ECRONICON OPEN ACCESS
EC EMERGENCY MEDICINE AND CRITICAL CARE
Literature Review

Malaria Parasite Detection-An Approach

Shawni Dutta[1]* and Samir Kumar Bandyopadhyay[2]

[1]Lecturer, Department of Computer Science, The Bhawanipur Education Society College, Kolkata, India

[2]Academic Advisor, The Bhawanipur Education Society College, Kolkata, India

***Corresponding Author**: Shawni Dutta, Lecturer, Department of Computer Science, The Bhawanipur Education Society College, Kolkata, India.

Received: April 27, 2020; **Published:** May 20, 2020

Abstract

Malaria has been a well-researched domain over the last 2 decades. While clinical research pertaining to curing of Malaria has been majorly conclusive, detection of malaria parasite in digitized thin blood smear image still stands open for investigation. Manual investigation of malaria parasite in thin blood smear image stands tedious and is liable to be affected by human bias. Herein lies the need for an automated CAD system. The proposed system is aimed to serve as a trustworthy aid to pathologists in malaria parasite detection and subsequent stage/species interpretation. The system has two broad phases, namely, the initial screening phase and main phase. The initial screening phase uses 40X magnification based images and is a precursor decision making phase to estimate whether the concerned digital image contains infection or not. If the image contains infection, the image is further investigated at 100X magnification in the main phase. The main phase encapsulates the infection detection phase along with stage/specie classification. So, the main phase is an independent, standalone phase in itself tailored as per WHO specification for malaria parasite detection. The main phase extends the initial screening phase by incorporating texture features for stage/specie of infection prediction. The initial screening phase was tested on 250 images (MaMic database) that were handpicked by the consultant pathologist. The main phase was built and evaluated on two datasets, namely, the MaMic and Government Hospital dataset (total 2730 images) using 10 fold cross validation to prevent overfitting of the data model. The prediction Accuracy, Sensitivity and Specificity of the main phase was recorded as 0.9917, 0.9948, and 0.9892 respectively using an Adaboost Classifier. Both phases use a strategic amalgamation of rule based and machine learning algorithms. The algorithm developed works on variedly stained, digitized thin blood smear slides. The performance of the algorithms developed was comparable with the state of the art methods in terms of prediction infallibility and time complexity. Cost benefit analysis for the proposed system revealed enhanced profitable in terms of time, labour and money as opposed to the standard practice of diagnostic laboratories.

Image processing and image analysis is a core area of application in computer science. Several research work has contributed and consolidated this domain of research. With the progression and advancement of digital imaging, contribution towards meaningful assessment of images have gained importance. An image captured by some imaging modalities needs to be understood and deciphered to gain insight on the image/s and hence extract pertinent information from the captured image/s. Images can be captured using digital camera, scanners, telescopes, microscopes and other electro-optical devices used in different industries and domains. Similarly, images are also generated by medical imaging hardware like Computed Tomography Scanners (CT-Scans), Magnetic Resonance Imaging (MRI) machines, X-ray machines and digital Whole Slide Imaging Scanners (WSI).

The images obtained require processing so that meaningful information can be retrieved and collated. Computerized algorithms play a significant role to achieve this goal. Manual interpretation of image may vary based on the perception of details and reasoning capability of the operator. Human bias and fatigue are two key elements that impede the performance of manual recognition and interpretation of images. However, the advent of computerized image processing tools has helped to overcome these impediments by assisting the operators in image analysis.

Computerized Image processing or Digital image processing is often referred to as Computer Vision and found its initial application in the study of satellite imagery captured by spy satellites to track enemy movement. Though initial applications of Computer Vision and Pattern Analysis was in defence related research but now this has found applications in every walks of life. Digital image processing have wide range of application ranging from industrial manufacturing, Geographical Information System (GIS) and also in Medical imaging. Digital image processing and automated analysis of images have significantly contributed to the betterment of several processes that affect human lives.

This research work particularly contributes towards medical image processing. Medical images are generated by several imaging modalities. Such images are interpreted and analyzed by some medical professional. Based on the source and imaging modalities the professional may be a radiologist (in case of X-rays, CT, MRI imaging modalities) or a pathologist (microscopic).

