozoites.

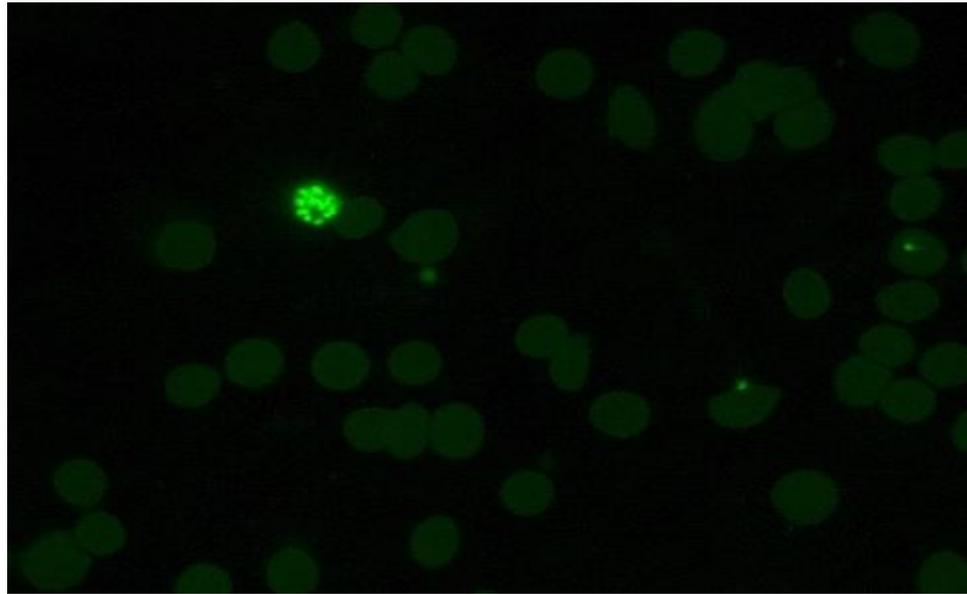


Figure 7.13: Microscopic image of an infected peripheral blood smear stained with SYBR Green 1 nucleic specific dye.

The existing studies conducted on image analysis of fluorescent images perform usual segmentation procedure to detect the stained pixels. In [87], Moon et.al. describes an inductive strategy of detecting the highlighted components followed by estimating the RBCs. The parasitic signals are detected using the local maxima estimation method. The algorithm also utilise information from the local mean and variance of the image. An edge detector is used to locate the boundaries of the individual infected cells. In [35], Vink et.al. describes a vision-based malaria diagnosis on fluorescent image using trained rule-based classifiers. In this study, an image analysis algorithm using modified ARR method has been used to detect the infected cells.

The method employs the ARR method for the detection of blood cells. In normal Giemsa stained images, the RBC appearance is circular with a light centre patch as shown in Figure 7.14(a). Hence morphological filtering is used with a concentric ring structuring element to remove the lighter centre patch within the RBC. However, fluorescent imaging gives a uniform solid pixel intensity distribution which simplifies the pre-processing by avoiding this morphological operation. This increases the speed of operation of the system compared to Giemsa stained image processing.

Automated Malaria Diagnosis Using Mobile Phones

For the fluorescent analysis, the ARR transform method is applied onto the complement of the given fluorescent image. Figure 7.15(a) shows the complement of a fluorescent image. In Giemsa-stained image processing, the ARR transform method is normally applied to the closed gray-scale image, since the method works well with darker objects against a lighter background. In fluorescent-image analysis, however, the negative of the image is used for ARR transform method processing because of its striking similarity with a closed gray level image as demonstrated in Figure 7.15 (a) and (b). Hence by applying the ARR transform method directly to the complement of the original image, the method eliminates the necessity of a morphological closing which would have otherwise needed for bright-light microscopic images. Figures 7.16, 7.17 and 7.18 show the results obtained using this algorithm.

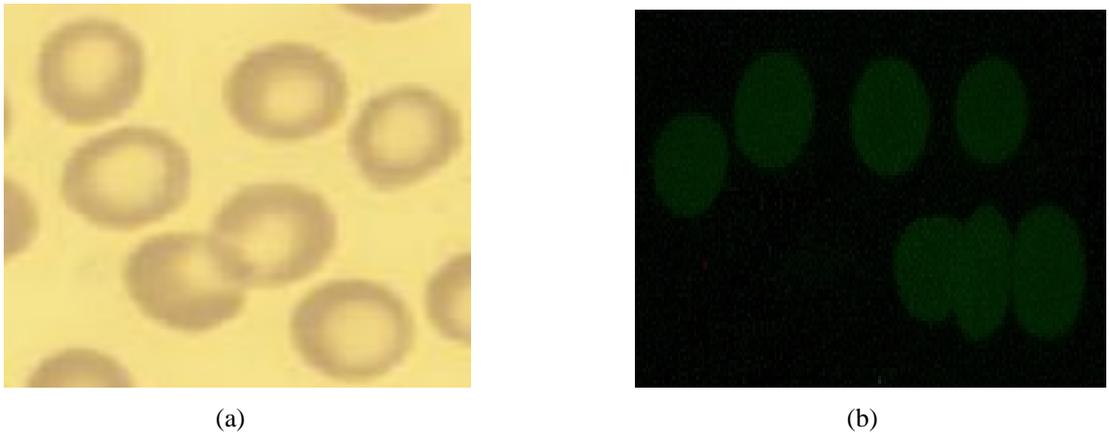


Figure 7.14: RBC morphology in Giemsa stained (a) and Fluorescent images (b).

The RBCs appears with lighter centres in (a) whereas the intensities are distributed uniformly within the RBCs of fluorescent images in (b).

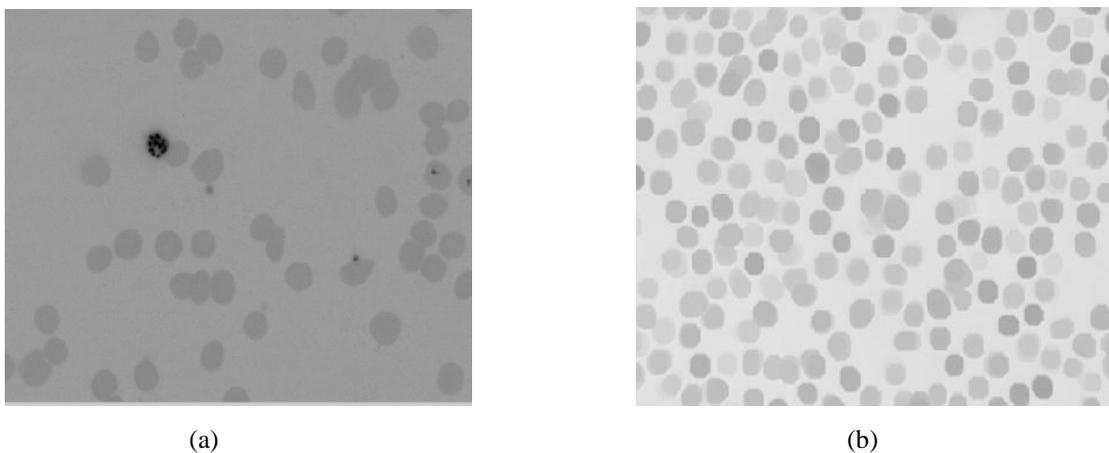
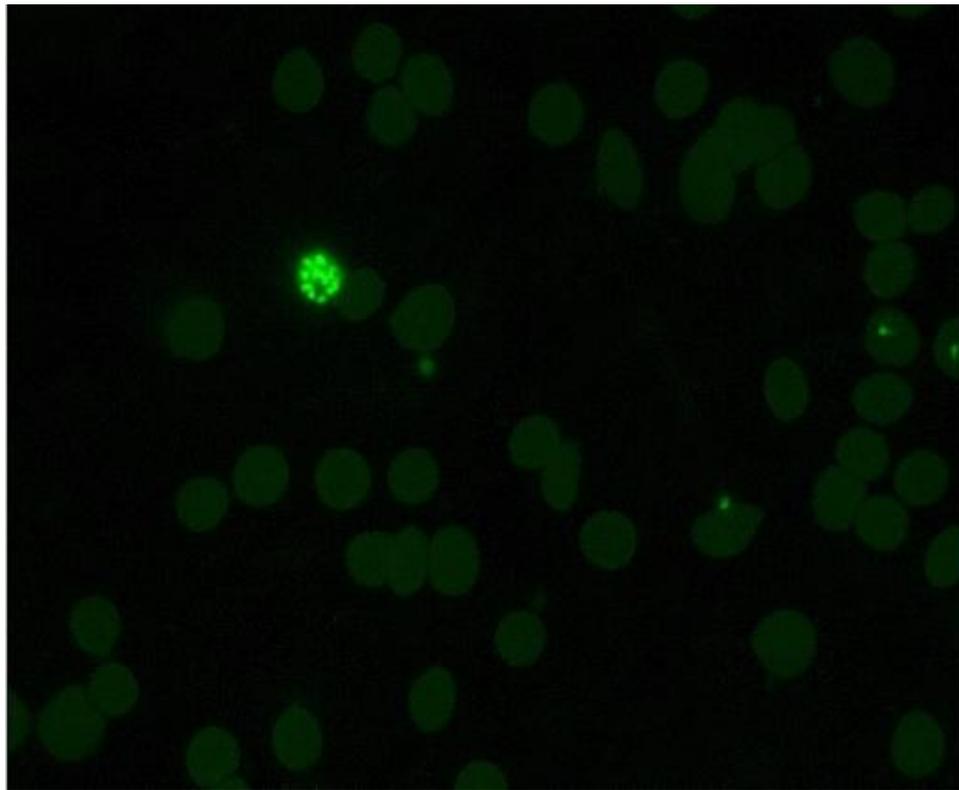
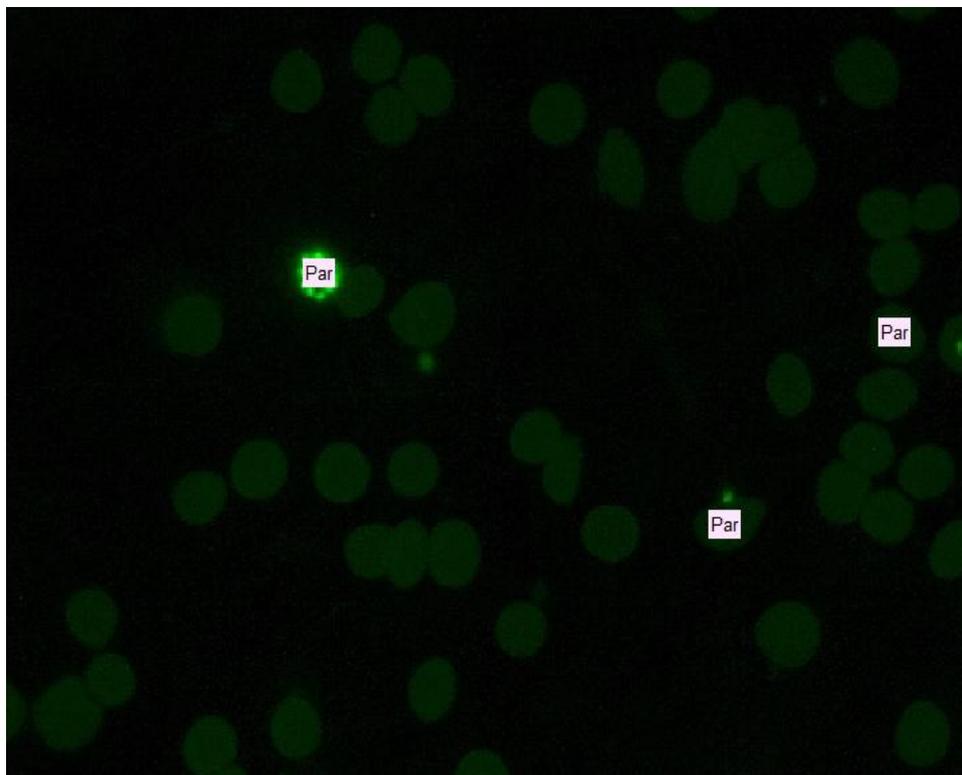


Figure 7.15: Demonstration of the similarities in the appearance of the cells in the complement of a fluorescent image in (a) and a morphologically closed image in (b).



