Image Processing Methodology for Blood Cell Counting via Mobile Devices

Hugo Miguel Felgueira de Andrade

Supervisor: Jaime Santos Cardoso, PhD
Supervisor at Fraunhofer: Luís Filipe Caeiro Margalho Guerra Rosado, MSc

Master in Bioengineering

June, 2015
Faculty of Engineering of University of Porto

Image Processing Methodology for Blood Cell Counting via Mobile Devices

Hugo Miguel Felgueira de Andrade

Dissertation submitted to Faculty of Engineering of University of Porto to obtain the Master degree in

Bioengineering

June, 2015
Abstract

Nowadays, malaria is one of the most severe public health issues in developing countries, affecting millions of people every year and contributing immensely to child mortality. The limitation of health and human resources as well as the lengthy microscopy-based blood diagnosis are impediments to timely diagnosis and population screening. Thus, Fraunhofer AICOS proposed MalariaScope, a research project aiming to develop a cheaper and faster alternative which, besides a mobile-phone adaptable microscope, includes an automatic analysis system. This system will be able to assess the parasite density and the cellular components in the bloodstream, the latter being this thesis focus. In fact, some computer vision approaches have already been developed in the blood cellular components detection and identification field. Even though several methodologies have been proposed on the literature with interesting results, their efficiency in low-quality images is limited. In this work, a methodology for red blood cell segmentation and counting, and for detection and recognition of white blood cells in mobile-acquired blood smear images is proposed.

The segmentation methodologies of the cellular components, besides the automatic threshold selection techniques, includes methods for refinement of the resulting binary images, such as anisotropic morphological operations and distance-transform-based watershed for red blood cells, and gradient-restricted morphological reconstruction operations for white blood cells. The estimation of red blood cells is assessed through regression analysis of the area and number of objects segmented, while the assigning of white blood cells as mono- or polymorphonuclear is based on classification methods after the detection of shape and texture features from the leukocytes nucleus. The classification results reveal that this methodology is capable of giving estimations of the number of red blood cells, extracting white blood cells candidates with high precision, and accurately distinguishing them in either mono- or polymorphonuclear with the considered features.

Finally, the integration of this methodology in mobile devices, particularly in the MalariaScope prototype application, was possible, despite of the speed-limiting high computational cost. The results from this study are very encouraging for ultimately integrating this methodology in a product conceived for aiding the population screening in malaria-affected regions.

Keywords: Malaria, White Blood Cells, Red Blood Cells, Counting, Recognition, Image Processing, Machine Learning
Resumo

Atualmente, um dos maiores problemas de saúde pública nas regiões em desenvolvimento é a malária, afetando milhões de pessoas anualmente e contribuindo imensamente para a mortalidade infantil. A limitação em termos de recursos humanos e equipamentos de saúde, assim como os demorosos diagnósticos baseados em microscopia do sangue são obstáculos tanto ao diagnóstico atempado como ao rastreio da população. Surgiu assim o projecto MalariaScope da Fraunhofer AICOS, cujo intuito consiste em desenvolver uma alternativa barata e rápida que, além de um microscópio adaptável a dispositivos móveis, inclui um sistema de análise automático. Este sistema será capaz de aferir a densidade de parasitas assim como os componentes celulares do sangue, sendo este último o foco desta tese. De facto, abordagens em visão computacional já têm sido desenvolvidas neste campo da deteção e identificação de componentes celulares. Apesar de já terem sido publicadas diversas metodologias com resultados interessantes, a eficiência das mesmas em imagens de baixa qualidade é limitada. Neste trabalho é proposta uma metodologia para segmentação e subsequente contagem de glóbulos vermelhos e deteção e reconhecimento de glóbulos brancos em imagens de esfregaços de sangue em gota fina e gota espessa adquiridas por dispositivos móveis.

As metodologias de segmentação dos componentes celulares, além de técnicas de seleção automática do threshold, incluem métodos de refinamento das imagens binárias resultantes, como operações morfológicas anisotrópicas e de watershed baseada na transformada de distância para os glóbulos vermelhos, e operações de reconstrução morfológicas com restrições impostas pelo gradiente para os glóbulos brancos. A estimativa do número de glóbulos vermelhos é realizada através de análise regressiva da área e número de objetos segmentados enquanto que a designação dos glóbulos brancos em mono- ou multimorfonucleados baseia-se em métodos de classificação após obtenção de características de forma e textura dos núcleos. Os resultados revelaram que esta metodologia é capaz de estimar com precisão o número de glóbulos vermelhos, de extrair candidatos a glóbulos brancos com elevada exatidão e distinguí-los em mono- ou multimorfonucleados com as características consideradas.

Por fim, a integração destas metodologias em dispositivos móveis, particularmente no protótipo da aplicação MalariaScope, foi possível, apesar do elevado custo computacional limitador da rapidez da mesma. Os resultados deste estudo são bastante encorajadores quanto à integração final desta metodologia num produto passível de ser utilizado nas zonas de impacto da malária, de forma a auxiliar o rastreio da população.

Keywords: Malária, Glóbulos Brancos, Glóbulos Vermelhos, Contagem, Reconhecimento, Processamento de Imagem, Aprendizagem Automática
As long as you leave everything out on the basketball court, you can look at the mirror and feel good about yourself

Allen Iverson
# Contents

<table>
<thead>
<tr>
<th>List of Figures</th>
<th>xii</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xv</td>
</tr>
</tbody>
</table>

1 Introduction 1
   1.1 Motivation and Context 1
   1.2 Blood Cells Counting 2
   1.3 Objectives 2
   1.4 Contributions of this work 2
   1.5 Structure of the document 3

2 Malaria Characterization 5
   2.1 Biology of Malaria 5
   2.2 Pathophysiology of the Malaria 8
      2.2.1 Hepatic Stage 8
      2.2.2 Erythrocytic Stage 8
   2.3 Malaria Diagnosis 9
      2.3.1 Routine Laboratory Diagnosis 10
   2.4 Blood Cells Morphology 11
      2.4.1 Haematopoiesis 11
      2.4.2 Red Blood Cells 13
      2.4.3 White Blood Cells 13
      2.4.4 Relevant Remarks 15

3 Literature Review 17
   3.1 Technology Review 17
      3.1.1 Mobile Operating System 17
      3.1.2 Computer Vision libraries 18
   3.2 Related Work 19
      3.2.1 Red Blood Cells Detection 19
      3.2.2 White Blood Cells Detection 21
   3.3 Image Processing Techniques 24
      3.3.1 Image Enhancement 25
      3.3.2 Morphological Image Processing 27
      3.3.3 Image Segmentation 29
      3.3.4 Automatic Thresholding - RATS method 30
      3.3.5 Image Features 31
3.4 Machine Learning Techniques .............................................. 34
  3.4.1 Feature Selection ..................................................... 34
  3.4.2 Model Evaluation Methods ......................................... 35
  3.4.3 Classification Methods .............................................. 36
  3.4.4 Regression Methods .................................................. 37
3.5 Relevant Remarks .......................................................... 38

4 RBC counting in Thin Blood Smear Images .................. 41
  4.1 Image Dataset Construction ........................................... 41
       4.1.1 Image Dataset Requirements ................................... 42
       4.1.2 Image Dataset Construction .................................... 42
  4.2 RBC Segmentation ......................................................... 43
       4.2.1 Algorithm Overview .............................................. 43
       4.2.2 Preprocessing and Image Enhancement ......................... 44
       4.2.3 Automatic Threshold Selection ................................ 46
       4.2.4 Directional Gradient ............................................. 48
  4.3 Regression Algorithm .................................................... 50
  4.4 Segmentation Results ..................................................... 51
  4.5 Regression Results ....................................................... 54

5 WBC recognition in Thick Blood Smear Images .......... 55
  5.1 Image Dataset Construction ........................................... 55
       5.1.1 Image Dataset Requirements ................................... 55
       5.1.2 Image Dataset Construction .................................... 56
  5.2 WBC Candidates Extraction ............................................ 56
       5.2.1 Algorithm Overview .............................................. 56
       5.2.2 Two-Step Thresholding Method ................................ 57
       5.2.3 WBC Candidates Segmentation ................................ 58
  5.3 Feature Vector Construction ........................................... 59
  5.4 Classification Methods .................................................. 60
  5.5 Segmentation Results ..................................................... 61
  5.6 Classification Results .................................................. 63

6 WBC recognition in Thin Blood Smear Images ........ 67
  6.1 Database Construction .................................................. 68
       6.1.1 Database Requirements ........................................... 68
       6.1.2 Database Construction ........................................... 68
  6.2 WBC Candidates Extraction ............................................ 68
       6.2.1 Algorithm Overview .............................................. 68
       6.2.2 WBC Nucleus Segmentation ..................................... 69
       6.2.3 WBC Cytoplasm Segmentation ................................... 69
  6.3 Feature Vector Construction ........................................... 72
  6.4 Classification Methods .................................................. 72
  6.5 Segmentation Results ..................................................... 73
  6.6 Classification Results .................................................. 75
## CONTENTS

7 **Integration in the MalariaScope Application**  
  7.1 MalariaScope Application Overview .......................... 79  
  7.2 Application Requirements ...................................... 80  
  7.3 Image Processing Module Integration .......................... 81

8 **Conclusions and Future Work**  

References  

ix 87
List of Figures

2.1 The life cycle of Plasmodium falciparum. ........................................ 6
2.2 Erythrocytic Invasion by Merozoite. .............................................. 7
2.3 Virulence clusters caused by infected erythrocytes due to changes in surface adhesion properties. .................................................. 8
2.4 Infected erythrocyte sequestration in microvascular endothelium. .......... 9
2.5 General model of haematopoiesis. .................................................. 12
2.6 Giemsa stained blood smear. ......................................................... 13
2.7 Erythrocyte in a Giemsa stained thin blood film. ................................. 14
2.8 Five different types of leukocytes. ................................................. 14
2.9 Neutrophilic series. ................................................................. 14
2.10 Eosinophilic series. ................................................................. 14
2.11 Basophilic series. ................................................................. 15
2.12 Monocyte maturation series. ...................................................... 15

3.1 Piecewise transformation function. .................................................. 26
3.2 Two $3 \times 3$ smoothing filter masks. .............................................. 26
3.3 $3 \times 3$ Sobel and Prewitt masks. ................................................ 28
3.4 K-fold cross-validation model. .................................................... 35
3.5 Nested K-fold cross-validation model. .......................................... 35
3.6 High-quality and low-quality thin blood smear images. ....................... 40

4.1 Pipeline for RBC counting in thin blood smear images. ....................... 41
4.2 Illustration of the image acquisition assembling. ................................ 43
4.3 FOV mask segmentation. ............................................................ 44
4.4 Separation of the 3-channel thin blood smear image and respective histogram. ................................................ ..................... 45
4.5 Linear contrast-stretching. .......................................................... 46
4.6 Illustration of the proposed moving-window method. .......................... 47
4.7 Anisotropic structural elements. .................................................. 49
4.8 Directional Dilation demonstration. .............................................. 50
4.9 Global vs Local thresholding application, prior to the application of the watershed method based on the distance transform. ..................... 52
4.10 Segmentation results when directional dilation is applied, no dilation is applied and isotropic dilation is applied. ........................... 53

5.1 Pipeline for WBC recognition in thick blood smear images. .................. 55
5.2 Separation of the 3-channel thick blood smear image. .......................... 57
5.3 WBC Candidates Segmentation. .................................................. 58
5.4 Illustration of the bounding-boxes intersection-based methodology for lobes inclusion in WBC candidates. ........................... 59
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>WBC candidates extraction results.</td>
<td>63</td>
</tr>
<tr>
<td>5.6</td>
<td>Normalized feature weight by the linear SVM algorithm (C = 64).</td>
<td>64</td>
</tr>
<tr>
<td>5.7</td>
<td>Optimized decision tree representation.</td>
<td>65</td>
</tr>
<tr>
<td>5.8</td>
<td>Results of the sequential forward feature selection method employed on k-NN, LDA and Naive Bayes.</td>
<td>65</td>
</tr>
<tr>
<td>6.1</td>
<td>Pipeline for WBC recognition in thin blood smear images.</td>
<td>67</td>
</tr>
<tr>
<td>6.2</td>
<td>WBC nucleus candidates extraction results without and with the Jaccard-based lobe-merging methodology.</td>
<td>69</td>
</tr>
<tr>
<td>6.3</td>
<td>Two examples of the results from automatic thresholding.</td>
<td>70</td>
</tr>
<tr>
<td>6.4</td>
<td>Two examples of the gradient-restricted morphological reconstruction methodology.</td>
<td>71</td>
</tr>
<tr>
<td>6.5</td>
<td>Two examples of the results of the cytoplasm segmentation methodology.</td>
<td>72</td>
</tr>
<tr>
<td>6.6</td>
<td>WBC candidates extraction results.</td>
<td>74</td>
</tr>
<tr>
<td>6.7</td>
<td>WBC cytoplasm segmentation.</td>
<td>74</td>
</tr>
<tr>
<td>6.8</td>
<td>Normalized feature weight by the linear SVM algorithm (C = 32).</td>
<td>76</td>
</tr>
<tr>
<td>6.9</td>
<td>Optimized decision tree representation.</td>
<td>76</td>
</tr>
<tr>
<td>7.1</td>
<td>Use-case Screenshots of the MalariaScope application, with particular focus on the methodology integration.</td>
<td>82</td>
</tr>
</tbody>
</table>
List of Tables

4.1 Thin blood smear images dataset constitution ........................................ 42
4.2 Cell count results following the distance-transform based segmentation algorithm. 52
4.3 Parameter Optimization and subsequent performance ............................. 54

5.1 Thick blood smear images image dataset constitution ............................. 56
5.2 Classifiers used for WBC classification in thick blood smear images, and the processes used in order to optimize the classification performance of each learning algorithm .............................................................. 60
5.3 WBC candidates extraction results .......................................................... 62
5.4 Parameter Optimization of the parametrized classification algorithms with the entire feature vector ............................................................ 63

6.1 Parameter combination of the optimization by grid search process ............ 73
6.2 WBC candidates extraction results .......................................................... 73
6.3 Parameter Optimization of the parametrized classification algorithms with the entire feature vector ............................................................ 75
# List of Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASG</td>
<td>Average Squared Gradient</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer-Aided Diagnosis</td>
</tr>
<tr>
<td>CBC</td>
<td>Complete Blood Count</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>FEUP</td>
<td>Faculdade de Engenharia da Universidade do Porto</td>
</tr>
<tr>
<td>FN</td>
<td>False Negative</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of View</td>
</tr>
<tr>
<td>FP</td>
<td>False Positive</td>
</tr>
<tr>
<td>GLCM</td>
<td>Grey Level Co-Occurrence Matrix</td>
</tr>
<tr>
<td>INSA</td>
<td>Instituto Nacional de Saúde Dr. Ricardo Jorge</td>
</tr>
<tr>
<td>kNN</td>
<td>k-Nearest Neighbours</td>
</tr>
<tr>
<td>LDA</td>
<td>Linear Discriminant Analysis</td>
</tr>
<tr>
<td>MW</td>
<td>Moving-Window</td>
</tr>
<tr>
<td>NO</td>
<td>Nitric Oxide</td>
</tr>
<tr>
<td>NRMSE</td>
<td>Normalized Root Mean Squared Error</td>
</tr>
<tr>
<td>RATS</td>
<td>Robust Automatic Threshold Selection</td>
</tr>
<tr>
<td>RBC</td>
<td>Red Blood Cell</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Squared Error</td>
</tr>
<tr>
<td>SDK</td>
<td>Software Development Kit</td>
</tr>
<tr>
<td>SFFS</td>
<td>Sequential Forward Feature Selection</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>SVR</td>
<td>Support Vector Regression</td>
</tr>
<tr>
<td>TP</td>
<td>True Positive</td>
</tr>
<tr>
<td>UI</td>
<td>User-Interface</td>
</tr>
<tr>
<td>WBC</td>
<td>White Blood Cell</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Motivation and Context

This research work is included in an internal research project at Fraunhofer Portugal, the MalariaScope, which aims to develop a cheaper and faster but still effective alternative approach to malaria diagnosis.

In fact, according to the World Health Organization, malaria is a major public health problem, mainly in developing countries. In 2012, there were 627 000 deaths caused by malaria, 90% of them in the sub-Saharan Africa and, more disturbingly, 77% of them were children under 5 years of age. The health resources in the most affected regions are very limited, most times leading to undiagnosed deaths. Even those who manage to be diagnosed are not properly treated due to misdiagnosis or a late diagnosis. Current diagnosis methodologies confine blood sample analysis mainly through microscopic examination and this procedure is performed by highly qualified health professionals. However, in developing countries an expertise deficiency is noticed. Furthermore, the need for relatively sophisticated laboratory equipment, which are scarce in those regions, results in both lengthy and costly diagnosis.

Therefore, the Malaria Scope is being developed which aims to be a less costly alternative to current microscopic-based diagnosis while maintaining the effectiveness. This mobile-based framework comprises both a mobile-phone-adaptable cheap microscope and an automatic analysis system. The latter consists on automatically assessing the cellular blood components of blood smears as well as the number of malaria parasites, which would replace the time-consuming expertise-dependent standard procedure. Variations from normal blood components concentration could be detected, which may be symptoms of not only malaria, but also several other types of diseases. Furthermore, in malaria cases, the number of parasites in the bloodstream is calculated as a ratio of number of parasites and white blood cells (WBC). Hence, blood cell counting is crucial for malaria diagnosis. Both components of the mobile-based framework compromise the quality of the blood smear images, thus the automatic analysis system must be robust enough to handle low-quality images a provide a medically acceptable blood assessment.
1.2 Blood Cells Counting

The cardiovascular system consists of a full-body network responsible for connecting all tissue and organs of the body, which communicate through molecular signals carried by the blood. Therefore, most dysfunctions, such as a malfunctioned or unregulated tissue or organ, induce fluctuations from the normal blood components’ levels. Hence, laboratory tests of blood samples prove to be useful diagnostic tools whose simplicity relies on a standardized blood extraction procedure across health professionals.

The doctors’ expertise is required for a proper blood analysis. The most common blood test is the complete blood count (CBC). This test evaluates and counts all cellular blood components, more specifically red blood cells (RBC), white blood cells and platelets, through microscopic examination. However, this technique is time-consuming, subjective and requires a trained laboratory technician, which are scarce in malaria-affected regions.

As aforementioned, the CBC test counts and assesses the morphology of blood cells. An abnormal feature or a significant deviation from normal numbers of RBCs, platelets or WBCs, either globally or each WBC subtype, may be the symptoms of a particular disease.

The subjectivity of this test can be removed through automatic blood cell counting systems. Additionally, the blood analysis is more rapidly obtained while relieving from labour-intensive work.

In fact, since the blood analysis is repetitive labour-intensive work that relies mainly on visual examination, a growing interest in computer vision approaches has been noticed. These systems are much faster than manual approaches and have a repeatability factor that erases the technician subjectivity from the equation.

1.3 Objectives

During the development of this project, the proposed objectives are summed up below:

- Develop a robust automated image processing methodology capable of estimating the number of red blood cells in low-quality smartphone-acquired thin blood smear images;
- Develop a robust image processing methodology capable of locating white blood cells and identify as mono- or polymorphonuclear in low-quality smartphone-acquired thin or thick blood smear images;
- Integration of the image processing methodologies into mobile devices.

1.4 Contributions of this work

The proposed work had two main contributions. Firstly, it proposed two image processing methodologies for blood cell counting in low-quality blood smear images: a regression-based approach for red blood cell counting, and a methodology for detection and classification of white blood cells. Secondly, less computationally demanding, but still robust, versions of these methodologies
were integrated in the prototype application of the MalariaScope project suitable for a wide-range of mobile devices and camera resolutions.

1.5 Structure of the document

Aside from this introductory chapter, this dissertation includes 6 other chapters. In Chapter 2, the contextualization of key aspects is presented, particularly the characterization of the cellular blood components and the depiction of the microscopic diagnosis of blood films. In Chapter 3, the image processing and machine learning fundamentals of interest in this thesis work are described in detail. Chapters 4 describes the methods and discusses the subsequent results for RBC segmentation and RBC estimation count. The methods used for WBC detection and classification and the results obtained in both thick and thin blood smear images are explored in Chapters 5 and 6, respectively. In Chapter 7, the proposed methodology integration in the MalariaScope Android application is addressed. Finally, in Chapter 8, the conclusions of this dissertation work and a few guidelines for future work in the development of this project are presented.
Chapter 2

Malaria Characterization

In this chapter, the contextualization of key biological aspects relevant to this thesis is presented. Initially, the pathophysiology description of malaria will allow a full understanding of the health problem in question. Subsequently, the current practices in malaria diagnostic strategies will be presented, with particularly focus on microscopic examination of blood films. Finally, the visual features of cellular blood components, specifically the WBCs and RBCs, are summarized in order to serve as starting point for the development of the image processing methodology.

