Comparison of Different Classification Techniques Using Knowledge Discovery to Detect Malaria-infected Red Blood Cells

Jameela Ali Alkrimi, Sherna Aziz Tome, Loay E. George and
Azizah Bet. Suliman

October 14, 2019

Comparison of Different Classification Techniques Using Knowledge Discovery to Detect Malaria-infected Red Blood Cells

Jameela Ali Alkrimi
College of Dentistry, University of Babylon
Babylon, Iraq
Email: jameela_ali65 @yahoo.com

Loay E. George
Department of Computer Science, College of Sciences
Baghdad University, Iraq
loayedwar57@scbaghdad.edu.iq

Sherna Aziz Tome
College of medicine, Baghdad University
Baghdad, Iraq
shernaaziz @yahoo.com

Azizah Bet. Suliman
College of Information Technology, University
Tenaga National, Malaysia
azizah57@uniten.edu.my

Abstract— **Malaria is an infectious disease which poses a major threat to the global health field. The objective of this research paper is to present an analysis on the main machine-learning algorithms for the classification techniques used for malaria-infected red blood cells (MRBCs) and determine the best techniques by comparing classification accuracy. This study uses knowledge discovery to analyses the accuracy achieved by each algorithm and their ability to forecast and predict classification results. The system that demonstrates the computerized methods of image analysis generally involves three main phases. Firstly, data collection, pre-processing and feature extraction are conducted based on the characteristics of normal and MRBCs. Secondly, artificial neural network (ANN) and support vector machine (SVM) classification algorithm are used to classify the data set of 1,000 MRBCs. We use ten-fold cross-validation methods to estimate three prediction models to compare their performance fairly. Thirdly, we investigate the main parameters of numeric and graphical performance to evaluate the classification algorithms. The results indicate that ANN is the best predictor with 94% accuracy in the holdout sample. Where, the model evaluated prediction with a 92.9% ability to distinguish positive and negative classification. Additionally, the model has high reliability at 93%.**

Keywords— ***malaria, knowledge discovery, blood cells, microscopic images, classification, feature extraction, ANN, SVM***

I. INTRODUCTION

Knowledge discovery is the main concept that contributes to data mining. It is a sequential process that extracts information from a data set and transforms this information into comprehensible data to aid users in decision-making (Cheng, 2018), (Li, 2019). Presently, data mining is becoming common in the healthcare field because of the need for an efficient analytical approach in detecting unknown and valuable health data (Chaurasia, 2018). Presently, data mining is becoming popular in the healthcare field because of the need for an efficient analytical approach in detecting unknown and valuable health data. Data mining analyses large observational data sets to find unsuspected relationships and summarise the data in novel ways that are understandable and useful. Knowledge discovery involves the non-trivial extraction of implicit, In real life application, data mining can be used to predict potential diseases that an individual may have based on his/her health records (Akter, 2018), (Olayinka, 2019), (Witten, 2016).

Nowadays, data mining has increasing applications in medical science (Pandey, 2016). Data mining allows doctors to provide necessary treatments, thus serving patients effectively at affordable costs. As a result, data mining has become popular in the medical field. In pathology, data mining has been established as a strong technique in dealing with enormous amounts of information (Abouelmehdi, 2018). Three major data mining techniques exist, namely, regression, classification and clustering (Ngai, 2011).

Malaria is a deadly disease caused by a blood parasite called Plasmodium. This disease is the leading cause of death in numerous developing countries. In 2015, an estimated 214 million cases of malaria were recorded, causing approximately 438,000 deaths (Fernando, 2018), (Report, 2015). Malaria develops when the protozoan parasites of Plasmodium are transmitted through the bites of infected female Anopheles mosquitoes, which infect the red blood cells (Poostchi, 2018). Malaria parasites attacks RBCs when they enter the bloodstream. In malaria, parasitemia is an important step to segment RBCs from blood images and classify them as either parasite-infected or normal. Furthermore, the morphology of cells can be clearly observed in thin blood images (Olayinka T. C., 2019), (Varma, 2019) (Savkare, 2011). Moreover, Plasmodium has five species that can cause human infection, namely, P. falciparum, P. vivax, P. ovale, P. malariae and P. Knowles; P. falciparum is responsible for most malaria fatalities (Poostchi et al, 2018), (Mahdieh, 2019).

