Recent advances of Malaria parasites detection systems based on mathematical morphology techniques

Andrea Loddo[1], Cecilia Di Ruberto[1], and Michel Kocher[2]

[1] *Department of Mathematics and Computer Science, Via Ospedale 72, 09124, Cagliari, Italy*
{*andrea.loddo, dirubert*}*@unica.it*
[2] *Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
michel.kocher@heig-vd.ch

Abstract

This paper investigates existing mathematical morphology based techniques applied for performing malaria parasites detection and identification in both Giemsa and Leishman stained blood smears images. Malaria is an epidemic health disease and a rapid, accurate diagnosis is necessary for proper intervention. Generally, pathologists visually examine blood stained slides for malaria diagnosis; this kind of visual inspection is subjective, error-prone and time consuming. In order to cope with such issues, computer-aided methods have been increasingly evolved for abnormal erythrocyte detection, segmentation and semi/fully automated classification. The aim of this paper is to present a review on the methods that are based on mathematical morphology techniques for malaria parasite detection.

Keywords— malaria, red blood cells segmentation, mathematical morphology, medical image analysis

1 Introduction

Haematology is the branch of medicine concerned with the study, diagnosis, monitoring, treatment, and prevention of diseases related to the blood and blood-forming organs. Haematology studies the blood in health and pathological conditions, firstly to identify the patient's health condition and, secondly, to predict how the bone marrow may have contributed to reach that condition.

Thus, haematology studies the relationship between the bone marrow and the systemic circulation. In fact, there are many diseases, disorders, and deficiencies that can affect the number and type of blood cells produced, their function and their lifespan. Usually, only normal, mature or nearly mature cells are released into the bloodstream but certain circumstances can induce the bone marrow to release immature and/or abnormal cells into the circulation. One of the most frequently ordered test to monitor the proportion of the cell

components into the blood stream is the Complete Blood Count (CBC), that offers various haematologic data represented by the numbers and types of cells in the peripheral circulation. The cells percentage is compared with the reference ranges in order to determine if the cells are present in their expected percentage, if one cell type is increased, decreased or if immature cells exist. Reference ranges for blood tests are sets of values used to interpret a set of diagnostic test results from blood samples. Since it is difficult to prove that healthy-considered subjects may not have infections, parasitic infection and nutritional deficiency, it is more feasible to talk about reference ranges rather than normal ranges. A reference range is usually defined as the set of values in which 95% of the normal population falls within. It is determined by collecting data from vast numbers of laboratory tests result from a large number of subjects who are assumed to be representative of the population. With automatic counters or the flow cytometry an automated CBC can be performed quickly. However, if the results from an automated cell count indicate the presence of abnormal cells or if there is a reason to suspect that abnormal cells are present, then a blood smear will be collected [18]. A blood smear is often used to categorize and/or identify conditions that affect one or more types of blood cells and to monitor individuals undergoing treatment for these conditions. The results of a blood smear typically include a description of the cells appearance, as well as any abnormalities that may be seen on the slide. The manual analysis of blood smears is tedious, lengthy, repetitive and it suffers from the presence of a non-standard precision because it depends on the operator's skill. The use of image processing techniques can help to analyse, count the cells in human blood and, at the same time, to provide useful and precise information about cells morphology.

Peripheral blood smears analysis is a common and economical diagnosis technique by which expert pathologists may obtain health information about the patients. Although this procedure requires highly trained experts, it is error-prone and could be affected by inter-observer variations. Moreover, blood cells images taken from microscope could vary in their illumination and colouration conditions. Typical blood cells images contain three main components of interest: the platelets (or thrombocytes), the red blood cells (or erythrocytes) and the white blood cells (or leukocytes). It is worth considering that blood cells exist with different shapes, characteristics and colourations, according to their types.

Many tests are designed to determine the number of erythrocytes and leukocytes in the blood, together with the volume, sedimentation rate, and haemoglobin concentration of the red blood cells (blood count). In addition, certain tests are used to classify blood according to specific red blood cell antigens, or blood groups. Other tests elucidate the shape and structural details of blood cells and haemoglobin and other blood proteins. Blood can be analysed to determine the activity of various enzymes, or protein catalysts, that either are associated with the blood cells or are found free in the blood plasma. Blood also may be analysed on the basis of properties such as total volume, circulation time, viscosity, clotting time and clotting abnormalities, acidity (pH), levels of oxygen and carbon dioxide, and the clearance rate of various substances. There are special tests based on the presence in the blood of substances characteristic of specific infections, such as the serological tests for syphilis, hepatitis, and human

immunodeficiency virus (HIV, the AIDS virus) [1].