2.1 Biology of Malaria

Malaria is an infectious parasitic disease caused by parasitic protozoans of the Plasmodium type, particularly Plasmodium malariae, Plasmodium ovale, Plasmodium vivax, and Plasmodium falciparum, being the latter the most deadly human plasmodium. The malaria parasite is most commonly transmitted by a mosquito bite.

The Malaria parasite has a complex life-cycle due to the necessity to infect more than one host species (Figure 2.1). In fact, during its life-cycle, the Malaria parasite is transmitted from the mosquito to the human body, and vice-versa, while the mosquito is feeding from human blood. Consequently, the malaria plasmodium is exposed to several different environments, in the vertebrate and mosquito. Thus, its apicomplexan-type life-cycle is suited to survive under this circumstances by undergoing morphological and biochemical changes as well as different types of replication along the several stages of its life-cycle [1].

The parasite is transmitted in its sporozoite stage into the bloodstream and rapidly infects liver cells, hepatocytes [1]. Even though as little as 15 to 25 sporozoites are injected each time by a mosquito [2, 3], the development of malaria disease can still occur. Thus, considering also the fact that the injection site is distant from the host cells, the efficiency of the parasite invasion is evident. Additionally, the circulatory system direct contact with hepatocytes through the sinusoidal endothelial cells fenestrations, induces a rapid infection of sporozoites into the hepatocytes [4]. The faster the sporozoite invades the hepatocyte, the less time it is subjected to the human immune system response. Several surface proteins, particularly circumsporozoite and thrombospondin-
Malaria Characterization

Figure 2.1: The life cycle of Plasmodium falciparum. Once transmitted into the bloodstream, the plasmodium invades rapidly the hepatocytes. The plasmodium develops into a merozoite and are released into the bloodstream, rapidly invading erythrocytes. The merozoites undergo endo-erythrocytic replication, burst from the erythrocytes and rapidly reinvade previously uninfected erythrocytes. The merozoites can differentiate into gametocytes, which are taken up by bloodfeeding mosquito, where they fertilize and develop to sporozoites that migrate to the salivary glands. Image taken from [6].

related adhesive protein, play a major role in subtract-dependant motility across the sinusoidal cell layer and sporozoites adhesion and invasion of host cells, primarily due to specific recognition of surface proteins at the target site [3, 5].

Once the sporozoites invade the hepatocytes, encapsulated within a parasitophorous vacuole, the plasmodium undergoes exo-erythrocytic intrahepatic shizogonic development and henceforward reproduces asexually with no cell segmentation. In other words, a shizogonic cell cluster is formed inside the hepatocytes, which eventually mature into first generation merozoites and the host cells bursts, releasing them into the bloodstream [7].

Subsequently, each merozoite can infect a RBC and the erythrocytic stage of the malaria parasite is followed. The erythrocytic invasion requires complex and dynamic biomolecular events, specifically between the merozoite surface and the RBC membrane, which ultimately lead to parasite-host cell junction, invagination and malarial vacuole formation within the erythrocyte [8].

The merozoites invasion process of erythrocytes involves five steps, as illustrated in Figure 2.2. Firstly, parasite binding antigenic, namely the erythrocyte-binding antigenic and the reticulocyte-
2.1 Biology of Malaria

Figure 2.2: Erythrocytic Invasion by Merozoite. Merozoite surface receptors bind to the erythrocyte surface (A), merozoite reorientation until apical end is adjacent to erythrocyte (B), formation of the parasite-erythrocyte cell junction (C), cell junction moves posteriorly along the merozoite surface from the apical end (D), and vacuole sealing (E). Image taken from [8].

binding antigenic, located in any part of the merozoite surface [9], bind to erythrocyte receptors [10]. Then, merozoite reorientation occurs until its apical end is directed to the erythrocyte. Ligands located on the apical end bind to erythrocyte surface receptors, belonging to either the Duff binding-like or the reticulocyte homology family of ligands. The third step consists on the parasite-erythrocyte cell junction formation. Some proteins move to the merozoite surface and others are secreted into the intermembrane space in order to promote parasite-erythrocyte surface proteins binding and, thus, strengthen parasite-host cell junction. Once the junction is completed, the parasite invasion step, fourth step, starts. Basically, the junction moves from the anterior to the posterior end of the merozoite along the parasite membrane due to actin-myosin stimulation. Meanwhile, a vacuole forms around the merozoite throughout the junction movement. Finally, the vacuole is sealed, the erythrocyte membrane is resealed and the junction surface molecules remain on the vacuolar membrane [5].

The higher redundancy of invasion pathways of Plasmodium falciparum results in a greater percentage of RBCs it can invade, thus its deadlier nature [9, 11].

Once within the erythrocyte, the malarial merozoites proceed with the erythrocytic cycle, de-differentiating into round-shaped tropozoite prior to asexually replicating. The ring-shaped merozoite metabolizes up to two thirds of the haemoglobin content. The host erythrocyte ultimately bursts and the replicated merozoites invade other non-infected erythrocytes.

Eventually, a small portion of the parasitic cells differentiate into male and female gametocytes and remain within the RBC until it is transmitted back to a mosquito. The parasite is transmitted back to the mosquito when bloodfeeding and takes up the gametocytes. Once they leave the RBC, they differentiate into gametes and rapidly the male gamete fertilizes the female equivalent, producing a diploid zygote which develops into an ookinete. Thereafter, prior to the meiosis stage, the ookinete migrates into de basil lamina and forms an oocyst. The diploid oocyst undergoes multiple nuclear divisions into haploid sporozoites. The sporozoites invade the salivary glands of
the Anopheles mosquito, completing thereby its lifecycle [5].

2.2 Pathophysiology of the Malaria

As aforementioned, when within the human, two stages of the Malaria development occur, hepatic and erythrocytic, in chronological order.

2.2.1 Hepatic Stage

After an infected mosquito bite, the malaria sporozoites migrate rapidly through the bloodstream and invade hepatocytes. Here, sporozoites replicate, mature into merozoites and are released to the bloodstream after causing the host cells to burst.

Even though malaria causes hepatocyte infection and disruption, no significant liver dysfunction is evident in an otherwise healthy organ. Solely in cases of existing liver pathologies, its condition is worsen due to the malaria. The small number of sporozoites infected through the mosquito bite, in addition to the merozoites incapability to re-enter hepatocytes, leads to non-pathologic hepatic condition.

2.2.2 Erythrocytic Stage

Summing up the erythrocytic stage of malaria plasmodia development, the merozoites, once released to the bloodstream from the hepatocytes, infect RBCs. Here, the plasmodium asexually replicates and periodically leaves the host cells to invade new RBCs. Therefore, in addition to the plasmodium metabolism consisting on digestion of erythrocytes organelles, particularly haemoglobin, the positive-feedback replication and infection of RBCs leads to several pathologic events.

The parasite causes changes on the infected erythrocyte surface. Several adhesion proteins are exclusively expressed on the infected RBCs surface during the asexual replication stage of the plasmodium [9, 12, 13, 14]. One surface protein with particular importance to parasitic survival and overall pathogenesis is PfEMP1 [11]. The capability of this protein to bind to several host receptors leads to formation of many virulence clusters, such as resetting structures, when binding

Figure 2.3: Virulence clusters caused by infected erythrocytes due to changes in surface adhesion properties. Parasite-infected erythrocyte binding to uninfected erythrocytes (Rosetting), to platelets (Clumping) and to dendritic cells. Image taken from [11].
2.3 Malaria Diagnosis

Figure 2.4: Infected erythrocyte sequestration in microvascular endothelium. Image taken from [11].

to uninfected erythrocytes [15, 16], clumping structures, when binding to infected erythrocytes through platelets [17], and dendritic structures, when this immune cells act as binding site to other infected erythrocytes [18] (Figure 2.3). The latter structure downregulates the host immune response. This adhesion protein, associated to the knobs emergence on the erythrocyte surface, leads to changes in cytoadherence and, consequently, to erythrocyte sequestration in microvascular endothelium as either one of the above-described structures or by itself solely [19, 20, 21]. Infected erythrocytes adhere to microvascular endothelium since the blood flow and blood pressure are lowest in these blood vessels structures. The primary effect of this event is to avoid filtering of these cells by the spleen [9, 19], thus increasing virulence. Moreover, sequestration of infected erythrocytes disrupts blood flow, causing tissue hypoxia and lactic acidosis [9], particularly in the brain (cerebral malaria) and placenta, as illustrated in Figure 2.4.

Eventually, infected erythrocytes burst, releasing both its content and merozoites into the bloodstream. Peaks in erythrocytes bursting causes anaemia, arising symptoms of fever and rigours [9].

Additionally, free haemoglobin resulting from haemolysis contributes to intravascular endothelial injury and dysfunction. These pathological effects are due to haemoglobin catalysing oxidative damage and consuming nitric oxide (NO), a molecule that maintains endothelial homoeostasis [22]. Additionally, free haemoglobin binds to haptoglobin, a protein which inhibits the oxidative activity of the haemoglobin, and the resulting complex is consumed by macrophages. The metabolism of this complex results in carbon monoxide and ferrous iron, the latter causing proinflammatory effects [23].

2.3 Malaria Diagnosis

Effective and accurate diagnostic strategies are crucial for managing pathologies, and malaria presents no exception. In the malaria-high-incidence resource-limited areas aforementioned, interest is in rapid but still effective diagnostic techniques that do not demand expertise, while in developed countries the methodologies adopted do not have these constraints.
2.3.1 Routine Laboratory Diagnosis

Currently, the microscopic examination of blood films is the most adopted diagnostic methodology. The blood, usually stained with Giemsa stain, is obtained from the finger or earlobe since this microvasculature/capillary-rich area is where malaria plasmodium are highly concentrated [24, 25]. Depending on the diagnosis interest, whether it is desired to identify the plasmodium species or to obtain a more sensitive estimation of plasmodium density within the blood, two types of blood film can be used: thin blood film and thick blood film, respectively. Nonetheless, the non-standardized blood film preparation procedure across technicians limits comparability of parasitemia estimations [26].

2.3.1.1 Thin Blood Film

The overall preparation procedure is very straightforward and is completed in few minutes. Briefly, one drop is dispersed over the slide and, after blood film fixation with methanol, the slide is stained with either Giemsa or Wright’s stain. Emphasizing the stains role in microscopic examination, their use allows differentiation of different cellular structures, particularly in malaria diagnosis the plasmodium insertion on RBCs. The standard use of 100x objective coupled with 10x eyepieces effectively limits sensibility to estimate the parasite density, even though percentage of infected red blood cells can still be estimated. However, thin blood films enhance specificity of plasmodium identification, which is crucial for preventing and foreseeing complications resulting from the specific type of plasmodium infected, particularly the deadlier species, Plasmodium falciparum [27].

2.3.1.2 Thick Blood Film

The blood film preparation procedure of thick blood film differs slightly from the thin analogues, particularly on its essence. Since better sensibility on parasitemia estimation is desired, the blood drop is not dispersed over the slide, the film is not fixed and the stain is diluted. Consequently, the increase concentration of RBCs per optic field enables easiness to see and count the parasites, even with equal magnification as the thin blood film examination. Usually, 100x objectives are used in the microscopic visualization of thick blood film. However, unfixed erythrocytes locate in different focal planes, thus, plasmodium species microscopic identification is not conceivable for this preparation type.

The number of parasites estimation is calculated as a ratio of number of parasite and WBCs, on the assumption of WBCs standard density of 8000 units/µL or a previous assessed WBC count. A major disadvantage, however, is the RBCs lysis during staining due to not being fixed to the slide. Therefore, in addition to disposal of several layers, approximately 6 - 20 layers thick, in this type of blood smear preparations, the impossibility of assessing the RBC count is evident and this is the reason for measuring solely WBCs in thick blood films [27].
2.4 Blood Cells Morphology

Mary Louise Turgeon, in her published book [28], described in detail the haematopoiesis, the formation of cellular blood components process. Additionally, the author describes the morphology of RBCs and WBCs (neutrophils, eosinophils, basophils, lymphocytes and monocytes).

2.4.1 Haematopoiesis

All types of blood cells derive from hematopoietic stem cells located in the bone marrow within long bones. This pluripotent stem cell may either self-regenerate by replicating without differentiating, or mature along the different cell lines, depending on the conditioning stimuli and mediators. The haemotopoiesis is modelled in Figure 2.5. There are two hematopoietic cell lines: lymphoid and myeloid cell line.

In the lymphoid cell line, the pluripotent stem cells undergo lymphopoiesis, originating one type of WBCs, the lymphocytes. The lymphoid precursor cells differentiate into cells capable of either expressing cell-mediated immune responses or secreting immunoglobulins, and it depends immensely on the maturation location. In case the lymphoid precursor cells differentiate in the bone marrow, they give rise to B cells, whose main function is to produce antibodies against antigens. In the other hand, if the lymphoid precursor travels to the thymus gland, specific molecular signals lead to differentiation into T cells, which play a major role in cell-mediated immunity. Additionally, there are morphologic differences between both lymphoid cell types. Generally, T cells tend to be smaller lymphocytes while the larger ones are B cells.

In the myeloid cell line, the hematopoietic stem cell differentiates into myeloid progenitor cell line, common to RBCs as well as several types of WBCs. The myeloid progenitor cells may either differentiate into granulocyte-macrophage progenitor or megakaryocyte–erythroid progenitor cells. The former may either undergo granulopoiesis deriving in polymorphonuclear leukocytes (neutrophil, eosinophil and basophil), or monocytopoiesis leading to monocytes. On the other hand, megakaryocyte–erythroid progenitor cells may either undergo trombopoiesis, leading to the formation of platelets, or erythropoiesis, producing RBCs.

During the monocytopoiesis, the monocyte progenitor cells undergo through a series of morphological modifications until it matures into monocytes. Particularly, the large and round nucleus and unorganized segmented nucleoli turn to a smaller kidney-shaped nucleus with chromatin network consisting on fine linear threads. In addition, it size range diminishes from 15-25 µm down to 14-18 µm in diameter. The cytoplasm may also be vacuolated in the mature monocytes. As above-mentioned, granulopoiesis originates three diverse polymorphonuclear leukocytes: neutrophils, eosinophils and basophils. The myeloid progenitor cells differentiates into myeloblasts, a round shaped 20-25 µm diameter cell which is the precursor in the granulocytic series. The nucleus is large and round with dispersed chromatin within it. Subsequence segmentation lead to progressively differentiated cells, differing in eosinophilic, basophilic and neutrophilic granulocytes, described below. Moreover, myeloid progenitor cells undergo erythropoiesis, forming RBCs. The large and round nucleus becomes progressively smaller and the initially dispersed and indistinct
Figure 2.5: General model of haematopoiesis. All blood cells derive from the hematopoietic stem cells (HSC), which can either replicate or differentiate into two progenitor cells: common myeloid progenitor (CMP) or common lymphoid progenitor (CLP). The former eventually differentiates into platelets, erythrocytes, monocytes and the granulocytes (neutrophils, eosinophils and basophils), while the latter differentiates into lymphocytes (T Cells, B Cells and N Cells). Image taken from [28].
nucleoli becomes denser, until it is extruded in the reticulocyte stage of the erythrocyte maturation. The cytoplasm becomes basophilic and the haemoglobin accumulates within the anucleated cell. Reticulocyte ultimately mature into erythrocytes, with 7-8 µm diameter and biconcave shape. Finally, thrombopoiesis originates platelets, blood cells which play a major role in coagulation. Megakaryocyte–erythroid progenitor cells, particularly megakaryocytes, duplicate with segmentation. The mature multiple-nuclei megakaryocyte ultimately sends cytoplasmic projections and shed platelets into the bloodstream.

Figure 2.6: Giemsa stained blood smear. WBCs marked with colorant: basophil (b), eosinophil (e), lymphocyte (l), monocyte (m), and neutrophil (n); arrows indicate platelets; others elements are RBCs [28].

2.4.2 Red Blood Cells

Red blood cells are the most numerous in the blood, accounting for 4.5–5.9 \( \times 10^6/\mu L \). These non-nucleated blood cells are flattened and biconcave-discoid shaped with a 7.0-8.0 µm diameter and 1.7-2.4 µm thickness. In Giemsa stained smears, the erythrocytes appear circular, due to the flattened surface being faced up, with a pale staining in the central area occupying one-third of the cell area, corresponding to the indented regions, as illustrated in Figure 2.7.

2.4.3 White Blood Cells

White blood cells are a group of nucleated cells in the bloodstream (about 5’000-10’000 WBC/µL) responsible for the body’s defence against both infectious disease and foreign invaders. They derived from hematopoietic stem cells located in the bone marrow and differentiate into 5 distinct cell types with distinct morphologies and functional roles: neutrophils, eosinophils, basophils, lymphocytes and monocytes, illustrated in Figure 2.8.
Figure 2.7: Erythrocyte in a Giemsa stained thin blood film. Image acquired at Instituto Nacional de Saúde Dr. Ricardo Jorge, using a smartphone coupled to a microscope with the Sky Light adapter.

Figure 2.8: Five different types of leukocytes: neutrophil (A), eosinophil (B), basophil (C), monocyte (D) and lymphocyte (E). Images taken from [28].

The first three types of WBC are polymorphonuclear leukocytes. They have an unique lobed nucleus, small granules in the cytoplasm and differ according to their staining reactions:

- **Neutrophils**: this cell type has 10-12 µm diameter and its lobed nucleus stains purple. The light pink stained cytoplasm has pink granules. 2’000-7’500 neutrophils/µL. The neutrophilic series is illustrated in Figure 2.9;

- **Eosinophils**: this leukocyte type has 12-14 µm diameter and its two-lobed nucleus stains a little paler than the neutrophils’ nucleus. Moreover, the granules in the cytoplasm are larger than those of neutrophils and have a round/oval shape. 40-400 eosinophils/µL. The eosinophilic series is illustrated in Figure 2.10;

- **Basophils**: this cell type has 10-12 µm diameter and its nucleus has a kidney-like shape. In addition, the large granules stain purple/blue. 20-200 basophils/µL. The basophilic series is illustrated in Figure 2.11.
Lymphocytes and monocytes are mononuclear leukocytes. In other words, they have a one-lobed nucleus and its cytoplasm is absent of granules.

- **Lymphocytes**: this mononuclear leukocyte has 7-10 µm diameter with a large round nucleus eccentrically located in the cell. Furthermore, the nucleus, which occupies most of the cell, stains deep-purple and the cytoplasm observed stains pale blue. 2'500 lymphocytes/µL;

- **Monocytes**: this cell type has 14-18 µm in diameter with a large and centrally located nucleus, which is ‘horseshoe’ shaped and stains pale violet. In addition, the cytoplasm stains pale blue. 700-1’500 monocytes/µL. The monocyte maturation series is illustrated in Figure 2.12.

**2.4.4 Relevant Remarks**

As seen in Figure 2.6, the cellular blood components have distinctive features in blood smear images, which proves to be useful in identification of the cell type in computer vision approaches. While the platelets are identified by the significantly smaller size, the RBCs are segmented by their red staining, pale red in the centre and the absence of nucleus. The remaining cellular components are WBCs, which are clearly differentiated within the WBC types by the different size and shape of the nucleus and granularity of the cytoplasm. However, as seen in Figure 2.8 through Figure 2.12, the different types of mature WBCs have common progenitors, which may cause incorrect identification of the cell type. Moreover, the agglomeration and formation of cell clusters may hamper the segmentation and identification methodology in computer vision approaches.
Chapter 3

Literature Review

3.1 Technology Review

This thesis aims to develop an image processing methodology for blood cell counting which can be suited to mobile devices. Thus, some technologic aspects need to be considered during the development of the thesis, which are described in more detail below.

3.1.1 Mobile Operating System

The image processing methodology aims for mobile operating systems. These operating systems combine features of personal computer operating systems with many other features such as sensors, GPS, Wi-Fi, mobile communication system and a user-interface consisting solely on touchscreen. In this Master thesis context, the camera features are particularly important. In this subsection, the three most widely used mobile operating systems are succinctly described: Android, iOS and Windows Phone.