Pediatric malaria is still endemic to certain parts of the world, especially in less-developed and developing countries. The lack of adequate medical experts to diagnose the disease and prescribe treatment is largely responsible for pediatric mortality. To decrease mortality from pediatric malaria, a fast and effective detection method should be applied. Furthermore, the occurrence of pediatric malaria must be stopped urgently. Data mining can be an appropriate tool to help general practitioners detect malaria earlier by obtaining knowledge and information from the patient's health records (Poostchi et al, 2018), (Olayinka, 2019). Determining the stages of malaria is still a challenge. Without the help of pathologists, detection at this stage becomes crucial in diagnosing the disease because each stage consists of numerous characteristics, such as shape, structure, number of chromatin dots and size occupancy in the RBCs. (Olugboja, 2017)) (Loddo, 2018) (Jamil, 2019). During such hematologic disorder, changes occur in the structure and viscoelastic properties of individual RBCs. Thus, studying their physical properties can greatly contribute to the understanding and possible discovery of new treatments for this disease. New advanced mathematical computing techniques are valuable for determining and distinguishing between healthy and infected RBCs (Purnima,2016. (Das DK, 2015).

Classification is one of the data mining methodologies used to predict and classify the predetermined data for specific classes. Different algorithms for classification have been proposed by researchers (Chakraborty, 2015) (Vanaja, 2015). The most common algorithms are artificial neural network (ANN) and support vector machine (SVM) (Devi, 2018).

This study focuses on comparing the most frequently used classification algorithms in malaria and choosing the most suitable method based on the accuracy and performance these algorithms.

II. RELATED WORK

In the past few years, several study found that used machine learning for automatic detection of malaria infection from microscopic images of stained blood cells to avoid human interpretation error, most of the machine learning methods such as Naive Bayes, ANN, and SVM applied, where accuracy much higher than other machines learning algorithems (Saiprasath et al., 2019). Sajana et al. have surveyed and reported that numerous conventional machine learning and data mining algorithms can be applied to classify malaria-infected RBCs (MRBCs). Several of them use ANN and SVM to provide high accuracy (Rajaraman, 2019) (Sajana T. &., 2018,) (Vanaja, 2015). Consequently, various methods have been developed, proposed and analysed to detect and classify RBCs in malaria. Purnima Pundit et al. have used ANN technology to detect the infection of malaria RBCs after analysing digital holographic interferometric microscopic images (Olugboja, 2017). Khot et al. have proposed a method for the detection of malaria parasites basedon KNN with ANN techniques (Sajana T, 2018). Jan et al. review on automated diagnosis of malaria parasite in microscopic blood smears images using ANN and SVm (Jan et al., 2018). Diaz et al. have used multilayer perceptron and SVM techniques to classify malaria parasites in RBCs (Gloria Diaz, 2009) (S.T. Khot, 2014). Sharma et al. have applied SVM and ANN to predict malaria outbreaks, which is the key to controlling malaria morbidity (Sharma V, 2015). Kapor and Rani have employed an efficient decision three algorithm using J48 to reduce error pruning, Bayesian and SVM use the best set of their discriminating features to provide high classification accuracy (Kapor P, 2015). Narayanan et al. have addressed the development of the computer-assisted malaria parasite characterisation and classification through machine learning approach. Moreover, based on light microscopic images of peripheral blood smears, the characterisation uses SVM as a classifier to determine whether the patient is affected by malaria (Narayanan, 2019). Singla et al. have used the pretrained convolutional neural networks (CNNs) classification of malaria-infected stages with limited labelled data size (Singla, 2019).

This research paper analyses the different classification techniques that are applied to MRBCs and compares the performance of classification accuracy in choosing the best techniques.

III. Automated classification of Malaria RBCs

The system for automated cell microscopy usually implements a sequence of key processing steps. Each step includes techniques or algorithms. Therefore, each of the subsections will focus on one specific aspect of the processing task. To enhance performance, hematological data analysis is divided into three phases: malaria data collection, classification of algorithms and evaluation of results and performance. These phases are shown in Figure 1.