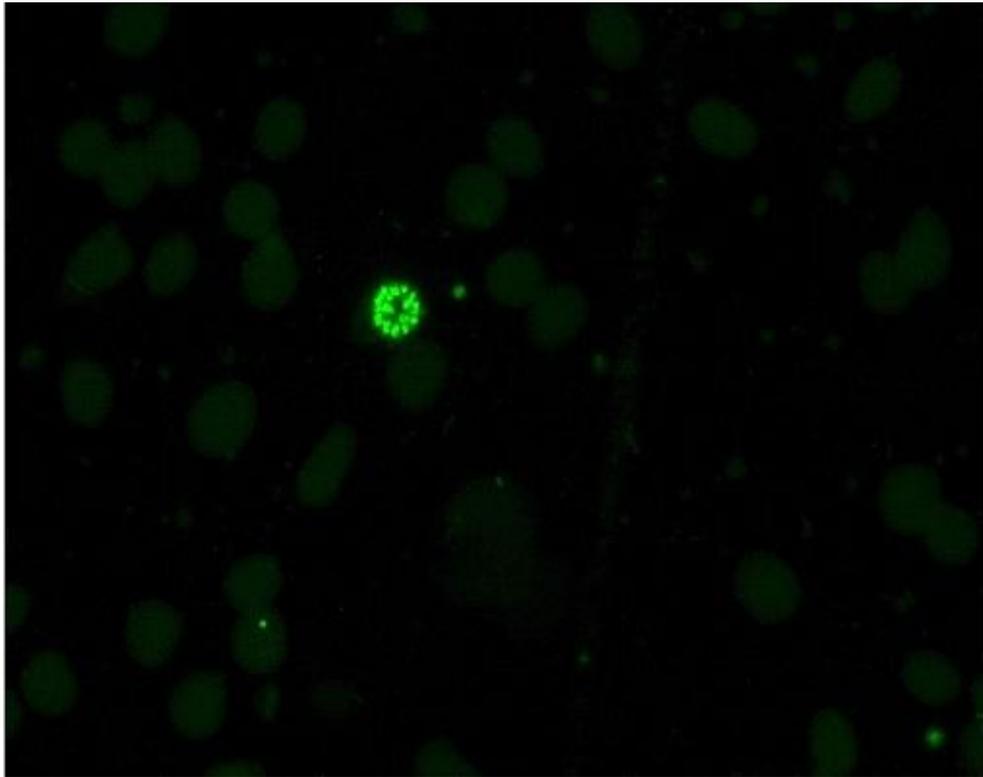
(a)



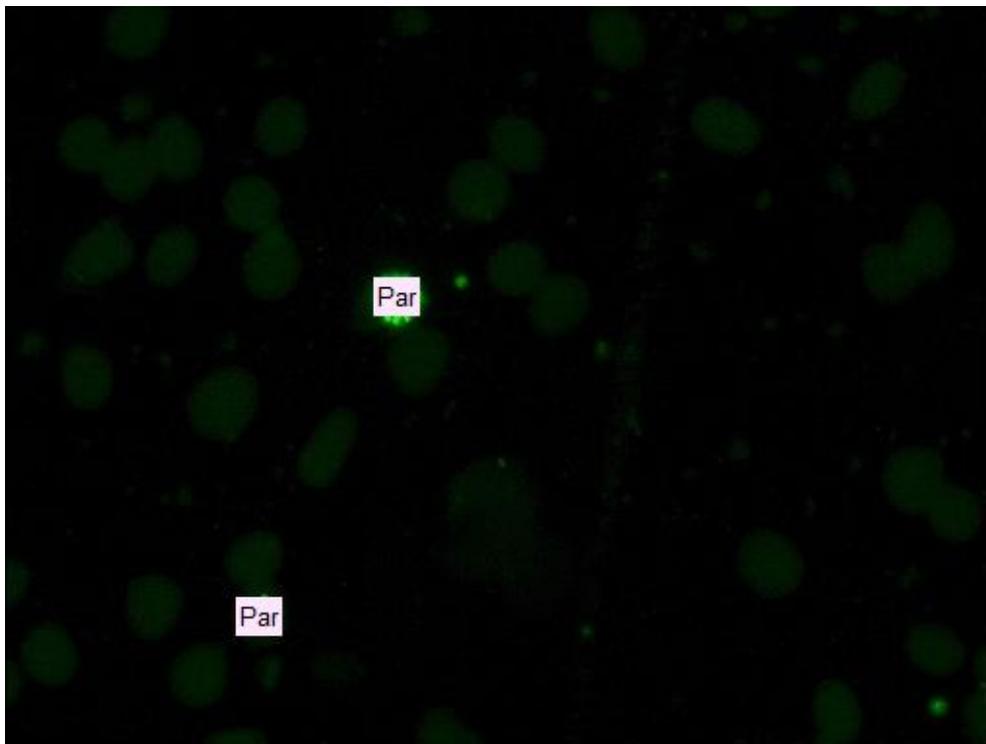
(b)

Figure 7.16: Fluorescent image processing.

(a) Original image. (b) Processed images with parasites detected and labelled.



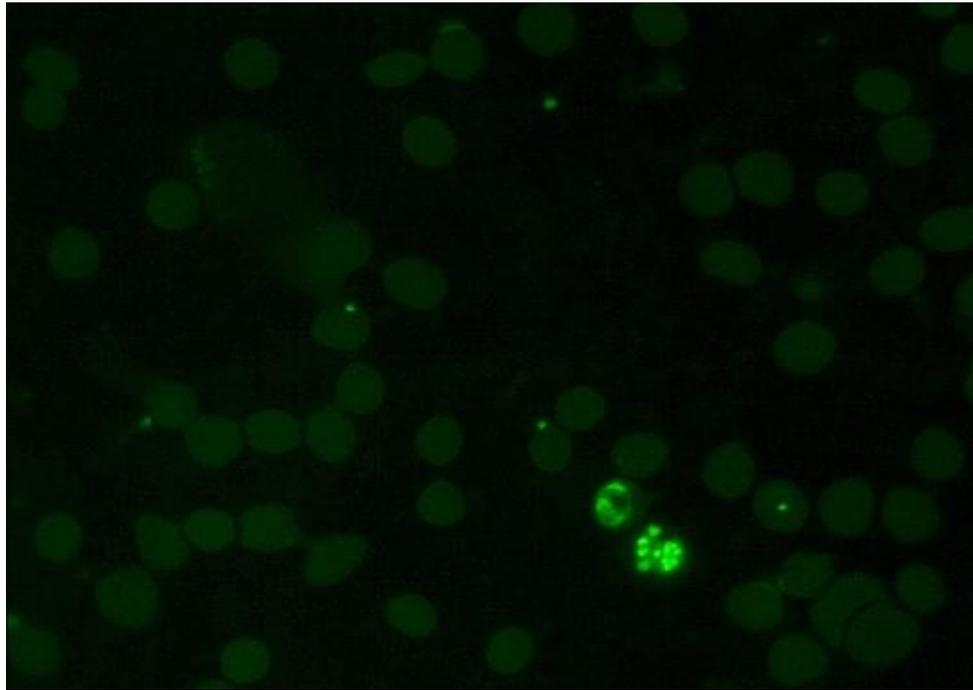
(a)



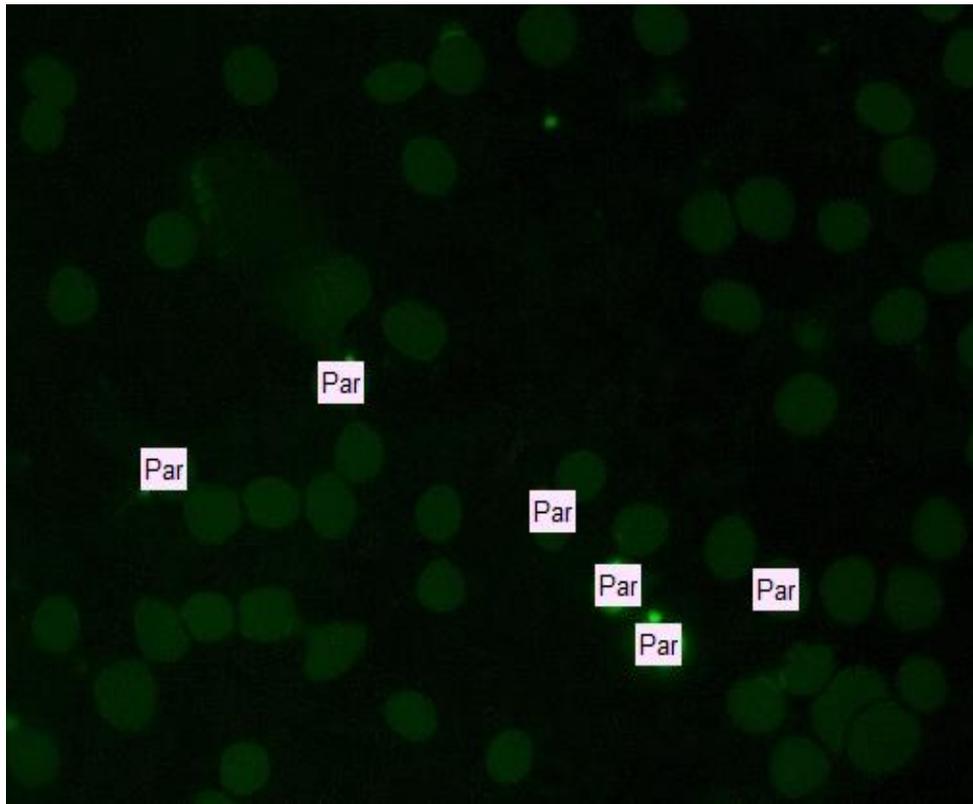
(b)

Figure 7.17: Fluorescent image processing.

(a) Original image. (b) Processed images with parasites detected and labelled.



(a)



(b)

Figure 7.18: Fluorescent image processing.

(a) Original image. (b) Processed images with parasites detected and labelled.

7.2.1 Summary and Discussion

The Section 7.2 described a novel image processing tool to detect malaria parasite in fluorescent images. Nucleic-acid-specific fluorescent markers were used to highlight the nucleated particles in the image. The method provides promising results with a given set of images. However along with the parasite detection, this process will also detect the WBCs if present as the extent of brightness is same for both objects. This is one of the challenges in the fluorescent image processing since the nucleic acid specific dyes will highlight both the WBCs as well as the parasite nuclei. However the WBCs and infected cells can be differentiated using the mean and standard deviation method as explained in Section 7.1 with the size and intensity of the cells as the selection attributes.

The process is less complex; however, the selection process of fluorescent markers for appropriate imaging has yet to be investigated. The proposed method was presented and published in the Pathology Visions Conference, 2012, conducted by the Digital Pathology Association.

The fluorescent images were supplied and the results were verified by the Centre for Medical Electronics, Anna University, Chennai, India. Due to the limited number of images supplied, no performance validation was carried out. In addition, the slow advancements in thin blood film processing and image acquisition equipments for fluorescent imaging of malaria parasite lead to the scarcity of images. An extended literature survey has to be carried out in order to conduct more experiments in this field.

7.3 Conclusion

In this chapter two studies conducted during the research were discussed that are extended application of the ARR transform method. In Section 7.1 a post treatment malaria diagnostic tool for *P.falciparum* detection was explained. A fluorescent image analysis algorithm for malaria diagnosis was described in Section 7.2. Both methods use a modified form of the ARR algorithm described in Chapter 4. The morphological filtering and closing operations that were used for the parasite detection in Giemsa-stained blood image is not required for both the applications described in this chapter. In the gametocyte detection algorithm the ARR method was performed directly with the gray scale image and

Automated Malaria Diagnosis Using Mobile Phones

in fluorescent analysis, the ARR method is performed directly to the negative of the image. By eliminating the morphological filtering operations prior to the ARR transform method, the method is faster and less complex than existing methods. Also both methods do not require any image pre-processing procedures for noise filtering, illumination correction or contrast enhancement of the stained pixels. In addition, unlike other segmentation methods the proposed method stores the located coordinates of the components. This enables the estimation of the number of components in the image.

The future work on the area will be concentrating on conducting further extensive studies of the application for species identification. Also a comparative study on bright field and fluorescent microscopy for malaria screening is also under process. A trial using different fluorescent stains and their impact identification of the blood components will be undertaken in the future.

Chapter 8

Malaria Parasite Detection Using Mobile Phones

The ultimate aim of this research is to develop a mobile phone camera enabled malaria diagnostic set-up which has enormous potential in the qualitative and quantitative analysis of malaria in the field of digital pathology. Foreseeing the necessity of a flexible and computationally less complex architecture to be integrated on to a mobile phone, this research explored the possibilities of developing an effective malaria diagnostic tool with image processing algorithms simpler than the existing ones. As demonstrated in the previous chapters, the research succeeded in developing effective image-processing software which is fundamentally different from the existing ones. These algorithms were first tried in Matlab before developing as a mobile application tool. This chapter describes the implementation of a mobile phone application for automated malaria diagnosis by processing the microscopic image of a thin blood smear using ARR transform method. This work was carried out as a collaborative venture as part of the student exchange programme between University of Westminster and École Nationale Supérieure d'Electronique, Informatique, Télécommunications, Mathématique et Mécanique de Bordeaux (ENSEIRB-MATMECA).

8.1 Literature Review

The recent advances in smart phone technology and its wide use even in remote areas of the world have increased the potential of these devices for medical diagnosis. Several studies in the field are being carried out since the smart phone revolution took off in the early 2000. Woodward et al. [88] describe a GSM based cellular telemedicine signal processor design which transmits the patient data to the monitoring unit using mobile phones integrated to the patient body. Similar literature can also be found in [89] and [90]. However, the application of mobile phones for image data transmission were first introduced in 2005 and numerous studies on medical-image data transfer for applications such as ultrasound, neurosurgery and urology were reported in the literature [114], [115]

Automated Malaria Diagnosis Using Mobile Phones

and [116]. A cell-phone-based medical imaging concept was reported by Granot et.al. in [91] which developed the mobile phones as a part of medical diagnostics to collect and display data. In 2007, Freaan [98] reported his study that was conducted on transmission of microscopic images using mobile phones by placing the mobile phone camera lens on to the eyepiece of a microscope. Similar studies were also conducted with fluorescent microscopy which is one of the popular subjects of studies in digital pathology [92], [93]. Advancement in the field of mobile-based microscopy was initiated by University of California (UCLA), Berkley with the introduction of CellScope which was reported in numerous literatures [94], [95]. Similar studies related to health care applications of mobile phones can also be found in [96] and [97].

The work cited above describes the application of mobile phones in medical diagnostics, mainly for data collection and transmission and remote sensing. For medical imaging, the cell phones are used as image acquisition units which capture the data leaving it to the experts to analyse. The study conducted in this research, however, has developed an application tool for malaria diagnosis which not only captures the image but also analyses them to provide a diagnosis.

8.2 Mobile Phone Camera Enabled Malaria Diagnosis

With the advancements in mobile technology and broadband communications, smart phones plays a momentous role in facilitating access to some diagnostic services even to people living in the most remote areas. Due to the scarcity of clinical workstations in malaria endemic countries such as India and Sub-Saharan Africa, there has been a wide interest in the field of diagnostic pathology to incorporate modern telecommunication technologies in telemedicine [96]. The remote sensing of malaria by accessing microscopic images via mobile phone is an upcoming pathological diagnostic procedure in the modern world. With the advancement of the processor power and memory storage of mobile phones, a complete diagnostic platform using a cell phone which performs image acquisition, processing and analysis is possible which enables a fast and cost effective diagnosis without human intervention and laboratory procedures. One such work has been conducted as part of the research to implement the malarial diagnosis algorithm described

Automated Malaria Diagnosis Using Mobile Phones

in Chapters 3-6, which has been tested and evaluated on a MATLAB platform, into an Android mobile phone.