3.1.1.1 Android

Google developed an operating system, Android, designed for touchscreen mobile devices. It has been regularly updated for improvements on the operating system itself, addition of new features or bug-fixing. According to a study conducted by mobiForge (Global Mobile Statistics 2014 Part A: Mobile Subscribers; Handset Market Share; Mobile Operators), Android operating system had a 78.6% market share in the smartphone market in 2013 and is expected to remain the leading smartphone operating system through to 2017. The Android’s dominating position in the mobile devices market is mostly due to its open source code and licensing. It grew interest from technology companies, developers and enthusiasts to whom were made easier to develop and market applications for mobile devices. Even though the Android Software Development Kit is the development environment most usually used by developers, there are other development platforms which run on Linux, Windows or Mac operating systems, adding up to the Android’s interest...
from developers. However, the limited hardware capability of current smartphones restrains the applications’ complexity.

3.1.1.2 iOS

Apple Inc. developed a mobile operating system, iOS, exclusively for Apple devices and it was unveiled when the iPhone was launched in 2007. Moreover, it is updated annually. According to a study conducted by mobiForge (Global Mobile Statistics 2014 Part A: Mobile Subscribers; Handset Market Share; Mobile Operators), iOS had a 15.2% market share in the smartphone market in 2013. Third-party applications can be developed exclusively on Xcode, the development environment for the iOS Software Development Kit, and are mostly written in Objective-C. Even though the iPhone, and subsequently the iOS, played a major role in the smartphone revolution, the restraints to third-party applications development led to the loss of the leading position in the smartphone market. Such restraints include paying an iPhone Developer Program annual fee for loading an application onto the iDevices, paying 30% of the application revenue to Apple and strict criteria for approval of the applications.

3.1.1.3 Windows Phone

Microsoft developed Windows Phone, an operating system for smartphones, particularly Nokia devices. According to a study conducted by mobiForge (Global Mobile Statistics 2014 Part A: Mobile Subscribers; Handset Market Share; Mobile Operators), Windows Phone had a 3.3% market share in the smartphone market in 2013. The small market share is mostly due to the late entry into the smartphone market. However, the excellent development tools are favourable to the developer community. Third-party applications are written in C#, Visual Basic.NET or C++ and can be designed exclusively on Windows operating systems, either within Visual Studio or Windows Phone Developer Tools. Nevertheless, the applications undergo an approval process and 20

3.1.2 Computer Vision libraries

Both image processing and machine learning algorithms can be very complex to implement and most were already established before. Thus, in order to avoid the time-consuming task of coding these algorithms, computer vision libraries are available. OpenCV and FastCV are succinctly described in this subsection.

3.1.2.1 OpenCV

OpenCV is an open source library of image processing functions, whose goal is real-time computer vision. Even though OpenCV is written almost solely in C++, it still is platform independent since it has interfaces in C++, C, Python, Java and MATLAB and supports Windows, Linux, Mac OS, iOS and Android.
3.2 Related Work

3.1.2.2 FastCV

FastCV is a computer vision library which aims real-time computer vision applications in mobile devices, specifically Android operating system. Future versions are expected to support iOS and Windows Phone smartphones. Additionally, FastCV consists on low level computer vision function optimized for use by ARM architectures, hence, improving the functions performance and speed.

3.2 Related Work

Computer vision methodologies have been developed to help improve not only image quality of blood smears but also the effectiveness of its analysis. Particularly, image processing proves to be advantageous to blood cell counting since it provides an experience-independent and less time-consuming analysis.

Image segmentation is an inevitable step in blood cell counting, specifically in detecting the types of blood cells of interest. Additionally, computer-aided image processing systems often require a preprocessing step before image segmentation is performed. The success of the designed system depends immensely on the efficiency of the preprocessing step and the type of preprocessing approach must be congruent to the following image segmentation methodology. In other words, preprocessing consists on enhancing the overall image quality and specific features of interest, the latter which would be the focus in the subsequent process. In blood cell counting, all cellular blood component, and even some key blood compounds (haemoglobin) are regarded. Firstly, the state of the art of RBCs detection, segmentation and counting methodologies is addressed. Subsequently, recent advances in WBCs segmentation, and ensuing differentiation, is presented.

3.2.1 Red Blood Cells Detection

In the preprocessing step, studies follow different approaches for image enhancement, particularly noise-removal and illumination inhomogeneity. Haider Adnan Khan et al. [29] followed Contrast-Limited Adaptive Histogram Equalization to get a nonuniform-illumination rectified image. Other approaches are as basic as mean or median filtering applied to either the RGB or gray-scale images, in order to remove spectral noise and other artefacts followed by colour conversion from RGB to YCbCr to overcome illumination issues. Morphological operations were also used to remove salt and pepper noise, specifically opening and closing operations [30].

Then, red blood cell segmentation follows this first step.

3.2.1.1 Red Blood Cells Segmentation

Several challenging occurrences must be accounted during the development and selection of the image processing methodology to use, such as clusters of overlapping erythrocytes or the existence
of other blood cell types (leukocytes or platelets) in the blood smear. In the literature, several diverse methodologies were followed.

A. Sai Prasad [31] proposed a straightforward distance-transform based methodology for erythrocyte segmentation. Initially, the pre-processing step consists on transforming the input image to gray-scale, followed by Canny edge detection operator, a very consensual edge detection algorithm within the community, with the intent of finding the contours of the RBCs. Then, morphological operations were processed in order to remove noise while maintaining the essential shapes of the objects. Following the preprocessing step, the author’s goal was to separate the platelets and WBCs from the RBCs. The extraction of geometrical features, mainly the area of the blood cells, and relative comparison to the blood cell with the highest value, leukocyte, enabled the separation of the three different types of blood cells. Finally, the clumped erythrocytes are split through using euclidian distance transform followed by watershed segmentation.

However, despite the implementation simplicity, two major drawbacks are observed. Firstly, the separation of the blood cells in their three types is prone to failure, particularly when the WBCs in the image do not have significantly larger areas than the other cellular blood components. Moreover, the possibility of round erythrocytes clusters having equivalent areas to WBCs was not addressed, but from the implementation standpoint, similar failures would occur [31].

Hassan Khajehpour et al. developed a distance-transform based methodology for erythrocyte segmentation which accounts for the oversegmentation problem obtained through implementation of the previous image segmentation system [32]. A mean filtering on gray-scale image, followed by a morphological operation sequence erosion-dilation-reconstruction, results in an enhanced image which preserves the blood cells margins while removing the bright spots inside them as well as other artifacts. Then, after computing the distance transform on the binary images, a line operator was applied for detection of optic disc structure whose centre is brighter than in the surroundings. In other words, the brightness of the structure is gradually reduced from its centre to surrounding. Basically, the optic disc structure is composed of 20 line segments in various direction, and each accounts for the image variation along the line. The direction which has maximum and minimum variation can be useful in cell centres detection. In this way, an "orientation map" image is obtained. After conversion to a binary image, the centre areas are obtained through convolution with a mask suitable for peak detection. Finally, to eliminate full connections between cell centres that might still appear, a first-order derivation of Gaussian filter and binary operations are applied. The resulting cell centres are superimposed on the resulting image from a previous step and act as local minima point in the watershed algorithm, resulting in segmented erythrocytes. Even though this proposed method resulted in lower oversegmentation and increased accuracy and sensitivity, it is computationally slow and still has some limitations in overlapping cells. In addition, it did not distinguish erythrocytes from the other blood cell types [32].

Haider Adnan Khan et al. proposed a novel method, "Distance Mapping", that combines both the distance transform map and the template-based pattern matching [29]. After an initial histogram equalization, obtaining a highly-contrasted nonuniform-illumination-rectified image, fol-
3.2 Related Work

Followed by global thresholding using the Otsu’s method, the distance transform is applied on the binary image. Then, in order to identify the cell centroids, template matching with a distance-transformed circular disk template was performed using normalized cross-correlation. The resulting image was complemented and, finally, cell segmentation achieved with watershed transform. In fact, this proposed method resulted in higher specificity in cell detection than distance transform methods. However, the authors did not account for the existence of other cell types within the blood smear images that may make this methodology vulnerable [29].

Another approach very commonly used in cell segmentation is the Hough Transform. Venkatalakshmi Balakrishnan et al. develop a Hough transform based method for RBC segmentation [33]. After image pre-processing, consisting on analysing the saturation component of the HSV-converted image, double thresholding logical operation was applied in order to extract the RBCs from the other blood cell types and background. Then, the circular Hough transform is used to find circular patterns within the image, corresponding to the RBCs borders, and the resulted image highlights the cell centres. Despite the simplicity of this method, there are major drawbacks. Firstly, the RBCs separation from the other cell types is very rudimentary, proving to be ineffective in extracting the WBCs with similar diameter. Additionally, any partial circular boundary results in circle detection, thus, background spaces circled by RBCs are false segmented [33]. In fact, Nasrul Humaimi Mahmood et al. developed a conceptually similar methodology to the above-stated [34]. Mausumi Maitra et al. developed a cell segmentation methodology which differed only in the pre-processing step. However, it had alike drawbacks in cell segmentation [35].

Antonio Garrido et al. developed a much more robust and complex methodology. Firstly, a generalized variation of the Hough transform is used in order to get a larger region in which the cell centre is located. Then, a deformable circle, approximated as an octagon disk, is used and iteratively deformed in order to fit the cells contours [36].

3.2.2 White Blood Cells Detection

As abovemented, in CBC, it is of major importance to not only identify and count WBCs but also differentiate within its 5 types: neutrophil, monocyte, eosinophil, basophil and lymphocyte. Thus, most cell segmentation methodologies, whose purpose is to count WBCs differentially, remove background and RBC information at first, and subsequently segment each leukocyte type. The latter consists on computerized pattern processing which confines determination of the region of interest, feature extraction either in the frequency or spatial domain, and a pattern recognition algorithm. Preprocessing techniques can be as removing artifacts by applying median/mean filters as well as morphological operations, such as opening operator. Moreover, cropping the regions of interest is of major importance for the following feature extraction step. Most notably, edge detection algorithms, such as Canny edge detection and Marr and Hildreth edge detection algorithms, are used to extract these regions.
3.2.2.1 White Blood Cells Segmentation

Cell segmentation proceeds the pre-processing step, and enables not only to segment the WBCs and, if desired, their immediate surroundings, but also exclude other blood cell types, specifically erythrocytes and platelets, from further analysis. Additionally, WBCs segmentation enables to extract meaningful features. Several approaches can be followed in this step. Due to the higher intensities of the leukocytes nucleus, when converted to grayscale, authors have applied a simple global thresholding. The major complexity of this method is to correctly select the proper threshold [37]. Additionally, three different grayscale values range can be identified in blood smear images (background, cytoplasm and nucleus), which may assist selecting a threshold value that segments both the nucleus and the cytoplasm. Techniques which may be used are histogram analysis for detection of the three different peak values or multiple gray level thresholding.

Van der Heijden [38] developed a WBC segmentation method whose major drawback was erythrocyte segmentation due to having similar pixel values to cytoplasm of WBCs. However, this issue can be overcome. Due to the nucleus presence in WBCs and easiness to segment, the cytoplasmic regions should only be considered in its immediacy. Another methodology, developed by Firdaus Ismail Sholeh [39], showed promising results in nucleus and cytoplasm segmentation of WBCs. After noise reduction and elimination of RBCs, platelets and cytoplasm, the WBCs nucleus are segmented, through thresholding the green channel of the original RGB image and morphological opening. Then, a region growing algorithm enables reconstruction of the cytoplasm of the respective nucleus. Finally, the contours are smoothed through morphological closing and opening operations. Vincenzo Piuri et al. [40] performed leukocytes detection through a set of steps which included nuclei detection and opening morphological operations, before thresholding.

3.2.2.2 Feature Vector Construction

Once the leukocytes are segmented, features must be obtained so that the WBCs can be differentiated. Such features include geometric properties of the binarized images, colour, gradients and texture features. As abovemented in the previous chapter, the main differentiation feature of WBC types is the nucleus, specifically its size, shape and colour. Furthermore, the cytoplasm also differs in colour and ratio with the total leukocyte area. In fact, Xubo Song et al. [41], in the feature vector construction step of their developed methodology for leukocyte differentiation, emphasized exactly on these features, particularly nucleus and cytoplasm areas and areas ratio, nucleus perimeter, colour and other texture-based features, indicative of the granularity of the cytoplasm of the segmented cells. The authors claimed above-80% WBC classification accuracy. Vincenzo Piuri et al., in his work [40], extracted biological relevant morphological features of leukocytes, particularly from the membrane, cytoplasm and nucleus.

3.2.2.3 Feature Selection

In order to obtain an efficient classification system, only the relevant features should be considered. In fact, the number of features that can be extracted from the segmented image is usually
very large, which, if the entire set is used, leads to computationally complex systems and poor classification results. While the former is an obvious consequence, the latter results from the inclusion of redundant features, which may either provide irrelevant information or may be highly correlated with other features. In addition, too many features result in overfitting classifiers. In other words, the classification system exaggerates minor fluctuations in the training set, leading to a poorer predictive performance (worse generalization). This erroneous behavior occurs when there are many features relative to the number of training examples. Thus, feature selection and feature extraction prove to be a crucial step in designing an accurate classification system.

Feature extraction consists on reducing the initial features vector by removing redundant information, and thus obtaining a reduced representation of the initial set with relevant information. There are many dimensionality reduction techniques and the most popular is Principal Component Analysis. This technique converts a set of possibly correlated variables into a set of linearly uncorrelated variables by using an orthogonal transformation based on the calculation of the eigenvectors of the covariance matrix of the data.

Feature selection consists on selecting a subset of relevant features for designing the classification system. There are three different feature selection methods: filter methods, embedded methods and wrapper methods. While the former, filter method, consists on ranking each feature individually and independently, the latter two methods are computationally more exhaustive since they require a classification model along with an evaluation measure. Feature selection in embedded methods is performed as part of the classification model construction model. The Recursive Feature Elimination is an example of an embedded feature selection method commonly used with Support Vector Machine. In the other hand, wrapper methods consist on iteratively finding a subset of features which results in the best classification performance using a specific classifier. The performance of the predictive model is usually assessed through cross-validation and the most common feature selection methods are exhaustive search, where all possible feature subset combinations are evaluated, sequential forward selection, where an unselected feature which most improves the accuracy is added to the feature subset, and sequential backward selection, where features are iteratively removed from the feature set that cause the most accuracy improvement.

In the literature, Vincenzo Piuri et al., in his work [40], carried out feature selection with the forward selection technique, using the nearest neighbor classifier and the performance was evaluated with the leave-one-out method, a variant of K-Fold Cross-Validation method where K equals the number of samples of the dataset.

3.2.2.4 Classification

Finally, once defined the feature vector, pattern recognition algorithms classify the WBCs. These algorithms rely on past experience to classify a certain input instance, in this case a leukocyte. Thus, pattern recognition algorithms must previously "learn" from supervised examples to classify WBCs. This process is designated as "Supervised Learning" which requires a classified dataset and consists of three main steps: division of the dataset into a training set and a test set, train the algorithm by producing a classifier using solely the training set, and use the resulting classifier on
the test set. The performance of the classifier is translated as the number of correct predictions of the samples of the test set. The selection of the algorithm must be appropriate according to the developer’s objectives. In WBC differentiation, the usual approach consists on classification algorithms as well as clustering algorithms since the aim is to label the leukocytes.

Vincenzo Piuri et al. developed an automatic morphological method to identify the leukocytes from other blood cells and then classify them [40]. In his work, different classifiers were tested, namely the k-nearest neighbours, feed-forward neural networks for each class and radial basis function neural networks with different centres. The classifiers’ performance was assessed through mean miss-classification error, using a cross validation technique for evaluation of the system accuracy. In fact, feed-forward neural network had the best classification performance [40]. In a subsequent study, Fabio Scotti tested the linear Bayes Normal classifier but the feed-forward neural network still had the lowest mean miss-classification error [42].

Another author who, in the conceptualization of an automated bone marrow WBC classification system, compared the performance of two classifiers was Nipon Theera-Umpon [43]. In his work, the classification problem was to label the WBCs in myelocytic series and the authors based solely on features extracted from the nucleus. However, some considerations can still be withdrawn for this thesis. Two classifier were used, Bayes classifier and neural network classifier, and the latter resulted in better classification performance. The performance was assessed through a cross-validation technique and all features extracted were used. Additionally, the author tested naive Bayes and decision tree classifiers but resulted in worse classification performances [43].

Another classifier, whose interest in WBC segmentation has recently grown, is Support Vector Machine. Daniela Mayumi Ushizima et al. investigated the use of Support Vector Machine classifiers to classify WBCs [44]. After segmentation and subsequent feature extraction of leukocytes, several strategies of multiclass Support Vector Machine algorithms were trained and tested on a previously expertise-labelled dataset. All Support Vector Machine strategies produced similar accuracy rates. They are also compare to Naive Bayes and tree-based classifiers and, while the former produced significantly lower accuracy rates, the latter produced similar confidence levels to SVMs classifiers. In fact, the tree-based technique did not favor one or more classes, in contrast to the SVMs classifiers, and generate a more compact multiclass classifier, important requirement for real time systems [44].

### 3.3 Image Processing Techniques

Image processing for computer vision consists on modifying the images, using mainly mathematical operations, in order to enhance them or extract a set of characteristics or parameters of interest. Through this section, some theoretical fundamentals of image processing operations used in the state-of-art, described in Section 3.2, and crucial for the development of this work are presented in this section.
3.3 Image Processing Techniques

3.3.1 Image Enhancement

Image enhancement is a pre-processing operation which aims to improve the image data, particularly enhance features useful in subsequent operations and suppress undesired ones. There are two types of image enhancement approaches depending on the domain the enhancement is processed. In the spatial domain, methods directly manipulate the pixels in the image while, in the frequency domain, methods consist on modifying the frequency-transformed image, most commonly using the Fourier transform.

Briefly, regarding image enhancement in the frequency domain, methods consist mainly on applying filters to the Fourier transform image. The filter is designed according to the aim of the preprocessing. Smoothing images is achieved through suppressing the high frequencies while sharper images are obtained when using filters which attenuate the low frequencies.

In the spatial domain, the methods are depicted in Equation 3.1. In this expression, $T$ is the operation applied to the original image that transforms the pixel value $f(x,y)$ of the original image to $g(x,y)$. The several spatial image enhancement techniques differ on the transformation operation.

$$g(x,y) = T[f(x,y)] \quad (3.1)$$

The most simple techniques consist on gray-level transformation functions which transform the input pixel value into an output pixel value. These operations are typically used for contrast enhancement and are elected according to the specific application.

Logarithmic or fractional power-law transformations are used for expanding lower gray pixel values into a wider range, while compressing the higher values. The opposite occurs when inverse logarithmic or power-law, with and higher-than-1 exponent, transformations are used.

At first, linear operations may seem redundant. However, the interest in these operations arise when used in piecewise approaches. This approach consists on splitting the entire gray-level range into $N$ pieces and independently apply a transformation operation to each piece, according to the aim of the pre-processing operation. The complexity of this method is highly variant since it comprises operations from as simple as two-piece linear stretching to complex multiple-piece with logarithmic or power-law transformations. This type of approaches are mostly used for contrast-stretching purposes and for increasing the dynamic range of the input image.

Considering a 3-piecewise transformation of linear stretching functions, depicted in Figure 3.1, each piece is transformed as described in Equation 3.2.

$$I_N = (I - \text{Min}) \frac{\text{newMax} - \text{newMin}}{\text{Max} - \text{Min}} + \text{newMin} \quad (3.2)$$

This procedure is useful for contrast enhancing a particular pixel values range while attenuating, but still maintaining, the remaining ones. If the interest is exclusively in a particular brightness range, this is stretched to the full range of available grayscale values while the meaningless ranges can be completely suppressed by setting them to either the minimum or maximum values. The most simple method is the minimum-maximum linear contrast stretching operation that uniquely
stretches the input dynamic range to the full grayscale available values. On the other hand, the percentage linear contrast stretching operation uses specified minimum and maximum values according to the either the mean or maximum of the histogram. This method may lead to suppression of some brightness values of the input image.

Another histogram-based image enhancement operation is histogram equalization. Following the computation of the image histogram, the transformation function is obtained which is the accumulated normalized histogram of the image. This method also enhances the global image contrast by provided a better distribution of the brightness values and, unlike the previously described operations, does not require user inputs (such as the percentage of the mean of the histogram).

Finally, other image enhancement operations worth mentioning in this thesis context are achieved through spatial filtering. In these cases, the transformation operator $T$ in Equation 3.1 is defined over some neighbourhood of $(x,y)$. The mask, centred at $(x,y)$, moves from pixel to pixel, the operator $T$ is applied over that subimage of equal dimensions as the mask and the output value is set at $g(x,y)$. The output is given by the sum of products of the mask coefficients and the corresponding subimage pixels. The type of mask, and consequently its coefficient values, depend on the application purpose.

