ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

SOFTWARE DEVELOPMENT FOR DETECTING MALARIA TROPIKA ON BLOOD SMEARS IMAGE USING SUPPORT VECTOR MACHINE

Sri Widodo [*], Pipin Widyaningsih

[*] Medical Records and Health Information Academy APIKES Citra Medika Surakarta, Central Java, Indonesia

Department of Informatic Systems, STMIK Duta Bangsa Surakarta, central Java, Indonesia

ABSTRACT

Tropical malaria is a disease caused by protozoan parasites of genus Plasmodium that infect red blood cells of patients. Tropical malaria falciparum is the most severe form of malaria. Process is performed by physicians to diagnosing tropical malaria diseases in blood smear with directly observed using a microscope. This procedure is certainly not effective. Purpose of research is to develop a software that can automatically detect tropical malaria of blood smear image as a diagnostic aid. This study consists of three steps. First step is to create a user interface. Second step is to develop software to perform segmentation of blood smears image using active contour method, and the third step is development of software for malaria detection by classification of blood smears images suspected containing plasmodium malaria using support vector machine (SVM). The results of testing that has been done for normal blood obtained accuracy 100 % , thropozoit phase with accuracy 100 %, Schizont phase with accuracy 85 %, and gamethozit phase with accuracy 95%.

KEYWORDS: active contour, tropical malaria, support vector machine.

INTRODUCTION

Malaria is a disease caused by protozoan parasites of genus plasmodium that infect red blood cells of patients. Parasites enter human body through bite of a female Anopheles mosquito. Plasmodium species that infect humans, namely plasmodium falcifarum, plasmodium vivax, plasmodium ovale, and plasmodium malariae. Plasmodium vivax causes tertian malaria, plasmodium malaria is cause of kuartana malaria, plasmodium ovale causes ovale malaria, while plasmodium falciparum causes tropical malaria [1].

Falciparum malaria in tropical or tropical malaria is the most severe form of malaria. Characterized by irregular heat, anemia, splenomegaly, and frequent parasitaemia complications. Incubation period is 9-14 days. Tropical malaria attack all forms of erythrocytes. Caused by Plasmodium falciparum. Shape of this Plasmodium is ring or small ring. Has a diameter 1/3 of normal diameter of erythrocytes, and only species that has 2 core chromatin [2]. Diagnosis can be done if malaria parasite found in blood of patients. Method is performed by physicians in diagnosing diseases of tropical malaria on blood smears is to observe directly through use microscope. From observation with microscope will be known who is image of blood smears of healthy red blood cells, red blood cells in which there is plasmodium falciparum with different phases, such as thropozoit phase, schizont phase, and gametozide phase located outside the red blood cells [2]. Detection manually will consume lot of time, other than that doctors may differ in defining image of a normal blood preparations and infected. Thus detection of tropical malaria automatically by observing shape and pattern of image of blood smears will be very helpful. Previous studies related to detection of tropical malaria automatically through blood smears image has been done [3]–[7].

Purpose of research is to develop software to automatically detect tropical malaria of blood smears image using support vector machine (SVM). There are three main steps undertaken in this research. The first step is to build a user interface using matlab programming language. The second step is to develop software to perform a blood smear image segmentation using active contour method [8] which

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

resulted in the suspected area as red blood cells, and third step is to develop software for detection of tropical malaria disease by performing classification on image of suspected disease, by using support vector machine (SVM) [9]. Support vector machine method used in this study are based on binary svm one against all [10]. Features used are statistics and Gray Level Co-occurrence Matrix (GLCM [11].

MATERIALS AND METHODS

Material

Input image used in this study is image of blood smears taken from www.dpd.cdc.gov. Consists of two data types, namely image used for training data and test data. Image of blood used for training data consists of 120 images, to images representing each class number 30, each of which has a size of 50x50 pixels.