The work aimed to implement the algorithm onto the mobile platform with minimal memory footprint, loss of information integrity and acceptable processing speed and duration when processing within the mobile phone, however a complete hardware implementation was beyond the scope of the research and an Android based mobile application tool was developed. Figure 8.1 illustrates the objective of the research. The mobile phone acts as a complete diagnostic tool with an added advantage of storing the results and images in remote locations such as PC and/or clinical server for future referencing.

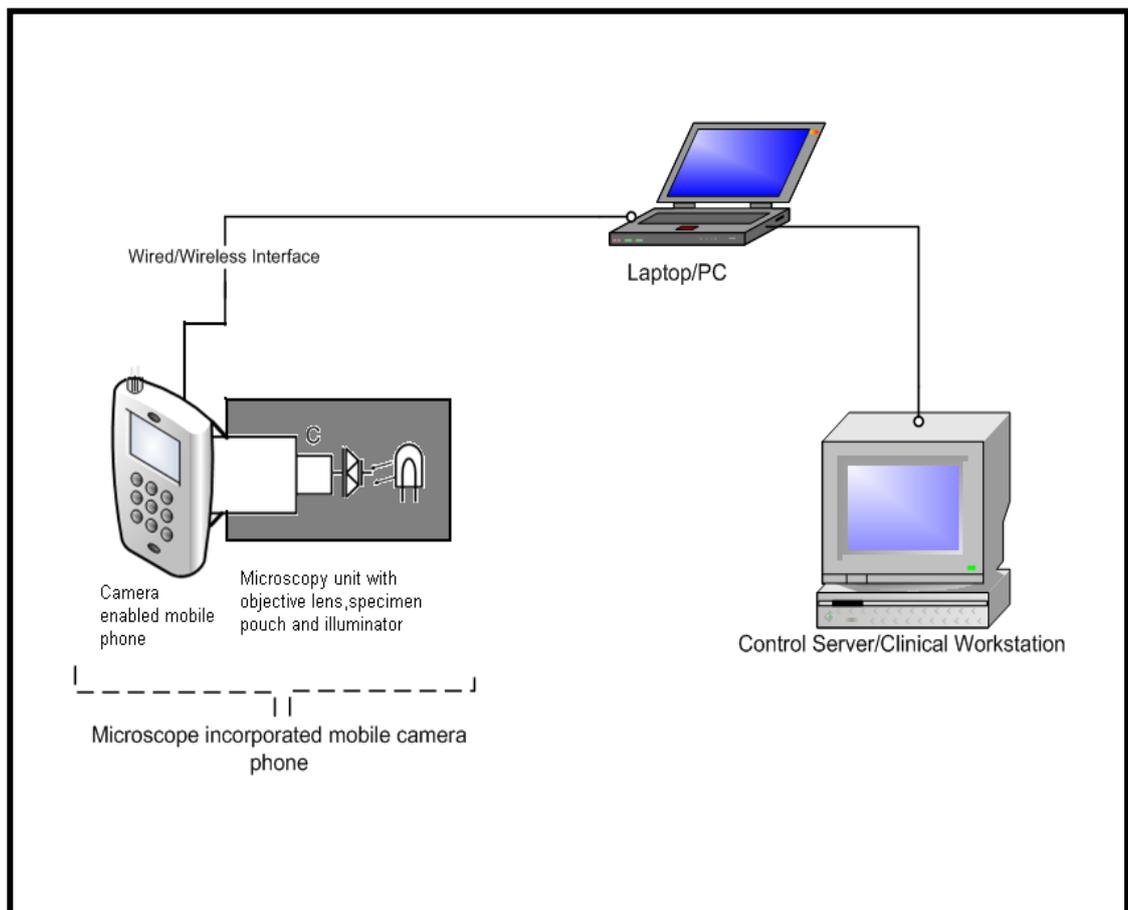


Figure 8.1: Schematic of the objective of Malaria diagnosis using mobile phones

Automated Malaria Diagnosis Using Mobile Phones

8.2.1 System Overview

The image processing algorithm for malaria diagnosis uses (a) the ARR transform algorithm to segment the blood components, (b) WBC differentiation algorithm using size of the cells and mean intensity values to distinguish WBCs and RBCs and (c) Hue-ARR algorithm for infected cells identification and life-stage recognition. In order to implement these algorithms into the mobile phone, a mobile platform (Operating System-OS) has to be chosen. The three popular mobile application platforms are Windows 7, Apple iOS and Android. For this study, the Android platform was selected due to its flexibility to customize, its use of open-source Linux kernel and the availability of a virtual simulator for mobile software design [99]. In addition, Android-based smart phones are popular in malaria-endemic countries and are comparatively cheaper than the others which makes the application affordable compared to existing laboratory diagnostic procedures.

The implementation of the application within the Android Operating System (AOS) needs specialist development software for Android and the Java development tool. Eclipse is an Integrated Development Environment (IDE) which contains the tools and frameworks for designing, installing and managing the software [100]. Android applications are developed in a Java language environment and can be open sourced using Java programming. Hence along with Eclipse, the essential tools required to develop and deploy the application are the Java Development Kit (JDK) which is a bundle of software for Java development and Java Runtime Environment (JRE) to run the application in a virtual environment [101]. In addition, an Android Software Development Kit (SDK) with relevant Application Program Interfaces (API) has to be added to the Eclipse IDE with Android Development Tools (ADT) to streamline the development [102]. An Android mobile device with a minimum requirement of Android Version 2.2 is also essential.

Like the Matlab software, the application developed in Android also uses the radius of the RBCs in pixels as the scaling parameter. In order to avoid zero padding at the edges of the image, which affect the results while performing morphological filtering, dynamic masking is used in which the neighbourhood used for filtering at the edges will contain only the pixels carrying the image information. Like the Matlab version, the mobile application was also developed using floating-point arithmetic.

Automated Malaria Diagnosis Using Mobile Phones

The application was tested on an Android emulator and has been successfully run on a Samsung Galaxy Mini, Samsung S2, Samsung S3, Samsung Galaxy Note and Samsung Galaxy Tablet. Figures 8.2 -8.10 demonstrate the working of the application. Screen shots of the application and the result obtained are shown in Figures 8.2, 8.3 and 8.4 respectively. A demonstration performed using a Samsung Galaxy S3 Mini is shown in Figures 8.5 and 8.6.



Figure 8.2: Screenshot of Samsung mobile phone with the application for malaria diagnosis. The black arrow points to the icon for the malaria application.

Automated Malaria Diagnosis Using Mobile Phones

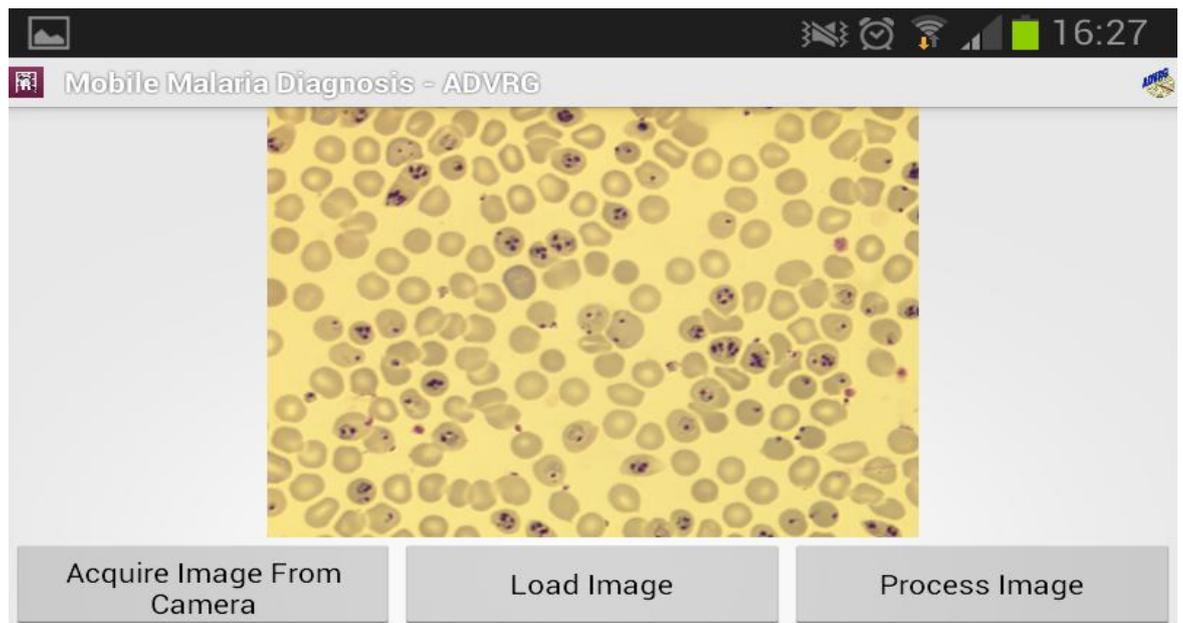


Figure 8.3: Screenshot of the malaria diagnosis application installed on to a Samsung mobile phone.

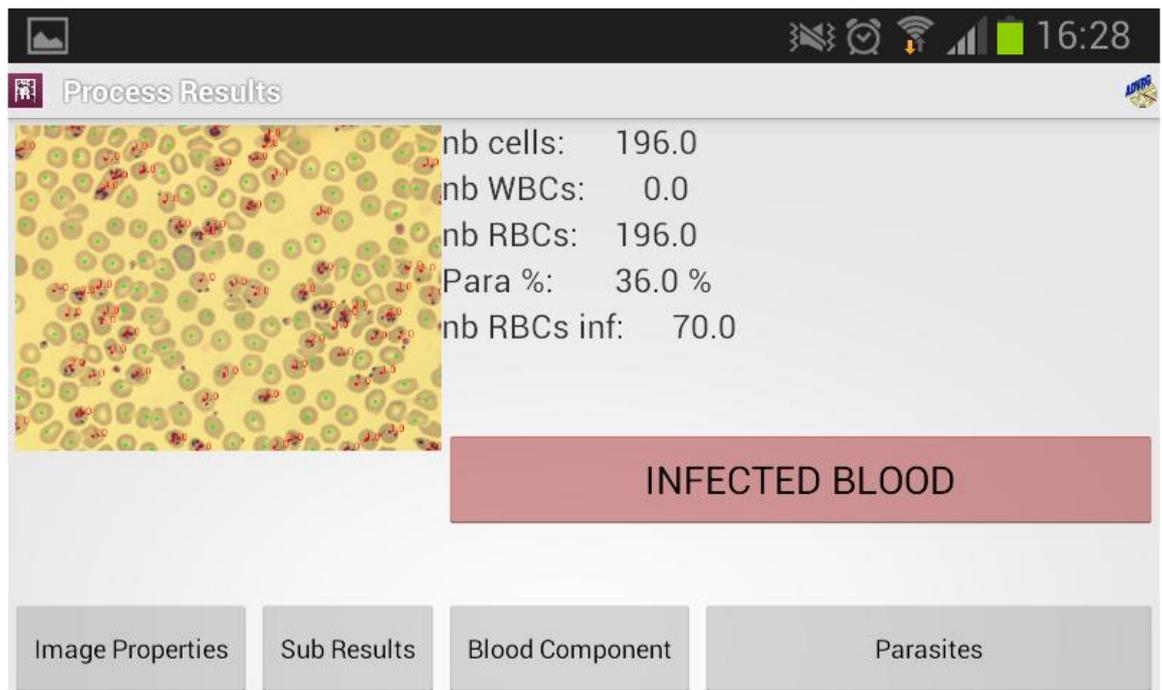


Figure 8.4: Screenshot of the malaria diagnosis application installed on to a Samsung mobile phone. The figure shows the diagnosis after processing the image.

Automated Malaria Diagnosis Using Mobile Phones

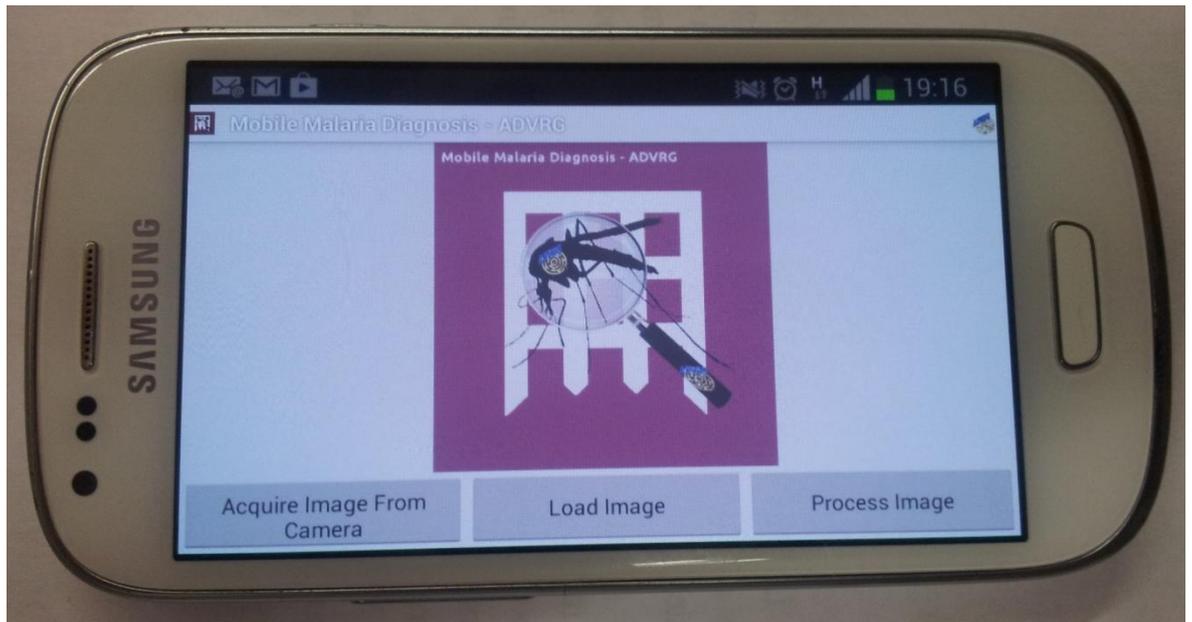


Figure 8.5: Demonstration of the mobile phone malaria diagnosis application installed on to a Samsung mobile phone.