Smoothing masks are used most commonly for noise reduction. This can be achieved by either using linear or nonlinear mask. The former leads to image blurring and consequently loss

![Figure 3.2: Two 3 × 3 smoothing filter masks. (a) Average of the grayscale values under the mask; (b) Weighted average of the grayscale values under the mask](image-url)
of sharpness of the edges. Such examples are depicted in Figure 3.2. Therefore, the degree of blurring and noise reduction must be thoroughly balanced. On the other hand, nonlinear filters, such as the order-statistics filter, are suitable for noise reduction, particularly of the salt-and-pepper type, while maintaining most of the image sharpness. The most commonly used is the median filter which consists on the ranking the neighbourhood pixels and setting $g(x, y)$ to the median value.

On the other hand, edges and other graylevel sharp transitions can be enhanced through sharpening filters, which consist on spatial differentiation of the pixels within the neighbourhood. The output value of the sharpening operation using derivative kernels is proportional to the discontinuity within the set neighbourhood. There are several sharpening masks within each of the following types: first-order image derivatives and second-order image derivatives.

The first order derivatives is of great interest in this thesis context. It consists on the implementation of the gradient on a discrete domain and is defined as a two-dimensional vector as shown in Equation 3.3.

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

(3.3)

The magnitude $G$ is given by Equation 3.4 and the orientation is given by 3.5.

$$G = \sqrt{G_x^2 + G_y^2}$$

(3.4)

$$\theta = \tan^{-1} \frac{G_y}{G_x}$$

(3.5)

Due computationally expensive nature of implementing Equation 3.5, the gradient magnitude is commonly approximated summing the absolute values of both gradient components or using the highest absolute value.

There are several first order derivative masks, of both direction, used for image sharpening spatial filtering. A few examples of the most commonly used masks are shown in Figure 3.3. This operation highlights edges and small sharp details of the input image while suppressing the constant or slowly variant regions across the image.

### 3.3.2 Morphological Image Processing

Mathematical morphology is a technique useful for image analysis and processing, particularly of geometrical structures which was initially aimed at binary images. Morphological transformation has similarities with spatial filtering described in Section 3.3.1 since these operations consist on the relation of each image pixel with a set neighbourhood defined by a structuring element. Unlike the kernels used in spatial filtering, the coefficients of the structuring images are either 0 or 1. Morphological operations have several applications such as image filtering by removing small and irrelevant details, segmentation of objects and enhancing structures.
The two lower-level morphological operations are dilation and erosion. The common dilation operation uses isotropic structuring elements for expanding the shapes uniformly in all directions. This operation consists on the dilation of the input image, $A$, by the isotropic structuring element, $B^i$, and is mathematically described in Equation 3.6.

$$A \oplus B = \{z \in E \mid (B^i)_z \cap A \neq \emptyset\} \quad (3.6)$$

The symmetry of the isotropic structuring element is assured according to Equation 3.7

$$B^s = \{(x,y) \in E \mid (-x,-y) \in B\} \quad (3.7)$$

The dilation of $A$ by $B$ is the set of all $x$ displacements such that the reflection of $B$ (180° rotation of $B$ around the origin) and $A$ overlap by at least one element. This operation can be visually comprehended as setting to foreground every pixel covered by the structuring element $B$ when its centre is located on a foreground pixel inside image $A$.

On the other hand, the erosion operation aims to compress the shapes uniformly in all directions. The erosion of the input image, $A$, by the isotropic structuring element, $B^s$, is mathematically described in Equation 3.8. The erosion of $A$ by $B$ can be visually comprehended as setting to background every pixel when $B$ centred in it does not cover exclusively foreground pixel.

$$A \ominus B = \{z \in E \mid (B^s)_z \subseteq A\} \quad (3.8)$$

The opening and closing operators are two other important morphological operators which derive from the lower-level operations of erosion and dilation. The opening of the input image $A$ by the structuring element $B$ is the erosion of $A$ by $B$ followed by the dilation of the eroded image by $B$, and is mathematically described in Equation 3.9. This operation leads to three structural
changes in shapes included in image $A$: the foreground structures smaller than $B$ are eliminated, narrow foreground connector are broken, and the contours of the objects are smoothed. While the former two effects are direct results of the erosion step of the opening operation, the latter is a result of the dilation step, as the eroded structures expand back to approximately the original size.

$$ A \circ B = (A \ominus B) \oplus B $$ (3.9)

On the other hand, the closing operation of the input image $A$ by the structuring element $B$ is the dilation of $A$ by $B$ followed by the erosion of the dilated image by $B$, and is mathematically described in Equation 3.10. Contrarily to the opening operation, the closing of an input image leads to filling holes and contour gaps smaller than $B$, fusing of narrow breaks and adjacent structures, and smoothing of objects contours.

$$ A \bullet B = (A \oplus B) \ominus B $$ (3.10)

As aforementioned, the mathematical morphology was initially aimed at binary images. However, another mathematical representation of the dilation and erosion operations enables to use this operators in grayscale images. Therefore, dilation and erosion can be also defined as the maximum or minimum operation, respectively, in the neighbourhood covered by the structuring element $B$, and these operation are depicted in Equation 3.11 and Equation 3.12.

$$ (A \ominus B)(s,t) = \min\{A(s-x,t-y) - b(x,y)|(s-x),(t-y) \in D_A : (x,y) \in D_b\} $$ (3.11)

$$ (A \oplus B)(s,t) = \max\{A(s-x,t-y) + b(x,y)|(s-x),(t-y) \in D_A : (x,y) \in D_b\} $$ (3.12)

Flat structuring elements are most commonly used in grayscale morphological operations, mathematically described in Equation 3.13.

$$ b(x,y) = \begin{cases} 0, & (x,y) \in B, \\ -\infty, & \text{otherwise} \end{cases} $$ (3.13)

Thus, dilation and erosion operations with flat structuring elements can be viewed as particular cases of nonlinear order-statistic filters whose result is the top and lowest value respectively. The opening and closing operations are also applicable in grayscale images.

### 3.3.3 Image Segmentation

Image segmentation is the process of subdividing an image into meaningful objects or regions, depending on the problem to be solved. It may also be viewed as labelling every pixel in an image so that equally-labelled pixels share certain characteristics. Segmentation algorithms can be categorized as those based on the feature domain of the image and those based on the spatial domain.
Within the latter category, algorithms are generally based on either the discontinuity or similarity of the brightness values and are categorized as boundary-based or region-based, respectively.

The majority of feature-domain based segmentation algorithms are histogram dependent. In other words, the partition of the image into regions is performed due to the measurement and analysis of the histogram of the image, most commonly the intensity histogram. The most simple method is thresholding. It consists in the selection of a threshold value that divides the image in two regions. This method is ideal for segmenting bimodal images - those whose histogram has two clear data peaks. There are several automatic thresholding methods, such as the Otsu’s method which is described in detail in Section 3.3.3.1. Other feature-domain based algorithms consist on clustering approaches.

In images with brightness variance across them, the effectiveness of the global thresholding is compromised. This fact justifies the need of exploring more robust approaches, such as adaptive/local thresholding. This technique consists on dividing the image into smaller subimages and computing the automatic threshold selection algorithm in each subimage. Then, in order to avoid discontinuity between consecutive windows, the implementation of Gaussian-weight or bilinear interpolation methods enables to compute a threshold value for each pixel of the image.

3.3.3.1 Automatic Thresholding - Otsu’s Method

The Otsu’s method is a commonly used bimodal clustering-based thresholding method. This method iterates through all possible threshold values aiming to select the threshold value which minimizes the intra-class variance and maximizes the inter-class variance. The intra-class variance is defined as the weighted sum of variances of the two classes separated by the set threshold value and can be computed through Equation 3.14, after computing each class probability and variance, \( \omega_i \) and \( \sigma^2_i \) respectively.

\[
\sigma^2_w = \omega_1(t)\sigma^2_1(t) + \omega_1(t)\sigma^2_1(t) \tag{3.14}
\]

3.3.4 Automatic Thresholding - RATS method

Another automatic threshold selection method ideal for bimodal images is the Robust Automatic Threshold Selection (RATS). This method computes the optimal threshold as the gradient-weighted sum of the intensity values which is computed through Equation 3.15.

\[
T = \frac{\sum \omega(x,y)p(x,y)}{\sum \omega(x,y)} \tag{3.15}
\]

Using thresholded edge strength, \( \omega \), enables to ignore pixels whose edge strength, \( e(x,y) \), is caused by image noise. After computing the horizontal and vertical components of image gradient, \( G_x \) and \( G_y \) respectively, by using Sobel kernels of both orientations, the edge strength is computed using Equation 3.16, a more computationally efficient alternative to gradient magnitude.

\[
e(x,y) = 0.5|\max(G_x(x,y))| + 0.5|\max(G_y(x,y))| \tag{3.16}
\]
3.3.5 Image Features

In this section, the various methods employed in order to obtain features from the segmented WBC candidates are presented. In using either supervised or unsupervised machine learning techniques, the data presented to these algorithms is commonly designated as feature vector and is preferably represented in numerical form, enabling to build a statistical-based model of the example data.

In the scope of this thesis, the differentiation of the candidates as either mono- or polymorphonuclear WBCs depends on the features of the nucleus and cytoplasm. Such features include geometric properties of the binarized images, colour, gradients and texture features. As above-mentioned in the Section 2.4.4, the main differentiation features of WBC types is the nucleus, specifically its size, shape and colour, as well as the cytoplasm, particular its colour and area-ratio with the nucleus. The geometrical features are computed from the binary segmented objects while the textural features are obtained through the grey level co-occurrence matrix method (GLCM) which is based on grayscale spatial variations. In the subsequent subsections, the computation of both features are described in detail.

### 3.3.5.1 Shape Based Features

The shape of an object can be represented as a binary function and its description can be obtained through several contour-based methods, i.e. the shape features are obtained through assessing the boundary points, particularly their relative distribution in the space domain, of the object in question. Perimeter and other perimeter-resultant measurements, such as circularity or solidity, are shape features obtainable from contour-based methods. Several shape features, withdrawn from Mingqiang Yang et al. work [45], are described below.

- **Area**: measurement of the area within the contour.
- **Perimeter**: measurement of the contour length.
- **Convex Hull Area**: a region R is convex if and only if for any two points $P_1, P_2 \in R$, the entire line segment $P_1, P_2$ is inside the region. The convex hull of a region is the smallest convex region including it. The convex hull area is the area within this region.
- **Solidity**: describes the extent to which the shape is convex or concave and it is defined by:
  \[
  \text{Solidity} = \frac{A_{\text{shape}}}{A_{\text{ConvexHull}}}
  \]
- **Waviness Shape Factor**: is defined as the perimeter ratio of the convex hull over the object contour.
  \[
  \text{Convexity} = \frac{P_{\text{ConvexHull}}}{P_{\text{Shape}}}
  \]
- **Rectangularity**: represents how rectangular a shape is, i.e. how much is fills its minimum bounding rectangle.
  \[
  \text{Rectangularity} = \frac{A_{\text{shape}}}{A_{\text{MinimumBoundingRect}}}
  \]
- **Equivalent Diameter**: is the diameter of the circle whose area is same as the shape.
  \[
  \text{EquivalentDiameter} = \sqrt{\frac{4 \times A_{\text{shape}}}{\pi}}
  \]
- **Circularity**: describes the shape similarity to a circle. This shape feature has several definitions and the expression used in the thesis scope was:
Circularity = \frac{p_{\text{shape}}^2}{4\pi A_{\text{shape}}}

**Thinness Ratio:** is the inverse of circularity feature.

\[ \text{ThinnessRatio} = \frac{4\pi A_{\text{shape}}}{P_{\text{shape}}} \]

**Euler Number:** describes the relation between the number of contiguous parts (S) and the number of holes (H) on a shape.

\[ \text{EulerNumber} = S - N \]

**Holes Total Area:** measurement of the total area of all holes in a shape.

**Holes Area Ratio:** describes the relation between the total area of all holes in the shape and its area

\[ HAR = \frac{A_{\text{holes}}}{A_{\text{shape}}} \]

**Maximum Axis Length:** the length of the major axis of the minimum bounding box containing every point in the shape (\(\lambda_{\text{majoraxis}}\)).

**Minimum Axis Length:** the length of the minor axis of the minimum bounding box containing every point in the shape (\(\lambda_{\text{minoraxis}}\)).

**Bounding Box Area:** measurement of the area of the minimum bounding box containing every point in the shape.

\[ \text{Area}_{\text{Bounding Box}} = \lambda_{\text{MinorAxis}} \times \lambda_{\text{MajorAxis}} \]

**Eccentricity:** measurement of the length ratio of major axis to minor axis, of the minimum bounding box containing every point in the shape.

\[ \text{Eccentricity} = \frac{\lambda_{\text{MinorAxis}}}{\lambda_{\text{MajorAxis}}} \]

**Diameter of Minimum Enclosing Circle:** as the designation suggests, this feature is the diameter of minimum circle enclosing every point in the shape.

**Elongation:** is defined as the square root of the ratio of the two second moments of the particle around its principal axes.

\[ f_{\text{elong}} = \sqrt{\frac{\mu_2}{\mu_1}} \]

### 3.3.5.2 Texture Features

The texture of a region can be defined as the degree of brightness variation. This spatial perception is unobtainable in a common histogram visualization since it does not take into account the spatial properties of the image. Grey Level Co-occurrence Matrix methodology is the commonly used texture operator that measures how often different combinations of pixel intensity values occur in a specific spatial displacement (\(\Delta x, \Delta y\)). The co-occurrence matrix is mathematically described in Equation 3.17.

\[ C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^{n} \sum_{q=1}^{m} \begin{cases} 1, & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{otherwise} \end{cases} \quad (3.17) \]

The occurrence of the pixel located at \((p, q)\), with a brightness value of \(i\), and the pixel located at \((p + \Delta x, q + \Delta y)\), and with brightness value of \(j\), is record in the co-occurrence matrix \(C_{\Delta x, \Delta y}\) at the position \((i, j)\).
The use of an offset vector with set spatial displacement parameters makes the GLCM rotational variant. In order to circumvent this drawback, some methods have been employed, such as using a set of offsets spanning every direction, particularly the unitary directions of an 8-connected neighbourhood. Once the GLCM is computed, several features can be obtained, either contrast order or descriptive-related, such as the ones shown below.

**Contrast**: Measurement of the local variations in the grey level co-occurrence matrix calculated using as follows.

\[
Contrast = \sum_{i,j} |i - j|^2 p(i, j) \tag{3.18}
\]

**Energy**: Returns the sum of squared elements in the GLCM and is calculated using the following equation.

\[
Energy = \sum_{i,j} p(i, j)^2 \tag{3.19}
\]

**Entropy**: Measurement of the uncertainty associated with a random variable and is as follows.

\[
Entropy = \sum_{i,j} -p(i, j) \log(p(i, j)) \tag{3.20}
\]

**Correlation**: Measure of how correlated a pixel is to its neighbourhood over the whole image and is calculated using the following relation.

\[
Correlation = \sum_{i,j} \frac{p(i,j)(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} \tag{3.21}
\]

**Homogeneity**: Measurement of how close the distribution of elements in the GLCM is to the GLCM diagonal and is calculated using the following relation.

\[
Homogeneity = \sum_{i,j} \frac{p(i,j)(i - \mu_i)(j - \mu_j)}{\sigma_i \sigma_j} \tag{3.22}
\]

**Mean for** $i$ (same for $j$):

\[
\mu_i = \sum_{i,j} i p(i, j) \tag{3.23}
\]

**Variance for** $i$ (same for $j$):

\[
\sigma_i^2 = \sum_{i,j} p(i, j)(i - \mu_i)^2 \tag{3.24}
\]

**Maximum probability**: The maximum element in the co-occurrence matrix and calculated through the following equation.

\[
MaxProb = \max(p(i, j)) \forall i, j \tag{3.25}
\]
3.4 Machine Learning Techniques

The most critical stage in developing robust computer-aided systems (CAD) is machine learning. This field consists on finding patterns and generally involves algorithms which build a statistical-based model from example data presented to them, and are able to make predictions or decisions on previously given data. Besides learning algorithms, this field also comprises methods for optimization of data dimensionality and model validation techniques. In this section, the techniques and methods used in the scope of this work, and also used in the state-of-art, are described in detail.

3.4.1 Feature Selection

Feature selection is the process of reducing the data dimensionality through selecting a subset of features for designing the classification system. This technique not only simplifies the complexity and training times of the developed model, but also enhances the generalisation capability of the resulting classification algorithm. Unlike feature extraction techniques, the data dimensionality reduction is not attained through creation of new features, thus the original meaning of the features is maintained. In fact, feature selection techniques aim to include the most relevant features in the feature vector, while excluding the redundant or irrelevant ones without loss, or even enhancement through avoiding overfitting, of the classifier performance. There are three different feature selection methods: filter methods, embedded methods and wrapper methods. While the former, filter method, consists on ranking each feature individually and independently, the latter two methods are computationally more exhaustive since they require a classification model along with an evaluation measure. Feature selection in embedded methods is performed as part of the classification model construction model. The Recursive Feature Elimination is an example of an embedded feature selection method commonly used with Support Vector Machine. On the other hand, wrapper methods consist on iteratively finding a subset of features which results in the best classification performance using a specific classifier. The performance of the predictive model is usually assessed through cross-validation.

The most common method within feature selection techniques is the sequential search, which consists on minimizing a classifier performance criterion, over the sequential features subsets. Furthermore, while the exhaustive feature subset search is infeasible, particularly in sizeable feature data, due to its computationally expensive nature, the most commonly used methods are the sequential forward selection and the sequential backward selection. These antagonist methods are similar in their essence since the former consists on sequentially adding features to the subset until the performance criterion does not decrease while the latter consists on sequentially removing features from the feature vector until the performance criterion does not increase.
3.4 Machine Learning Techniques

3.4.2 Model Evaluation Methods

Model validation techniques are crucial to evaluate the generalization capability of a learning algorithm, i.e. they estimate how the resultant predictive model performs to independent data. Theses techniques are used for both regression and classification machine learning algorithms and can be used to evaluate the learning algorithms performance or intermediate steps in their optimization such as in feature selection methods or parameter optimization. The most commonly used model validation techniques are briefly described below:

- **K-fold Cross-Validation**: This model validation method consists on partitioning the dataset into \( K \) equally sized subsets. The process of retaining one subset as the testing data, while the remaining \( K - 1 \) are used to train the algorithm, is repeated until all subsets are used once as testing data. Then, the algorithm performance is given as the average performance from each iteration.

- **Nested K-fold Cross-Validation**: This model validation method consists on partitioning the dataset into \( K \) equally sized subsets. Then, one of the subsets is retained in order to provide the optimized classifier performance. This optimization process is performed using the remaining \( K - 1 \) folds by computing \((K - 1)\)-fold cross-validation for each parameters combination, i.e. for each combination the algorithm is computed \( K - 1 \) times. The optimum parameters are attained from the parameters combination whose average performance was better.
3.4.3 Classification Methods

Classification algorithms, also designated as classifiers, are methods which are capable of assigning a class to an unlabelled observation on the basis of a model built from a training set of data whose class is known. The learning process can be either supervised, in which the algorithm is trained using labelled data in order to build a statistical model, or unsupervised, in which the algorithm builds a model from unlabelled data using clustering methods. Additionally, there are semi-supervised methods which use both labelled and unlabelled data for training. In the subsequent subsections, several supervised learning algorithms are briefly described, such as Decision-Tree, Support Vector Machine, Bayesian Classifiers and Linear Discriminant Analysis as well as the k-Nearest Neighbours, the most widely used clustering classifier.

3.4.3.1 Decision Tree

Decision Tree algorithm create a tree-shape model able to predict the classes of a set of input features. Each leaf represents a class label and a set of input variables is given that particular label if it goes through the path from the root to the leaf. The in-between path from the root to the leaf is constituted by branches which represent conjunctions of features that lead to those class labels.

Decision Trees are generated by recursive partitioning of the value of the features, i.e. the learning process of decision tree consists on splitting the source set into subsets based on an attribute value test, and this process is repeated recursively on each derived subset until splitting no longer adds value to the predictions. The most commonly selected attribute value test is information gain which is based on the concept of entropy from information theory. The information gain is the entropy reduction of a particular features achieved by learning the classification.

3.4.3.2 Bayesian Classifiers

Bayesian classifier aims to minimise the probability of misclassification by using the concept of conditional distribution, in which the label of the training data is used as the "given" part. The probability distributions of each of the classes are mapped into the feature space and this classifier assigns to a unlabelled observation the class whose probability distribution at that particular point is bigger.