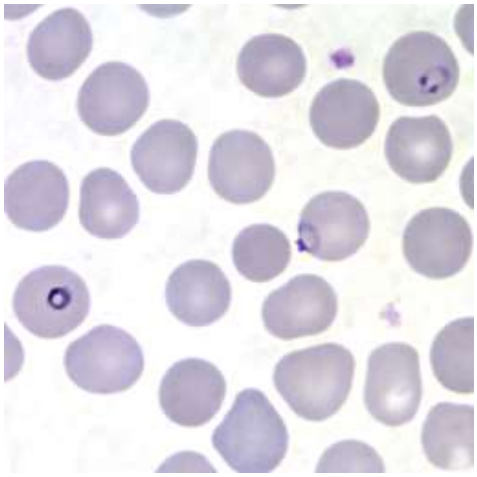


Figure 1 Example of Blood Smear

USER INTERFACE DEVELOPMENT

The steps of software development for detection of tropical malaria can be explained in Figure 2. The first step of software development for detection of tropical malaria disease is development of a user interface using Matlab programming language. User interface consists of three parts. The first part is display screen, second is segmentation of blood smear, and third part is detection of tropical malaria.

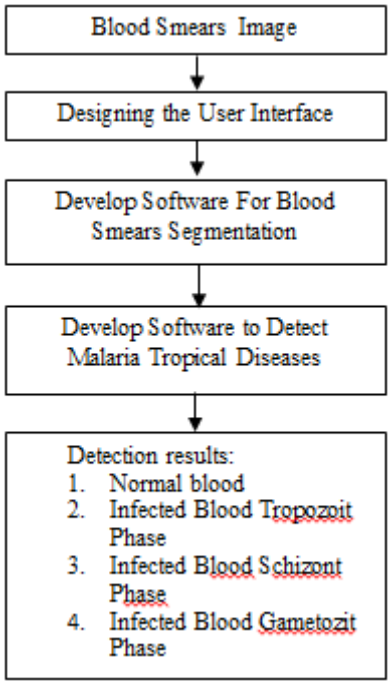


Figure 2 Steps of the Proposed Method

User interface is developed is shown in Figure 3.

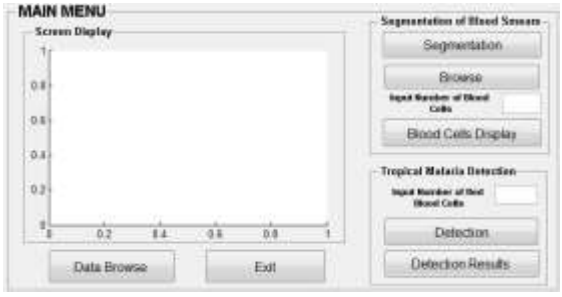


Figure 3 User Interface

SOFTWARE DEVELOPMENT FOR BLOOD SMEARS SEGMENTATION.

This step is a step to develop software for segmentation blood smear image. The purpose of this step is to divide blood smear image into several parts called blood cells, making it easier to detect. In the process of segmentation blood smear image with size of 256x256 pixels produce images of blood cells to be used in the detection process with size of 50x50 pixels. Process segmentation of blood smears image using active contour without edges [12-13]. In this paper we do not discuss active contour method, as we have discussed in the previous paper [14].The process of segmentation is shown in Figure 4.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

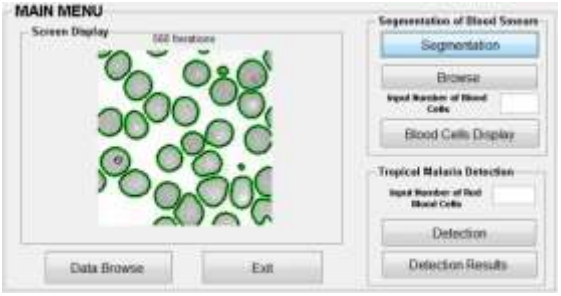


Figure 4 Segmentation Process

Results of segmentation process are blood cells that will be used for detection process. Segmentation results can be seen in Figure 5.