Among the several available blood tests, the most common are certainly the blood cells counts, e.g., a CBC is a measure of the haematologic parameters of the blood. Included in the CBC is the calculation of the number of red blood cells (red blood cell count) or white blood cells (white blood cell count) in a cubic millimetre (mm^3) of blood, a differential white blood cell count, a haemoglobin assay, a hematocrit, calculations of red cell volume, and a platelet count. The differential white blood cell count includes measurements of the different types of white blood cells that constitute the total white blood cell count: the band neutrophils, segmented neutrophils, lymphocytes, monocytes, eosinophils, and basophils. A specific infection can be suspected on the basis of the type of leukocyte that has an abnormal value.

Human malaria infection is not strongly related to cells count but it needs different tests in order to be identified. It can only be caused by parasitic protozoans belonging to the Plasmodium type. The parasites are spread to people through the bites of infected female Anopheles mosquitoes, called "malaria vectors". There are five parasite species that cause malaria in humans and two of these species, Plasmodium falciparum and Plasmodium vivax, constitute the greatest threat. Plasmodium ovale, Plasmodium malariae and Plasmodium knowlesi are the three remaining species which are less dangerous in human [37], as shown in fig.1. All five species may appear in four different life-cycle stages during the infection phase in peripheral blood: ring, trophozoite, schizont and gametocyte. Some examples are shown in fig.2. The life-cycle-stage of the parasite is defined by its morphology, size and the presence or absence of malarial pigment. The species differ in the changes of infected cell's shape, presence of some characteristic dots and the morphology of the parasite in some of the life-cycle-stages [32].

Computer vision techniques for malaria diagnosis and recognition represent a relatively new area for early malaria detection and, in general, for medical imaging, able to overcome the problems related to manual analysis, that is performed by human visual examination of blood smears. The whole process requires an ability to differentiate between non-parasitic stained components/bodies (e.g. red blood cells, white blood cells, platelets, and artefacts) and the malarial parasites using visual information. If the blood sample is diagnosed as positive (i.e. parasites present) an additional capability of differentiating species and life-stages (i.e. identification) is required to specify the infection. Numerous methods of automatic malaria diagnosis have been proposed so far, in order to overcome the issues before mentioned. The aim of this paper is to review and analyse the works of different researchers who in particular have used mathematical morphology as a tool for computer aided malaria detection and classification.

1.1 Mathematical morphology

Mathematical morphology (MM) can be defined as a theory for the analysis of spatial structures. It is called morphology because it aims at analysing the shape and form of objects. It is mathematical in the sense that the analysis is based on set theory, integral geometry, and lattice algebra. MM is not only

[1] https://www.britannica.com/topic/blood-analysis

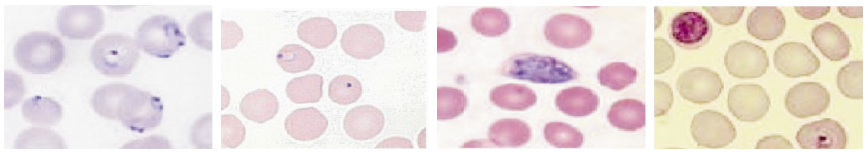

Figure 1: Types of human malaria: from left to right, P. Falciparum, P. Vivax, P. Ovale, P. Malariae [16].

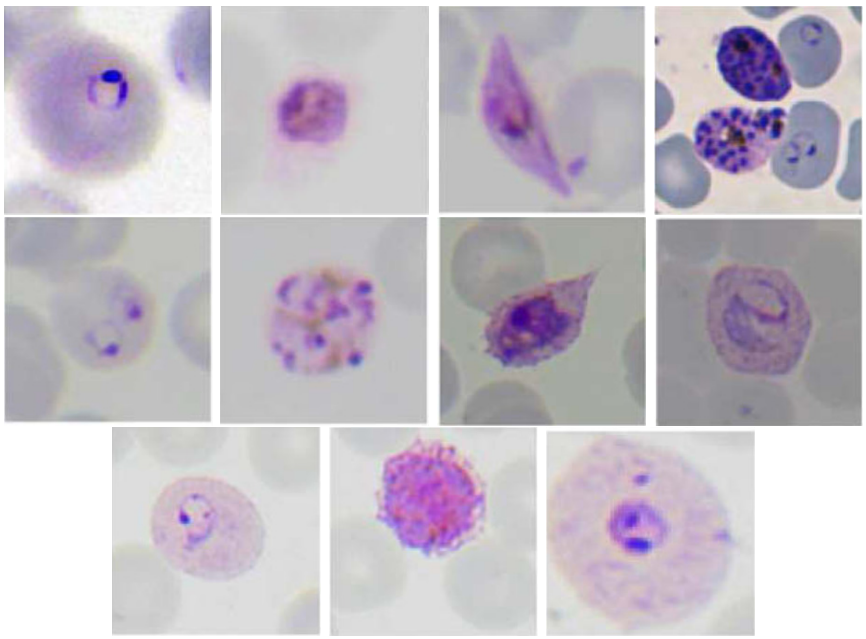

Figure 2: Examples of malaria parasite stages. From top left: P.falciparum ring, trophozoite, gametocyte, schizont; P.malariae ring and schizont; P.ovale and P.vivax trophozoites; P.vivax ring and gametocyte; P.vivax ring [36]

a theory, but also a very powerful image analysis technique [29]. It was introduced by Matheron as a technique for analysing geometric structure of metallic and geologic samples. It refers to a branch of non-linear image processing and analysis that concentrates on the geometric structure within an image. The morphological filter, which can be constructed on the basis of the underlying morphological operations, are more suitable for shape analysis than the standard linear filters since the latter sometimes distort the underlying geometric form of the image. Some of the salient points regarding the morphological approach are as follows [12]:

1. Morphological operations provide for the systematic alteration of the geometric content of an image while maintaining the stability of the important geometric characteristics.

2. There exists a well-developed morphological algebra that can be employed for representation and optimization.

3. It is possible to express digital algorithms in terms of a very small class of primitive morphological operations.

4. There exist rigorous representations theorems by means of which one can obtain the expression of morphological filters in terms of the primitive morphological operations.

Dilation and erosion are the basic morphological processing operations. They are defined in terms of more elementary set operations, but are employed as the basic elements of many algorithms. Both dilation and erosion are produced by the interaction of a set called structuring element with a set of pixels of interest in the image. The structuring element has both a shape and an origin.

In this paper we present a review of computer-aided methods oriented to malaria parasites detection and segmentation by mathematical morphology based techniques. The rest of the paper is organised as follows. Section 2 introduces the adopted materials and methods for malaria-infected peripheral blood cells images analysis. It has been structured in preprocessing, segmentation, feature extraction, feature selection and classification subsections, consistently with a typical pipeline of a computer-aided image analysis process. All the considered works make use of morphological operators in at least one of the phases of image analysis. Section 3 contains an overall discussion about the methods and the conclusions are expressed in section 4.

2 Materials and methods

This section presents a review of some of the main recent studies existing in literature regarding the analysis of malaria infected blood smears using mathematical morphology. A typical approach usually comprises five different image processing and analysis tasks, as follows:

1. Pre-processing.

2. Segmentation.

3. Feature extraction.

4. Feature selection.

5. Classification.

Since morphological techniques have been used in the first three phases, the reviewed works have been divided into the following sub-sections: pre-processing, segmentation and feature extraction. Each sub-section contains description about methods that cope with malaria parasites (MP) stained components analysis, both on thin and thick blood smears, without distinction.

Extensive search of articles has been made in PubMed and Google Scholar search engines based on the keywords: "malaria, mathematical morphology, automated malaria diagnosis" up to October 2017. The search includes papers published in English and titles and abstracts of potentially relevant studies were selected and presented from the most recent ones. Thereafter, the full texts of these studies were evaluated as per the exclusion criteria.

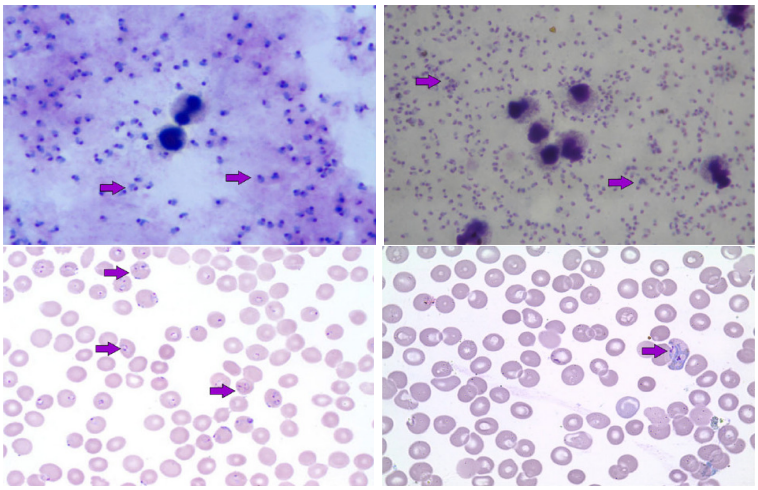

Figure 3: Malaria infected blood smear images. From top left, clockwise: thick smear with Giemsa stain, thick smear with Leishman stain, thin smear with Leishman stain, thick smear with Giemsa stain [7]. Arrows in top images indicate chromatin dots while, in the bottom ones, they show the infected erythrocytes.

Two main factors are generally considered if we refer to staining techniques: the type of colouration, in which Giemsa and Leishman are the most common, and the thickness of blood slide, which may be thick or thin. The majority of studies have been employed on thin blood smear images (over two-third of the total count) while only a few have used thick blood smear images. Typically, thin smears permit the identification of specific parasitic stage and quantification of malaria parasite; on the other hand, thick smears are better if the target is to perform an initial identification of malaria infection using blood pathology. Some examples are shown in fig. 3. Giemsa stained blood smear is considered in most of the analysed literatures whereas Leishman stain is considered in few studies, as showed in table **??**.