3.4.3.3 Support Vector Machines (SVM)

Support Vector Machines is a function-based learning algorithm which aims to compute the maximum margin hyperplane separating two classes. This supervised classifier maps the labelled data onto the feature space and computes the large margin decision boundaries which maximizes the separability between the classes. In order to avoid overfitting, the computation of the large margin decision boundaries uses only the data point that lie closer to the boundaries.

Some dataset may not have an hyperplane which can perfectly split the data. The soft margin method allows for mislabelled data and, thus, the hyperplane optimization becomes a trade-off
between large margin and an error penalty for misclassification of the data. In this way, the hyperplane would split the data as cleanly as possible while maximizing the distance to the nearest cleanly split data.

3.4.3.4 k-Nearest Neighbours (kNN)

The k-Nearest Neighbours algorithm is a widely used lazy learning algorithm. An unlabelled test example is classified by assigning the label which is most frequent among the k training samples nearest to it. Firstly, the classifier computes a distance metric (for example, Euclidian, Mahalanobis or Hamming) between the feature vectors of the unlabelled observation and the data in the training dataset. Apart from the number of “votes” from the k nearest observation, a more robust approach consists on assigning “weighted votes” which takes into account the distance from the unlabelled example to each of its k nearest neighbours. The weight is proportional to the inverse of the distance.

In “unweighed” k-NN, the k parameter should be an odd value in order to avoid draw-voting cases. In addition, a few drawbacks of k-NN are the biased classification due to unbalanced training dataset and the computational complexity for large dataset since the classifier is required to store all training examples.

3.4.3.5 Linear Discriminant Analysis (LDA)

Linear discriminant analysis is a generalization of Fisher’s linear discriminant, a method used to find a linear combination of features that separates two or more classes. It aims to maximize the difference between the means by looking for a projection where the intra-class data is close to each other (low variance) and the classes means are as farther as possible from each other. The computation process consists on calculating the sample mean of each class and, then, computing the sample covariance, by first subtracting the sample mean of each class from the observations of that class, and taking the empirical covariance matrix of the result.

3.4.4 Regression Methods

Regression algorithms build a statistical model correlating the independent variables with the dependent variable. The resulting regression function or regression model is able to provide estimation or prediction values given a set of independent variables. Two regression methods are described in the subsequent subsections: linear regression - a widely used regression algorithm -, and support vector regression (SVR) - a regression version of the robust classifier algorithm SVM.

3.4.4.1 Linear Regression

Linear regression builds a model in which the relationship between a scalar dependent variable and a set of independent variables is fitted by a linear equation.
In general terms, the linear relationship between the independent variables $y$, defined as an dependent variable vector, and the dependent variable $X$, the design matrix which stores the independent variables, is mathematically described as:

$$y = X\beta + \epsilon$$

where $\beta$ is a regression-coefficients vector, also interpreted as the partial derivatives of the dependent variable, of length equal to the number of independent variables and $\epsilon$ is called the error term which considers other factors which influence the relationship.

Several regression coefficients estimation methods have been developed, such as the Akaike information criterion. The preferred model is obtained through minimization of the Akaike information criterion value, a criterion which balances the likelihood of the a set of parameters and a penalty value which contributes to avoiding overfitting.

### 3.4.4.2 Support Vector Regression (SVR)

Despite a few minor differences, Support Vector Regression uses the same principles as the SVM: minimizing error while maximizing the hyperplane margin. The main difference is in regards to the penalty definition. In the SVM hyperplane optimization, the penalty is for misclassification, i.e. if an example is on the wrong side of the hyperplane. In SVR, the penalty is in function of the distance of the example from the predicted line, while if the example data is within an $\epsilon$-distance no penalty is given.

This algorithm has several parameters, namely the kernel type, C and, depending on the kernel type used, $\gamma$. Intrinsically to SVR, whether it is linear or nonlinear, the $\epsilon$ and C parameters must be accounted for. The former consists on the width of the zone in which the training data is considered to build the model, while the latter is the penalty weight which determines the trade-off between the model complexity and the generalization power of such algorithm. For example, a carefulness high C leads to overfitting of the training data. Nonlinear regression approaches enable to fit the hyperplane in a feature space. This is achieved by using kernel functions on the training data and, instead of learning fixed set of parameters, the nonlinear regression algorithms learn the weight corresponding to each point in the feature space. One particular kernel type is the Gaussian radial basis function and the $\gamma$ parameter is related to it. It defines how far the influence of a single training example reaches and can be intuitively related to the inverse of the radius of influence. Lower $\gamma$ values leads to a further influence of a set training sample on the feature space.

### 3.5 Relevant Remarks

For RBC and WBC detection and WBC classification via an image processing system, some considerations about the blood smear images should be noted. RBCs can exclusively be analysed in thin blood films since blood components are disposed in a single layer. On the other hand, WBCs identification and classification is feasible in both thin and thick blood smear images. Additionally, the low-quality nature of the acquired images should also be considered, particularly in the selection of the image processing methods.
A critical analysis of the methodologies described in Section 3.2 is presented below, in order to delineate possible directions to follow during the development of this work.

For RBC segmentation, the preprocessing methods used in [31], along with the distance-transform based methodology, led to efficient segmentation results. However, since the blood smear images to be used in this Master thesis are lower quality, the pre-processing step should be more robust and the RBC segmentation methodology should be suited for the presence of erythrocyte clusters. For these reasons, some considerations should be taken from [29] since the method proposed resulted in higher specificity in cell detection than distance transform methods. Finally, Hough transform-based methods should be disregarded due not only to false positive detection proneness, but also to poor segmentation contours that are not aligned with the RBCs contours.

For WBC segmentation, most methodologies rely on the contrastive intensities of the WBC nucleus. Thresholding methods are applied in order to segment the nucleus, followed by either region growing [39] or morphological operations [40] in order to segment each nucleus respective cytoplasm. For WBC classification, the features most commonly considered include geometric properties of the segmented cells (particularly the nucleus), colour, gradients and texture features (to determine granularity). In Vincenzo Piuri al. work [40], biological relevant morphological features of leukocytes, particularly from the membrane, cytoplasm and nucleus were considered.

Regarding the classifiers used, authors compared several different classifiers in their respective works and the better classification results were obtained with neural networks. In Vincenzo Piuri al. work [40], different classifiers were tested, namely the K-nearest neighbours, radial basis function neural networks with different centres and feed-forward neural networks for each class. The latter feed-forward neural network had a better classification performance. In two different studies [42, 43], this classifier performed better than the linear Bayes Normal classifier. Another classifier which led to good classification results was Support Vector Machine (SVM), which produced better results than Naive Bayes and tree-based classifiers [44].

It should be noted that the aforementioned studies relied on high-quality blood smear images either obtained from private databases, such as the ALL-IDB [46] provided by Fabio Scotti, from public databases, such as the MaMic [47], or even their own constructed databases. Since this work is aimed at low-quality images, no public database was found, thus, an appropriate image dataset has to be constructed. The influence of the image quality is more evident in thin blood smear images for several reasons. Firstly, the contrast between cellular structures is poorer. Then, not only are the cell borders inferiorly defined, but also the texture information is loss in low-quality images, particularly in white blood cells. Regarding thick blood smear images, the influence of the image quality is not as crucial for the segmentation and recognition of blood cells since in these images the number of cellular structures is lower (only the nucleus of white blood cells are visually detectable) and due to the impracticality of extracting meaningful textural information for discriminating the white blood cells in five types (only in two broaden ones) in high-quality images. The differences between high and low-quality blood smear images can be visualized in Figure 3.6.

Finally, OpenCV was the computer vision library used throughout this work. Even though
FastCV is faster than OpenCV performance-wise, the number of image processing functions are still limited and it does not include machine learning functions, two requirements particularly crucial for the development of this Master thesis. In addition, the fact that OpenCV supports Android makes this library ideal for the integration of the image processing methodologies in a mobile-phone application.
Chapter 4

RBC counting in Thin Blood Smear Images

In this chapter, the proposed methodologies for RBC segmentation and counting are explained in detail. Firstly, in section 4.1 the processes followed for the creation and annotation of the used image dataset are presented. In the subsequent sections, the proposed automatic threshold selection and segmentation algorithms are presented while the feature vector construction and supervised learning approaches for regression purposes are elucidated. The proposed methodology for RBC counting in thin blood smear images is schematized in Figure 4.1.

Figure 4.1: Pipeline for RBC counting in thin blood smear images.

4.1 Image Dataset Construction

According to the objectives of this chapter, a dataset constituted exclusively of thin blood smear images was created. This dataset proved to be essential in the development of both the RBC segmentation and regression methodologies. The partnership with INSA (Instituto Nacional de Saúde Dr. Ricardo Jorge), in particular Dr. José Manuel Costa from the Research and Development Unit for Infectious Diseases, was crucial in acquiring these images, not only in the construction of this particular dataset but also the ones created for the subsequent chapters. In fact, INSA supplied blood smears of 4 distinct people, with two thick blood smear samples and one thin blood sam-
ple for each person. The image dataset used in the work needed to meet several requirements, enumerated in the following section.

4.1.1 Image Dataset Requirements

The image dataset for RBC counting should meet the following requirements:

- The thin blood smears must be prepared by a specialized doctor;
- The thin blood smears collected should be obtained from various patients;
- The thin blood smear images should be acquired with 1000x magnification;
- The thin blood smear images should be acquired with a low-cost commercial microscope using a smartphone device with a high-resolution camera;
- The image dataset should include equal number of blood smear images from each patient;
- The number of RBCs should be annotated by a specialist;
- The total number of images for each dataset should be at least 200.

4.1.2 Image Dataset Construction

The acquisition of the thin blood smear was very straightforward since, prior to the start of the thesis work, the samples had already been prepared and available at Fraunhofer AICOS. According to the requirements stated in the Section 4.1.1, the dataset constitution should be equally distributed among images from each patient, thus ideally the dataset should include 50 images from sample.

The images were acquired with the smartphone LG-Nexus 5 coupled to the low-cost microscope Bresser Microscope - 5102000 - Erudit DLX 20x-1000xa with the Sky Light adapter. The abovedescribed assembling is illustrated in Figure 4.2. The smartphone had camera resolution of 8MP(3264 x 2448). The acquired images are compressed to .jpeg format. It should be noted that the low quality nature of the smartphone-acquired images is enhanced even more by image compression due to the file format, thus affecting the segmentation and machine learning algorithms proposed in this work.

<table>
<thead>
<tr>
<th>Blood Smear Id</th>
<th>Blood Smear Type</th>
<th>Images Acquired</th>
<th>Images to Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>171-I</td>
<td>Thin Film</td>
<td>82</td>
<td>44</td>
</tr>
<tr>
<td>0195-I</td>
<td>Thin Film</td>
<td>45</td>
<td>29</td>
</tr>
<tr>
<td>306-I</td>
<td>Thin Film</td>
<td>60</td>
<td>31</td>
</tr>
<tr>
<td>307-I</td>
<td>Thin Film</td>
<td>71</td>
<td>70</td>
</tr>
<tr>
<td>Other Dataset</td>
<td>Thin Film</td>
<td>55</td>
<td>22</td>
</tr>
</tbody>
</table>

At first, an excess number of images from different locations within each sample were acquired, according to Table 4.1. Subsequently, the acquired images were analysed, and those with an extremely high concentration of RBCs were discarded. In fact, a correct area of a specimen from a patient with a normal RBC count should have 200-250 RBCs per 100x oil-immersion field [48]. Despite this standardized value, images with RBC concentration up until 300 RBCs
per field of view were still included in order to improve the adaptability and robustness of the methodology. Moreover, those images with acceptable RBC concentration yet unfocused were still discarded. Finally, the different sized subsets of focused and RBC-concentration-acceptable images were merged as depicted in Table 4.1.

Finally, the number of erythrocytes in each image was annotated. The data annotation process is labour-intensive, time-consuming and, especially, prone to error. In this field of cell counting, the subjectiveness lies mostly on overlapping cells and whether to consider cells touching the field of view borders or not, which differs from technician to technician. On top of it, even though the specialist provided some guidelines for the annotation of the number of RBCs in this work, it was not performed by him, thereby, increasing the error susceptibility for the ground truth annotation.

4.2 RBC Segmentation

4.2.1 Algorithm Overview

The proposed method is divided in three main distinctive steps. Initially, the preprocessing methodology consists on using a linear contrast-stretching method which not only enhances the contrast of the RBCs and background but also converts the image to bi-modal by saturating the WBCs and platelets to zero.

An automatic threshold selection algorithm is then applied. In this step, two different methods were tested: automatic gradient-weighted threshold selection, based on RATS (Robust Automatic Threshold Selection), and the Otsu’s method. In addition, two different approaches were applied
to each method: global application of the algorithms and local application through overlapped Gaussian-weighted moving-window (MW) method. Subsequently, an anisotropic morphological dilation is applied to the segmented images with the purpose of closing existing interruptions of the ring-shaped segmented RBCs before applying a holes-filling algorithm.

Finally, the overlapping RBCs are segmented individually through a distance-transform-based method in order to obtain a more precise count on the number of RBCs in the image. The number of RBCs in addition to the total segmented area, obtained from both automatic thresholding methods, are then used in the regression methodology described in Section 4.3.

4.2.2 Preprocessing and Image Enhancement

The preprocessing step was divided in two steps: field of view (FOV) segmentation and image enhancement.

The interest in FOV segmentation is that within this area resides the region of interest for the entire image processing methodology. Through region growing, the non-FOV area is segmented. Region growing is a region-based image segmentation method. It is an iterative process which requires the selection of initial seed points. In each iteration, the pixels adjacent to the segmented region are added depending on set criterion, which in this work was as presented in Equation 4.1.

\[
src(x',y') - loDiff \leq src(x,y) \leq src(x',y') + upDiff
\]  

(4.1)

In Equation 4.1, \( src(x',y') \) is the value of one of the neighbouring pixels already belonging to the segmented region, and \( loDiff \) and \( upDiff \) are the lower and upper brightness difference between the current pixel \( src(x,y) \) and \( src(x',y') \). The adjacency criterion used was 4-connected neighbourhood, or in other word the pixels which share and edge to the one being considered. Four seed points were used, the corner pixels of the original image, in order to avoid cases which the FOV fills the image width and the segmented region would reach the bottom non-FOV.

Figure 4.3: FOV mask segmentation: (a) original image; (b) segmented FOV image.
4.2 RBC Segmentation

At last, an median filter is applied in order to remove small unsegmented points, and an opening operation is performed in order to smooth the FOV borders. The output of this FOV segmentation operation is illustrated in Figure 4.3.

Subsequently, image enhancement of the FOV region is performed. Firstly, the conversion of the 3-channel image to a single-channel was performed. Literature suggests that the green channel is the optimal colour channel for segmentation purposes since the contrast between the structures within the image is higher. This statement can be confirmed through assessing each channel and respective histogram, as illustrated in Figure 4.4. In fact, through comparing the histograms of each channel, the bi-modality is evident in the green-channel image, as one modal distribution is due to the background area while the other is due to the red blood cells.
The contrast between the RBCs and background should still be enhanced and this is achieved though percentage linear contrast stretching. This operation consists on linearly stretching the original pixel values in order to use the full range of 0-255 scale system, as depicted in Equation 3.2. Recalling this particular histogram-stretching method, depicted in detail in Section 3.3.1, it sets a percentage of the maximum value of the histogram as the new minimum and new maximum values.

In fact, typical thin blood smear images present three very distinct regions, each with a particular brightness distribution. Platelets and WBCs have similar intensity values. On the other hand, while RBCs and the background are visually dissimilar and their intensity values are considerably higher than platelets and WBCs, their brightness distribution still overlap. Thus contrast should be enhanced for further segmentation procedures purposes. The ideal percentage would set the minimum value to lay in-between the platelets-WBCs and RBCs-background brightness values which would enable both contrast enhancement and thresholding of platelets and WBCs’ nucleus. The latter is then added to the FOV mask as illustrated in Figure 4.5.

4.2.3 Automatic Threshold Selection

The inclusion of the WBCs and platelets in the mask image turns the contrast-enhanced image into a bimodal image whose intensity levels can be clustered as foreground (RBCs) and background.

A commonly used bimodal clustering-based thresholding method is the Otsu’s method which was previously described in Section 3.3.1. In the same Section, another automatic threshold method suited for bimodal images, the Robust Automatic Threshold Selection (RATS), is also described.

The low-quality images present brightness variance across them, compromising the effectiveness of the described automatic threshold selection algorithms. This fact justifies the need of exploring more robust approaches. Local applications of the Otsu’s method and RATS algorithms were elected in this work and were achieved through a Gaussian-weighted moving-window approach. While benefits from local thresholding techniques are evident, some challenges arise and must be accounted for when designing the algorithm. As above-stated both automatic threshold selection algorithms, Otsu’s method and RATS, assume that the histogram of the image is bimodal.
However, a particular region may contain no foreground object, thus the algorithm, assuming a bimodal histogram, would erroneously split a unimodal histogram into two clusters.

Regarding the RATS method, detecting the presence of an object, in this case an edge, in a certain window is a straightforward process, consisting on inspecting the denominator of Equation 3.15. If the sum of thresholded edge strength values is below a set threshold value, it is indicative of the absence of edges in that particular window and the computed threshold value is decrease by a set offset value in order to force undersegmentation in that given region. Regarding the Otsu’s method, evidence of unimodal histogram and consequential absence of objects is unclear. The computation of the thresholded edge strength was used for this purpose solely, despite being completely unrelated to the Otsu’s method itself. Subsequently, the same procedure of offsetting the threshold value was adopted when the sum is lower than a set value.

In this work, in order to address discontinuity between consecutive windows, it is proposed an alternative and less-computationally-demanding moving-window method to those which require to compute the threshold value for each pixel. This method, all pixels within a window are assigned with the computed threshold value. Then, the window slides several pixels to a new region and a new threshold value is computed, as illustrated in Figure 4.6. Two drawbacks are evidenced in the straightforward application of this method. Pixels from the window edges are assigned with the threshold value most suitable to central pixels i.e. if the window was centred in these edge pixel, the threshold value computed would most likely be different. On the other hand, pixels within overlapping regions are assigned multiple values.

![Figure 4.6: Illustration of the proposed moving-window method.](image)

In order to overcome both problems, a Gaussian-weighted method was proposed. This method consists on giving weights to pixels according to their distance to the window centre \( p(x_i, y_i) \). Each window is associated with an Gaussian kernel of the same size computed using Equation 4.2. This
kernel type was preferred due to its isotropic nature.

$$G_i(x, y) = \frac{1}{2\pi\sigma^2}e^{-\frac{(x-x_i)^2+(y-y_i)^2}{2\sigma^2}}$$

(4.2)

Therefore, pixels closer to the window centre are given higher weights than distant ones. Hence, pixels from overlapping regions are assigned different threshold values, each one with its respective weight. To elucidate this concept, a given pixel $p(x, y)$, which is in the overlapping region between window 1, $w_1$, and window 2, $w_2$, is considered. Both threshold values are computed using Equation 4.3, as well as both Gaussian values, $G_i(x, y)$.

$$T_i(x, y) = \frac{\sum_{w_i} \omega(x, y)p(x, y)}{\sum_{w_i} \omega(x, y)}$$

(4.3)

Finally, the threshold value for $p(x, y)$ is computed using Equation 4.4. If $p(x, y)$ is closer to $w_1$’s centre than $w_2$’s, the threshold value from $w_1$ is given higher weight in the calculation of the threshold value assigned to $p(x, y)$. Additionally, this method enables to overcome discontinuity problems that might occur otherwise.

$$T(x, y) = \frac{\sum_i G_i(x, y)T_i(x, y)}{\sum_i G_i(x, y)}$$

(4.4)

### 4.2.4 Directional Gradient

As abovestated in Section 2.4.2, the flattened surface of RBCs faces up in thin blood smear images, corresponding to the biconcave regions. Consequentially, the central area of RBCs presents similar intensity values as the background, leading to these regions being unsegmented. The expected outcome of the previously described thresholding algorithms is a ring-shaped binary object, as the hole would represent the biconcavity region.

Flood-filling algorithms would fill the unsegmented holes of the binary image, resulting in the complete segmentation of RBCs. However, the low-quality nature of the acquired thin blood smear images leads to interruptions of the binary rings. In these cases, the flood-filling algorithm would prove inefficient, compromising the RBCs segmentation algorithm since the segmented RBC borders would disappear in later opening and erosion morphological operations. In order to avoid this undersegmentation issue caused by gapped ring-shaped binary objects, a morphological operation is performed.