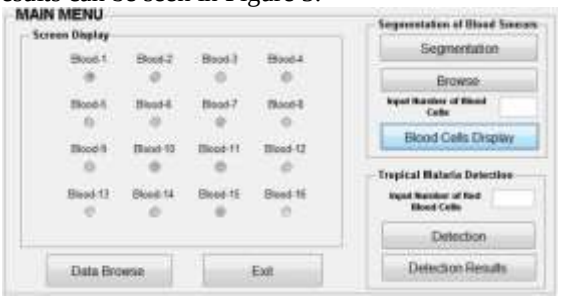


Figure 5 Display of Segmentation Results

Feature Extraction

In feature extraction step of image data from blood smears is process to get accurate information so that identification process can be done. Classification process of image based on texture analysis generally requires a feature extraction step, which consists of three kinds of methods, statistical methods, spaktral methods and structural methods. In our study using statistical methods that include: mean, standard deviation, smoothness, third moment, uniformity, and entropy. Statistical method consists of first-order feature extraction and feature extraction of second order. First-order feature extraction is done through the image histogram. To calculate the above features are used following equation.

$$Mean = \sum_{i=0}^{L-1} Z_i \; P(Z_i) \quad (1)$$

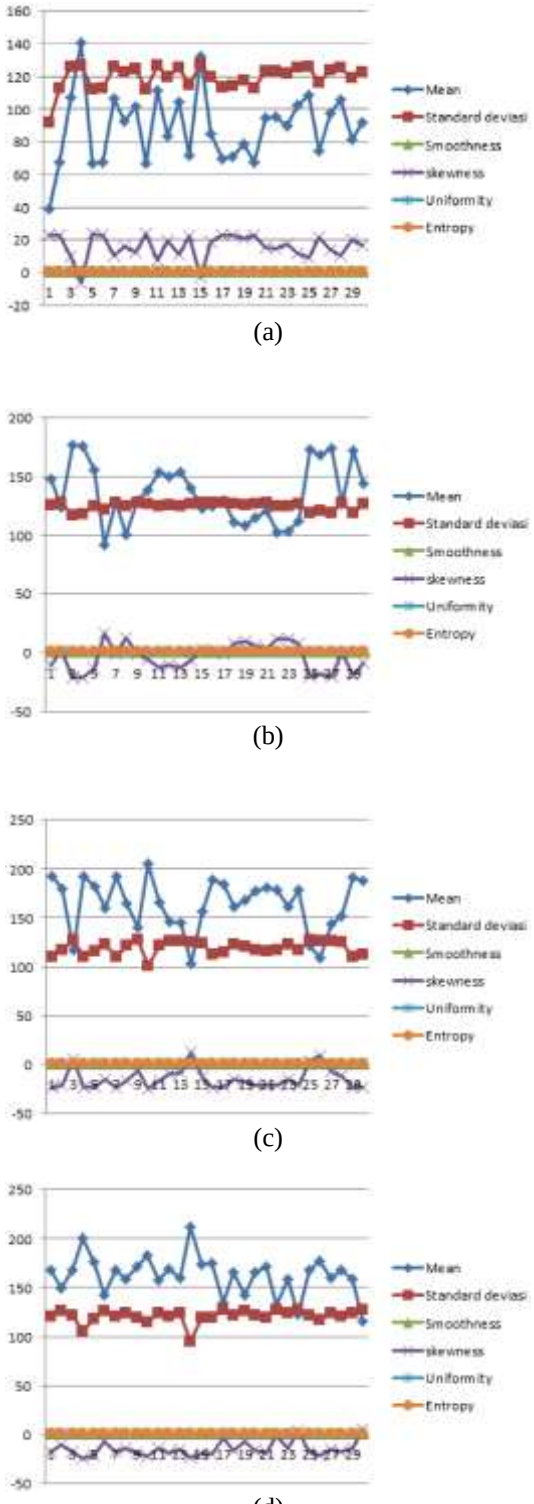
$$Standar\ Deviasi = \sigma = \sqrt{\sum_{i=0}^{L-1}(Z_i - m)^2 p(Z_i)} \quad (2)$$