It is reported that Leishman stain has bigger sensitivity for parasite detection than Giemsa [17] and is superior for visualization of red and white blood cell morphology. [26]. On the contrary, Giemsa stain highlights both malaria parasites and white blood cells and, therefore, it is an additional issue to deal with. Giemsa stain is much costly and also time-taking procedure than Leishman. Moreover, magnification of 100X by using an oil immersion objective is used for capturing microscopic images of thin blood smear for identification of specific parasites and their infected stage.

2.1 Preprocessing

In image analysis field, especially when we refer to complex computer-aided pipelines, pre-processing methods are particularly used in order to improve the image data by suppressing unwanted noise or enhancing some image features for further processing. It is worth to mention pre-processing methods because they are an important step regarding image analysis field but, for what concerns

the malaria-affected blood image analysis, in our review we particularly found methods which operate for illumination correction and noise filtering purposes. Generally speaking, digital microscopy images can be acquired in different lighting conditions, with several types of acquisition devices or from blood smears stained with various staining protocol and, consequently, the features of similar images could differ a lot. Different techniques for illumination correction have been suggested to reduce such variation, e.g., a lot of authors work with grayscale-converted images as an illumination correction method On the other hand, noise filtering aims to remove the noise introduced by mishandling the slides and/or the camera settings.

The system proposed in [21] is based on image processing, artificial intelligence techniques and an adapted face detection algorithm to identify Plasmodium parasites. The algorithm uses the integral image and haar-like features concepts, and weak classifiers with adaptive boosting learning. The search scope of the learning algorithm is reduced in the preprocessing step by removing the background around blood cells.

In [24] an algorithm that uses morphological operations, the watershed method, the Hough transform and the clustering method of k-means to detect overlapped RBCs is presented. In the pre-processing stage white blood cells and platelets are removed before the segmentation task. During this step, some noise, the WBC cytoplasm and platelets still remain on the image. Therefore, the small objects are removed using a morphological opening and then the image is dilated with a disc-shaped structuring element.

An edge-based segmentation of erythrocytes infected with malaria parasites using microscopic images is proposed in [30]. A fuzzy C-means clustering is applied to extract infected erythrocytes, which is further processed for the final segmentation. The experimental results showed that the proposed method can gain 98%, 93.3%, 98.65% and 90.33% of sensitivity, specificity, prediction value positive and prediction value negative, respectively. A morphological erosion is used to erase some small noises and spots before the segmentation and holes inside the infected erythrocytes are filled using a morphological hole filling operation for the final segmentation. The work [23] proposes a new algorithm for morphological filtering of the blood images as a pre-processing tool for segmentation. Conventional morphological closing on blood images removes the unwanted components but also useful information. On the opposite the proposed method preserves the necessary information of foreground components while removing noise and artefacts. The paper in [14] illustrates a morphological approach for blood cell identification and uses the image features such as intensity, histogram, relative size and geometry for further analysis. Before the identification of blood cells, the authors propose a novel morphological filtering based on the size of RBC for platelets and/or artifacts elimination. A dilation is performed by a concentric ring structuring element and erosion by disk shaped structuring element. The radius of the structuring element depends on the radius of the RBC, so that all the components smaller than the RBCs are removed. In [36] a complete framework to detect and identify malaria parasites in images of Giemsa stained thin blood film specimens is proposed. Also, the system is able to identify the infecting species and life-cycle stages. The preprocessing step of the proposed method is applied to reduce the variations in the observed size, intensity, and colour of the cells and stained objects before the detection and classification steps. The aim is to correct the non-uniform illumination

in the images. The estimation is based on a morphological closing operation using a sufficiently large structuring element. The "sufficiently" large size for an input image is determined automatically with respect to its average cell size computed from the area granulometry distribution. In the method proposed in [28] the first phase is the stain object extraction which extracts candidates objects that can be infected by malaria parasites by using intensity and colour. Before extracting the stained objects the method firstly extracts the foreground and enhanced images. Foreground image is a binary image which is produced after applying morphological hole filling on such pixels which have lower intensity value than average intensity value of green layer. After the stained objects extraction process, a series of morphological operations is also employed after every part in order to eliminate small components and complete the final stained objects.

2.2 Segmentation

Segmentation is a key step in image analysis because it permits the identification and separation of the regions that compose an image, according to certain criteria of homogeneity and separation. Its main target is to divide the image into parts that have a strong correlation with objects or areas of the real world contained in the image. The commonly used segmentation methods essentially operate considering characteristics such as the brightness value, colour and reflection of the individual pixels, identifying groups of pixels that correspond to spatially connected regions. As for many problems of image processing, there is no standard solution valid in general, so different segmentation techniques can be applied, according to the characteristics of the images to process and of the objects to segment. Medical images segmentation is typically performed using two main strategies: the first level aims to separate whole cells or tissues from the background and the second one aims to separate the tissue structure in different regions or the cell in their components as the nucleus from the cytoplasm or intracellular parasites. The latter case is commonly used in applications in which the cell class depends on the morphological characteristics of its components.