The figure shows the initial stage of loading the image.

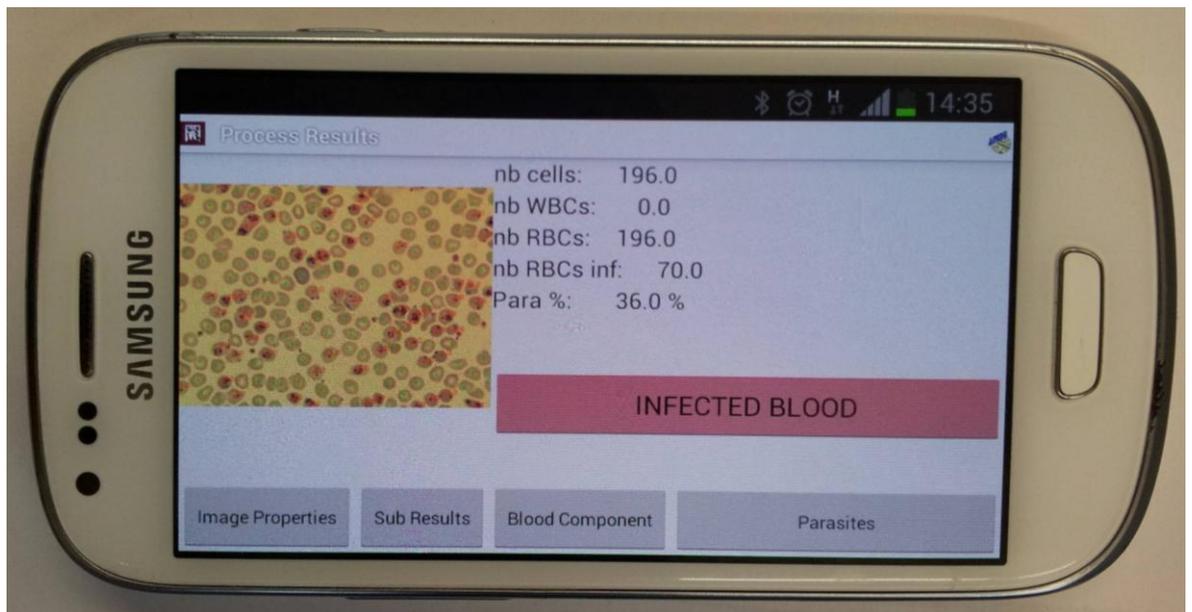


Figure 8.6: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung mobile phone.

The figure shows the diagnosis after processing the image.

Figures 8.7, 8.8, 8.9, 8.10, 8.11 and 8.12 show different stages of the results provided by the application when installed and processed on a Samsung Galaxy Tablet.

Automated Malaria Diagnosis Using Mobile Phones

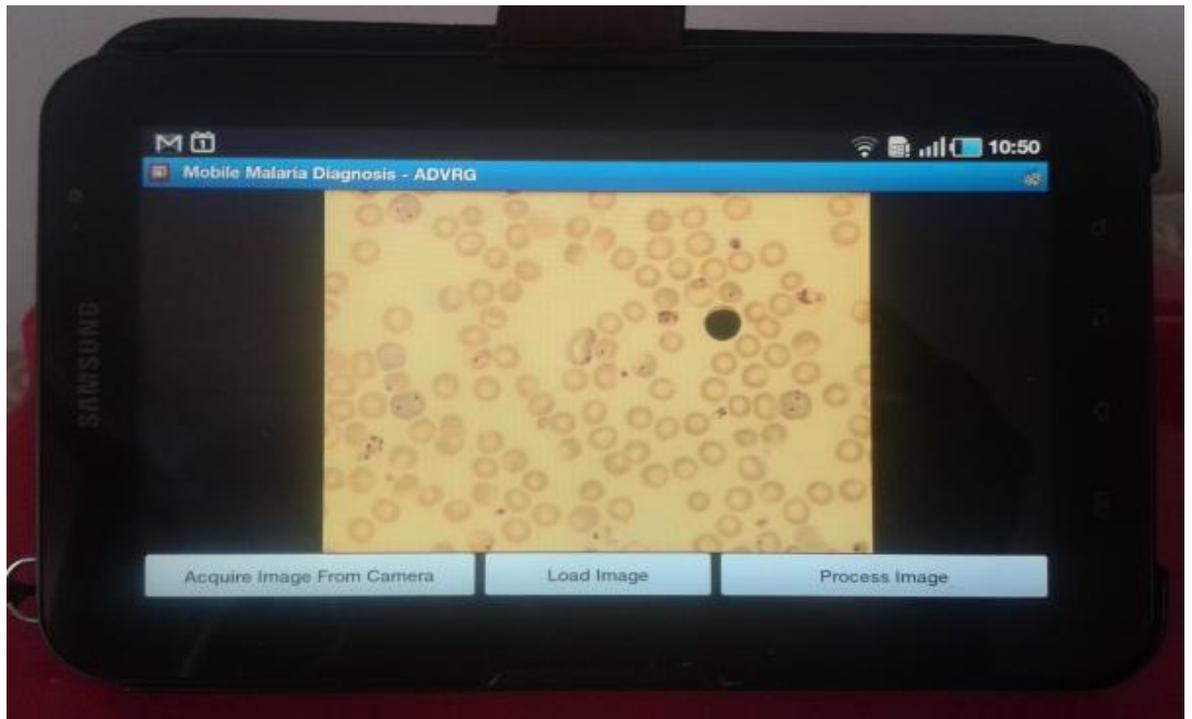


Figure 8.7: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.

The figure shows the initial stage prior to processing the image.

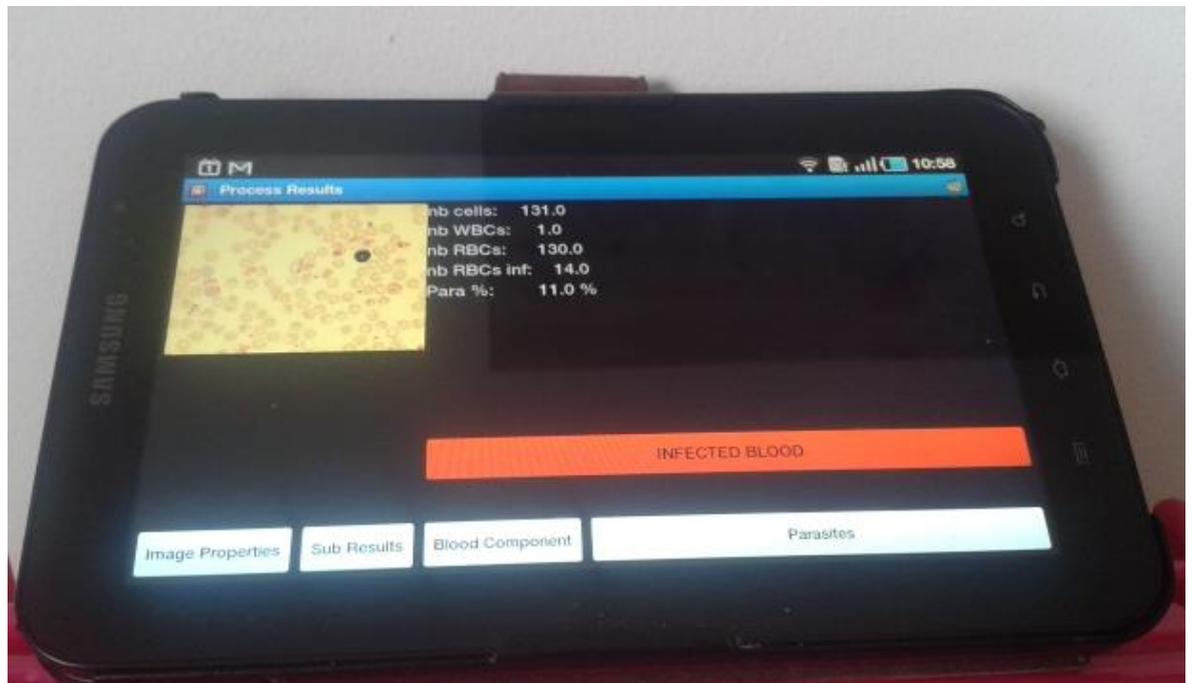


Figure 8.8: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.

The figure shows the diagnosis after the processing of the image.

Automated Malaria Diagnosis Using Mobile Phones

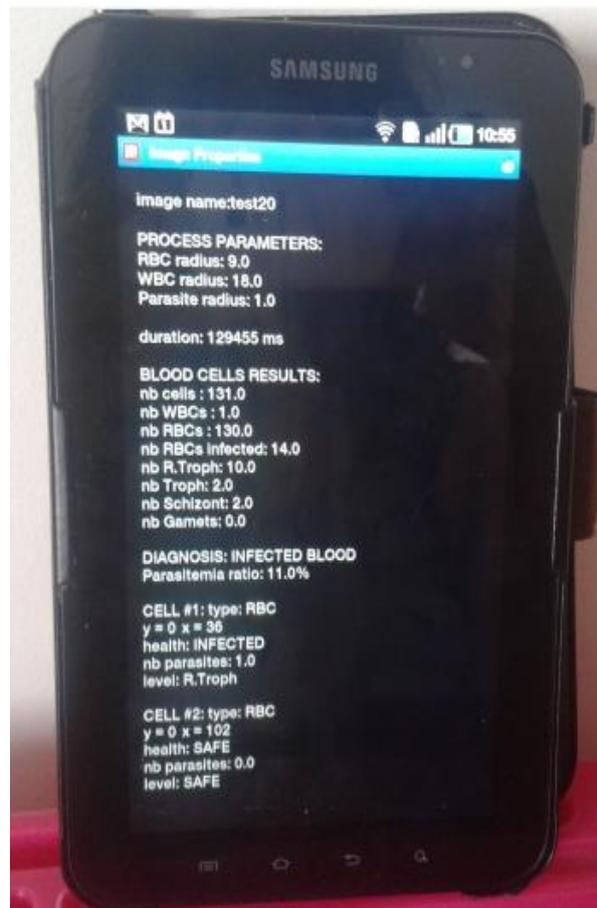


Figure 8.9: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.

The figure shows the text file generated after processing the image containing information on the population of RBCs, WBCs and infected cells and the centre pixel coordinate value of each components identified.

Automated Malaria Diagnosis Using Mobile Phones

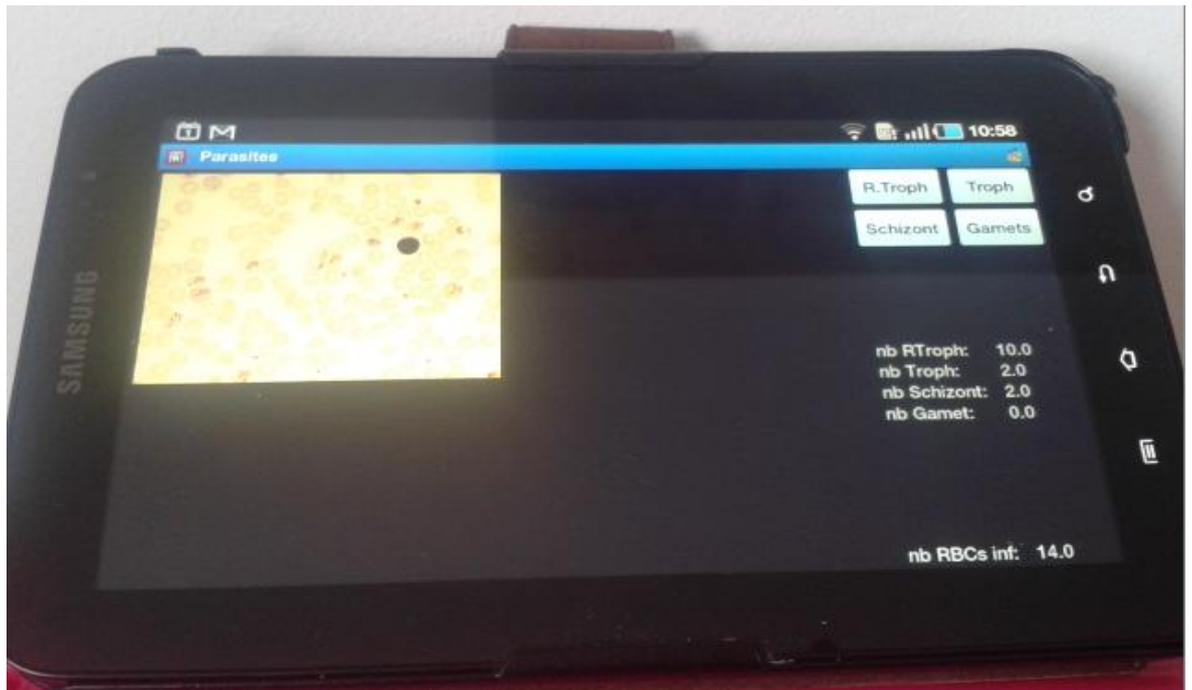


Figure 8.10: Demonstration of the mobile phone malaria diagnosis application installed on a Samsung Galaxy Tablet.

