Micron 45 (2013) 97–106

Contents lists available at SciVerse ScienceDirect

Micron

journal homepage: www.elsevier.com/locate/micron

Machine learning approach for automated screening of malaria parasite using light microscopic images

Dev Kumar Das[a], Madhumala Ghosh[a], Mallika Pal[b], Asok K. Maiti[b], Chandan Chakraborty[a,*]

[a] School of Medical Science and Technology, IIT Kharagpur, India
[b] Department of Pathology, Midnapur Medical College & Hospital, Midnapur, West Bengal, India

ARTICLE INFO

Article history:
Received 24 November 2011
Received in revised form 3 November 2012
Accepted 6 November 2012

Keywords:
Malaria parasite
Erythrocyte
Texture
Bayesian classifier
Machine learning

ABSTRACT

The aim of this paper is to address the development of computer assisted malaria parasite characterization and classification using machine learning approach based on light microscopic images of peripheral blood smears. In doing this, microscopic image acquisition from stained slides, illumination correction and noise reduction, erythrocyte segmentation, feature extraction, feature selection and finally classification of different stages of malaria (*Plasmodium vivax* and *Plasmodium falciparum*) have been investigated. The erythrocytes are segmented using marker controlled watershed transformation and subsequently total ninety six features describing shape-size and texture of erythrocytes are extracted in respect to the parasitemia infected versus non-infected cells. Ninety four features are found to be statistically significant in discriminating six classes. Here a feature selection-cum-classification scheme has been devised by combining *F*-statistic, statistical learning techniques i.e., Bayesian learning and support vector machine (SVM) in order to provide the higher classification accuracy using best set of discriminating features. Results show that Bayesian approach provides the highest accuracy i.e., 84% for malaria classification by selecting 19 most significant features while SVM provides highest accuracy i.e., 83.5% with 9 most significant features. Finally, the performance of these two classifiers under feature selection framework has been compared toward malaria parasite classification.

1. Introduction

Malaria is one type of parasitic infectious disease caused by Plasmodium species viz. *Plasmodium falciparum* (*P. falciparum*), *Plasmodium vivax* (*P. vivax*), *Plasmodium malariae* (*P. malariae*) and *Plasmodim ovale* (*P. ovale*) (Greer et al., 2009). This parasite exhibits a complex life cycle involving an insect vector (mosquito) and a vertebrate host (human). Malaria is common in Asian and Sub African populations (Frean, 2010) and is responsible for 1.5–2.7 millions of death per year (Raviraja et al., 2006). In the Indian population, the incident rate is higher in *P. vivax* infection cases than that of *P. falciparum*. It has been observed that ~50–60% of malaria patients are affected by *P. vivax* while ~40–50% is affected by *P. falciparum* in India (NVBDCP, 2010-2011).

Like other diseases, it is well understood that early detection of malaria infection leads to prevention and cure by means of providing treatment and management. Red blood cells or erythrocyte blood cells are mainly affected by the malaria parasites. In human blood, three life stages viz., trophozoite, schizont and gametocyte are cycled for the parasite. These infection stages viz. trophozoite, schizont and gametocyte are visible under light microscope using peripheral blood smears. The trophozoite stage is often known as ring stage (WHO, 2010). In case of *P. falciparum* trophozoite and gametocyte stage are visible under microscope but schizont stage is rarely visible because it remains in capillaries and bone marrow (Cuomo et al., 2012). In case of *P. vivax* infection, all three stages and in *P. falciparum* infection, two stages (trophozoite and gametocyte) are visible under microscope during peripheral blood smear screening. Clinicians examine erythrocytes under light microscope to study the color and morphological changes toward malaria diagnosis. Evaluation accuracy mostly depends on the expert's clinic-pathological understanding. In effect, such procedure involves humanistic error in terms of subjectivity, which leads to inconsistent as well as less diagnostic accuracy. To increase diagnostic precision by minimizing such subjectivity, developing a computer assisted malaria parasite detection tool has given importance in modern pathological services where a clinician will get assistance in order to quickly make better decision toward malaria diagnosis.

In modern diagnostic system development, machine learning techniques have enormous contributions for achieving higher diagnostic precision in medical imaging informatics like microscopy, ultrasound imaging, MRI, and CT. Microscopic image analysis is the most important as well as highly informative tool toward

* Corresponding author. Tel.: +91 3222 283570; fax: +91 3222 282221.
E-mail address: chandanc@smst.iitkgp.ernet.in (C. Chakraborty).

0968-4328/$ – see front matter
http://dx.doi.org/10.1016/j.micron.2012.11.002

pathological evaluation of different diseases viz. hematological disorder, oral cancer, breast cancer, cervical cancer, etc. Like others microscopic image analysis, peripheral blood smear screening is one of the essential diagnostic techniques to identify hematological disorders (anemia, thalassemia, etc.) and parasitic infection (malaria, filaria) in the blood. In case of malaria detection, pathologists frequently use light microscope to detect infection in erythrocytes based on color as well as morphological changes of the erythrocyte.