The common dilation operation uses isotropic structuring elements for expanding the shapes uniformly in all directions. This operation consists on the dilation of the input image, $A$, by the isotropic structuring element, $B^i$, and is mathematically described in Equation 3.6.

However, in this thesis context, particularly in the gap-closing step of segmented RBC rings, the isotropic dilation is unsuitable. The expansion of the ring-shaped objects in other directions but solely the gap direction impairs the entire segmentation methodology since adjacent RBCs would
connect to each other and to small artefacts present in the binary image. Thus, the approach taken in this work was to use morphological dilation whose structuring elements would be spatial-variant based on image-dependent gradient fields, a simplified variation of the methodology proposed by the Verdú-Monedero et al [49].

The estimation of the orientation of the structures is obtained by using the average squared gradient (ASG) method consisting on squaring and averaging the gradient vectors in a set neighbourhood defined by window $W$. The window dimensions used in this study was 21x21 pixels for FOVs with $\sim 2100$ pixels diameter. In fact, the 1-100 width-to-diameter proportion led to successful gap-closing results and this rule can be used for images with different resolutions as long as the 1000x magnification is used. The horizontal and vertical components of image gradient, $G_x$ and $G_y$ respectively, are computed by using Sobel kernels of both orientations and, subsequently, the ASG matrix is obtained through using Equation 4.5.

$$
\overline{G_s} = \left[ \frac{G_{sx}(x,y)}{G_{sy}(x,y)} \right] = \frac{\sum_W (G_x^2(x,y) - G_y^2(x,y))}{\sum_W (2G_x(x,y)G_y(x,y))}
$$

(4.5)

The vectorial directional field is converted to angle, which is computed using Equation 4.6, and is in the range of $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

$$
d(x,y) = \arctan \left( \frac{G_{s,y}(x,y)}{G_{s,x}(x,y)} \right)
$$

(4.6)

As above-stated, this is a simplified implementation of the ASG method proposed by Verdú-Monedero et al [49]. The major divergence is the regularization of the ASG, which is absent in this work. The purpose of this step is to extend the orientation information to homogeneous regions where the average squared gradient is residual. However, in the context of this thesis, there is no need for the regularization of ASG step since this is method is used solely on binary image, unlike Verdú-Monedero et al [49] whose method is suitable for grayscale images as well. Thus, the spatially-variant anisotropic dilation operation proposed is mathematically described in Equation 4.7.

$$
A \oplus B = \{ z \in E | B_z \cap A \neq \emptyset \}
$$

(4.7)

The structuring element $B$ is linear with a set length $\lambda$ and its orientation $\theta$ is pixel-to-pixel dependent, as described in Equation 4.8. Through experimental observations, the length was also set at 21 pixels, value which enabled to close wider RBC gaps. The 1-100 length to FOV diameter
Figure 4.8: Directional Dilation demonstration: (a) Gapped ring-shaped RBC, result from (b) directional dilation (SE size of 21) and (c) isotropic dilation (SE size of 21)

proportion can also be used across image resolutions for 1000x magnification. The orientation values are obtained through the ASG matrix.

\[ B(x) \equiv L_{\lambda}^{\theta(x)} \]  

An exemplification of the directional gradient method implementation is illustrated in Figure 4.8. Figure 4.8(a) shows the segmentation result of the automatic thresholding algorithm which, in these specific case, displays a ring-shaped object with a gap in its structure. Figures 4.8(b) and 4.8(c) illustrate the differences between the directional and isotropic dilation, respectively. Even though the isotropic dilation solved the gapping issue, it led to merging of adjacent RBCs. In contrast, the directional dilation method was able to close the gaps while causing minimum issues related to merging of adjacent RBCs.

4.3 Regression Algorithm

Recalling the objective of this chapter, the CAD system must provide a prediction on the number of RBCs present in the thin blood smear images. Thus, regression algorithms must be used, which necessarily implies the use of supervised learning techniques. The data annotation was previously described in Section 4.1.2. Then, the regression algorithm builds a statistical model relating the independent variables with the dependent variable, in this case the number of RBCs annotated. The independent variables provided to the regression algorithms tested were the number of RBCs segmented using the RATS method and Otsu’s method, and the ratio between the segmented area and the FOV area using the RATS method and Otsu’s method. The segmented area is provided as a ratio of the total field of view area for applicability reasons. In other words, the described ratio is independent of the image resolution, thus being suitable for multiple smartphone devices with different camera resolutions.

Two different regression algorithms were used: Linear Regression and Support Vector Regression. The theory behind each algorithm is documented in more detail in Section 3.4.

The performance of the Linear Regression algorithm was evaluated using the 10-fold cross-validation method on the root mean squared error (RMSE), the most commonly used regression
4.4 Segmentation Results

metric, which is calculated through Equation 4.9, as \( P \) is the predicted value and \( M \) is the ground truth value. The sampling of the subsets followed a stratified approach in order to ensure that the data distribution, in this case the mean ground-truth cell-count, in the subsets is approximately the same in all folds and in the complete dataset. The normalization of RMSE results in a scale-indifferent metric, enabling a more intuitive understanding of the performance values. This normalization is achieved through Equation 4.10.

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i} (P_i - M_i)^2} \quad (4.9)
\]

\[
NRMSE = \frac{RMSE}{M} \quad (4.10)
\]

Regarding the Support Vector Regression, the methodology is more complex due to the need to optimize the parameters of the algorithm, namely the kernel type, \( C \) and, depending on the kernel type used, \( \gamma \). In this work, the linear SVR and the nonlinear SVR with the radial basis function kernel were used.

The parameter optimization and subsequent performance evaluation was performed using the nested 10-fold cross-validation method on the RMSE, with the stratified sampling method. The kernel types used, linear and radial basis function, were evaluated separately. In other words, the process described in the above paragraph was performed twice with each kernel type fixed. In addition, the \( \varepsilon \) parameter was also set at 0.001 for both. In the linear support vector regression algorithm, \( C \) was the single parameter to be optimized and was grid searched from the following set of logarithmic increasing values:

\[
C \in \{2^1, 2^2, 2^3, ..., 2^{12}\}
\]

When using the radial basis function kernel, both \( C \) and \( \gamma \) were optimized using grid search as well for the following set of values:

\[
C \in \{2^1, 2^2, 2^3, ..., 2^{13}\}
\]

\[
\gamma \in \{10^{-9}, 10^{-8}, 10^{-7}, ..., 10^2\}
\]

Then, the parameters combination with the highest NMSE was used to building an SVR model, using the 9 folds as the training set, and the algorithm was evaluated by the initially retained subset.

4.4 Segmentation Results

In this section, the segmentation step is assessed, particularly the suitability of the RATS method as an automatic thresholding technique, the effectiveness of the local application of automatic threshold selection algorithms in low quality blood smear images, and the effect of the directional dilation operation as a gap-closure morphological operation. The predicted number of segmented objects measured just after the distance-transform-based step are evaluated against the ground-truth data through the RMSE and NRMSE metrics. The performance of the various methodologies followed in the segmentation step are depicted in Table 4.2.
Table 4.2: Cell count results following the distance-transform based segmentation algorithm.

<table>
<thead>
<tr>
<th>Segmentation Methodology</th>
<th>RMSE</th>
<th>NRMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW RATS</td>
<td>12.92</td>
<td>7.80%</td>
</tr>
<tr>
<td>MW Otsu’s method</td>
<td>12.79</td>
<td>7.72%</td>
</tr>
<tr>
<td>Global RATS</td>
<td>21.34</td>
<td>12.89%</td>
</tr>
<tr>
<td>Global Otsu’s method</td>
<td>21.69</td>
<td>13.10%</td>
</tr>
<tr>
<td>MW RATS noDirDilation</td>
<td>26.18</td>
<td>15.81%</td>
</tr>
</tbody>
</table>

As expected, adaptive approaches improve notably the segmentation process over global application of automatic threshold selection techniques. In this work, this performance improvement is evidenced by the 40% decrease of RMSE, either using the RATS or the Otsu’s method. This event is illustrated in Figure 4.9.

Figure 4.9: Global vs Local thresholding application, prior to the application of the watershed method based on the distance transform. From left to right, top row: Moving-Window Gaussian-Weighted RATS (left) and Otsu’s method (right) thresholding; bottom row: Global RATS (left) and Otsu’s method (right) thresholding.
4.4 Segmentation Results

Figure 4.10: Segmentation results when (a) directional dilation (SE size of 21) is applied (b), no dilation is applied and (c) isotropic dilation (SE size of 21) is applied

In fact, illumination uniformity across the field of view is crucial in order to assure optimal conditions in quantitative optical imaging, often nominated as "flat field". However, microscope manufactures acceptance criteria consists on maximum intensity variation of 5% to 10% along the horizontal and diagonal axis, as stated in [50]. Several factors intrinsic to the microscope lead to illumination variance. The illumination source usually has a brighter central spot that requires homogenization often with a light diffuser or an optical scrambler, even though it is not completely attained. Even when using lasers, a coherent and uniform radiation, as illumination sources, the remaining microscopic components, such as lenses and relay optics, lead to illumination inhomogeneity. While human experts can intuitively detect this phenomenon without impacting significantly the quantitative assessment of the microscopic images, it must be accounted for when developing the CAD system, particularly during the segmentation process.

Regarding the RATS method as an automatic threshold selection algorithm, it proved to be as robust as the universally use Otsu’s method. In fact, the insignificantly RMSE variation of <1% in using either one of the automatic thresholding method at the expense of the other evidences the suitability of the simplified version of the RATS method developed in this work.

Finally, the impact of the gaps in the ring-shaped segmented RBCs was assessed, through carrying out the regular RBC segmentation pipeline using the adaptive RATS thresholding methodology, but excluding the directional dilation step. As noticed in Table 4.2, its exclusion leads to significantly negative effects on RBC segmentation, as the NRMSE metric increased by more than 8 percentage points. In fact, Figure 4.10 illustrates this event as a few RBCs are left unsegmented when no dilation is applied. Furthermore, in Figure 4.10, the unsuitability of isotropic dilation is evidenced since some gaps within RBCs clusters are included in the segmented area.

As final remarks in regards to the RBC segmentation step, even though individual RBCs can be almost perfectly detected and segmented, splitting and detecting overlapping RBCs proved to be a rough task, despite the efficiency of the distance-transform based algorithm. This is due to the different overlapping degrees, as a RBC may either just touch the border of one of its neighbouring RBCs, or overlap by almost 50% of its area. Therefore, the robustness of this cell count prediction system may be enhanced by performing a regression analysis with not only the results of the number of objects but also considering the total segmented area, using both the RATS and Otsu’s
methods.

4.5 Regression Results

In this section, the results of the regression analysis are presented. In order to predict the number of RBCs, the independent variables (attributes) used to train the regression algorithm were the number of segmented objects and the ratio between the segmented and the FOV area, using the MW RATS method and the MW Otsu’s method. Additionally, two regression algorithms were used: SVR and Linear Regression. While the later is a nonparametrized model, the former requires to optimize several parameters, specifically the kernel type, the penalty weight $C$ and the distance of influence of the kernel $\gamma$. The performance of the used algorithms are summarized in Table 4.3

As described in detail in Section 4.3, for the SVR algorithm, the optimum parameters and subsequent performance evaluation were assessed through combining parameter combination grid search and nested 10-fold cross-validation method. On the other hand, the performance of the nonparametrized algorithm, Linear Regression, was evaluated through 10-fold cross-validation.

The results in Table 4.3 unveil the linear relationship between the independent variables - area segmented and number of objects detected in the segmentation step - and the number of RBCs, since the regression algorithms with the best RMSE performance were Linear Regression, 11.70, and $\varepsilon$-SVR using a linear kernel, 10.24. The linearity between both variables is intuitively evident and the regression analysis solely provides a better-fitting model capable of more accurately predicting the number of RBCs than simply considering the number of objects segmented when using only one thresholding method. In fact, even when using non-linear kernels in SVR algorithm, the sheer fact that more attributes are weighted in in the predicting model leads to an enhanced performance, 11.96 and 12.26 for sigmoid and RBF kernels respectively, than using merely the output result of the segmentation step, 12.92.

Table 4.3: Parameter Optimization and subsequent performance

<table>
<thead>
<tr>
<th>Regression Type</th>
<th>Parameter Optimization</th>
<th>RMSE</th>
<th>Final RMSE</th>
<th>Best Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$-SVR: Linear</td>
<td>11.62</td>
<td>10.24</td>
<td>C = 1024</td>
<td></td>
</tr>
<tr>
<td>$\varepsilon$-SVR: Sigmoid</td>
<td>11.96</td>
<td>11.96</td>
<td>C = 4096 ; $\gamma = 0.01$</td>
<td></td>
</tr>
<tr>
<td>$\varepsilon$-SVR: RBF</td>
<td>11.24</td>
<td>12.26</td>
<td>C = 2048 ; $\gamma = 0.1$</td>
<td></td>
</tr>
<tr>
<td>Linear Regression</td>
<td>-</td>
<td>11.70</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5

WBC recognition in Thick Blood Smear Images

In this chapter, the proposed methodologies for segmentation of WBC candidates and subsequent classification, as either multimorphonuclear or mononuclear, are described in detail. Firstly, in section 5.1, the creation and annotation of the image dataset is elucidated. In the subsequent sections, the proposed automatic extraction method of WBC candidates and the machine learning methods, including feature vector construction, feature selection and classification, are elucidated. Finally, the results from previously described methods are presented. The proposed methodology for WBC recognition in thick blood smear images is schematized in Figure 5.1.

Figure 5.1: Pipeline for WBC recognition in thick blood smear images.

5.1 Image Dataset Construction

The image dataset used in this chapter was constituted exclusively of thick blood smear images. As in Section 4.1.2, the partnership with INSA was critical in the acquisition of these images, since this institution supplied two thick samples from each blood smear from different people. The requisites the image dataset must meet are enumerated in the following section.

5.1.1 Image Dataset Requirements

The image dataset used in this section should meet the following requirements:
The thick blood smears must be collected by a specialized doctor;
- The thick blood smears collected should be obtained from various patients;
- The thick blood smear images should be acquired with 1000x magnification;
- The thick blood smear images should be acquired with a low-cost commercial microscope using a smartphone device with a high-resolution camera;
- The image dataset should include an approximately equally distributed number of blood smear images from each thick film of each patient;
- The WBCs type should be annotated by a specialist.

5.1.2 Image Dataset Construction

Parallel to the previous section, the thick blood smear samples had already been prepared and available at Fraunhofer AICOS. Besides, images were acquired with the low-cost Bresser Microscope - 5102000 - Erudit DLX 20x-1000x and the smartphone LG-Nexus 5 as well. Once again, an excess number of images were acquired from different locations within each sample, according to Table 5.1. Subsequently, the acquired images were qualitatively assessed and those acceptably focused were included in the image dataset, as depicted in Table 5.1.

Table 5.1: Thick blood smear images image dataset constitution

<table>
<thead>
<tr>
<th>Blood Smear Id</th>
<th>Blood Smear Type</th>
<th>Images Acquired</th>
<th>Images to Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>171-I</td>
<td>Thick Film 1</td>
<td>76</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Thick Film 2</td>
<td>70</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Thick Film 1</td>
<td>53</td>
<td>25</td>
</tr>
<tr>
<td>306-I</td>
<td>Thick Film 2</td>
<td>58</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Thick Film 1</td>
<td>70</td>
<td>25</td>
</tr>
<tr>
<td>307-I</td>
<td>Thick Film 2</td>
<td>93</td>
<td>25</td>
</tr>
</tbody>
</table>

Finally, for each image, the WBCs type was annotated, in addition to its location with marking a bounding box surrounding each WBC. Again, the prone-to-error data annotation process was not performed by a specialist. However, the manual detection and location of WBCs is intuitive, meaning that the annotation by a specialist is not as essential as, for instance, discerning their type. Moreover, due to cytoplasm lysis in thick blood smear preparations, WBCs can only be differentiated as mono- or polymorphonuclear. Thus, in annotating the WBCs type, even though it was not entirely performed by a specialist technician, Dr. José Manuel Costa classified a few demonstrative WBCs and provided some guidelines for an "unsupervised" annotation of the remaining cells.

5.2 WBC Candidates Extraction

5.2.1 Algorithm Overview

The proposed method for WBC segmentation in thick blood smear images is divided in three main distinctive steps. Analogously to the previous section, the initial step was FOV segmentation using
5.2 WBC Candidates Extraction

Figure 5.2: Separation of the 3-channel thick blood smear image. From left to right, top row: original RGB image, gray-scale image; bottom row: red-channel image, green-channel image, blue-channel image.

the same region-growing methodology. The importance of this step boils down to consider solely the region of interest in subsequent image processing methods. Recalling the theory described in Section 2.4.4, typical thick blood smear images present two very distinct regions: platelets and WBCs have low brightness values in contrast to the high intensity values of the background. Thus, preprocessing consisted simply on selecting the RGB channel in which the contrast between the cellular structures and background is more evidenced, as illustrated in Figure 5.2, and applying a Gaussian filter in order to attenuate possible existing artefacts.

Subsequently, a two-step thresholding method, in addition with Jaccard-index-based merging process of detached lobes, enables the extraction of WBC candidates. These two methods are described more thoroughly in the following sections.

5.2.2 Two-Step Thresholding Method

In thick blood smear images, the modal distribution of the background pixels in the histogram overshadows the lower brightness values of the cellular structures, particularly WBCs. Thus, in addition to illumination variance across the image, the straightforward application of automatic thresholding techniques may imperil proper WBC segmentation. Thus, a two-level thresholding method was proposed.

Firstly, similarly to the percentage linear stretching algorithm implemented in Section 4.2.2, the threshold value is set as a percentage of the maximum value of the histogram. Most background
region are left unsegmented as illustrated in Figure 5.3. Following an erosion morphological operation, whose purpose was to remove small binary objects corresponding to artefacts and platelets, the remaining objects segmentation were refine. This threshold refinement consisted on using the RATS method, described in Section 4.2, on the pixels within the bounding box enclosing each object. It is interesting to note that all objects connected to the complement FOV mask, i.e., touch the FOV borders, are discarded because part of its structure is not displayed in the image, which imperils the correct classification of such WBC candidates. The output of the proposed WBC segmentation methodology is shown in Figure 5.3.

However, in some cases, the nucleus of WBCs, particularly of the polymorphonuclear type, have well-marked lobes which, either due to the staining procedure or the microscopy visualization / image acquisition, separate from each other. Therefore, considering the binary objects as WBC candidates may lead to classifying a single detached lobe as mononuclear WBC or excluding it from the classification step due to its small area. Thus, further processing is required in order to circumvent this issue and the proposed methodology is described in detail in the following section.

### 5.2.3 WBC Candidates Segmentation

Briefly, the proposed methodology designed to address the issue of detached lobes belonging to the same WBC consists on identifying the objects of small size as lobe-candidates and, through expansion of the bounding box of every object, detect if the bounding box of the lobe-candidate intersect with the bounding box of either another lobe-candidate or a WBC-candidate. In intersecting cases, the lobe both bounding box are merged and the enclosing objects regarded as a WBC candidate. Otherwise, the lobe-candidates are discarded from classification consideration.

In the proposed methodology, objects with size inferior than 800 pixels per FOV area (10^9 pixels) are considered as lobe-candidates. This value was attained from experimental observations of perfectly segmented WBCs, whose size was 1120 (± 257) pixels per FOV area (10^9 pixels). Likewise, the bounding-box expansion value was also obtained from averaging the distance between detached lobes belonging to the same WBC.

As above-stated, the inclusion of a lobe-candidate in a WBC candidate depends on the intersection of its expanded bounding box with a neighbouring one, as illustrated in Figure 5.4.

---

**Figure 5.3: WBC Candidates Segmentation.** (a) original thick blood smear image; (b) segmentation result of the first thresholding (c) segmentation result following thresholding refinement.
5.3 Feature Vector Construction

Figure 5.4: Illustration of the bounding-boxes intersection-based methodology for lobes inclusion in WBC candidates: (left) two segmented objects, as the right object is considered a lobe-candidate; (right) bounding-boxes intersection within neighbouring objects.