Several other authors attempted to use thresholding as segmentation method in their computer-aided systems. In [25] the authors propose a method which provides a positive or negative diagnosis of malaria and differentiates parasites by species. The segmentation step relies on a thresholding strategy which aims to identify and segment potential parasites and erythrocytes from the image background after a six steps threshold selection. Mathematical morphology has been used for parasite size estimation, erythrocytes reconstruction and cells bigger than erythrocytes removal. The authors in [15] used the Annular Ring Ratio (ARR) transform method. Before applying it, a pre processing phase for removing platelets, parasites and other artefacts in the image has been performed. In the proposed method, the image after being converted to grayscale undergoes a morphological opening similar to closing. Unlike conventional closing (dilation followed by erosion) which uses the same structuring element, two different structuring elements are used, a concentric ring for dilation and a disk for erosion. The inner and outer diameter of the dilation ring is set to 35% and 70% of RBCs size, respectively and the erosion disk has the same diameter. Therefore, considering that fixed manually defined parameters are used for this strategy,

the results may substantially differ depending on the image resolution. This approach results in locating only the stained components in the image instead of all the cells and hence will not only speed up the operation but reduces the complexity. Anggraini et al. [2] proposed a method for separating blood cells, parasites and other components from background in a microscopic field of a thin blood smear. They applied several global thresholding methods and visually compared the results to qualitatively determine which technique yields the best result. The binary image was then subjected to hole filling morphological operator and applied as marker to label blood cells. From each identified cell (RBC and WBC), constituents of the parasite (nucleus and cytoplasm) were extracted using multiple threshold. Savkare et al. [27] worked on thin blood films with Giemsa staining and used global threshold and Otsu threshold [22] on grayscale enhanced image (green channel) for separating foreground from background. Hole filling has been performed on identified cells and morphological operators have been used to identify overlapping cells. Then, watershed transform has been applied for separating overlapped cells. In [11], authors used divergence based threshold selection in order to segment P.vivax parasites from Leishman-stained thin blood films. This method is based on Cauchy membership function [**?**] and is applied to the C channel of CMYK colour space. Morphological operators of opening and closing have been used for artefacts removal. Tek et al. [36] addressed the problem of parasite detection and identification on thin blood film. They performed a pre-processing phase, in which morphology has been used with a granulometry-based cell size estimation, considering the peak index of the granulometry distribution as the average cell area. Then, two different segmentation steps followed: the foreground-background segmentation has been realized using morphological area top-hats (using the average cell area value) and morphological double thresholding; the stained pixels segmentation has been performed by modelling the stained and unstained pixel distribution with RGB space histograms and used the probability density function to determine if a pixel is stained or not. This work also addresses non-uniform illumination problem, removing it by using a pre-recorded illumination image or using a morphological closing operation. Authors in [3] worked on thick blood films and proposed a method that uses an adaptive thresholding based scheme, which also allows an effective classification of pixels. This means that the election of whether a pixel belongs to the background or to the signal (parasites and white blood cells) is only established by the pixels around it, that is its neighbourhood. Then, morphological methods are applied to evaluate the area of connected components, labelling those belonging to parasites and counting their number. The following strategy has been used:

1. an algorithm based on run-length encoding has been used for analysing the connected components of the binary masks obtained in segmentation phase;

2. preliminary tagging, keeping the label equivalences or connected regions with different labels in a table of local matches. Resolution of the equivalences is, then, performed;

3. measurements on the remaining connected components: WBCs and MP region areas are checked and WBCs connected components are removed with a closing operation.

Also in the method proposed in [31] the segmentation of the infected parasites is based on thresholding. The segmentation is achieved in two stages by maximizing between-class variance of an original image and consequently by an iterative threshold selection from a stage-one threshold image with suitable stopping criteria. The segmented results are post processed to improve the accuracy of the detection of malaria parasites by morphological operators (erosion and closing).