The figure shows the display option on the life stages present in the image. It also displays the total number of infected cells belonged to each life stage.

The application displays the results of the analysis along with statistical information of the total number of RBCs, WBCs, infected RBCs, life stages and the location of the cells which aids to the measurement of the parasitemia. Separate display option for closer analysis of detected RBCs, WBCs and infected cells are also provided along with displays of intermediate results of the morphological operations such as the morphologically closed and ratio-transformed images. The user also has the flexibility to visualise the potential life stages present in the image. For future reference, the results are stored in a folder, with the same name of the input RGB picture, in the mobile phone SD card to be available at any time.

The application was tested on different versions of Android. The following Table 8.1 explains the functioning characteristics of the application when performed on different mobile phone processors. The processing speed of the application varies with the type and performance of the mobile phone processors as well as the size of image. The application was tested using images already available in the UoW image data set. A standard size of 512X381 pixels images were used for processing. However, in real time the images will be

Automated Malaria Diagnosis Using Mobile Phones

acquired on the site using the camera of the mobile phone. Depending on the processor, the larger images cause overloading of the application. However, this problem was rectified by setting the phone camera resolution to 640x480 pixels.

Table 8.1: Mobile phone description and process duration [112]

Materials	Tested Android mobile phones		
	<i>Samsung Galaxy S3 Mini</i>	<i>Samsung Galaxy Mini S5570</i>	<i>HTC One SV</i>
Android version:	4.1 Jelly Bean	2.2	4.1.2
RAM	1 GB	384 MB	1 GB
Camera Resolution	5 Mpixels	3 Mpixels	5 Mpixels
Autofocus	Yes	Yes	Yes
Processor	STE U8420	Qualcomm MSM7227	Qualcomm MSM8230
Core	Dual Core	Dual Core	Dual Core
Processor frequency	1 GHz	600 MHz	1.2 GHz
Processor Frequency (GHz)	1.0	0.6	1.2
Process duration (s)	77.2	52.53	55.765

8.3 Image Acquisition and Processing Using Mobile Phones

The application was at first tested by processing the images that were already available in the dataset. However, to prototype a complete diagnostic set up, the application should run in real time with image acquisition and processing performed on the spot. Hence, a suitable microscope set up was explored in order to integrate the camera phone to acquire the picture. At first the experiment was conducted by placing the camera of the mobile phone over the eyepiece of a normal light microscope. However, there were focusing issues due to the mismatch of the field curvatures of the camera lens and eyepiece. In addition, the limitations of the microscopic hardware such as condenser adjustments and variable iris worsen the illumination issues. Due to the time limitations, rather than correcting the focus and illumination issues, alternatives were explored.

Automated Malaria Diagnosis Using Mobile Phones

A portable Newton Nm1 series microscope was tried to analyse the smear and capture the image. The Nm1 series is an inverted folded optics, monocular microscope with a built in white LED illumination source and a portable battery powered for up to 300 hours [103]. The microscope works under three different magnifications, 400(x10 N/A 0.25 and x40 N/A 0.65), 600(x10 N/A 0.25, x40 N/A 0.65, x60 N/A 0.80) and 1000(x10 N/A 0.25, x40 N/A 0.65, x100 oil N/A 1.25) where N/A stands for numerical aperture. The microscope has adaptive mounting for both Apple I phone as well as Android phones. The mobile phones when placed at the eyepiece of the microscope with the help of an eyepiece fitment provided promising results. The image acquisition set-up using the portable Newton microscope is shown in Figures 8.11-8.13.

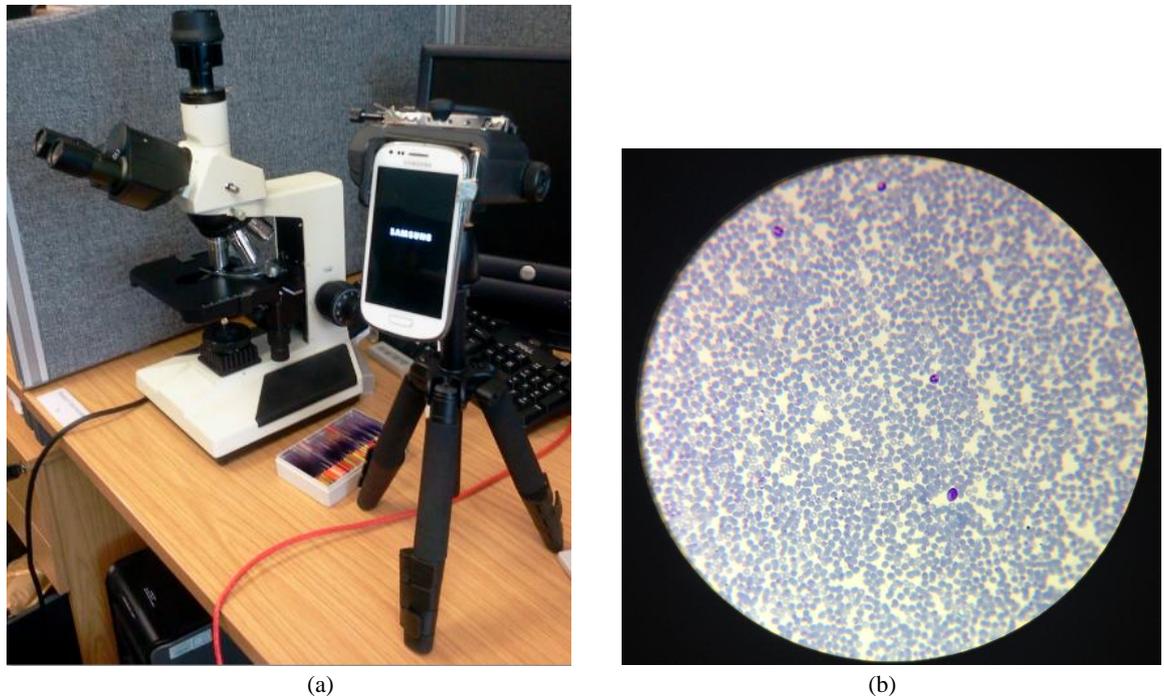


Figure 8.11: Demonstration of mobile phone microscopy.

(a) A Samsung Galaxy 3 Mini is mounted on to the eye piece of a portable Newton Nm1 microscope using the Android phone adapter. (b) The captured image with 40x magnification.

Figures 8.12 and 8.13 demonstrate the real-time field selection and capturing of the image which is then loaded directly to the application and the result of processing and diagnosis.

Automated Malaria Diagnosis Using Mobile Phones



Figure 8.12: Demonstration of mobile phone microscopy.

Upon selecting the field to capture and the application prompts the user to decide if the field is suitable for diagnosis. An good monolayer area is extremely important for accurate diagnosis.



Figure 8.13: Demonstration of mobile phone microscopy.

The result of diagnosis is displayed with all the necessary information on total number of RBCs, WBCs, infected cells, life stages and parasitemia.

Automated Malaria Diagnosis Using Mobile Phones

The processing was carried out using a 100x objective magnification with oil-immersion lenses. The memory limitations of the phone used for these experiments meant that the full-resolution images of the camera could not be processed so the resolution was reduced to 640 x 480 pixels. The user has the flexibility to choose the appropriate field to be processed and the variable zoom feature of the phone aids the location of a suitable zone prior to acquisition. The camera lens and eyepiece integrated well so that no focusing issues have been experienced so far.

8.4 Discussion and Summary

This chapter successfully demonstrated a user-friendly mobile-malaria diagnostic system which not only diagnose the infection but give a statistical data on the estimated number of RBCs, WBCs, and infected cells. The application uses morphological image processing algorithm and the intermediate results of the processing are also provided. The results obtained are saved in the SD card memory of the mobile phone which is available for later use or analysis for clinical experts. This application was successfully demonstrated on several versions (2.2 and above) and types of Android mobile phones and tablets. The application allows the user to either acquire fresh images directly from the camera or load the image from the gallery. This provision indeed accelerates the flexibility of the system to be used as both on-the -spot field trials and/ or remote clinical analysis where a set of images can be captured and stored and can later be used for qualitative analysis.

This work has explored the prospective of mobile phones in medical imaging and has paved the way to many potential studies which utilise its enormous potential. Even though the porting of the algorithm to a mobile phone is at an early stage, it has shown very promising results. The final aim is to set up a complete image acquisition system integrated to the mobile phone. However, the current real-time processing set-up using the mobile phone camera attached to the eyepiece of the portable Newton microscope has great potential for use as a complete diagnostic set up with minimal laboratory requirements and/or manual intervention. Since the application also estimates the RBC and WBC population, it has potential in the quantitative analysis of the blood images and is not limited to malaria diagnosis. The research will focus on the benefits it can provide for the

Automated Malaria Diagnosis Using Mobile Phones

successful diagnosis of malaria, and the supportive treatment and further studies in even remote environments without specialist clinical support.

This work has the potential for a major breakthrough in the field of mobile microscopy and medical diagnostics and a major contribution to global health care. It has been accepted for publication at the upcoming International Symposium for Circuits and Design, 2014 [112]. Future work will concentrate on being able to handle larger resolution images as well as to speed up the algorithm to reduce the diagnostic time. The work can be extended to other operating systems used in the mobile phones such as Apple iOS and Windows 7. A hardware realisation of the algorithm to be incorporated on to the Graphics processing unit (GPU) of the mobile phones is also under consideration.

Chapter 9

Conclusions

Morphological image processing approaches for blood image analysis and its subsequent applications in malaria diagnosis have been investigated in this thesis. Even though the research is application oriented, it also addresses the general issues related to microscopic image analysis. This chapter reviews what has been achieved, reflects on the learning experience and suggests further improvements and extensions to the research.

9.1 Learning Achievements

The research involved the development of many novel algorithmic techniques as well as the successful implementation of the core process onto a mobile phone. The problems associated with automated malaria diagnosis were extensively investigated during the course of the study. Chapter 3 described the development of a tool to determine the optimum RGB weights in the polychrome to monochrome conversion which maximised the contrast in the resultant gray-scale image. It was discovered that the blue component provided very little information and so the optimum lied along the red-green axis but the exact ratio of red and green weights was not critical. Chapter 3 also described a novel morphological filtering approach to remove artefacts, platelets and other noise in the image. The method uses a concentric ring structuring element for dilation which is then followed by morphological erosion using a disk shaped structuring element. This eliminated the problem of the light centre patch of the RBC attenuating the overall RBC intensity distribution when more conventional disc-based dilation is performed.

Chapter 4 described a novel approach to the identification of blood cells. Key to this is the Annular Ring Ratio (ARR) image transform. The ARR method involves converting a gray-scale image into a ratio-flattened image and a peak detection algorithm to detect the bright peaks in the ratio-transformed image. The ratio-transform method works with a ratio of intensities in the immediate vicinity of the cell and, consequently, is independent; not only of the average level of illumination but also any variation of illumination or film

Automated Malaria Diagnosis Using Mobile Phones

thickness across the image. It, therefore, avoids the necessity of pre-processing, colour normalisation and calibration techniques. The only calibration required for different imaging equipment is the determination of the RBC radius in terms of pixels as all the other constraints are scaled to this parameter. The output of the peak detection algorithm is not an image but a list of coordinates, each one identifying the location of a blood cell. This makes this approach significantly different to other existing and conventional segmentation algorithms. Another advantage of the ARR transform method is that it facilitates the counting of the number of RBCs in the sample in order to calculate the parasitemia. Unlike per-image estimation, the method can analyse as many image fields on a slide, counting the cells of clearly independent and evenly distributed areas until it reach a count of 1000 RBCs, to get the percentage. This is an advantage where the slide has been poorly prepared as would be the case in the field where there is a shortage of skilled laboratory technicians.

Chapter 5 explored the possibilities of various object extraction techniques to distinguish the RBCs from the WBCs that are located by the ARR transform method. An effective RBC and WBC differentiation tool, based on the mean intensity and area of the components, was developed. Instead of processing the gray-scale image, the mean intensity of closed (morphologically dilated and closed using different structuring elements) image is used for processing. The processing of only the pixels in the vicinity of the cell speeds up the operation compared to other methods which use the localised mean intensity of all the pixels in the image.

The detection of infected cells is a key stage in the diagnostic procedure. Chapter 6 investigated different stained pixel extraction tools for parasite identification and pattern recognition techniques for life-stage recognition. Foreseeing the necessity for computationally simple methods to be incorporated into a mobile phone, the research focused on possible alternatives to pattern recognition and stained pixel classification methods. The chapter proposes a novel Hue-ARR algorithm to detect the infected cells which also estimate the total number of nuclei within each cell. This, in fact, aids the identification of potential life stages within the cells by counting the total number of nuclei. The chapter also describes the performance measurement experimented on a set of objects in an image data set plus the probabilities associated with the identification of a

Automated Malaria Diagnosis Using Mobile Phones

single cell as being parasitized and its extensions to a number of cells from a single patient as well as a number of patients.

Chapter 7 discussed two parallel investigations that were carried during the course of the research. One is an effective *P.falciparum* gametocyte detection tool which also differentiates the WBCs present in the image. The method has great potential in post-treatment malaria diagnosis where the presence of gametocytes is often found due to their resistance towards anti-malarial drugs. The chapter also introduced an effective fluorescent image analysis for malaria diagnosis. Unlike other fluorescent imaging techniques, the method addresses the issue of false identification of WBCs, which are also highlighted by the nucleic acid specific fluorescent dyes, and propose a possible solution based on the mean intensity, variance and area of the cell to differentiate them. However due to lack of fluorescent images carrying WBCs, a thorough investigation remains for future work. These two studies contribute to the applications of ARR transform method which has the scope for several future works in the field of blood image analysis.