Keywords: *Image Processing; Malaria; Thin Blood; GIS*

Introduction

Image processing and image analysis is a core area of application in computer science. The images required processing so that meaningful information can be retrieved and collated. Computerized algorithms play a significant role to achieve this goal. Manual interpretation of image may vary based on the perception of details and reasoning capability of the operator. However, the advent of computerized image processing tools has helped to overcome these impediments by assisting the operators in image analysis. Automated computer application or Computer Aided Diagnosis (CAD) systems can provide adequate analysis that can maximize the limited human resource for correct analysis. Effective, timely and accurate analysis contribute to correct inference is vital for diagnosis of disease.

This paper contributes to the knowledge domain of medical image processing. The contribution is aimed at developing automated CAD applications for Pathology. Malaria disease identification is an age-old problem that significantly affect a huge population. The research work presents an approach for automated analysis of Malaria infected slides.

A Pathologist determines the cause of a particular disease conditions based on tests (chemical/clinical/microscopy. Most of these test are conducted by automated equipment using body fluids/tissue sample extracted from the patient. Microscopy plays a vital role in disease determination for cases of parasitic invasion within tissues and to locate abnormality in histological/cytological body samples.

Microscopic examination of cellular and histological samples are widely used as a basis for disease detection. Handling of glass slides across the labs is cumbersome and susceptible to loss of the slide or decreased quality of the specimen. The associated turnaround time (from the sample collection to report generation) is time consuming. With the increase in reliability of digital equipment like digital imaging technologies, computer hardware and software are accepted Digital Pathology in the medical community [1].

Conversion of a biological specimen in a glass slide to an image is referred to as Virtual Microscopy. This is the first step towards digital pathology. High resolution Whole Slide Imaging Scanners (WSI) perform this task at resolutions of 40-60X magnification. A high resolution image obtained of the whole slide can be extended to the size of a tennis court when projected/printed at 300 dpi resolution [2]. The information can be archived for training and technical education purpose, telemedicine applications, primary/secondary diagnosis or for a second opinion, review of consultation and for quality assurance mechanisms [3]. The consulting pathologist will view the slide images on high resolution monitors that are specialized for medical purpose. The pathologist will be able to analyse remotely at the time of his preference without the sample being affected/destroyed or stained sample getting discoloured. The images can be distributed among consultants for double review or expert review and can result in faster workflow in pathological laboratories.

Digital pathology using WSI has been granted certification by Food and Drug Administration (FDA) [4] in 2017 for its application in primary diagnosis of disease. Medical images obtained from different equipment/vendors and of different modalities have been standardized by the Digital Imaging and Communications in Medicine (DICOM) standard. The DICOM images and Picture Archiving and Communication system or PACS system are being evolved to accommodate WSI for medical diagnosis. A working group, WG-26, was established by DICOM for this very purpose in 2005 [5]. Several studies indicate that the performance of WSI and glass slide is similar [6].

The use of digital imaging has also opened a new technological dimension for pathology. The images can be processed with Artificial intelligence and machine learning algorithms in CAD system. Such a system can be used to identify abnormalities independent of human intervention and is referred as the third revolution in pathology [7].

Data sets

For computerized algorithms identification and differentiation of infected and normal RBC is vital. The possible hindrance in the segmentation process are the presence of White Blood Cells (WBC) or leukocytes that are the primary defence system of the body against infections. Figure 1 shows all the different types of WBC that are found within a smear slide image.