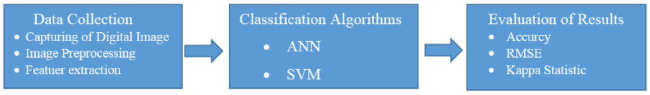


Fig 1. The hematological data analysis phases

A. Data Collection

1. Capturing of Digital Image

Peripheral blood smear slides of anemic cases were obtained from the hematology unit of the Pathology Department, Faculty of Medicine in Serdang Hospital from March 2012 to August 2012. All peripheral blood slides related to hematological cases were stained by experts with May Grünwald–Giemsa under a light microscope with 10×100 magnification. No loss of parasite was observed during the staining of thick blood smear; the artifacts and parasites were observed in their natural location. To diagnose malaria under a microscope, a drop of the patient's blood is applied to a glass slide, which is then immersed in a staining solution to make parasites more visible under a conventional light microscope, usually with a 100× oil objective. Two different types of blood smears, namely, thick and thin smears, are typically prepared for malaria diagnosis.

A thick smear is used to detect the presence of parasites in a drop of blood. Thick smears allow more efficient detection of parasites than thin smears. However, thin smears, which is a result of spreading a drop of blood across a glass slide, still have advantages. They allow the examiner to identify malaria species and recognise parasite stages more easily. Figure 2 shows the blood smear images of patients infected with malaria.

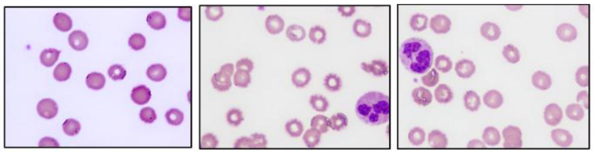


Fig 2. blood smear images infected malaria

2. Image Preprocessing

- Preprocessing Blood Smear Images

The captured raw images require several preprocessing steps to isolate each individual MRBC. Each step includes the algorithm, methods and results. The result of each process is input into the second process, and the algorithms improve the quality of the preprocessing. The first step is the conversion of the image from colored to grey. The image uses spectral analysis to obtain the grey stretched image then apply the new image using gamma mapping method to address the distortion in the image edges. To smooth out the MRBCs, the image uses mean filter algorithm. Thereafter, banalization process uses the morphological tiling operation algorithm to convert the grey image to a binary image. The isolation of an individual RBC image is achieved using threshold algorithm. The four types of malaria-infected RBCs are grouped and named in Figure 3.

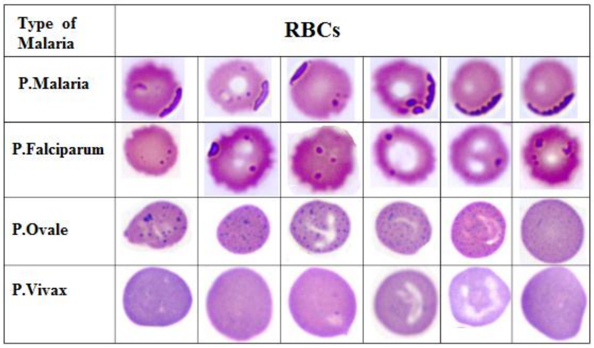


Fig 3.The type of Malaria RBCs

3. Feature extraction

Feature extraction is the process of defining a set of features, or image characteristics, which will represent important information for analysis and classification. Ideally, healthy RBCs are doughnut-shaped, whereas infected RBCs become spherical due to the growth of parasites within the RBC. In our work, the different classification techniques of healthy and infected RBC are compared using textures and geometric and statistical features. The set of features are based on the characteristics of normal and MRBCs. The geometric features include the irregularity, size and circularity of RBCs. The textures' features are based on statistical measures which include variance, contrast and moment. This set of features holds the following subsets:

a) The histogram moments of up to the fifth order for the colour components and the corresponding grey image.
b) The spatial moments of the colour components and grey values. The pivot point is the centre of mass.
c) The contrast histogram moments for each colour and grey component.
d) The gradient histogram moments for each colour and grey component.

Thus, the new and accreted MRBC data set includes 70 attributes and 1000 instances. It also consists of 200 samples from each of the four malaria types and healthy RBCs.