Chapter 8 demonstrated a mobile application tool developed on an Android platform that has been tested in various portable devices. The application contains the image processing algorithms developed in this thesis, including morphological dilation and erosion, the ARR transform method, the WBC and RBC differentiation algorithm and the Hue-ARR algorithm for infected cell identification. A real-time image acquisition and processing scenario has been demonstrated using a Samsung Galaxy S3 Mini which was mounted on the eyepiece of a portable, monocular Newton Nm1 microscope [103]. The experiments showed enormous potential for development of a pathological set up for not only malaria diagnosis but also to differentiate and estimate other blood components.

9.2 Learning Experience

This research is of interdisciplinary nature and is application oriented. It is equally important to achieve and contribute the outcomes to both healthcare and image-processing fields. Hence the research simultaneously focused on two paradigms- (a) Medical aspects of malaria diagnosis (b) Image analysis and processing. As explained in Chapters 1 and 2, the medical aspects of malaria diagnosis involve microscopic examination of the blood smear to detect the parasites in the RBC. To familiarise with this process, the author attended a course on ‘Updates in blood and tissue parasitology’ which taught the laboratory procedures required for malaria diagnosis, starting from smear preparation through to parasitemia estimation. Attending this course was a major advantage since it enabled to achieve a better understanding on smear analysis and aided in the decision making throughout the research.

Image analysis for malaria diagnosis is an essential part of the automated malaria diagnosis. A machine based learning approach is adapted in the research which mimics the manual microscopic analysis. Hence the primary investigation was based on the nature of the images used. Unlike those images taken under natural lighting conditions and acquisition settings, the input image to be processed for malaria diagnosis is a microscopic image and its quality mainly depends on the slide preparation (and thereby film thickness) and variations in the lighting and staining conditions. The images used for the study were obtained from different sources and hence were highly diverse. This in fact raised serious issues while considering different image processing algorithms to process them.

Another challenging part was to deal with the variety and randomness of the components present in the image. Other than the RBCs, the presence of all the other components including WBCs, platelets, parasitic nuclei and other artifacts are unpredictable even though can be foreseen. As a result, trained algorithms used to be the common choice in conventional blood image analysis algorithms as in [21], [62]. Furthermore, the components have varying size and staining concentration and hence extracting them requires highly sophisticated procedures. Existing image processing algorithms dealt with the above problems separately. The illumination and intensity variations were eliminated using illumination correction and contrast enhancement

Automated Malaria Diagnosis Using Mobile Phones

techniques respectively and the object extraction was performed using conventional segmentation and stained pixel extraction tools [21]. Initially time was spent investigating these conventional procedures that are mainly based on the size and colour of the components. However, exploring alternative approaches, which deals with the variation in intensities of different regions, led to the development of the ARR transform method which calculates the ratio of intensities within a circular region. The initial motivation for developing the ARR transform method was to deal with non-uniform illumination. However, the ARR method turned out to be a major breakthrough in the research programme. The ARR method is fundamentally different from other segmentation techniques in that the output of the method is a list of co-ordinates rather than an image. In fact, it can be considered as an object locating tool rather than a segmentation tool. It turned out to be pivotal in that it impacted the processing of the image both before and after the transform stage. Before the transform, it eliminated the necessity of any additional pre-processing procedures for and colour and illumination normalisation and contrast enhancement. After the transform, the processing stages to identify the type of cell and the degree of any parasitemia were simplified in that they only need to examine the image in the vicinity of each cell rather than the whole image so reducing the energy and time required to process an image. It is worth noting that the ARR method could also be used in other applications, outside malaria, where the location of each cell is required.

The ARR transform method becomes the backbone of this research and further progress was based on the information obtained from it. The list of coordinates obtained located the centre pixel of the foreground components (RBCs and WBCs). Differentiating them was the next challenge. The two major distinguishing factors were the size and colour of these components and the possibility of a differentiation tool based on these parameters was first considered. However, further studies proved that these factors were unreliable for a heavily infected blood image since the level of infection can change the appearance of the RBC and thereby impair the diagnosis. The research continued to concentrate on the outcome of the ARR method and eventually developed an effective RBC and WBC differentiation tool based on the average intensity and area of the cellular region in the morphologically closed image as described in Chapter 5.

Automated Malaria Diagnosis Using Mobile Phones

Progress was slow during the development of infected cell identification and life-stage recognition algorithm. Various existing pattern recognition algorithms and neural network architectures were considered but finding a workable algorithm with reduced memory requirement, architectural complexity and better feasibility to implement on a mobile device was difficult. The arbitrary nature of the images, coupled with the unpredictability of the occurrence of the disease/levels of infection and inconsistencies in the colour characteristics of the parasites, made it difficult to develop an untrained, generic algorithm applicable to a wide variety of images. The necessity for a stable system with its own image acquisition tool became apparent during this period. However, an extensive literature survey along with consecutive experimentation using the location information obtained from the ARR transform method led to the design and development of a novel Hue-ARR algorithm which utilises the visual perception characteristics of Hue-Saturation-Value (HSV) colour space combined with the ARR transform method to detect the infected cells and thereby estimate the number of parasitic nuclei present in the RBC. The work done by [19] was the key sign post to consider the HSV colour space and the modified ARR method enables the parasitic nuclei to be identified reliably thereby allowing recognition of the life stages and measurement of the parasitemia. This in fact eliminates the requirement of an integral image acquisition set-up. (However, the process does require an initial calibration to take into account the changes to the resolution of the image.)

Overall, the research experience consisted of one challenge after another. Having solved one problem, a new different problem soon presented itself. This could be disheartening and there were times when it seemed that the current problem was insurmountable. However, the acknowledgements received during the presentation at Pathology Visions Conference, IEEE Pan American Health Care Exchange conference and British Society for Parasitology exhibition provided the encouragement to persevere and move forward to eventual success. The wide variety of image processing considered, developed, adopted or discarded contributed to a very rich learning experience which was further enhanced by the biological and medical understanding gained.

9.3 Improvements and Future Work

The work done in this thesis has lot of advantages compared to existing diagnostic algorithms. However, there remain a number of challenges and extensions still to be undertaken.

[1] Cell overlapping remains a serious issue which sometimes result in off-centred peaks in the ratio transform. The overlapped cells leave a shaded effect in the image which will be picked up by the ARR transform method and can lead to false identification as infected cells. By carefully selecting the appropriate field with evenly distributed cells, this issue can be eliminated. The overall accuracy would be improved with the introduction of an algorithm to automatically identify and eliminate regions where cells are overlapped. Furthermore, an automatic process for locating the mono-layer on the slide would eliminate this currently manual process. The studies by Angulo and Flandrin [104] and Xion et.al [105] explains automated working-area classification by cell clumping detection using morphological processing of connected pixels and trained classifiers respectively. As an extension of this research a fast and cost effective working area estimation tool to detect and capture an evenly distributed monolayer needs to be developed.

[2] The diagnostic set up also requires an effective image capturing unit which will in turn eliminate the need to recurrent calibration based on the type and size of the images. The measured specificity of 97% and overall accuracy of 96% obtained so far were based on a variety of slides and conditions, including some that were very poor. There is an urgent need to measure the performance of the overall process under more controlled and consistent conditions using a fixed optical system.

[3] Extending the image processing algorithm to species identification. Since each species of genus plasmodium exhibit distinctive characteristics, an identification strategy for four species needs to be developed. However, certain life stages of *P.vivax* and *P.ovale* shows similar features such as enlargements of the cells and scruffy stained dots on the membrane of the RBC, a machine vision analysis will be

Automated Malaria Diagnosis Using Mobile Phones

difficult [63]. A thorough investigation on the feature relevancies and classification schemes would be required.

- [4] Development of a speed-up algorithm to increase the efficiency of the overall diagnostic process. Fast closure and ARR methods are possible and would improve the processing time. Parallel processing and vectorisation are common procedures performed in image processing algorithms to decrease the processing time. At the moment, the approximate time taken to perform the algorithm in Matlab is 28 seconds on a 2 GHz quad processor PC and in mobile phone application it is considerably longer depending on the size of the image. In order to enhance the algorithm to be a strong diagnostic tool, computational efficiency is extremely important.
- [5] A hardware realisation of the algorithm onto the mobile Graphics Processing Unit (GPU). The image processing algorithms described here were developed in Matlab using floating-point arithmetic. The quantisation of the algorithm for fixed-point implementation could cause potential problems including multiple peaks within a single cell in the ratio transform. However this will not be required if the algorithm realisation is performed on the GPU of the mobile phone. Unlike a Central Processing Unit (CPU), the GPU of the mobile phone can be used to implement image processing algorithms with floating-point arithmetic and is computationally more effective [NITI2011]. An extensive study should be conducted to investigate the hardware realisation of the algorithm and its optimisation for extended applications.
- [6] Future work should also concentrate on mobile microscopy being able to handle larger resolution images as well as to speed up the algorithm to reduce the diagnostic time. The work can be extended to other operating systems used in the mobile phones such as Apple iOS and Windows 7.

Automated Malaria Diagnosis Using Mobile Phones

As mentioned earlier the method developed in the research is fundamentally different from any image segmentation and parasite diagnostic techniques existing so far. Apart from the morphological filtering, the algorithm does not require any pre-processing techniques and hence speeds up the application and saves memory. Even though a calibration procedure to determine the optimum weights in the colour to gray-scale conversion has been derived, our experience to date suggests that it is not very critical. The algorithm is efficient in locating (and hence counting) cells and locating components within each cell. This in fact widens the application to be used as a blood analysis tool which estimates the components in the image.

The work has the potential to be extended to other image processing applications involving object identification and extraction of satellite images as well as other medical image data analysis involving muscular and lymphatic tissues and animal blood analysis.

Automated Malaria Diagnosis Using Mobile Phones

REFERENCES:

- [1] WHO, "World Malaria Report 2012 Fact Sheet", http://www.who.int/malaria/publications/world_malaria_report_2012/wmr2012_factsheet.pdf, Date Accessed July 2013.
- [2] Malaria No More, <http://malarianomore.org.uk> Date Accessed August 2013.
- [3] Roll back Malaria, "www.rollbackmalaria.org", Date Accessed July 2013.
- [4] Centers for Disease Control and Prevention, "CDC's Malaria Program", Division of Parasitic Diseases, http://www.cdc.gov/malaria/resources/pdf/fsp/cdc_malaria_program.pdf, Date Accessed July 2013
- [5] WHO Publication, "<http://www.who.int/malaria/publications/en/>", Date Accessed August 2013
- [6] Malaria Journal "<http://www.malariajournal.com/content>", Date Accessed, August 2013
- [7] WHO, "Malaria World Report 2009 Fact Sheet", http://www.who.int/malaria/world_malaria_report_2009/factsheet_wmr2009.pdf, Date Accessed, July 2009.
- [8] WHO Malaria Global Program, "World Malaria Report 2010", http://whqlibdoc.who.int/publications/2010/9789241564106_eng.pdf, Date Accessed July 2013.
- [9] WHO Global Malaria Program, "World Malarai Report 2011 Fact Sheet", http://www.who.int/malaria/world_malaria_report_2011/WMR2011_factsheet.pdf, Date Accessed August 2013.
- [10] H.M Gilles, "Management of Severe and Complicated Malaria: A Practical Handbook", ISBN 92-4-154436-8 , 1991.
- [11] Sean C. Murphy and Joel G. Breman, " Gaps in the Childhood Malaria Burden in Africa: Cerebral Malaria, Neurological Sequelae, Anemia, Respiratory Distress, Hypoglycemia, and Complications of Pregnancy", The Intolerable Burden of Malaria: A New Look at the Numbers: Supplement to Volume 64(1) of the American Journal of Tropical Medicine and Hygiene ,January 2001
- [12] Malaria, Encyclopaedia Britannica" <http://www.britannica.com/EBchecked/topic/359534/malaria>", Date Accessed August 2013.
- [13] Division of Parasitic Diseases, "Parasites and Health-Malaria" <http://www.dpd.cdc.gov/dpdx/HTML/Malaria.htm>, Date Accessed September 2013.

Automated Malaria Diagnosis Using Mobile Phones

- [14] Noppadon Tangpukdee, Chatnapa Duangdee, Polrat Wilairatana, and Srivicha Krudsood, "Malaria Diagnosis: A Brief Review", Korean Journal for Parasitology; Pages: 93–102. Published online 2009 May 26. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2688806/>, June 2009
- [15] Garcia, "Determination of Parasitemia" Preanalytical Considerations, Determination of Parasitemia Protocol, http://www.med-chem.com/pages/lab_procedures/pdf/determination_of_parasitemia.pdf, Date Accessed September 2013.
- [16] NCCLS. "Laboratory Diagnosis of Blood-borne Parasitic Diseases; Approved Guideline." NCCLS document M15-A (ISBN 1-56238-401-5). USA, 2000, Date Accessed September 2013.
- [17] Chansuda Wongsrichanalai, Mazie J. Barcus, Sinuon Muth, Awalludin Sutamihardja, and Walther H. Wernsdorfer, "A Review of Malaria Diagnostic Tools: Microscopy and Rapid Diagnostic Test (RDT)", Defining and Defeating the Intolerable Burden of Malaria III: Progress and Perspectives: Supplement to Volume 77(6) of American Journal of Tropical Medicine and Hygiene. Northbrook (IL): American Society of Tropical Medicine and Hygiene, December 2007.
- [18] Stephanie P. Johnston, Norman J. Pieniazek, Manipheth V. Xayavong, Susan B. Slemenda, Patricia P. Wilkins, and Alexandre J. da Silva, "PCR as a Confirmatory Technique for Laboratory Diagnosis of Malaria", Journal of Clinical Microbiology, 44(3): 1087–1089. PMID: PMC1393165, March 2006;
- [19] C. Di Ruberto, Andrew G. Dempster, S. Khan, and B. Jarra, "Analysis of infected blood cell images using morphological operators", Image and vision computing, vol 20, no. 2, pp. 133-146, February 2002.
- [20] K.N.R.M. Rao, "Application of Mathematical Morphology to Biomedical Image Processing," PhD thesis, University of Westminster, 2004
- [21] Boray Tek, Andrew G. Dempster, Izzet Kale. "Computer Vision for microscopy diagnosis of malaria". Malaria Journal, BioMed Central, doi:10.1186/1475-2875-8-153, Online Access : <http://www.malariajournal.com/content/8/1/153>, July 2009.
- [22] W. I. Sherman, "Malaria: Parasite Biology, Pathogenesis and Protection", Washington DC: ASM Press, 1998.
- [23] S.H Lee, U.A Kara, E Koay, M.A Lee, S Lam, D Teo. "New strategies for the diagnosis and screening of malaria". International Journal of Hematology, 76(suppl 1):291-293, August 2002;.
- [24] Anna Rosanas-Urgell, Dania Mueller, Inoni Betuela, Céline Barnadas, Jonah Iga, Peter A Zimmerman, Hernando A del Portillo, Peter Siba, Ivo Mueller and Ingrid Felger, "Comparison of diagnostic methods for the detection and quantification of the four sympatric *Plasmodium* species in field samples from Papua New Guinea" Malaria Journal, Volume 9, www.malariajournal.com, 2010.