$$Smoothness \; = 1 - \frac{1}{1+\sigma^2} \quad (3)$$

$$Skewness \; = \sum_{i=0}^{L-1}(Z_i - m)^3 \; p(Z_i) \quad (4)$$

$$Uniformity \; = \sum_{i=0}^{L-1} p^2 \; (z_i) \quad (5)$$

$$Entropy = -\sum_{i-0}^{L-1} p(z_i) \; log_2 \; p(z_i) \quad (6)$$



(a)

(b)

(c)

(d)

Figure 6. Extraction of Statistical Features of Blood Image.
(a) Normal, (b) thropozoit, (c) Schizont, (d) gamethozit

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

In the first order statistical feature extraction in Figure 6 there are features that overlap each other, meaning that its value are among other features. As a result, the characteristics of first-order can not be used to recognize the difference between the image accurately.

In cases like this, we would require a second order statistical characteristics. While the second-order statistical feature extraction performed by Co-occurrence matrix or Gray Level Co-occurrence Matrix (GLCM). Co-occurrence matrix is a matrix that represents the adjacency relationship between pixels in the image at different orientation direction and spatial distance. Co-occurrence means joint events, is the number of occurrences of one level of neighboring pixel values with one level another pixel value in the distance (d) and the orientation angle (θ) specific. The distance is expressed in pixels and orientation expressed in degrees. Orientation is formed in a four-way corner with an angle of 45 ° intervals, ie 0°, 45°, 90°, and 135°. While the distance between pixels is usually set at 1 pixel, 2 pixels, 3 pixels and so on. Co-occurrence matrix is a square matrix with the number of elements as the square of intensity level number of pixels in image. Each point (i, j) in Co-occurrence matrix oriented, contain opportunities incident pixels valuable i neighboring with pixel-value j at a distance d as well as orientation and (180-θ). For example, 5 × 5 matrix has GLCM matrix 45^0. Co-occurrence matrix will be calculated with the value of d = 1 and θ = 45. The number of times the emergence of the pair (i, j) is calculated for the whole matrix. Number of occurrence loaded on GLCM matrix at positions corresponding cell. Examples kookurensi matrix with distance 1 and the angle 45^0 is shown in Figure 7.

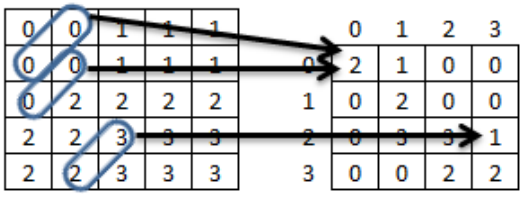


Figure 7. Co-occurrence Matrix

Software Development to Detect Malaria Tropical.

The next process is to develop software to detect malaria by performing classification using Support Vector Machine (SVM) [15]. Concept of SVM can be explained simply as an attempt to find the best hyperplane that serves as a separator of two classes in input space. Hyperplane in a d-dimensional vector space is an affine subspace dimension d-1 which divides vector space into two parts, each of which corresponds to a different class. Classification problem can be translated to business of finding a line (hyperplane) that separates two groups. Various alternative dividing line (discrimination boundaries).

The best separation hyperplane between two classes can be found by measuring hyperplane margin, and seek its maximum extent. Margin is distance between hyperplane to closest pattern of each class [16]. In this study using multiclass SVM one against all.