B. *Classification*

The classification task uses MRBCs with two machine learning algorithms: ANN and SVM. Initially, the dataset is pre- processed to remove the noise and fill the missing values. The pre-processing step prepares the classification of the dataset. Figure 3 depicts these steps. The dataset is split into 30% training and 70% testing samples. In the training phase, the pre-determined data and its associated class label are used for classification. In the testing step, the test data are used to estimate the accuracy of the classifier rule. If the accuracy of the classifier rule on test data is acceptable, the rule can be applied for further classification of unseen data. The most important factor in medical diagnosis is the accuracy of the classifier. Figure 4show the classification of MRBCs data set steps.

A. *Artificial neural networks (ANN)*

ANN is a type of generative modelling in which machine learning occurs in an unsupervised learning context. It is also a broad area of machine learning where models learn probability distribution P(X) and generate samples from that distribution. We can use these models to create new unseen images by training them on a dataset of images. The ANN structure comprises three layers: input, hidden and output. The first layer is a set of inputs that present 70 features. An input vector is used for all radial basis functions, each with different parameters. The second layer presents the hidden layer with a non-linear RBF activation function and a linear output layer.

The hidden units and linear activation functions on input and output units will generate a network that is equivalent to a linear classifier. Then, larger numbers of hidden units will be necessary to obtain good classification results.

B. *Support Vector Machines (SVM)*

SVM algorithm is based on a learning system which uses statistical learning methodology. These algorithms are widely used for classification. In SVM, the optimal boundary, known as the hyperplane, of two sets in a vector space is obtained independently on the probabilistic distribution of training vectors in the set. This hyperplane locates the most distant boundary from the vectors near the boundary in both sets. The vectors that are placed near the hyperplane are called supporting vectors. If the space is not linearly separable, no separating hyperplane is present. Sequential minimal optimisation (SMO) is used for training a support vector classifier through polynomial or RBF kernels. SMO replaces all the missing values and transforms nominal attributes into binary ones. A single layer of a hidden neural network uses exactly the same form of model as an SVM (Keerthi, 2001).

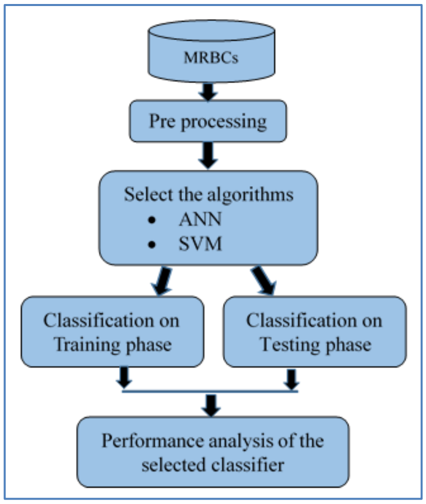


Fig 4. Classification of MRBCs data set steps

B. *Evaluation Classification Algorithms*

Performance evaluation is an important aspect of the machine learning process. The performance of a machine learning classification model can be evaluated in various ways as classification accuracy alone cannot be trusted to select a wellperforming classification model (Das, 2013), (Sokolova, 2006). Thus, classification can be conducted in two ways. Firstly, the features used by the machine learning models are their own measurements. Secondly, learned models use performance evaluation (Flach, 2019), (Bbosa F, 2016). At first, cross-validation and accuracy are applied. Then, confusion matrix is applied in order to accurately evaluate the machine learning classifiers.

C.1 Cross-validation

Witten et al. have asserted that the stratified ten-fold cross-validation technique is the acceptable method in validating the classifiers. In machine learning, the testing dataset must have sufficient size, which is usually much larger than the training dataset to avoid an overlap between the two datasets (Jiao Y. a., 2016) In bioinformatics research, the available samples are usually limited. Meanwhile, in machine learning, cross-validation methods are considered as a compromise solution when the number of available samples is very limited. Given the number of available samples in bioinformatics, cross validation methods become the most recommended evaluation methods. In a cross-validation test, all the data are used as training and testing datasets (Danjum, 2015). In an n-fold cross-validation test, the entire dataset is randomly partitioned into equal-sized n parts. In this study, we have used the most common partitions: (3,5and 10-fold). Figure 5 shows the accuracy with different k-folds for both algorithms.