Now-a-days there are various techniques for malaria diagnosis available in the market (Tangpukdee et al., 2009) but conventional microscopic technique remains the gold standard for malaria diagnosis. Other methods are not cost effective and also these require further improvement for diagnostic precision. Few literatures have suggested computer vision approach to detect malaria infection based on digital microscopic images of peripheral blood smear. Color histogram based malaria parasite detection (Tek et al., 2006) has been carried out. Further, Diaz et al. (2009) showed quantification and classification of *P. falciparum* infected erythrocytes. Morphological and novel thresholding selection techniques for identification of erythrocytes were used by Ross et al. (2006). Malaria parasite in HSV (Hue, Saturation, and Value) color space was segmented (Makkapati and Rao, 2009). Erythrocytes infected by malaria parasites were detected by using statistical approach (Raviraja et al., 2008). Mathematical morphology and granulometry approaches (Dempster and Ruberto, 1999) and gray level thresholding (Toha and Ngah, 2007) for estimation of parasitemia were applied. Kumar et al. (2006) suggested clump splitting algorithm and rule base approach to segment out clump erythrocytes from peripheral blood smear images. Sio et al. (2007) applied rule based approach for *P. falciparum* infection detection purpose. Most of the literatures showed malaria classification based on cultured blood smear sample. But no comprehensive approach toward developing a pathological decision support system is still available for computerized detection of malaria parasitemia viz., *P. vivax* and *P. falciparum* using peripheral blood smear images.

In view of this, our study focuses on development of machine learning approach for discriminating five (three *P. vivax* and two *P. falciparum*) different stages of infected erythrocyte due to malaria infection and non-infected erythrocytes using color, textural and morphological information. Fig. 1 depicts systematic approach for

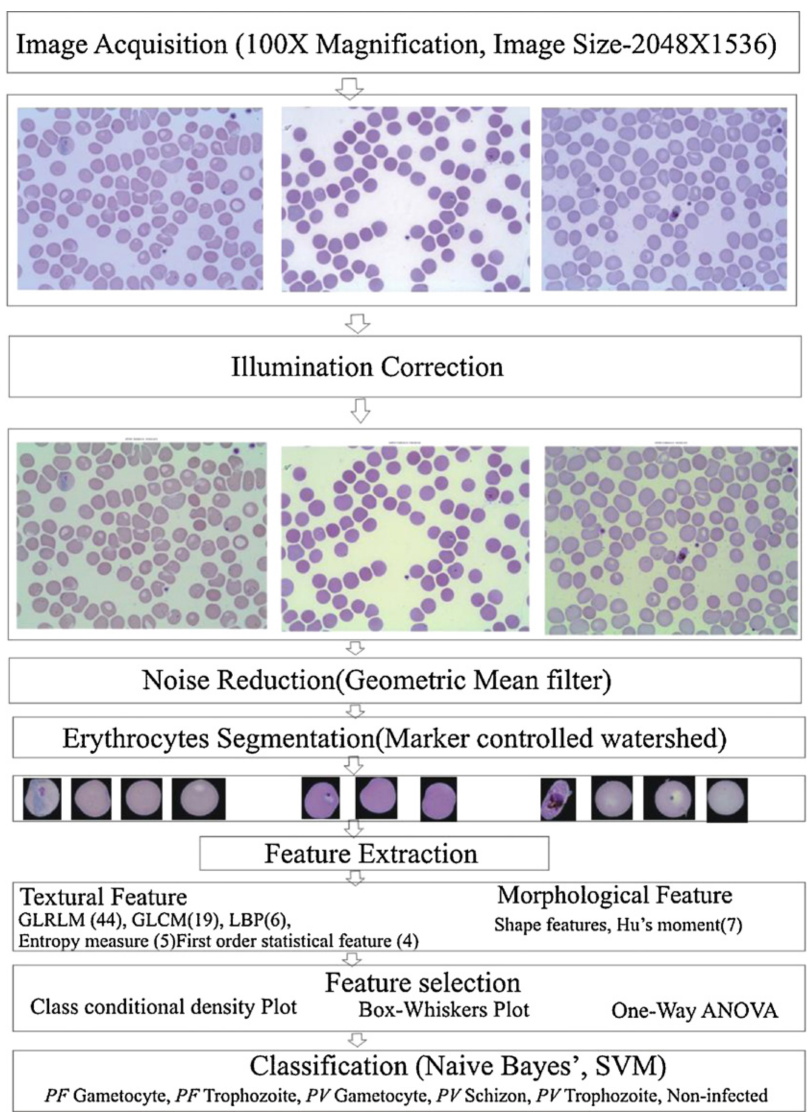

Fig. 1. Work flow diagram of the proposed methodology.