Mushabe et al. [20] used morphological and statistical classification to detect malaria in blood smears by identifying and counting red blood cells and Plasmodium parasites. Morphological operations and histogram-based thresholding are used to extract RBCs and boundary curvature calculations and Delaunay triangulation are used for splitting clumped RBCs. They worked on Giemsa-stained thin blood smears. The proposed automated method in [10] for parasite detection and identification worked on thin blood film acquired with Giemsa stain. The authors found that the G and B channels of the RGB colour are very good features to identify objects containing chromatin in Giemsa stained blood films, being not only considered highly discriminative but also almost independent of differences in illumination and staining intensity. They transformed the colour input image into a monochrome image *I(x,y)*, that highlights objects containing chromatin: $I(x,y) = arctan\frac{I_{green}(x,y)}{I_{blue}(x,y)}$. In this work, mathematical morphology has been used with a black top-hat operator to separate MP from both leukocytes and platelets, with a non-flat paraboloid structuring element of radius of 9 and a slope of 1 pixel. It should be taken into account that these fixed parameters might not be suitable for images with different pixel resolutions. The black top-hat operator is followed by a thresholding operation with a fixed threshold, which according to the authors is reliable given the independence of the G and B channels with regard to illumination and staining intensity. However, the authors do not define the value of this fixed threshold on the publication. Di Ruberto et al. [9] aimed to detect the parasites by means of an automatic thresholding based on a morphological approach applied to cell image segmentation, that is more accurate than the classical watershed-based algorithm. They have applied grey scale granulometries based on opening with disk-shaped elements, flat and hemispherical. They have used a hemispherical disk-shaped structuring element to enhance the roundness and the compactness of the red cells improving the accuracy of the classical watershed algorithm, while they have used a disk-shaped flat structuring element to separate overlapping cells. These methods make use of the red blood cell structure knowledge, that is not used in existing watershed-based algorithms. Soni et al. [33] performed segmentation of erythrocytes by using granulometry as well. The size and eccentricity of the erythrocytes are also required for the calculation of some feature values (as these can be indicative of infection). The shape of the objects (circular erythrocytes) is known a priori, but the image must be analysed to determine the size distribution of objects in the image and to find the average eccentricity of erythrocytes present. Gray-scale granulometries based on opening with disk shape elements are then used. Non flat disk shaped structural element are applied to enhance the roundness and compactness of the red blood cells and flat disk shaped structural element applied to segment overlapping cells. The object to be segmented differs greatly in contrast from the background image. Changes in contrast can be detected by operators that calculate the gradient of an image. The gradient image can be computed and a thresh-

old can be applied to create a binary mask containing the segmented cell. The binary gradient mask is dilated using the vertical structuring element followed by the horizontal structuring element. The cell of interest has been successfully segmented, but it is not the only object that has been found. Any objects that are connected to the border of the image can be removed. Also Airwhar et al. [1] based their approach on thresholding and granulometry. The histogram of the complemented, green component has been used and it is said to be a bimodal distribution in all the considered images. Then, both local and global thresholds are used, and the union of the two binary images is chosen as the parasite marker image. A morphological opening filter, using a disk-shaped SE with radius equal to the mean erythrocyte radius less the standard deviation, is applied to the grayscale morphologically filtered green component of the image to remove any objects smaller than an erythrocyte. The morphological gradient is then calculated using a diamond-shaped SE with unity length. The segmentation method is applied to each object in the reconstructed binary image of erythrocytes individually. Those objects that do not exceed the area of a circle with radius equal to the mean erythrocyte radius plus the standard deviation are regarded as being single cells, and are unmodified. On the other hand, the clumped cells are segmented as follows. First, the intersection of the morphological gradient image and the dilated cell cluster is taken. This image is then transformed to a binary image by thresholding any value greater than zero. A series of morphological operations, namely a closing operation, thinning, and spur removal are then applied to generate a contour of the segmented erythrocytes. The contours are filled, and the segmented mask is again reconstructed with the valid parasite marker image to result in a segmented mask of infected cells. Several authors used marker controlled watershed with morphological approach. Khan et al. [16], among other experimentations, used it in order to try to separate overlapping cells because, according to their statements, watershed transform can separate touching cells but it is not sufficient for overlapping cells. In [34] red blood cell segmentation has been performed by using marker-controlled watershed transformation based on the image gradient. Markers are computed as a combination of the binary mask of the red blood cells and centres of the cells which are computed using a similar algorithm that was utilized for the evaluation of the average cell radius. The binary mask is obtained by thresholding the grey-scale image with an automatically estimated threshold using Otsu method [22]. Das et al. [4], [5], [6], [7] segmented erythrocytes as aforesaid and then morphological operators are used to eliminate unwanted cells like leukocytes and platelets. To conclude, overlapping erythrocytes are segmented by using marker controlled watershed segmentation technique.

In the paper [8] the authors propose a computer assisted system for quantification of erythrocytes in microscopic images of thin blood smears. The proposed method consists of preprocessing, segmentation, morphological filtering, cell separation and clump cell segmentation. The performance of the system in classifying the isolated and clump erythrocytes by geometric features is evaluated for the different classifiers. Moreover, the clump erythrocytes are segmented using marker controlled watershed with h-minima as internal marker. Based on the experimental results, it may be concluded that the proposed model provides satisfactory results with an accuracy of 98.02% in comparison to the state of art.