Figure 8. Design Process of Malaria Detection

In classification process of training data, variable hyperplane for each classifier (classifier) gained will be saved and will be used as data for each classifier in testing process, in other words, the training data classification process is to seek support vector, alpha and bias of input data (in this case used quadratic programming). While in testing process, image of blood preparations used is blood smears which had never been trained in training process with a size of 256 x 256 pixels. Segmentation process is done using active contour cropping to get image of size of 50 x 50 pixels which will be used for classification process. For testing process using results of data feature extraction and classification results of data training. Results of this process is index value of the largest decision function, stating class of testing data. If a class in classification test match test data classes, classification is stated correctly. Final result of classification is image of blood that matches with index value of decision function using SVM one against all.

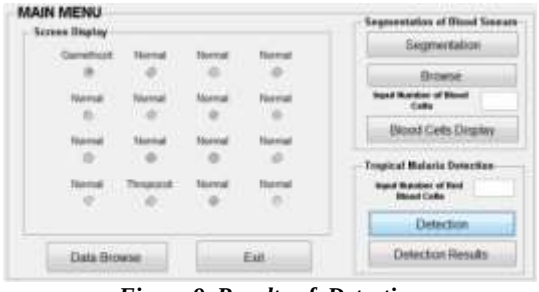


Figure 9. Results of Detection

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

In this paper, I do not discuss about the method of support vector machines in depth, as already discussed in my previous paper [14].

CALCULATE ACCURACY

To calculate accuracy, sensitivity and specificity using Receiver Operating Characteristics (ROC). First, results of classification will be done so that comparison will be obtained four values, each of which is a true positive, false negative, false positive, and true negative. True positive (TP) shows image of blood mears correctly identified according to the class (V). False positive (FP) is image of blood smears that should be identified with class turned out to be right in process of identifying wrong classification. True negative (TN) is an image that is not a member of class identified right is not a member of class (NV). False negative (FN) shows image of blood smears that should not members of class identified as a member of class.

EXPERIMENT AND RESULTS

Test of software for detection of malaria on blood smear images using SVM classification method used 120 blood smears images data with size 50x50 pixels. Each class of classification consists of 30 data. In a trial carried out a two-steps process. The first step is training phase, while second is test phase. The training phase is used to obtain coordinates of support vector, weight, bias and suppot distance vector, whereas testing phase is to use data other than training data to obtain the results of classification, so as to know level of accuracy. The testing process is done in four classes (normal blood, thropozoit blood, schizont blood and gametozit blood). Each group consists of 40 data.

Test results for normal blood cell, thropozoit, schizont and gametozit image based on two times of testing with training data and test that do not overlap can be seen in table 1.

Table 1. Accuracy of Classification Process

EXPERIMENT	ACCURACY			
	Normal (%)	Thropozoit (%)	Schizont (%)	Gametozit (%)
Testing I and II	100	100	85	95

CONCLUSION

From test results using a support vector machine, one against all method obtained an accuracy rate of 100% for normal blood, thropozoit 100%, schizont 85% and gametozit 95%. Accuracy is less related to some constraints, namely amount of training data is lacking, segmentation process can only be performed on images of blood preparations which do not overlap, extracted features are incomplete, making it less able to describe characteristics of image of blood.

ACKNOWLEDGEMENTS

Research on detection of tropical malaria is a study funded by Indonesian Education Ministry, through research grants in 2014

References

[1] U.S. Department of Health and Human Services National Institutes of Health National Institute of Allergy and Infectious Diseases, “Understanding malaria Fighting an Ancient Scourge”, 2007, Bethesda.

[2] Hariyanto P. Malaria. In: Sudoyo A, Sotiyohandi B, Alwi I, Simadibrata M, Setiati S (eds). Textbook of Medicine, Jakarta: Pusat Penerbitan Departemen Ilmu Penyakit Dalam Fakultas Kedokteran Universitas Indonesia. 2006; p. 1754-66.