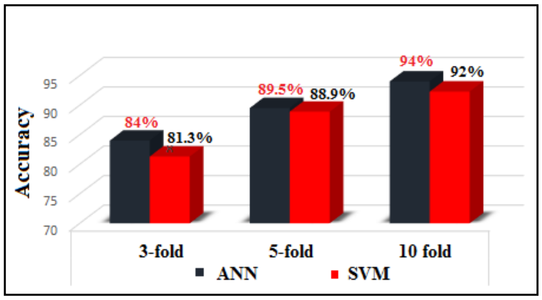


Figure 5 Comparison accuracy of ANN and SVM algorithms with different k-fold

C2. Accuracy

The classification of MRBC dataset results must be accurate. Accuracy analysis is based on the main parameters' root mean squared error (RMSE) and mean absolute error (MAE) as well as time. The MAE and RMSE are two of the most common metrics used to measure the accuracy of classification. MSE is used for the validation of models. Model validation is an important part of the machine learning process. Validation methods consist of creating models and testing them on a dataset (Jiao Y. a., 2016), (Cawley, 2010 July). Table 1 shows the results of comparative ANN and SVM algorithm in the classification of MRBCs.

Table 1.
The comparative results of ANN and SVM algorithm

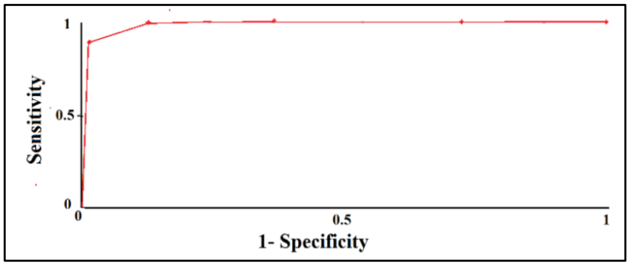
Algorithms	MAE	RMSE	Accuracy	Time sec.
ANN	9.0124	0.986	94	4.34
SVM	13.1367	0.1167	92.3	1.03

C3. Confusion Matrix

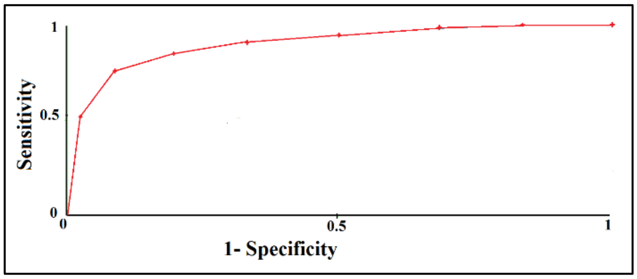
The metrics selected to evaluate the machine learning model is important as the choice influences how the performance of machine learning algorithms is measured and compared[Sunasra,2017]. A confusion matrix is a summary of prediction results on a classification problem. It shows the ways in which a classification model is misinterpreted when making predictions [Sammut,2017] [Ting,2017]. In this research, the numbers and graphs of the main parameters are used to compare the evaluation performance between classifiers and predictive ANN and SVM algorithms for the MRBC dataset. Sensitivity, specificity, f-measure and kappa statistic are used as numeric parameters. Also, receiver operating characteristics (ROC) are used as graphical parameters. Table 2 shows the evaluation performance of SVM and ANN algorithms, while Figure 6 shows the ROC curves of ANN and SVM.

Table 2
Prediction Performance Measures algorithms

Algorithm	Accuracy	Sensitivity	Specificity	F-measure	ROC	Kappa Statistic
ANN	94%	0.92	0.89	0.929	0.988	0.9308
SVM	92.30%	0.919	0.82	0.918	0.804	0.9113



a. ANN ROC curve



b. SVM ROC curves

Fig 6. ROC Curves of ANN and SVM

IV. COMPARATIVE ANALYSIS

Several machine learning classifiers have been used for MRBCs. Moreover, the literature in finding the best classifier for this application are SVM and ANN. Hence, this paper presents a comparison and analysis on the accuracy of these algorithms in the classification of MRBC datasets. These algorithms are evaluated based on the following;

a) Validation of the models

ANN algorithms give the best validation as shown in Figure 4, wherein ANN achieves higher accuracy than SVM, as well as MAE and RMSE.

b) Performance metrics

Confusion matrix is used to distinguish the best evaluation classification model. The matrix has found that the RBFANN model is better than the RBFSVM model, as shown in Table 3. Additionally, ANN gives significant prediction agreement with the actual class label.