Automated Malaria Diagnosis Using Mobile Phones

- [25] Roger S. Riley, G. Watson James, Sandra Sommer, Mary Jo Martin, "How to Prepare and Interpret Peripheral Blood Smear", Medical College of Virginia, Virginia Commonwealth University, <http://www.pathology.vcu.edu/education/PathLab/pages/hematopath/pbs.html>, Date Accessed August 2013.
- [26] Blood Cells," <http://legacy.owensboro.kctcs.edu/gcaplan/anat2/histology>", Date Accessed August 2013
- [27] M. Schaefer And R. M. Rowan, "The Clinical Relevance of Nucleated Red Blood Cell Count", Sysmex Journal International Vol.10 No.2, Series 11, 2000 .
- [28] TechnoClin," Trends in Pathology and Laboratory Medicine", <http://technoclin.com/articles/trends-in-pathology-and-laboratory-medicine>, Date Accessed September 2013.
- [29] Tang, H Laura, Gonen Mithat, Hedvat, Cyrus, Modlin, Irvin M, Klimstra, David S," Objective Quantification of the Ki67 Proliferative Index in Neuroendocrine Tumors of the Gastroenteropancreatic System: A Comparison of Digital Image Analysis with Manual Methods, American Journal of Surgical Pathology: Volume 36 - Issue 12 – Pages: 1761–1770, December 2012.
- [30] H R Ali, M Irwin, L Morris, S-J Dawson, F M Blows, E Provenzano, B Mahler-Araujo, P D Pharoah, N A Walton, J D Brenton and C Caldas,"Astronomical algorithms for automated analysis of tissue protein expression in breast cancer", British Journal of Cancer and Molecular Diagnosis, 108, Pages: 602–612.www.bjcancer.com, January 2013.
- [31] Anthony E Rizzardi, Arthur T Johnson, Rachel Isaksson Vogel, Stefan E Pambuccian, Jonathan Henriksen, Amy PN Skubitz, Gregory J Metzger and Stephen C Schmechel," Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring" Diagnostic Pathology 2012, <http://www.diagnosticpathology.org>, April 2012.
- [32] Electron Microscopy Sciences, "Focus on Emerging Technology", <http://www.emsdiasum.com/microscopy>, Date Accessed August 2013.
- [33] ScanScope OS Operating System, "Hematopathology and Microbiology Applications", <http://www.medicine.usask.ca/academic-units/support-units/itu/equipment-and-facility-resources/virtual-microscopy-facility/os.pdf>, Date Accessed August 2013.
- [34] Vipul Parkhi, Pooja Pawar, Archana Surve, "Computer Automation for Malaria Parasite Detection Using Linear Programming", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.2, Issue 5, http://www.ijareeie.com/upload/may/35_%20Computer.pdf, May 2013.

Automated Malaria Diagnosis Using Mobile Phones

- [35] J.P. Vink, M. Laubscher, R. Vlutters, K. Silamut, R.J. Maude, M.U. Hasan And G. De Haan, "An automatic vision-based malaria diagnosis system", *Journal of Microscopy*, Vol. 250, Issue 3, Pages: 166–178 doi: 10.1111/jmi.12032, 2013.
- [36] Keerthana Prasad, Jan Winter, Udayakrishna M. Bhat, Raviraja V. Acharya, Gopalakrishna K. Prabhu, "Image Analysis Approach for Development of a Decision Support System for Detection of Malaria Parasites in Thin Blood Smear Images", *Journal of Digital Imaging*, Volume 25, Issue 4, Pages: 542–549, August 2012.
- [37] Yashaswi Purwar, Sirish L Shah, Gwen Clarke, Areej Almigairi and Atis Muehlenbachs, "Automated and Unsupervised detection of Malarial Parasites in Microscopic Images", *Malaria Journal* 2011, 10:364, www.malariajournal.com, 2011
- [38] C. Di Ruberto, Andrew G. Dempster, S. Khan, and B. Jarra, "Automatic thresholding of infected blood images using granulometry and regional extrema." In *Proceeding International Conference on Pattern Recognition*, Barcelona, Spain, 2000.
- [39] K.N.R.M. Rao and A. Dempster, "Area-granulometry : an improved estimator of size distribution of image objects," *IEEE Letters*, vol. 37, Pages: 50-951, 2001
- [40] K.N.R.M. Rao and A. Dempster, "Modification on distance transform to avoid over-segmentation and under-segmentation," *Proceeding International Symposium on Video/Image Processing and Multimedia Communications*, Croatia, 2002.
- [41] Fred L. Metzger Jr., Alan Rebar, "Three minute peripheral blood film evaluation: Preparing the film", *Veterinary medicine*, www.vetmedpub.com, December 2004
- [42] Boray Tek, "Computerised Diagnosis of Malaria". PhD Dissertation, University of Westminster, September 2007.
- [43] A. Wetzel, P. Feineigle and J. Gilbertson, "Design of a high-speed slide imaging system for pathology", in *Proceeding IEEE International Symposium on Biomedical Imaging*, Washington DC, USA, 2002
- [44] N. Otsu, "A Threshold Selection Method From Gray-level Histograms", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, Pages: 62-66, 1979.
- [45] Qiang Chen, Quan-sen Sun, Pheng Ann Heng, De-shen Xia, "A double-threshold image binarisation method based on edge detector", *Pattern Recognition*, Volume 41, Issue 4, Pages: 1254–1267, April 2008

Automated Malaria Diagnosis Using Mobile Phones

- [46] Y.Liu, G.Liang, P.K.Saha, "A new multi object image thresholding method based on correlation between object class uncertainty and intensity gradient", *Medical Physics* ,39(1) Pages:514-32. <http://www.ncbi.nlm.nih.gov/pubmed/22225322>, January 2012
- [47] C.Saravanan,"Color Image to Grayscale Image Conversion.", IEEE Second International Conference on Computer Engineering and Applications, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5445596>, 2010.
- [48] L.Gooldman, A.Schafer,"The peripheral blood smear",*Cecil Medicine*, Saunders Elsevier, Ch.160,Philadelphia, 2011.
- [49] Mehdi Habibzadeh, Adam Krzyzak, Thomas Fevens,"Application of Pattern Recognition Techniques for the Analysis of Thin Blood Smear Images",*Journal of Medical Informatics and Technologies*, Vol 18, ISSN 1642-6037, 2011.
- [50] K.V.Kale, S.C. Malhotra, R.R Manza,"Computer Vision and Information Technology: Advances and Applications",I.K International Private Limited, ISBN-13: 978-9380026954, May 2010
- [51] Rafael.C.Gonzalez,Richard.E.Woods, "Digital Image Processing", Pearson International Edition, Third Edition, August 2007.
- [52] P.Soille,"Morphological Image Analysis", Second Edition, Springer-Verlag,ISBN 3-540-42988-3, 2003.
- [53] Anoraganingrum, D.," Cell segmentation with median filter and mathematical morphology operation", IEEE International Conference on Image Analysis and Processing, Venice, Pages: 1043 – 1046, 1999.
- [54] F. Boray Tek, Andrew G. Dempster and Izzet Kale," Blood Cell Segmentation Using Minimum Area watershed And Circle Radon Transformations", *Mathematical Morphology: 40 Years On Computational Imaging and Vision Volume 30*, Pages:441-454, 2005.
- [55] Pearl P. Guan , Hong Yan, "Blood Cell Image Segmentation Based On The Hough Transform And Fuzzy Curve Tracing", *Proceedings of the 2011 International Conference on Machine Learning and Cybernetics*, Guilin, 10-13 July 2011.
- [56] S.Kareem,R.C.S Morling, I Kale" A Novel Method to Count the Red Blood Cells in Thin Blood Films",*IEEE International Symposium on Circuits and Systems (ISCAS)*, Pages 1021-1024, 2011
- [57] J. M. Sharif, M. F. Miswan, M. A. Ngadi, Md Sah Hj Salam, Muhammad Mahadi bin Abdul Jamil," Red Blood Cell Segmentation Using Masking and Watershed Algorithm: A Preliminary Study", *IEEE International Conference on Biomedical Engineering (ICoBE)*, Pages:258 - 262 ,ISBN:978-1-4577-1990-5 , 2012.

Automated Malaria Diagnosis Using Mobile Phones

- [58] Valentín Osuna, Erik Cuevas, Humberto Sossa, "Segmentation of Blood Cell Images Using Evolutionary Methods", *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation II*, Advances in Intelligent Systems and Computing Volume 175, Pages: 299-311, 2013
- [59] L. B. Dorini , R. Minetto. ; N. J.Leite, "Semiautomatic White Blood Cell Segmentation Based on Multiscale Analysis", *IEEE Journal of Biomedical and Health Informatics*, Volume:17 , Issue: 1 , Pages:250 - 256 ,ISSN :2168-2194 , January 2013.
- [60] Faliu Yi ; Inkyu Moon ; Bahram Javidi ; Daniel Boss ; Pierre Marquet," Automated segmentation of multiple red blood cells with digital holographic microscopy", *Journal of Biomedical Optics*. 18(2), 026006; <http://dx.doi.org/10.1117/1.JBO.18.2.026006> , February 2013.
- [61] Wang Feng; Wang Xiuyou; Xu Lin, "An Improved Fingerprint Segmentation Algorithm Based on Mean and Variance," *Intelligent Systems and Applications, 2009. ISA 2009. International Workshop on* , vol., no., Pages:23-24 ,doi: 10.1109/IWISA.2009.5072716, May 2009.
- [62] H. Fleyeh, D. Jomaa, M. Dougherty, "Segmentation of low quality fingerprint images," *Multimedia Computing and Information Technology (MCIT), 2010 International Conference on* , vol., no., Pages:2-4,doi: 10.1109/MCIT.2010.5444844, March 2010.
- [63] M.Kettlehut, C.Rogers,"Updates in Blood and Tissue Parasitology", Short Course, Department of Life Science, University of Westminster, 2010.
- [64] M. Sheikhhosseini, H. Rabbani, M. Zekri and A. Talebi , "Automatic diagnosis of malaria based on complete circle-ellipse fitting search algorithm". *Journal of Microscopy*, 252: Pages: 189–203. doi: 10.1111/jmi.12081, 2013
- [65] Livestrong,"Abnormalities in Blood Cells" <http://www.livestrong.com/article/153256-abnormalities-in-blood-cells/> , Date Accessed September 2013.
- [66] K. Yabusaki , T. Faits, E. McMullen, J. L Figueiredo ,M. Aikawa, "A Novel Quantitative Approach for Eliminating Sample-To-Sample Variation Using a Hue Saturation Value Analysis Program". *PLoS ONE* ,9(3): e89627. doi: 10.1371/journal.pone.0089627 , 2014
- [67] Shamik Sural, Gang Qian and Sakti Pramanik," Segmentation And Histogram Generation Using The Hsv Color Space For Image Retrieval" *IEEE International Conference on Image Processing (ICIP)*, 0-7803-7622-6/02 ©2002 IEEE, 2002.
- [68] S.Kareem, I Kale, R.C.S Morling," Automated Malaria Parasite Detection in Thin Blood Films: - A Hybrid Illumination and Color Constancy Insensitive Morphological Approach", to be presented at *IEEE Asia Pacific Conference on Circuits and Systems, Kaohsiung, Taiwan, December 2012.*

Automated Malaria Diagnosis Using Mobile Phones

- [69] S.Kareem , I .Kale, R.C.S Morling , ”A Novel Fully Automated Malaria Diagnostic Tool Using Thin Blood Films” presented and published at Pan American Health Care Exchanges (PAHCE2013), in Madellin, Columbia, May 2013.
- [70] J.A Swets,”Measuring the accuracy of diagnostic systems”, Science.;240(4857):1285-93,PMID:3287615,June 1988.
- [71] Rajul Parikh, Annie Mathai, Shefali Parikh, G Chandra Sekhar, Ravi Thomas,”Understanding and using sensitivity, specificity and predictive values”,Indian Journal of Ophthalmology, 56(1) Pages : 45–50. PMID:PMC2636062, 2008.
- [72] Allisson D.Oliveira,Giordani Cabral, D.Lopez,Caetano Firmo, F.Zarzuela Serrat,J.Albuquerque,”A proposal for automatic diagnosis of malaria: extended abstract”, WWW’13 Companion Proceedings of the 22nd international conference on World Wide Web companion , pages 681-682, ISBN:978=1-4503-2038-2, 2013.
- [73] Hisashi Fujioka,Masamichi Aikawa,” Structure and Life Cycle: Malaria Parasites and Disease” Perlmann P, Troye-Blomberg M (eds): Malaria Immunology. Chem Immunol. Basel, Karger, vol 80, Pages 1–26, 2002.
- [74] D.C Warhurst and J.E Williams, ”Laboratory Diagnosis of Malaria”, Journal for Clinical Pathology,49(7)Pages : 533–538.PMCID: PMC500564, July 1996.
- [75] Catherine H Roberts, Margaret Armstrong, Ewa Zatyka, Samuel Boadi, Simon Warren, Peter L Chiodini, Colin J Sutherland and Tom Doherty ,” Gametocyte carriage in *Plasmodium falciparum*-infected travellers, Malaria Journal, Pages:12:31 doi:10.1186/1475-2875-12-31, 2013.
- [76] Teun Bousema, Lucy Okell, Seif Shekalaghe, Jamie T Griffin, Sabah Omar, Patrick Sawa, Colin Sutherland, Robert Sauerwein, Azra C Ghani and Chris Drakeley Bousema et al : “Revisiting the circulation time of *Plasmodium falciparum* gametocytes: molecular detection methods to estimate the duration of gametocyte carriage and the effect of gametocytocidal drugs”. Malaria Journal, 9:136, 2010.
- [77] Noppadon Tangpukdee, Srivicha Krudsood, Sriripun Srivilairit, Nanthaporn Phophak,Putza Chonsawat, Wimon Yanpanich, Shigeyuki Kano and Polrat Wilairatana,”Gametocyte Clearance in Uncomplicated and Severe *Plasmodium falciparum* Malaria after Artesunate-Mefloquine Treatment in Thailand”, Korean J Parasitol. Vol. 46, No. 2: 65-70, June 2008
- [78] MVI, “Fact sheet:*Plasmodium falciparum* malaria “, PATH, http://www.malariavaccine.org/files/FS_Pf_falciparum-Sept-2004_FINAL.pdf, Date Accessed September 2013.