However, in WBC clustering regions, a set lobe-candidate may intersect with several neighbouring bounding-boxes and its inclusion in a WBC candidate becomes ambiguous. In this work, the Jaccard index was the metric used as the tie-breaking criteria and it is mathematically described in Equation 5.1. In cases where the lobe-candidate is at equal distances from two neighbouring objects, its inclusion in the smallest object is prioritized. The reasoning is intuitive as the a lobe-detached WBC has smaller size. On the other hand, in cases where two equally-sized objects are at different distances from the lobe-candidate, the inclusion in the nearest one is prioritized. The sensing beneath this is also evident. Therefore, using the Jaccard index metric is ideal for this tie-breaking concept as it balances between the distance of the neighbouring objects to the lobe-candidate, which is reflected on the nominator of the Jaccard Index, and their size, prioritizing simultaneously the nearest and smallest object.

\[
J(A, B) = \frac{|A \cap B|}{|A \cup B|}
\]  

(5.1)

5.3 Feature Vector Construction

In this section, the feature vector construction process adopted is described in detail. This step is pivotal for subsequent implementation of machine learning techniques since it provides the data in which the learning algorithm produces the model able to map unlabelled input to one class.

In thick blood smear images, the main differentiating feature of WBCs as either mono- or polymorphonuclear is the nucleus, particularly its geometric properties and texture, reflecting the nucleus shape and granularity. The shape-based features are computed from the binary segmented objects obtained in Section 5.2.3 and the features used are described in Section 3.3.5.1.

On the other hand, the textural features are obtained through the grey level co-occurrence matrix method (GLCM) and are described in Section 3.3.5.2. In order to assure that GLCM is computed solely on the WBCs nucleus, the binary segmented objects obtained in Section 5.2 are used as masks, respectively. Additionally, the GLCM was computed on the grayscale images since
the nucleus brightness in the higher-contrast green-channel were undersaturated, therefore most of its textural information would be lost.

Once fully constructed, feature scaling was performed on the feature vector to range in [0, 1] according to the Equation 5.2. This step is important in order to assure proper functioning of the learning algorithms. Some algorithms lean on distance-based operations, while others lean on convergence through gradient-descent optimization algorithms, and, in these cases, feature scaling improves convergence speed and assures that a particular feature is not given more relevancy due to its larger scale.

\[ x' = \frac{x - \min(x)}{\max(x) - \min(x)} \]  
(5.2)

### 5.4 Classification Methods

In this section, the learning stage of this chapter is presented. This step aims to find patterns from the features data from which the learning algorithm can build a model capable of assigning the WBC candidates in thick blood smear images as mono- or polymorphonuclear. The supervised learning algorithms used were SVM, Decision-Tree, Naive Bayes and LDA, while the k-NN was the solely unsupervised learning algorithm tested. Their descriptions are explicit in Section 3.4.3. Furthermore, some of these learning algorithms not only are parametrized, thus they need to be optimized, but also do not embed feature selection, requiring to perform the optimization of the feature vector through reduction of its dimensionality. In the cases of those classifiers which embed feature selection as part of their overall learning operation, particularly the SVM and the Decision Tree, there is no need to optimize the feature vector. In other words, the SVM algorithm gives coefficient values to each dimension of the feature space, thus inferences about each feature relevancy can be taken from the respective coefficient value, while the Decision Tree algorithm is generated by recursive partitioning of the data, thus the features nearest to the root are more relevant to discriminate the data. The processes required for each learning algorithm are summarized in Table 5.2.

Sequential forward feature selection was employed and the criterion used was the accuracy of the resulting classifier model created on that specific feature subset, which in turn is evaluated through the 10-fold cross-validation method. In regards to parameter optimization, the perfor-

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Parameter Optimization</th>
<th>Feature Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supervised Learning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDA</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Naive-Bayes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SVM</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Unsupervised Learning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k-NN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 5.2: Classifiers used for WBC classification in thick blood smear images, and the processes used in order to optimize the classification performance of each learning algorithm
5.5 Segmentation Results

The performance was evaluated through the nested 10-fold cross-validation method with the accuracy of the classifier as the criterion.

Regarding the Support Vector Machine algorithm, the parameter optimization process aims to attain the kernel type, $C$ and, depending on the kernel type used, $\gamma$ which maximizes the classifier’s performance. The impact of these parameters on the learning algorithm performance, particularly the resulting complexity and generalization power, is previously described in Section 3.4.3.3. Two different kernel type were used - linear and Gaussian radial basis - separately, i.e. the parameters $C$ and $\gamma$ were optimized for each kernel type, while the $\varepsilon$ was set at 0.001. Moreover, the parameters combination followed a grid search approach. In the linear SVM, $C$, the single parameter to be optimized, varied from the following set of logarithmic increasing values:

$$C \in \{2^1, 2^2, 2^3, ..., 2^{12}\}$$

When using the radial basis function kernel, each parameter of the $C$-$\gamma$ combination ranged from following set of values:

$$C \in \{2^1, 2^2, 2^3, ..., 2^{13}\}$$
$$\gamma \in \{10^{-9}, 10^{-8}, 10^{-7}, ..., 10^{2}\}$$

The need for optimization of the parameters of the Decision-Tree, particularly the minimal leaf size and minimal size for splitting, may be the most intuitively comprehensible since using leaf of unitary size and low minimal size for splitting leads to clear overfitting of training data compromising the performance of the classifier. Again, grid search was the approach followed for finding the best parameters combination and ranged:

Minimal Leaf Size $\in \{1, 2, 3, ..., 25\}$
Minimal Size for Splitting $\in \{5, 10, 15, ..., 80\}$

The number of training samples to be considered closest to a particular point, $k$, is the lone parameter that requires optimization in the k-NN algorithm. Moreover, the type of measure used for finding the nearest neighbours was the Euclidean distance. The search of this value followed the ensuing set of values:

$$k \in \{1, 2, 3, ..., 10\}$$

Finally, k-NN required an optimization of the respective parameters and reduction of the feature vector dimensionality. In the scope of this thesis, the approach used was to check for variations of the optimal parameters throughout the sequential feature selection methods. In case the ideal parameters changed, the feature selection method proceeded with the new optimal parameters.

5.5 Segmentation Results

In this section, the WBC candidate extraction methodology is assessed, particularly the segmentation step and the merging of neighbouring lobe-candidates. The performance of the WBC candidates extraction methodology is depicted in Table 5.3.

Some inferences can be taken from the demonstrative results of the WBC candidate extraction, illustrated in Figure 5.5, regarding both the thresholding and the merging of the lobe-candidates.
Recalling the automatic thresholding methodology described in Section 5.2.2, the effectiveness of the double-thresholding approach is evidenced by the small number of unsegmented WBC structures, which is indicated by the low number of missed WBC candidates. The missed structures, such as illustrated 5.5, are mostly due to subsequent morphological operations, particularly erosion and opening, which discard small and thin WBC structures from the final segmentation image. Even though they contribute to number of missed WBC candidates, these morphological operations are essential to remove platelets which otherwise would eventually be merged to neighbouring WBC structures as a result of the subsequent step. Additionally, WBC structures with high brightness, which most commonly occurs near the FOV borders where the illumination is usually higher, were also unsegmented due to the harsh thresholding. However, the mere threshold value adjustment may not lead to improved WBC candidates extraction since more non-WBC objects would be segmented. This step is a trade-off between segmentation of non-WBC objects and missed WBC candidates, and since sensibility was set as the priority factor for the thresholding process, the threshold value was kept low.

However, the most significant drawback of the proposed methodology is the inability to individually segment agglomerated WBCs, i.e. neighbouring WBC nucleus, which either overlap or their borders touch, are considered as a unique WBC candidate. This shortcoming is illustrated in Figure 5.5. The combination of the highly-variant WBCs’ size and the indefinitition of the extent the borders of neighbouring WBCs touch is the major obstacle in splitting the WBC clusters, hampering the delineation of the splitting site.

On the other hand, the process of assigning small objects as lobe-candidates, searching for neighbouring WBC candidates or other lobe-candidates and ultimately merging them, enhanced the WBC candidates extraction performance. In fact, this process contributed for the correct extraction of 201 candidates whose nucleus’ structure presented detached lobes. Even though occasional undersized neighbouring nucleus were merged incorrectly as a unique WBC candidate, there were still cases of unmerged lobes since they did not filled the "small area" requisite. Therefore, analogously to the double-thresholding step, there is a trade-off between incorrectly merging undersized WBC nucleus and letting unmerged oversized lobes. The "small area" criterion defined in Section 5.2.3 was kept throughout this thesis work since it was based on the measurement of the average and standard deviation of perfectly segmented WBC nucleus.
5.6 Classification Results

The objectives outlined for this chapter are achieved in this section, i.e. the classification of WBC candidates as mono- or polymorphonuclear is addressed here. This process includes the optimization of both the parameters of the used classifiers and data dimensionality.

At first, the performance of learning algorithms was assessed, prior to any type of feature selection procedure. While the performance of the nonparametrized classifiers was evaluated through 10-fold cross-validation, the parametrized learning algorithms required the optimization of their respective parameters prior to the evaluation of their performance, which was assessed through nested 10-fold cross-validation method. Each classifier performance is shown in Table 5.4.

Table 5.4: Parameter Optimization of the parametrized classification algorithms with the entire feature vector.

<table>
<thead>
<tr>
<th>Classifier Algorithm</th>
<th>Optimization Accuracy</th>
<th>Final Accuracy</th>
<th>Best Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM: RBF</td>
<td>0.959</td>
<td>0.965</td>
<td>C = 256 ; ( \gamma = 0.1 )</td>
</tr>
<tr>
<td>SVM: Linear</td>
<td>0.965</td>
<td>0.971</td>
<td>C = 64</td>
</tr>
<tr>
<td>k-NN</td>
<td>0.958</td>
<td>0.959</td>
<td>k = 5</td>
</tr>
<tr>
<td>LDA</td>
<td>-</td>
<td>0.948</td>
<td>-</td>
</tr>
<tr>
<td>Bayes</td>
<td>-</td>
<td>0.943</td>
<td>-</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.950</td>
<td>0.918</td>
<td>Minimum Split Size = 35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minimum Leaf Size = 16</td>
</tr>
</tbody>
</table>

Figure 5.5: WBC candidates extraction results. From left to right, Top row: true WBC candidate, true WBC Candidate after merging lobe-candidates, missed WBC Candidate; Bottom row: WBC Cluster (of two WBC), unmerged WBC lobes, non-WBC object.
Figure 5.6: Normalized feature weight by the linear SVM algorithm (C = 64).

As shown in Table 5.4, the classification methods generated high accuracy, particularly the SVM algorithm with linear kernel. The high performance across all the different types of learning algorithms may be due to the relevancy of particular features, which was subsequently inquired more deeply through sequential feature selection and feature weight evaluation methods.

Recalling the theory of SVM, it consists on finding the maximal-margin separating hyperplane between two classes in the feature space. Interpreting the coefficients of the hyperplane enables to take some inferences about the relevancy of the features. Considering the example of a much higher hyperplane coefficient value on the $m$-dimension relative to the coefficient on the $n$-dimension, it can be inferred that the classes are more separable in the $m$-dimension than the $n$-dimension, thus the more relevancy of feature $m$. Figure 5.6 shows the normalized weight of the features used in the linear SVM. Analysing the features weight, it is revealed the more discriminative power of shape-based features.

Another classifier which includes feature-weighting in its conceptualization is Decision Tree. Decision Trees are generated by recursive partitioning of the value of the features, and the nearer the features are to the decision tree root, the more relevant are to the discrimination within classes. At each node, an attribute is selected to split depending upon the information gain criterion. Figure 5.7 is a representation of the optimized decision tree and the relevancy of the shape-based features is once again evidenced.

The remaining learning algorithms require optimization of the feature vector, which was achieved through reduction of the data dimensionality by sequential forward feature selection. This method consists on the evaluation of the classifiers trained with sequentially increasing feature subsets, and, at each iteration, including the feature which produces a performance increase. Figure 5.8 depicts this method for the k-NN, LDA and Naive Bayes algorithms.
5.6 Classification Results

Figure 5.7: Optimized decision tree representation.

Figure 5.8: Results of the sequential forward feature selection method employed on k-NN, LDA and Naive Bayes.
The features subset resultant from the SFFS on the k-NN was (by order of addition):

- Solidity
- Thinness Ratio
- Holes Total Area

The features subset resultant from the SFFS on the LDA were (by order of addition):

- Thinness Ratio
- Number of Connected Components
- Euler Number
- Elongation
- Holes Total Area
- Contour Area

The features subset resultant from the SFFS on the Bayes were (by order of addition):

- Solidity
- Entropy
- Thinness Ratio
- Euler Number
- Correlation
- Eccentricity
- Rectangularity

As final remarks, WBC classification as mono- or polymorphonuclear presented successful results, as every learning algorithm used presented >90% accuracy rate. Moreover, both the assessment of the features relevancy in SVM and Decision-Tree and the SFFS method revealed the discriminative power of shape-based features in detriment to texture features. This outcome is not unexpected since the low-quality nature of the image acquisition procedure in addition to the preparation of this type of blood smear, compromises the nucleus textural information. Finally, the features subsets resultant from SFFS, the 4-level decision tree and the multiple redundant and irrelevant features in linear-SVM eases the implementation of a less computationally expensive version of this methodology in a smartphone application. Not only would the complexity of the classifier be reduced, but less features would have to be computed during the feature construction step.
Chapter 6

WBC recognition in Thin Blood Smear Images

Analogously to Chapter 5, the proposed methodologies for segmentation of WBC candidates and subsequent classification as either mono- or multimorphonuclear are described in detail, this time in thin blood smear images. The initial goal outlined for this section was to classify WBCs as one of its 5 types mentioned in Section 2.4.3. However, the specialized doctor stated it was impractical to perform this type of annotation on low-quality images because the granularity of the cytoplasm, a crucial feature for such WBC differentiation, was unassessable. Nonetheless, the initial methodology of segmenting the cytoplasm was still followed. Firstly, in section 6.1, the creation and annotation of the database is elucidated. In the following sections, the proposed segmentation methodology of WBC nucleus and subsequently their respective cytoplasm is described. Thereafter, the machine learning methods for classification of WBC candidates as either mono- or multimorphonuclear are enlightened. Finally, the results from previously described methods are presented. The proposed methodology for WBC recognition in thin blood smear images is schematized in Figure 6.1.

Figure 6.1: Pipeline for WBC recognition in thin blood smear images.
6.1 Database Construction

The database used in this chapter was constituted exclusively of thin blood smear images which includes at least one WBC representative. Once again, the partnership with INSA was essential in acquiring blood smear samples. The requisites the database must meet are enumerated in the following section.

6.1.1 Database Requirements

The database used in this section should meet the following requirements:

- The thin blood smears must be collected by a specialized doctor;
- The thin blood smears collected should be obtained from various patients;
- The thin blood smear images should be acquired with 1000x magnification;
- The thin blood smear images should be acquired with a low-cost commercial microscope using a smartphone device with a high-resolution camera;
- Every thin blood smear image should contain at least one WBC;
- The database should include an approximately equally distributed number of blood smear images from each patient;
- The WBCs type should be annotated by a specialist.

6.1.2 Database Construction

In this section, the image acquisition process was identical to the one described in Section 4.1.2. The acquired images were assessed and only those which included at least one WBC were added to the database. Recalling the normal blood cells count mentioned in Section 2.4, the \(~600\)-1 ratio of RBCs to WBCs reveals the scarcity of WBC in blood smear images. Thus, in order to complete, it was required to turn to an already available database at Fraunhofer AICOS for additional thin blood smear images.

Analogously to Section 5.1.2, the WBCs type and location were annotated by a non-specialist. The manual location of WBCs is rather intuitive and the annotation of each WBC category (mononuclear or polymorphonuclear) followed a set of guidelines provided by the specialist doctor. At the end, the database constituted a total of 193 thin blood smear images with at least one WBC representative.

6.2 WBC Candidates Extraction

6.2.1 Algorithm Overview

The proposed method for WBC segmentation in thin blood smear images has a lot of similitudes with the method described in 5.2, particularly in regards to nucleus segmentation.

Again, preprocessing consisted solely on region-growing-based FOV segmentation and selection of the higher-contrast green-channel, followed by Gaussian blurring. The brightness-distinct
6.2 WBC Candidates Extraction

Figure 6.2: WBC nucleus candidates extraction results (left) without and (right) with the Jaccard-based lobe-merging methodology.

WBC nucleus to RBCs and background, particularly in green-channel images, eases the nucleus segmentation process. Thus, the same principles of a two-step thresholding method with a Jaccard-index-based merging process of detached lobes were used for WBC nucleus candidates extraction. Subsequently, a cytoplasm segmentation segmentation method was proposed based on automatic thresholding followed by a distance-transform-based method for splitting the cytoplasm from overlapping neighbouring RBCs.

6.2.2 WBC Nucleus Segmentation

As aforementioned, the WBC nucleus segmentation method was based on the two-step thresholding method described in Section 5.2. Firstly, benefiting from the brightness contrast of WBC nucleus relative to RBCs and background, an automatic global threshold value, which is set as a percentage of the maximum value of the histogram, was applied. Subsequently, after the application of an erosion morphological operation in order to remove small artefacts and platelets, the threshold value in each of the remaining segmented regions is refined through using the RATS method on the pixels within the bounding box enclosing each of them.

Subsequently, the lobes of the WBCs, particularly of the polymorphonuclear type, may appear detached from each other, but still belong as the same WBC. This issue is addressed through a merging methodology based on the overlapping of neighbouring expanded bounding-boxes enclosing the previously segmented objects, as described in Section 5.2.3. Exemplifications of the resulting WBC nucleus candidates are illustrated in Figure 6.2

6.2.3 WBC Cytoplasm Segmentation

In this WBC candidates extraction methodology, the segmentation of the cytoplasm is performed from the WBC nucleus extraction results. It is assumed that the robustness of the WBC nucleus segmentation is considerably superior than cytoplasm segmentation due to the higher contrast of the former structures to RBCs and background. In fact, the brightness of WBC cytoplasmic structures are extremely similar to RBCs in thin blood smear images. In simpler words, this method
assumes that WBC nucleus are correctly segmented and, thenceforth, the cytoplasm segmentation region has the restriction of being attached to the nucleus.

Firstly, the brightness value of the cytoplasm is estimated from the nucleus adjacent regions. The subtraction of the dilated by the original nucleus binary image results in the nucleus-adjacent-regions mask from where the brightness value is estimated. Once estimated, a thresholding operation is applied on the surrounding region of the respective WBC nucleus candidate in order to segment the cytoplasmic regions.

Assessing the images resulting from thresholding operation in Figure 6.3, the overlapping existence of RBCs and WBC cytoplasm compromises the straightforward application of the aforementioned automatic thresholding method. In order to circumvent this issue, a splitting method for touching cytoplasm and neighbouring RBCs was proposed and it is based on the morphological reconstruction operation with further restrictions in regards to the gradient direction of the distance-transform.

Morphological reconstruction involves two images, the mask-image \( F \) and a subset of the mask-image \( G \), the marker-image \( (F \subseteq G) \). This iterative process consists on the dilation of the marker-image while the mask-image constrains the transformation, i.e. the dilation of the marker-image is forced to fit within the mask-image, until the image values stop changing.

\[
h_{k+1} = (h_k \oplus B) \cap G, \text{ until } h_{k+1} = h_k
\]

This morphological operation is useful to extract marked objects or, in combination with an erosion operation, remove small objects while keeping the original shape of the larger objects.

The principles of the morphological reconstruction in binary images are withdrawn for the cytoplasm-RBCs splitting methodology. At first, the distance-transform is computed in the cytoplasm binary images. Subsequently, the markers are set through a subtraction of the dilated distance-transformed by the original image. This operation enables to find the distance-transform peaks which are located where the subtraction is null. However, since the cytoplasm image includes several objects (RBCs and WBC), multiple peaks result from the previously described
6.2 WBC Candidates Extraction

Figure 6.4: Two examples of the gradient-restricted morphological reconstruction methodology. From left to right: segmented objects; distance-transformed images; gradient direction of the distance-transformed images; original images with the markers in green; dilation fronts; output morphological reconstructed images.

operation which goes against the interest of having a sole peak respective to the WBC centre. This issue is evaded through considering the overlapping, or nearest, peak to the WBC nucleus.

In parallel, the gradient direction of the distance-transform image is computed according the mathematical equations depicted in Section 3.3.1. The resultant image acts as one further restriction in the morphological reconstruction operation applied on the marker-image.

Finally, once the gradient-direction and the marker-image are computed, a variant of the morphological reconstruction is applied, with the original cytoplasm segmentation image as the mask. The marker-image is iteratively dilated while it fits within the mask image and while the dilation front finds gradient-values pointing at congruent directions. In other words, in a particular iteration, the new pixels are included in the marker-image if the gradient-direction at that location is pointing towards previously marked pixels within a 8-connected neighbourhood. Figure 6.4 display two examples of the aforedescribed method.