V. CONCLUSION

In this paper, we have analysed and compared the most in-command classification algorithms that are applied in MRBCs in ANN and SVM. The performance of classification accuracy is compared to choose the best model. The experimental results show that ANN is more accurate than SVM, having 94% accuracy compared with SVM's 92.3%. The classifier's accuracy is evaluated using two points. Firstly, the features used by the machine learning models are their own measurements. Secondly, the performance of learned models are evaluated. In numeric parameters, the experimental results show that ANN achieves 92% sensitivity, 89% specificity, 92% f-measure and 93.08% kappa statistic. In graphical parameters, the ROC curves ANN's superior ability to classify the four types of MRBCs.

REFERENCES

[1] Abouelmehdi, K. B.-H. (2018). Big healthcare data: preserving security and privacy. Journal of Big Data, 5(1), 1.

[2] Akter, F. H. (2018). Classification of Hematological Data Using Data Mining Technique to Predict Diseases. Journal of Computer and Communications, 6, 76-83. https://doi.org/10.4236/jcc.2018.64007 Received.

[3] Bbosa F, R. W. (2016). Clinical malaria diagnosis: Ruled-based classification statistical prototype. Publisher: SpringerPlus. , 5:939.

[4] Cawley, G. C. (2010 July). On over-fitting in model selection and subsequent selection bias in performance evaluation. Journal of Machine Learning Research 11.J, 2079-2107.

[5] Chakraborty, D. D. (2015). Computational microscopic imaging for malaria parasite detection: a systematic review. Journal of microscopy, 260(1), 1-19.

[6] Chaurasia, V. P. (2018). Prediction of benign and malignant breast cancer using data mining techniques. Journal of Algorithms & Computational Technology,12(2), 119-126.

[7] Cheng, Y. C. (2018). Data and knowledge mining with big data towards smart production. Journal of Industrial Information Integration, , 9, 1-13.

[8] Danjuma, K. J. (2015). Performance evaluation of machine learning algorithms in post-operative life expectancy in the lung cancer patients. arXiv preprint arXiv:1504.04646

[9] Das DK, M. R. (2015). Computational microscopic imaging for malaria parasite detection: a systematic review. Journal of microscopy, 260(1), 1-19.

[10] Das, D. K. (2013). Machine learning approach for automated screening of malaria parasite using light microscopic images. Micron, 45, 97-106.

[11] Devi, S. S. (2018). Hybrid classifier based life cycle stages analysis for malaria-infected erythrocyte using thin blood smear images. . Neural Computing and Applications,29(8), 217-235.

[12] Fernando, M. B. (2018). Plasmodium Falciparum Malaria and Severe Dengue Coinfection in a Pulmonary Tuberculosis Patient: Case Report and Literature Review. Archives of Clinical and Medical Case Reports, 2(3), 82-89.

[13] Flach, P. (. (2019). Performance Evaluation in Machine Learning: The Good, The Bad, The Ugly and The Way Forward. In 33rd AAAI Conference on Artificial Intelligence.

[14] Gloria Diaz, F. A. (2009). A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. Elsevier - Journal of Biomedical Informatics., (42): 296–307.

[15] Jamil, M. M. (2019). Computational Automated System for Red Blood Cell Detection and Segmentation. In Intelligent Data Analysis for Biomedical Applications s., (pp. 173-189). Academic Pres.

[16] Jan, Z. K. (2018). A review on automated diagnosis of malaria parasite in microscopic blood smears images. Multimedia Tools and Applications, 77(8), 9801-9826.

[17] Jiao, Y. a. (2016). Performance measures in evaluating machine learning based bioinformatics predictors for classifications. Quantitative Biology 4.4, 320-330.

[18] K.M., T. (2017). Confusion Matrix. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning and Data Mining. Springer, Boston, MA.