In [38], Yunda et al. proposed a method for P.vivax parasites detection. The

segmentation phase is a combination of border and region detection that allows rejection of the image background and permits identifying each of the objects. Initially, the morphological gradient method is used to enhance the borders of previously found objects. This is followed by a threshold detection stage using the K-Median method. Furthermore, Laplacian operator was used to discriminate the pixels that are interior or exterior in relation to the regions of the images and then erosion operation followed by two dilations were applied to delete the pixels which did not make part of any object. In the end, Absence of Gradients and Nernstian Equilibrium Stripping (AGNES) and K-Median techniques were applied to assign the remaining number of pixels to each region, using the image regions previously identified as objects and background as the starting point. Tek et al. [35] proposed a classifier-based method, for the segmentation stage, which relies on a Bayesian pixel classifier to distinguish among stained and non-stained pixels. In particular, they used a non-parametric method based on histograms in order to produce the probability density functions of stained and non-stained classes. Stained pixels can belong to other components such as WBCs, platelets or artefacts, in addition to the parasites and so the detection procedure requires a further classification to distinguish among parasite and non-parasite pixels. However, the stained pixels have to be represented as connected sets, representing stained objects, to extract features for the classifier. Furthermore, top-hat extraction and infinite reconstruction were applied to find the regions that include the objects.

2.3 Feature extraction

Feature extraction has the target of reducing the computational complexity of the subsequent process and facilitating a reliable and accurate recognition for unknown novel data, considering that the input data to an algorithm could be too large to be processed and it could be redundant (e.g. repetitiveness of pixels patterns in an image). Moreover, the in-depth understanding of the domain-specific knowledge gained by human experts on the problem being addressed can be of extreme importance for the design of a reliable and effective feature extraction engine [13]. It starts from determining a subset of the initial features and this procedure is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data. Malaria parasite infection causes micro structural changes in erythrocytes. The microscopic features of the RBCs are usually specific to morphology, intensity and texture. They may also represent the differences that occur among healthy and unhealthy cells. Most of the studies have reported both textural and geometric features for describing malaria infection stages [7]. Generally speaking, features may be distinguished according to the following characteristics: morphological features and textural and intensity features.
It is a well known mathematical morphology approach to compute a size distribution of grains in binary images, using a series of morphological opening operations. It is the basis for the characterization of the concept of size. Some authors used area granulometry for pre-processing purposes in malaria characterization [36] even though it is certainly effective for extracting cells size features information [35], [19], [34].

References

[1] N. Ahirwar, S. Pattnaik, and B. Acharya. Advanced image analysis based system for automatic detection and classification of malarial parasite in blood images. *International Journal of Information Technology and Knowledge Management*, 5(1):59–64, 2012.

[2] D. Anggraini, A. S. Nugroho, C. Pratama, I. E. Rozi, Aulia Arif Iskandar, and Reggio Nurtanio Hartono. Automated status identification of microscopic images obtained from malaria thin blood smears. In *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*, pages 1–6, 2011.

[3] J.E. Arco, J.M. Gorriz, J. Ramirez, I. Alvarez, and C.G. Puntonet. Digital image analysis for automatic enumeration of malaria parasites using morphological operations. *Expert Systems with Applications*, 42(6):3041 – 3047, 2015.

[4] D.K. Das, M. Ghosh, C. Chakraborty, A.K. Maiti, and M. Pal. Probabilistic prediction of malaria using morphological and textural information. In *Image Information Processing (ICIIP), 2011 International Conference on*, pages 1–6, 2011.

[5] D.K. Das, M. Ghosh, M. Pal, A.K. Maiti, and C. Chakraborty. Machine learning approach for automated screening of malaria parasite using light microscopic images. *Micron*, 45:97 – 106, 2013.

[6] D.K. Das, A.K. Maiti, and C. Chakraborty. Textural pattern classification of microscopic images for malaria screening. *Advances in Therapeutic Engineering*, pages 419–446, 2012.

[7] D.K. Das, R. Mukherjee, and C. Chakraborty. Computational microscopic imaging for malaria parasite detection: a systematic review. *Journal of microscopy*, 260(1):1–19, 2015.

[8] S.S. Devi, J. Singha, M. Sharma, and R.H. Laskar. Erythrocyte segmentation for quantification in microscopic images of thin blood smears. *Journal of Intelligent and Fuzzy Systems*, 32:2847–2856, 2017.

[9] C. Di Ruberto, A. Dempster, S. Khan, and B. Jarra. Analysis of infected blood cell images using morphological operators. *Image and Vision Computing*, 20(2):133 – 146, 2002.

[10] M. Elter, E. Haßlmeyer, and T. Zerfaß. Detection of malaria parasites in thick blood films. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5140–5144, 2011.

[11] M. Ghosh, D. Das, C. Chakraborty, and A. K. Ray. Plasmodium vivax segmentation using modified fuzzy divergence. In *2011 International Conference on Image Information Processing*, pages 1–5, 2011.

[12] C.R. Giardina and E.R. Dougherty. *Morphological Methods in Image and Signal Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1988.