[3] Wahab, Iis Hamsir Ayub, “Identification of malaria parasites in the blood using a digital image segmentation and neural networks”, Image and Vision Computing, 2008, 20:133-146.

[4] Díaz, G., González, Fabio A., Romero, Eduardo,” A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images”, J. of Biomedical Informatics, 2009, 42:296–307.

[5] Di Ruberto, Cecilia, Dempster, Andrew, Khan, Shahid, Jarra , Bill (2002), “Analysis of infected blood cell images using morphological operators”, Image and Vision Computing, 2002, 20:133-146.

[6] Deepti Gangwar, Mridul K Kalita, Dinesh Gupta, Virander S Chauhan and Asif Mohmmed, A systematic classification of Plasmodium falciparum P-loop NTPases: structural and functional correlation, Malaria Journal 2009, 8:69 doi:10.1186/1475-2875-8-69, University of Liverpool.

[7] Noorhidayati Abu Seman, Nor Ashidi Mat Isa, Lim Chia Li, Zeehaida Mohamed, Umi Kalthum Ngah, Kamal Zuhairi Zamli, Classification Of Malaria Parasite Species Based On Thin Blood Smears Using

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 2.114

Multilayer Perceptron Network, International Journal of the Computer, the Internet and Management Vol. 16. No.1 (January-April, 2008) pp 46-52, School of Electrical & Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, 14300, Nibong Tebal, Pulau Pinang, Malaysia.

[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models" International Journal Computer Vision, 1988, vol. 1, pp. 321–331.

[9] Nugroho, A.S., Witarto, B.A., Handoko, D., (2003), Support Vector Machine Theory and Applications in Bioinformatics, Kuliah Umum Ilmu Komputer.com, 2003.

[10] Hsu, C.W., and Lin, C.J., "A Comparison of Methods for Multi-class Support Vector Machines", IEEE Transactions on Neural Networks, 2002, Vol. 13, No. 2, hal. 415-425.

[11] Mark S. Nixon A and Alberto S. Aguado, "Feature Extraction And Image Processing", Second Edition, AcademicPress is an imprint of Elsevier, 2008.

[12] Chan, Tony F., and Vese, Luminita A., "Active contour without edges", IEEE Transactions on Image Processing, 2001, vol. 10, no. 2.

[13] Mahabadi, K. Navid, "Region Based Image Segmentation Using Active Contours", 2008, Stockholm.

[14] Sri Widodo, Wijiyanto, "Texture Analysis To Detect Malaria Tropica In Blood Smears Image Using Support Vector Machine", International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163, Volume 1 Issue 8 (September 2014), pages 301-306, India.

[15] Burges, J.C., "A Toturial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, 1998, Vol.2, No. 2, hal. 955-974.

[16] Boser, B. E., Guyon, I. M. and Vapnik, V. N., "A training algorithm for optimal margin classifiers". D. Haussler, editor, 5th Annual ACM Workshop on COLT, 1992, pages 144–152, Pittsburgh.

Author Biblography

Sri Widodo
He was born in Sragen on October 28, 1972. He graduated Bachelor in Sekolah Tinggi Manajemen dan Informatika Komputer 'Akakom' Yogyakarta on majoring in Computer Science in 1998. post-graduate at the University of Muhammadiyah Surakarta in the management department in 2005. Field which concentrates on currently is image processing, neural networks, artificial intelligence, biomedical, and MATLAB. At this time author actively working as a computer lecturer in Department of Medical Records and Health Informatics, APIKES Citra Medika Surakarta, Cenral Java, Indonesia.

Pipin Widyaningsih
She was born in Boyolali on January 28, 1984. She was graduated bachelor in 2006 from Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Adi Unggul Bhirawa Surakarta in information system. She received the Master Degree in Information System in 2012 from Diponegoro University. Since 2007, she has lectured in the information system departement at STMIK Duta Bangsa Surakarta. Her research interests include computer programming, database system, system analysis and design, research methods, and strategic planning.