Automated Malaria Diagnosis Using Mobile Phones

- [79] S.Kareem, I Kale, R.C.S Morling,” Automated *P.falciparum* Detection system for Post-treatment Malaria Diagnosis using Modified Annular Ring Ratio Method”,IEEE 14th International Conference on Computer Modelling and Simulation ,Cambridge, UK, March 2012.
- [80] F.Kawamoto and P.F Billingsley,”Rapid Diagnosis of Malaria by Fluorescence Microscopy”,Trends in Parasitology,Volume 8 Issue 2, February 1992.
- [81] Rebecca Guy, Paul Liu, Peter Pennefather and Ian Crandall, “The use of fluorescence enhancement to improve the microscopic diagnosis of falciparum malaria”, Malaria Journal, 6:89, <http://www.malariajournal.com/content/6/1/89>, 2007.
- [82] Saad El-Din H Hassan, Abd Elrahium D. Haggaz, Ehab B Mohammed-Elhassan, Elfatih M Malik, and Ishag Adam,”Fluorescence microscope (Cyscope®) for malaria diagnosis in pregnant women in Medani Hospital, Sudan”,Diagnostic Pathology; 6: 88, doi: 10.1186/1746-1596-6-88, September 2011.
- [83] P. Caramello, A. Lucchini, D. Savoia, P. Gioannini,”Rapid diagnosis of malaria by use of fluorescent probes”.Diagnostic Microbiology of Infected Diseases,17, Pages:293-297, 1993.
- [84] D. Barkan, H. insburg, J. Golenser,”Optimisation of flow cytometric measurement of parasitaemia in *Plasmodium*-infected mice”.International Journal of Parasitology,30, Pages:649-653, 2010.
- [85] S.Kareem, I. Kale, R.C.S Morling,” Automated Malaria diagnosis in Fluorescent Thin Blood Image Using Annular Ring Ratio Method”, Presented at Pathology Visions Conference, Digital Pathology Association, October 28 – 31, Baltimore, Maryland, USA , 2012.
- [86] H. Jouin, W. Daher, J. Khalife, I. Ricard, O.M,Puijalon, M. Capron, D. Dive,”Double staining of *Plasmodium falciparum* nucleic acids with hydroethidine and thiazole orange for cell cycle stage analysis by flow cytometry”.Cytometry Article, 57,Pages:34-38, 2004.
- [87] S. Moon, S Lee, H. Kim, L.H Freitas-Junior, Kang, “An Image Analysis Algorithm for Malaria Parasite Stage Classification and Viability Quantification”, PLoS ONE 8(4): e61812. doi:10.1371/journal.pone.0061812, 2013.
- [88] B. Woodward, R.S.H Istepanian, C.I Richards ,”Design of a telemedicine system using a mobile telephone”. IEEE Transactions on Information Technology in Biomedicine, 5,Pages : 13–15, 2001.
- [89] S. Tachakra, X.H Wang, R.S Istepanian, Y.H Song ,” Mobile e-health: the unwired evolution of telemedicine”. Telemed J E Health, 9, Pages: 247–257, 2003.
- [90] R.P Braun, J.L Vecchietti, L. Thomas, C Prins, L.E French ,”Telemedical wound care using a new generation of mobile telephones: afeasibility study”. Arch Dermatol, 141: 254–258, 2005.

Automated Malaria Diagnosis Using Mobile Phones

- [91] Y. Granot, A. Ivorra, B. Rubinsky ,”A new concept for medical imaging centered on cellular phone technology”. PLoS ONE, 3: e2075, 2008.
- [92] N.V Hung, D.N Sy, R.M Anthony, F.G Cobelens, D. Van Soolingen,”Fluorescence microscopy for tuberculosis diagnosis”. Lancet Infectious Diseases, 7:238–239, 2007.
- [93] D. Jones, J. Nyalwidhe, L. Tetley, M.P Barrett,” McArthur revisited:fluorescence microscopes for field diagnostics”. Trends Parasitology, 23: 468–469, 2007.
- [94] D.N Breslauer, R.N Maamari, N.A Switz, W.A Lam, D.A Fletcher ,” Mobile Phone Based Clinical Microscopy for Global Health Applications”. PLoS ONE, 4(7): e6320. doi: 10.1371/journal.pone.0006320, 2009.
- [95] A. Tapley , N. Switz , C Reber , J.L Davis , C. Miller , J.B Matovu , W. Worodria , L. Huang , D.A Fletcher , A. Cattamanchi ,” Mobile digital fluorescence microscopy for diagnosis of tuberculosis”. Journal of Clinical Microbiology, 51(6):1774-8. doi: 10.1128/JCM.03432-12, June 2013.
- [96] C.S Godse, S Patkar, N.S Nabar, A.J Amonkar, R.A Vaidya, A.A Raut, A.B Vaidya. Mobile camera Microphotography: “A Simple but Elegant Technique for Tediagnosis of Malaria”, <http://www.jkscience.org/archive/volume103/new%20horizons/TELEPHONE.pdf>, 2008
- [97] Livia Bellina, Eduardo Missoni: “Mobile cell-phones (M-phones) in telemicroscopy: increasing connectivity of isolated laboratories”. Diagnostic Pathology, 4:19 doi:10.1186/1746-1596-4-19, 2009.
- [98] J. Frean ,” Mircoscopic Images Transmitted by Mobile Camera Phone,” Transactions Royal Society for Tropical Medicine and Hygeine ,101 (10): 1053. doi: 10.1016/j.trstmh.2007.06.008,2007.
- [99] Stanford University,”Mobile Image Processing “, http://www.stanford.edu/~dmchen/documents/EE3682_011Spring_GuestLecture_Part1.pdf, retrieved in September 2013.
- [100] Eclipse, “ What is Eclipse and eclipse Foundation” <https://www.eclipse.org/org/#about>, retrieved in September 2013.
- [101] Oracle documentation,”JDK 7 and JRE 7 Installation Guide”, <http://docs.oracle.com/javase/7/docs/webnotes/install/>, retrieved in September 2013.
- [102] Android,”Getting Android SDK”, <https://developer.android.com/sdk/index.html?hl=sk>, retrieved in September 2013.
- [103] Newton Microscopes,” Nm1 Specifications”, http://newtonmicroscopes.com/nm1_spec.html, retrieved in September 2013.

Automated Malaria Diagnosis Using Mobile Phones

- [104] J. Angulo , G. Flandrin ., ” Automated detection of working area of peripheral blood smears using mathematical morphology”. *Anal Cellular Pathology*,25(1):37-49, 2003.
- [105] W. Xiong , S.H Ong , J.H Lim , K.W Foong , J. Liu , D. Racoceanu , A.G Chong , K.S Tan , ” Automatic area classification in peripheral blood smears”, *IEEE Transactions on Biomedical Engineering*;57(8):1982-90. doi: 10.1109/TBME.2010.2043841. March 2010.
- [106] Nitin Singhal, Jin Woo Yoo, Ho Yeol Choi, In Kyu Park, ”Design and Optimization of Image Processing Algorithms on Mobile GPU”, *ACM Special Interest Group on Computer Graphics and Interactive Techniques* , Article no:21, ISBN: 978-1-4503-0971-4 doi>10.1145/2037715.2037741, 2011 .
- [107] Anthony K Akobeng, ” Understanding diagnostic tests : sensitivity, specificity and predictive values”, *Journal Compilation, Foundation Acta Pædiatrica/Acta Pædiatrica* , Pages. 338–341, 2006.
- [108] L.Newberg, ” Some Useful Statistics Definitions”, Date Accessed August 2012, <http://www.cs.rpi.edu/~leen/misc-publications/SomeStatDefs.html>.2005.
- [109] Tze-wei Loong, ” Understanding sensitivity and specificity with the right side of the brain”, *BMJ* 327 doi: <http://dx.doi.org/10.1136/bmj.327.7417.716>, September 2003.
- [110] Mariela Azul Gonzalez, Gustavo Javier Meschino, Virginia Laura Ballarin, ” Solving the over segmentation problem in applications of Watershed Transform”, *Journal of Biomedical Graphics and Computing*, Vol. 3, No. 3, 2013.
- [111] MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States, September 2012.
- [112] C.Dallet, S.Kareem, I.Kale, ”Mobile Malaria Diagnosis: Real time mobile diagnostic application on Android using Java” .*International Symposium for Circuits and Systems, ISCAS 2014*.
- [113] Jaynes, E.T. "Bayesian Methods: General Background." In *Maximum-Entropy and Bayesian Methods in Applied Statistics*, by J. H. Justice (ed.). Cambridge: Cambridge Univ. Press, 1986.
- [114] Blaivas M, Lyon M, Duggal S, ” Ultrasound image transmission via camera phones for overreading”, *The American Journal of Emergency Medicine*, Pages:433-435,<http://dx.doi.org/10.1016/j.ajem.2004.09.037>, 2005.
- [115] J. Piek , R. Hebecker,., M. Schütze , S. Sola , S. Mann , K. Buchholz, ” Image Transfer by Mobile Phones in Neurosurgery”, *Central European Neurosurgery*,Pages: 193-196, DOI: 10.1055/s-2006-94230, 2006.

Automated Malaria Diagnosis Using Mobile Phones

- [116] S. Razdan, J. Johannes, R. L. Kuo, D. H. Bagley, "The camera phone: A novel aid in urologic practice.", *Urology Journal*, Volume 67, Issue 4, Pages 665–669, <http://dx.doi.org/10.1016/j.urology.2005.10.005>, April 2006.
- [117] F. Meyer, S. Beucher, "Morphological Segmentation", *J. Vis. Comms. Image Representation*, vol. 1, pp. 21–46, 1990.
- [118] L. Najman, M. Schmitt, "Geodesic Saliency of Watershed Contours and Hierarchical Segmentation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 1163–1173, 1996.

Automated Malaria Diagnosis Using Mobile Phones

BIBLIOGRAPHY

- [1] Boray Tek, Andrew G. Dempster and Izzet Kale. "Malaria Parasite Detection in Peripheral Blood Images (2006)" Paper published retrieved from <http://www.bmva.ac.uk/bmvc/2006/papers/156.pdf>.
- [2] Andrej Trampuz, Matjaz Jereb, Igor Muzlovic, Rajesh M Prabhu. Clinical review: "Severe malaria. Critical Care" 2003;7:315-323 Available at <http://ccforum.com/content/7/4/315>.
- [3] Kim, Ph.D. and Clark D. Haass, MSEE. "A Next-based High Performance Image Computing Workstation for Biomedical Applications". Engineering in Medicine and Biology Society, 1990. Proceedings of the Twelfth Annual International Conference of the IEEE.
- [4] Lindsay W. MacDonald. "Digital heritage: applying digital imaging to cultural heritage" - Antiques & Collectibles – 2006.
- [5] Amitabh kumar. "Mobile TV: DVB-H, DMB, 3G systems and rich media applications". Technology & Engineering – 2007.
- [6] Julie.A.jacko."Human –Computer Interaction: New trends",13th International conference,HCI international 2009,San Diego.
- [7] Steven.W.Smith: "The Scientist and Engineer's guide to Digital signal processing", Chapter 25, pp 423-450.
- [8] Li.Yiran, "FPGA Implementation for Image Processing Algorithms"EEL6562 Course project report, December 2006.
- [9] Jinhong K. Guo1 and Matthew Y. Ma2 "A Low Resource Consumption Image Region Extraction Algorithm for Mobile Devices" Multimedia and Expo, 2007 IEEE International Conference.
- [10] Medical Imaging Implementation using FPGAs; Altera,White paper, WP-medical,April 2006,ver1.0.
- [11] Guerin PJ, Olliaro P, Nosten F, et al."Malaria: current status of control, diagnosis, treatment and proposed agenda for research and development". Lancet Inf Dis 2002, 2: 564-73.
- [12] Kaushal DC. "Recent approaches to malaria diagnosis. Current R&D highlights", 1995, Oct. Dec.:14-17.
- [13] Ebner C, Wurm EM, Binder B, et al. "Mobile teledermatology: a feasibility study of 58 subjects using mobile phones".J Telemed Telecare 2008;14(1):2-7.
- [14] Zhang P, Kogure Y, Matsuoka H, et al. "A remote patient monitoring system using a Java-enabled 3G mobile phone".
- [15] Wikipedia- Malaria Diagnosis www.wikipedia.org