[19] Kapor P, R. R. (2015). Efficient decision tree algorithm using j48 and reduced error pruning. International, Journal of Engineering Research and General Science. 3(3): 2091-2730.

[20] Khan, S. E. (2011). Hematology and Serum Chemistry Reference Values of Stray Dogs in Bangladesh.

[21] Li, J. S. (2019). A knowledge discovery and reuse method for time estimation in ship block manufacturing planning using DEA. Advanced Engineering Informatics, 39, 25-40.

[22] Loddo, A. D. (2018). Recent advances of malaria parasites detection systems based on mathematical morphology. Sensors, 18(2), 513.

[23] Mahdieh Poostchi, K. S. (2019). Image analysis and machine learning for detecting malaria. Translational Research, Volume 194, Pages 36-55.

[24] Narayanan, B. N. (2019). Performance analysis of machine learning and deep learning architectures for malaria detection on cell images. Applications of Machine Learning International Societ, (Vol. 11139, p. 111390W).

[25] Ngai, E. W. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature Decision support systems, 50(3), 559-569.

[26] Olayinka, T. C. (2019).). Predicting Paediatric Malaria Occurrence Using Classification Algorithm in Data Mining. Journal of Advances in Mathematics and Computer Science, 1-10.

[27] Olugboja, (2017). July). Malaria parasite detection using different machine learning classifier. In 2017 International Conference on Machine Learning and Cybernetics (ICMLC), (Vol. 1, pp. 246-250). IEEE.

[28] Pandey, S. C. (2016). Data mining techniques for medical data: a review. International Conference on Signal Processing, . Communication, Power and Embedded System (SCOPES) (pp. 972-982). IEEE.

[29] Poostchi, M. S. (2018). Image analysis and machine learning for detecting malaria. Translational Research, 194, 36-55.

[30] Rajaraman, S. J. (2019). Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. . PeerJ, 7, e6977.

[31] Report, W. H. (2015). URL: www.who.int/malaria/publications/world-malaria-report-2015.

[32] S.S. Keerthi, S. S. (2001). Improvements to Platt's SMO Algorithm for SVM Classifier Design. Neural Computation, 13(3), pp 637-649,

[33] S.T. Khot, R. (2014). Prasad. Optimal Computer Based Analysisfor Detecting Malarial Parasites. Advances in Intelligent systems and computing, :69-80.

[34] Saiprasath, G. B. (2019). Performance comparison of machine learning algorithms for malaria detection using microscopic images. IJRAR February 2019, Volume 6, Issue 1

[35] Sajana, T. (2018). Classification of Imbalanced Malaria Disease Using Naïve Bayesian Algorithm. International Journal of Engineering & Technology, 7(2.7), 786-790.

[36] Sajana, T. (2018,). Majority Voting Algorithm for Diagnosing of Imbalanced Malaria Disease. I. n International Conference on ISMAC in Computational Vision and Bio-Engineering (pp. 31-40). Springer, Cham.

[37] Sammut, C. a. (2017). Encyclopedia of machine learning and data mining. Springer Publishing Company, Incorporated.

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

[39] Sharma V, A. K. (2015). Malaria outbreak prediction model using machine learning. International. Journal of Advanced Research in Computer Engineering & Technology (IJARCET)., 4(12).

[40] Singla, N. &. (2019). Deep learning enabled multi-wavelength spatial coherence microscope for the classification of malaria-infected stages with limited labelled data size. arXiv preprint arXiv:1903.06056.

[41] Sokolova, M. N. (2006). Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. Australasian joint conference on artificial intelligence. Springer, Berlin, Heidelberg,

[42] Sunasra, M. (2017). Performance metrics for classification problems in machine learning. Medium

[43] Vanaja, S. &. (2015). Performance Analysis of Classification Algorithms on Medical Diagnoses-a Survey. Journal of Computer Science (JCS), 11(1), 30-52.

[44] Varma, S. L. (2019). Detection of Malaria Parasite Based on Thick and Thin Blood Smear Images Using Local Binary Pattern. In Computing. Communication and Signal Processing, Springer, Singapore., pp. 967-975).

[45] Witten, I. H. (2016). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.