[13] X. Jiang. Feature extraction for image recognition and computer vision. In *Proceedings - 2009 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009*, pages 1–15, 2009.

[14] S. Kareem, I. Kale, and R.C.S. Morling. Automated malaria parasite detection in thin blood films:- a hybrid illumination and color constancy insensitive, morphological approach. In *Proc. of the 2012 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, 2012.

[15] S. Kareem, R.C.S. Morling, and I. Kale. A novel method to count the red blood cells in thin blood films. pages 1021–1024, 2011.

[16] M.I. Khan, B. Acharya, B.K. Singh, and J. Soni. Content based image retrieval approaches for detection of malarial parasite in blood images. *International Journal of Biometrics and Bioinformatics (IJBB)*, 5(2):97, 2011.

[17] N.A. Khan, H. Pervaz, A.K. Latif, A. Musharraf, and Saniya. Unsupervised identification of malaria parasites using computer vision. In *2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 263–267, May 2014.

[18] A. Loddo, L. Putzu, C. Di Ruberto, and G. Fenu. A computer-aided system for differential count from peripheral blood cell images. In *2016 12th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pages 112–118, Nov 2016.

[19] L. Malihi, K. Ansari-Asl, and A. Behbahani. Malaria parasite detection in giemsa-stained blood cell images. In *2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP)*, pages 360–365, 2013.

[20] M.C. Mushabe, R. Dendere, and T.S. Douglas. Automated detection of malaria in giemsa-stained thin blood smears. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3698–3701, 2013.

[21] A.D. Oliveira, C. Prats, M. Espasa, and et al. The malaria system microapp: A new, mobile device-based tool for malaria diagnosis. *JMIR Research Protocols*, 6(4):1–11, Apr 2017.

[22] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

[23] S.K. Reni, I. Kale, and R. Morling. Analysis of thin blood images for automated malaria diagnosis. In *Proc. of the 5th IEEE International Conference on E-Health and Bioengineering - EHB 2015*, 2015.

[24] M.F. Romero-Rondon, L.M. Sanabria-Rosas, L.X. Bautista-Rozo, and A. Mendoza-Castellanos. Algorithm for detection of overlapped red blood cells in microscopic images of blood smears. *DYNA*, 83(198):187–194, 2016.

[25] N.E. Ross, C.J. Pritchard, D.M. Rubin, and A.G. Dusé. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Medical and Biological Engineering and Computing*, 44(5):427–436, 2006.

[26] S. Sathpathi, A.K. Mohanty, P. Satpathi, S.K. Mishra, P.K. Behera, G. Patel, and A.M. Dondorp. Comparing leishman and giemsa staining for the assessment of peripheral blood smear preparations in a malaria-endemic region in india. *Malaria Journal*, 13(1):512, 2014.

[27] S.S. Savkare and S.P. Narote. Automatic detection of malaria parasites for estimating parasitemia. *International Journal of Computer Science and Security (IJCSS)*, 5(3):310, 2011.

[28] M. Sheikhhosseini, H. Rabbani, M. Zekri, and A. Talebi. Automatic diagnosis of malaria based on complete circle ellipse fitting search algorithm. *Journal of Microscopy*, 252(3):189–203, 2013.

[29] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer, 2004.

[30] J. Somasekar and B.E. Reddy. Segmentation of erythrocytes infected with malaria parasites for the diagnosis using microscopy imaging. *Computers and Electrical Engineering*, 45:336–351, 2015.

[31] J. Somasekar and B.E. Reddy. A novel two-stage thresholding method for segmentation of malaria parasites in microscopic blood images. *Journal of Biomedical Engineering and Medical Imaging*, 4(2):31–42, 2017.

[32] J. Somasekar, B.E Reddy, E.K. Reddy, and C.H. Lai. Computer vision for malaria parasite classification in erythrocytes. *International Journal on Computer Science and Engineering*, 3(6):2251–2256, 2011.

[33] J Soni, N Mishra, and C Kamargaonkar. Automatic difference between rbc and malaria parasites based on morphology with frst order features using image processing. *Int. J. Adv. Eng. Technol*, 1:290–297, 2011.

[34] V. Špringl. Automatic malaria diagnosis through microscopy imaging. *FACULTY OF ELECTRICAL ENGINEERING*, page 128, 2009.

[35] F.B. Tek, A.G. Dempster, and I. Kale. Malaria parasite detection in peripheral blood images. pages 347–356, 2006.

[36] F.B. Tek, A.G. Dempster, and I. Kale. Parasite detection and identification for automated thin blood film malaria diagnosis. *Computer Vision and Image Understanding*, 114(1):21 – 32, 2010.

[37] WHO. Malaria fact sheet december 2016, 2016.

[38] L. Yunda, A. Alarcón, and J. Millán. Automated image analysis method for p-vivax malaria parasite detection in thick film blood images. *Sistemas & Telemática*, 10(20):9–25, 2012.