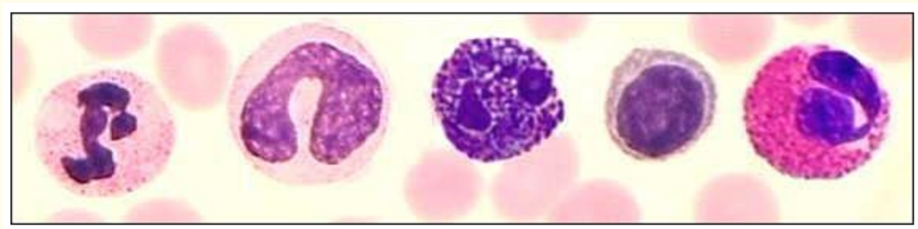

Figure1: *The figure shows five different types of WBC with Neutrophil at the leftmost position followed by Monocyte, Basophil, Lymphocyte and Eosinophil at the extreme right. This rare image was recorded by a Haematologist while scanning through smear slides.*

Early and effective treatment of malaria can prevent its rapid spread among the immediate geographic location. The diagnosis of malaria identifies the presence of malaria parasite cells, antigens and antibodies within the human blood. There are different malaria life-cycle stages and five types to identify from. Early diagnosis can prevent mortality rates. Diagnosis of malaria depends on the availability of appropriate diagnostic equipment and presence of adequate trained technicians.

The database that was acquired from MaMic (which is a publicly available database) pertains to snapshots taken from a whole thin blood smear slide scanned at 100X resolution of *P. falciparum* infection. The *P. vivax* dataset is not available publicly and was provided on request. Some of the infected slide scan areas were devoid of infection. This is due to low Parasitaemia for those slide scans. Some of these slides were taken into consideration to test the robustness of the proposed system.

The dataset is acquired from the Pathology Department of a Government Hospital in Kolkata under the supervision of Dr. D. Ckakraborty. The slides were prepared by the Ronald Ross Malaria Centre within the hospital campus.

A combined dataset of the images obtained from MaMic and Government Hospital in Kolkata.

Literature Review

The intervention of technology to assist pathologists and medical practitioners is vital for the fight against malaria and its early diagnosis to prevent mortality. The biologists and the chemists were busy discovering means to control the disease. Advancement in microscopy and computer technology has bolstered this effort. Several works can be found in the research domain that have significant contribution. Several authors have tried to compare different methodology both in terms of detection technology and computational methods to establish better ways to identify and diagnose malaria [8].

Research work related to the medical/clinical methods for Malaria parasite detection are dated to pre-historic times. However, research work on Malaria in context to medical image analysis is a relatively new concept. As already discussed with the advent of digital imaging the focus has shifted to image analysis as a method for detection of parasite within the thin/thick blood smear images. Quite a number of research work was focused on identifying Malaria parasites as well as the type of specie and lifecycle stages from image/s. Most of the research work conducted during the early years of digital pathology was to identify simple graphical software tools for better understanding of the images. Subsequently, the use of image processing algorithms for pre-processing and segmentation, as a part of image analysis, was also utilized for parasite identification. Most of these systems were based on rule-based image processing techniques. Image segmentation techniques like image binarization, edge detection and color-thresholding was widely used by authors for parasite segmentation. Morphological operators and Morphometry has also been effectively implemented for segmentation and parasite detection [9].

However, recent works have shown a shift towards machine learning approach with supervised and un-supervised learning approach in both image segmentation and parasite identification. Authors have derived host of features including color channels of different color model, textural and morphological features from the images. Using these features they have implemented supervised/unsupervised machine learning algorithms to segment/classify parasite and cellular components. Some of the research works are focused on identifying the infection, stage and/or species directly from the slide images without segregating the infected region as region of interest. CAD system should be able to screen for infection, extract region of interest, segregate from normal cells, extract the parasite region, differentiate between species and stage, quantify the infection regions and calculate Parasitaemia [10].

Proposed Method

The use of digital microscopy/ WSI scanners to obtain image data and further analysis by general purpose image processing software was the initial step towards automation. Subsequently, several scientists across the globe contributed towards the development of an automated disease diagnosis system that will detect the presence of parasites from the images obtained from digital slides. The search for the best possible method is still ongoing and remains a challenge in medical/computer science interdisciplinary research domain. The authors have broadly categorized the research work cited here into two categories. First one is 'rule based system' based on image processing and define rules set to achieve the objective. In contrast. The second category is 'machine learning based hybrid systems' utilizes image processing to train the system so that the system will derive its own rules for classification of the image accordingly. Figure 2 shows the generalized scheme of Malaria parasite detection used by both category of authors. Pre-processing of digital images are important feature of any image processing algorithm. Images obtained from digital capturing equipment often suffer from illumination and noise related issues. Before application of any automated algorithm it is pertinent to make adjustment to the images in terms of illumination correction and noise removal. The illumination of images and colour density may vary intra/inter dataset. This arises due to different staining methods, stains and drying time. Moreover, the image capturing medium also utilises different lighting condition giving rise to illumination differences among images. Moreover, color images vary based on staining methods used.