The purpose of this methodology is to morphological reconstruct the cytoplasm from the WBC centre up to the frontier between the WBC cytoplasm and the overlapping RBC, site where the gradient-direction of distance-transformed image turns towards the RBC centre. Results from this methodology are illustrated in Figure 6.5.
6.3 Feature Vector Construction

In this section, the feature vector construction process is described in detail, which was pivotal for subsequent machine learning techniques. Again, the main differentiating feature of WBCs as either mono- or polymorphonuclear is the nucleus, particularly its geometric properties and texture, reflecting the nucleus shape and granularity. The shape-based features were computed from the nucleus binary images attained in Section 6.2.2 and the features used are described in Section 3.3.5.1. The textural features were obtained through the GLCM of the grayscale images, computed solely on the WBC nucleus, i.e. binary nucleus image are used as masks.

6.4 Classification Methods

In this section, the learning stage of this chapter is presented. This step aims to attain learning algorithm models capable of assigning the WBC candidates in thin blood smear images as mono- or polymorphonuclear. The learning algorithms used were SVM, Decision-Tree, Naive Bayes, LDA and k-NN. Again, the learning algorithms which require either parameter optimization or feature selection, or even both, are summarized in Table 5.2 in Section 5.4.

The sequential forward feature selection was employed, with the accuracy of the classifier as the criterion, which is evaluated through 4-fold cross-validation. On the other hand, the parameter optimization process was assessed through nested 4-fold cross-validation and followed a grid search of parameters combination, as summarized in Table 6.1.
Table 6.1: Parameter combination of the optimization by grid search process.

<table>
<thead>
<tr>
<th>Classifier Algorithm</th>
<th>Parameter Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM: RBF</td>
<td>$C \in {2^1, 2^2, 2^3, ..., 2^{12}}$, $\gamma \in {10^{-9}, 10^{-8}, 10^{-7}, ..., 10^{2}}$</td>
</tr>
<tr>
<td>SVM: Linear</td>
<td>$C \in {2^1, 2^2, 2^3, ..., 2^{12}}$</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>Minimal Leaf Size $\in {1, 2, 3, ..., 25}$</td>
</tr>
<tr>
<td>k-NN</td>
<td>$k \in {1, 2, 3, ..., 10}$</td>
</tr>
</tbody>
</table>

6.5 Segmentation Results

In this section, the effectiveness of the WBC candidate extraction methodology for thin blood smear images is assessed, particularly the WBC nucleus extraction method and the ensuing cytoplasm segmentation. The performance of former is depicted in Table 6.2.

Evaluating Table 6.2 in conjunction with the qualitatively assessment of the resultant images, such as the ones illustrated in Figure 6.6, the effectiveness of the WBC nucleus candidates extraction method can be verified. Despite the 87.6% accuracy of true candidates, the high sensitivity (98.5%) of the proposed method insures an efficient study triage. Assessing the false negative results, i.e. WBCs not detected in the WBC nucleus extraction methodology, the missed WBCs consisted mostly on detached lobes which, through morphological erosion and opening operation, were discarded due to their small size, as illustrated in Figure 6.6. Still, these morphological operations are pivotal in removing artefacts in intermediate steps. On the other hand, the false positive results were mostly due to either platelets clustering (Figure 6.6, bottom-left image) or multiplicative non-WBC objects existent in the images as a result of the microscopic observation (Figure 6.6, bottom-middle image).

Regarding the cytoplasm segmentation methodology, some inferences can be withdrawn from qualitatively evaluating the resulting images, such as illustrated in Figure 6.7. Even though it proved to be efficient for isolated WBC (Figure 6.7, left image), the robustness of this methodology decreased in WBCs which overlapped with neighbouring RBCs. Apart from the inherent difficulty in segmenting the cytoplasm in these situations, the degree of overlapping affects the efficiency of the proposed methodology. In fact, in cases where the RBCs slightly touch the WBC (Figure 6.7, both middle images), the morphological-reconstruction-based methodology was capable of separating the cytoplasm from the neighbouring RBCs. However, it proved to be inadequate in situations with high degree of overlapping. This was mostly due to the fact that the large extent

Table 6.2: WBC candidates extraction results.

<table>
<thead>
<tr>
<th>Metric</th>
<th>No. of Cases</th>
<th>No. Cases per Thin Smear Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>True WBC Candidates</td>
<td>263</td>
<td>1.36</td>
</tr>
<tr>
<td>Missed WBC Candidates (FN)</td>
<td>4</td>
<td>0.02</td>
</tr>
<tr>
<td>Non-WBC objects (FP)</td>
<td>33</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Figure 6.6: WBC candidates extraction results. From left to right, top row: true WBC candidate, true WBC Candidate but deficient cytoplasm segmentation; Bottom row: non-WBC object, non-WBC object (clustering of platelets), missed WBC Candidate.

of RBC overlapping (\(~ \sim 50\%\) of the RBC border touches the WBC cytoplasm) leads to the non-existence of an inflection point in the gradient-direction of the distance-transformed image, i.e. , that would restrict the morphological reconstruction past the cytoplasm border. In order to obtain a performance value of the cytoplasm segmentation, a acceptably segmented cytoplasm is required to overlap with at least 80\% of the ground truth and less than 10\% of the non-cytoplasmic ground truth region, usually represented by neighbouring RBCs. Thus, the performance of this method on the true-positive WBC nucleus extracted on the previous step was 77\%, i.e. 202 of the 263 correctly extracted WBC nucleus had their respective cytoplasm correctly segmented according to the aforementioned criteria adopted.

Figure 6.7: WBC cytoplasm segmentation. Top row: binary images resultant from the cytoplasm automatic thresholding method; Bottom row: respective original images with both the nucleus (yellow) and cytoplasm (green) marked.
Table 6.3: Parameter Optimization of the parametrized classification algorithms with the entire feature vector.

<table>
<thead>
<tr>
<th>Classifier Algorithm</th>
<th>Optimization Accuracy</th>
<th>Final Accuracy</th>
<th>Best Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM: RBF</td>
<td>0.919</td>
<td>0.887</td>
<td>C = 32 ; γ = 0.00000001</td>
</tr>
<tr>
<td>SVM: Linear</td>
<td>0.905</td>
<td>0.981</td>
<td>C = 8</td>
</tr>
<tr>
<td>k-NN</td>
<td>0.900</td>
<td>0.924</td>
<td>k = 5</td>
</tr>
<tr>
<td>LDA</td>
<td>-</td>
<td>0.901</td>
<td>-</td>
</tr>
<tr>
<td>Bayes</td>
<td>-</td>
<td>0.796</td>
<td>-</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.901</td>
<td>0.943</td>
<td>Minimum Split Size = 70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minimum Leaf Size = 3</td>
</tr>
</tbody>
</table>

### 6.6 Classification Results

The classification of WBC candidates as mono- or polymorphonuclear is addressed in this section. Firstly, the performance of the learning algorithms was assessed, including parameter optimization for the parametrized classifiers, prior to any type of feature selection procedure. The performance evaluation was performed as described in Section 6.4 and the results are depicted in Table 6.3.

Analogously to the WBC classification process in thick blood smear images, the SVM algorithm with linear kernel had the highest accuracy performance across all classifiers. Figure 5.6 shows the normalized weight of the features used in the linear SVM. The weight of the features is based on the coefficients of the maximal-margin separating hyperplane on the feature space, revealing the relevancy of each feature. Similarly to Chapter 5, the higher discriminative power in WBC classification of shape-based features relative to textural features is evidenced.

Decision-Tree classifier also had high accuracy performance. This is another learning algorithm with embedded feature-weighting in its conceptualization. Figure 5.7 is a representation of the optimized decision tree and it demonstrates the higher relevancy of shape-based features. In fact, the simplicity of the decision-tree is an interesting characteristic which creates good perspectives for integration of this methodology in mobile devices.

Furthermore, the lower performance of the remaining learning algorithms may be caused by redundant features in the complete feature vector. Thus, the optimization of the feature vector was performed through sequential forward feature selection. This data-dimensionality-reduction method consists on iteratively increasing the feature subset by including the feature which leads to an increased classifier performance.

The features subset resultant from SFFS on the k-NN, with an accuracy of 0.927, was (by order of addition):

- Eccentricity
- Number of Connected Components
- Circularity

The features subset resultant from SFFS on the Bayes, with an accuracy of 0.909, were (by order of addition):
Figure 6.8: Normalized feature weight by the linear SVM algorithm (C = 32).

Figure 6.9: Optimized decision tree representation.
6.6 Classification Results

- Solidity
- Bounding Box Area
- Contour Area

The features subset resultant from SFFS on the LDA, with an accuracy of 0.924, were (by order of addition):

- Waviness Shape Factor
- Holes Total Area
- Entropy
- Energy
- Mean Column
- Euler Number
- Max Axis Length
- Number of Connected Components
- Correlation

In conclusion, WBC classification as mono- or polymorphonuclear in thin blood smear images produced efficient results as every optimized learning algorithm, either through optimization of its parameters or the feature data dimensionality, had >90% accuracy rate. Additionally, the higher relevancy of shaped-based features relative to texture features is confirmed through the assessment of the features weight in SVM and Decision-Tree and the predominant presence of shape-based features in the SFFS-optimized feature subsets in k-NN, LDA and Bayes. Again, the low-quality of the blood smear images hampers the extraction of the textural information of the nucleus. Analogously to Chapter 5, the reduced complexity of the optimized feature subsets and the assessment of the feature relevancy in SVM and Decision-Tree brings good perspectives towards implementing the WBC segmentation, feature vector construction and classification methodologies in a smartphone application.
Chapter 7

Integration in the MalariaScope Application

The ultimate goal of this thesis work is the integration of the developed image processing methodology in the MalariaScope project, particularly in the automatic analysis system included in the MalariaScope mobile-application. At first, an overview description of the MalariaScope application is presented. Subsequently, this chapter addresses the integration of image processing methodologies developed throughout this study in the application.

7.1 MalariaScope Application Overview

The MalariaScope prototype application has a built-in database which collects and stores the patients’ information, sessions’ record and blood smear images. This database relational model is very straightforward and it succinctly consists on the the one-to-many relationship between the "patient" and the "session" entities and between the "session" and the "sample" entity. In other words, the data is collected in the following fashion: each patient can potentially reference several sessions, which in turn can reference several images. For storage purposes, the images are saved in the mobile-device internal storage instead of the database for a few reasons. The file system storage is optimized in terms of computational cost and no accessibility advantages are obtained from saving in the database. In addition, even though it is more reliable to control access to the images if stored in the database, no privacy is violated if the images saved to the internal storage are set to private to the MalariaScope application, which warrants that nor the user nor other applications can access them. Thus, each "sample" attribute stores its respective file path.

Considering the activity-driven user-interacting architecture of Android applications, the MalariaScope activity map is analogous to its database relational model. After an opening splash-screen activity displaying the application logo, a ListActivity displays the currently database-stored patients which are selectable. Once selected, a subsequent ListActivity displays the sessions history for that particular patient, which once again are selectable. Selecting one of his sessions leads to a new ListActivity displaying the respective currently-stored images. Throughout these Lis-
There are options for the creation and deletion of patient profiles, sessions or samples respectively, and, regarding the addition of new samples, it may either be from the mobile-device gallery or acquired from the camera and subsequently stored in the mobile device. Additionally, the images can be analysed. At this stage of the prototype development, an external image processing module already addresses the parasitaemia levels of the images through a back-end server. Briefly, it consists on sending the images to the external server where the image processing module is included and, after remotely processing the images, sends back the output, in this case the parasitaemia levels as the number of trophozoites. This is processed in a working thread, enabling the user to continue using the application and, once the analysis is completed, the user is notified. The analysis is then displayed in a new activity in the form of a report.

7.2 Application Requirements

Recalling the objectives of the MalariaScope project, the main goal is to create a mobile-based solution that can provide an effective pre-diagnosis of malaria to be used in medically underserved areas. Thus, the technical and technological liabilities in these regions should be attentively considered during the development of the MalariaScope, particularly the automatic analysis system.

For the integration of the image processing modules developed in this work, the mobile application should meet three major requirements: it must be intuitive for medical use, should have a built-in database for storage of patient’s sessions and samples, and the image processing should be fast, while not disrupting the application flow. As described in the previous section, the MalariaScope application already has an built-in database and has already been tested and favourably regarded as intuitive for medical use. Therefore, the focus in this work was centred solely on integrating the image processing modules.

There are three major factors that impact the integration of the image processing methodology developed in this thesis: the promptness of the analysis, the computational cost and the availability.

The processing modules may be integrated locally in the application or in a remote server and serve as a back-end application which supports the MalariaScope application. Despite the evident benefits in terms of computational cost and, consequently, promptness of the analysis of the back-end integration, this system requires internet accessibility for transferring images and analysis results between the back-end and front-end applications, which is still very limited in these regions. On the other hand, the inclusion of the image processing methodology within the application arises several computational difficulties due to the ever-increasing but still limiting computational power of mobile devices, i.e. even though efforts have been employed in reducing the gap between the hardware of mobile devices and desktop computers in processing power, storage capacity and energy, the difference in computational power is still significant. In fact, computer vision applications are computationally demanding due to the large amount of data to process and the complexity of the algorithms, which additionally leads to faster power consumption. The major benefit of local integration is that the analysis results would still be obtainable in regions voided of internet accessibility.
Therefore, the methodology integration approach is up to the application developer to weight these factors and make a commitment between computational cost and availability of the analysis to the user. In the context of this work, the proposed processing methodologies are integrated locally in the mobile-device application, process which is detailed in the following section.

7.3 Image Processing Module Integration

Analogously to the parasitaemia analysis module, the image processing methodology is integrated as a selectable option in the "Sample" ListActivity, i.e. the methodology module can be called individually for each image listed. Thereafter, the results are displayed in the "Report" Activity.

Since the image processing methodology is a long-running Central-Processing-Unit (CPU)-intensive task, it is key to move this work off the User-Interface (UI) thread and into the background, which was done through IntentService, a class which handles each request in turn using a worker thread. Once the service is stopped, the user is prompted and directed to the "Report" Activity. Briefly, the information of the selectable option, such as the id of the user, session and sample, is sent from the "Sample" ListActivity to the IntentService. In this Android class, a new class is created which stores the OpenCV Mat structure of the respective selected image, and the processing modules are integrated in this class.

Regarding the image processing implementation in the prototype application, several modifications were considered. Firstly, there is the need to integrate the methodology in C++ using the Native Development Kit (NDK). Even though the task can be coded in Java programming language - the programming language used in Android -, and even though coding in C++ expands significantly the size of the application package because the native code needs to be built for every device architecture, using the NDK improves significantly the processing speed. In fact, the employment of several morphological operations and other methods which require direct access to pixels hampers the performance of the application when using the Java-based Software Development Kit (SDK), most noticeable in terms of processing speed.

In order to ensure the modules integration is universal across mobile-devices, particularly regarding the different camera resolutions of smartphones, some variables dependant of the image resolution, such as the structuring element size in morphological operations, vary as function of the FOV diameter.

Additionally, some methods and algorithms were simplified in order to increase computation speed and lighten the CPU burden, as well as deallocate temporary variables in order to relieve the device’s memory. Such simplifications include the FOV segmentation which was previously performed by region growing and the modified method is based on the Otsu’s method, the resize of the image matrices once their full detail is no longer required nor provides segmentation performance advantages, the reduction of intermediate morphological operations which do not lead to an increase in performance of the image processing modules, and the use of solely one thresholding method, either the MW Otsu’s or the MW RATS method, for the regression analysis. Figure 7.1 displays some use-case examples of the MalariaScope application.
Figure 7.1: Use-case Screenshots of the MalariaScope application, with particular focus on the methodology integration. From left to right, top row: MalariaScope logo on the opening splash-screen activity, ListActivity displaying the images respective to one session of a particular patient, Dialog displaying the options to be executed on the selected image, Dialog prompting the user that the module is be processed in the background; middle row: notification bar informing the current state of the processing module, Dialog notifying the user that the processing module is complete, "Report" activity displaying the processing results as well as the thumbnails of the original and processed images; bottom row: input thin blood smear image, image after RBC segmentation, input thick blood smear image, image after WBC detection and segmentation.
Chapter 8

Conclusions and Future Work

Blood tests are one of the most commonly requested diagnostic tests since most dysfunctions induce fluctuations from normal blood biochemical, molecular and cellular levels. In particular, the cellular information can be retrieved from complete blood count tests which nowadays are preferably performed through automated analysers instead of the labour-intensive manual count through microscopic analysis. Considering that this work is part of the broader project, MalariaScope, aiming to assess the malaria parasitaemia through mobile devices, not only are the automated analysers inadequate for the assessment of the parasites but also the malaria-affected regions lack such health resources. Thus, this work focus on image processing approaches for the automatic analysis of blood cells in microscopic images.

Interest in computer vision approaches as faster alternatives for labour-intensive time-consuming visual blood analysis has been growing. However, their limited efficiency for processing low-quality images and computational cost for implementation in a mobile-device framework makes them unsuitable for the project. In this work, an image processing methodology for blood cell counting and differentiation in low-quality images is presented and its integration in Android mobile devices is addressed.

In the chapter of literature review, even though existing studies relied on high-quality images during the development of their methodologies, several relevant considerations were withdrawn, particularly the characteristic relationship between brightness of each cellular components, the cellular features useful for recognition and the typical image processing methodology path, from the segmentation of each of the cell type until the classification.

One of the main goals of this work was to develop a robust image processing methodology for estimation of the number of red blood cells. The proposed algorithm consisted on the segmentation of the erythrocytes and subsequent estimation based on the both the number of objects and area segmented. The algorithms employed in segmentation step were capable of successfully detecting RBCs in low-quality illumination-variant images, with RMSEs of 12.92 and 12.79 for the local application of the RATS and the Otsu’s automatic thresholding algorithms respectively. Future improvements should focus on improving the segmentation of RBC clusters in their individual cellular constituents since the efficiency of the proposed distance-transform-based methodology
Conclusions and Future Work

decreased with the overlapping level and number of RBCs per cluster. Moreover, the regression analysis with the linear Support Vector Regression algorithm eased the shortcomings of the segmentation step and enabled to make accurate RBC count estimations, decreasing the RMSE to 10.24.

The remaining objectives were to detect WBCs and subsequent classification as mono- or polymorphonuclear. Even though it was initially proposed to classify WBCs in its 5 types, it was indicated by the specialized doctor that the low-quality images did not enable to assess the small intrinsic features, particularly the granularity, to manually assign the WBCs candidates as one of the five types. Moreover, these goals were employed in thin and thick blood images and, despite having some different traits, the methodologies had a common pipeline. The distinctive brightness of WBCs nucleus prompted their segmentation as the first step of the WBC extraction methodology. In addition, the merging process of neighbouring lobes improved significantly the efficiency of the WBC extraction step. Nonetheless, there is room for future improvements in the Jaccard-based merging method, particularly the stretched bounding-box approach used in this work could be replaced by a more suitable enclosing geometry that does not intrinsically prioritize some directions, such as stretched enclosing circle or stretch convex hull. In thin blood smear images, the segmentation of the cytoplasm follows a consequential path of the WBC nucleus segmentation since the methodology proposed assumes that the nucleus are correctly detected and segments the cytoplasm in the surrounding regions. The automatic thresholding technique in conjunction with the gradient-based morphological reconstruction method for splitting the cytoplasm from overlapping RBCs had adequate results. In fact, the low-quality nature of the images in association with high levels of overlapping RBCs compromised the proper cytoplasm segmentation. Future work should reside on improving the robustness of this step despite these unfavourable conditions. Regarding the classification step, the recognition of the WBCs candidates as mono- or polymorphonuclear was accurate across the several learning algorithms used in this study, particularly the Support Vector Machine, with 97.1% accuracy, and k-Nearest Neighbouring, with 95.9% accuracy, and in thick blood smear images. The feature relevancy analysis performed on the various learning algorithms, through feature selection processes or through the weight the algorithms embeddedly gave (in SVM and Decision-Tree) revealed the best discriminative power of shape-based features, particularly the solidity of the WBC nucleus.

As future work for both image processing methodologies, the algorithms developed should be tested in databases of high-quality images, and vice-versa (the performance of the state-of-art algorithms in low-quality images should be assessed), in order to make fair performance comparisons between these and previously developed methodologies. In addition, the ground truth annotation, and thereby results validation, should be completely performed by medical specialists in future work.

Finally, the integration of the proposed image processing methodologies in the MalariaScope prototype application was an important starting point for ultimately implementing them in the final prototype of the MalariaScope project. Despite the successful integration using NDK in the Android application, future efforts should focus on the optimization of this module in terms
of computation speed and processing cost, as the current version presents a few memory leak problems. In addition, the medical validation of the image processing results obtained from the application should be concerned in future work.
References

REFERENCES