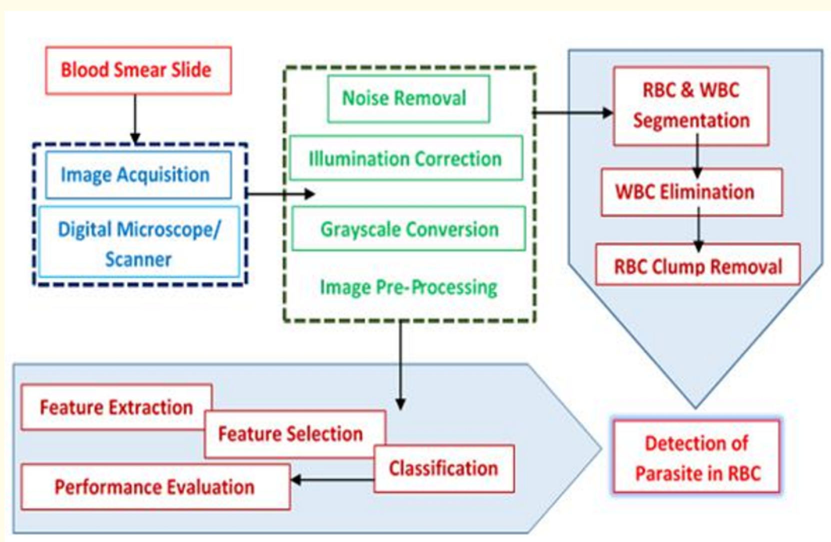

Figure2: *Generalized scheme for detection of malaria parasite.*

Conclusion

The contribution of the work is particularly two folds. The proposed method stands as a tool to assist medical practitioners for effective detection of malaria parasite. The CAD systems are designed to assist medical practitioners for making decisions.

Bibliography

1. M García-Rojo., *et al*. "Perspectives on Digital Pathology". Amsterdam: IOS Press (2012).
2. D Treanor and B Williams. "The Leeds Guide to Digital Pathology". The Leeds Teaching Hospitals NHS, University of Leeds (2019).
3. G Bueno., *et al*. "New Trends of Emerging Technologies in Digital Pathology". *Pathobiology* 83.2-3 (2016): 61-69.
4. L Pantanowitz., *et al*. "Twenty Years of Digital Pathology: An Overview of the Road Travelled, What is on the Horizon, and the Emergence of Vendor-Neutral Archives". *Journal of Pathology Informatics* 9 (2018): 40.
5. R Singh., *et al*. "Standardization in digital pathology: Supplement 145 of the DICOM standards". *Journal of Pathology Informatics* 2 (2011): 23.
6. TW Bauer and RJ Slaw. "Validating Whole-Slide Imaging for Consultation Diagnoses in Surgical Pathology". *Archives of Pathology and Laboratory Medicine* 138.11 (2014): 1459-1465.
7. M Salto-Tellez., *et al*. "Artificial intelligence-the third revolution in pathology". *Histopathology* 74.3 (2019): 372-376.
8. BS Kakkilaya. "The Challenge of Malaria" (2018).
9. R Neghina., *et al*. "Malaria, a Journey in Time: In Search of the Lost Myths and Forgotten Stories". *American Journal of the Medical Sciences* 340.6 (2010): 492-498.
10. TD Otto., *et al*. "Genomes of all known members of a *Plasmodium* subgenus reveal paths to virulent human malaria". *Nature Microbiology* 3.6 (2018): 687-697.

Volume 4 Issue 6 June 2020