

# Red Blood Cells Abnormality Classification: Deep Learning Architecture versus Support Vector Machine

Hajara Abdulkarim Aliyu<sup>1,2,\*</sup>, Rubita Sudirman<sup>2</sup>, Mohd Azhar Abdul Razak<sup>2</sup>,  
Muhamad Amin Abd Wahab<sup>2</sup>

<sup>1</sup>Jigawa State Polytechnic Dutse,  
Kiyawa Road, Jigawa State, 7040, Nigeria,

<sup>2</sup>Universiti Teknologi Malaysia,  
Faculty of Electrical Engineering, Johor Bahru, 81310, Malaysia

\* Corresponding Author

DOI: <https://doi.org/10.30880/ijie.2018.10.07.004>

Received 26 October 2018; Accepted 15 November 2018; Available online 30 November 2018

**Abstract:** The most common and dangerous defect of red blood cells (RBCS) is shape abnormality, The primary detection and confirmation of anaemic stage(shape abnormality) is based on haemoglobin level or manual microscopic examination of peripheral blood smears. The problem of classifying the abnormal cells manually under microscope is that it consumes time, working on huge number of sample manually is burdensome which leads to poor result quality with unnecessary medication leading to life trait to the patient and cause eye fatigue to the technicians. This paper proposed a method to classify Rbc's abnormalities based on deformed shaped RBCs image by using SVM and Deep learning in comparison on the RBCs cell Classification. Classifying normal cells of RBCs indicate a healthy patient and Classifying anemic abnormalities indicate presence of disease. And is very important in medical field to detect and classify disease in early stage because it saves and protects human lives. The patients waiting time for blood test is longer because the time taken to generate the result of the blood test is more due to high demand and less equipment. This lead to comparison of the two classifiers in order to predict the one that will best perform on RBCs in order to achieved maximum accuracy for the classification. This study shows that SVM classifier can classify the cells in all condition either small or large dataset while deep learning performs mainly on large and very large dataset which RBCs dataset will be generated in large amount in order to work successfully with the state of the earth on RBCs deformity.

**Keywords** — Red blood cells (RBCs); Deep Learning; SVM; Rbc's abnormality

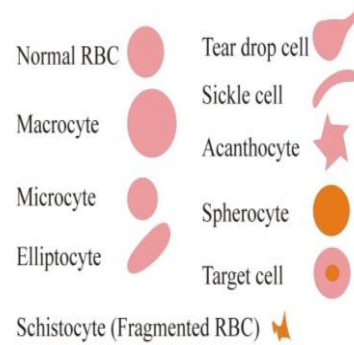
## 1. Introduction

The human blood cell is comprise of the three major components of blood cells that are white blood cell (WBC), platelets and red blood cell (RBC). The RBCs are majority of cells in human body and it has many functions in human body, like moving oxygen round the body, carrying waste and carbon dioxide products away from tissue and cells. The normal shape of RBCs are biconcave disk with 7 to 8µm in cell diameter and 2.2 µm thickness (Aliyu, 2017). The RBCs abnormal morphological nature of the cells gives anemia sign, hemoglobin reduction (the protein that bind with an oxygen molecule in RBCs), also the secondary effect of many other disorders. Considering medical perspective, the diagnosis of RBC gives more information on various related blood cell diseases. For example, the shape of the RBCs with its deformity has connection to the relevant disease more especially anemia and the secondary effect of several other disorder (Webster, & Cazzanti, 2004). Approximately 24.5% of the world population are affected with anemia and other related blood disorders. This makes most pathological laboratories to used visually inspection of the blood smear slide under the microscope. The method is expensive, time consuming, laborious, and need skilled technicians (Dalvi & Vernekar, 2016).

\*Corresponding author: [hajabdulkarim@jigpoly.edu.ng](mailto:hajabdulkarim@jigpoly.edu.ng)

2018 UTHM Publisher. All right reserved.  
[penerbit.uthm.edu.my/ojs/index.php/ijie](http://penerbit.uthm.edu.my/ojs/index.php/ijie)

The abnormal shaped RBCs differ greatly depending on the kind of blood disorder a patient is suffering from shown in Fig 1.



**Fig. 1- normal and abnormal rbc's (<https://bigpictureeducation.com/blood-cells-images>)**

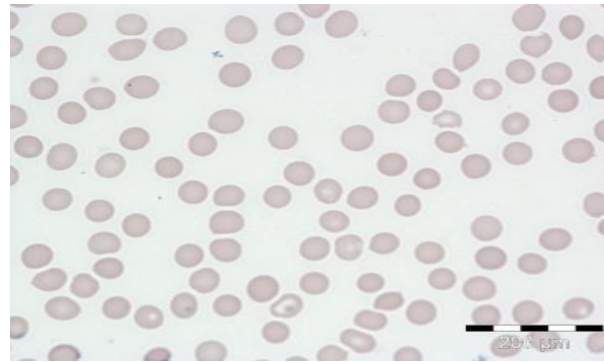
The major stages in image analysis are pre-processing, segmentation, feature extraction and classification. The important and challenging steps are feature extraction and classification to detect and classify the exact feature of the blood cells in time.

Many researchers have worked on similar research since some decades ago and still the area need more effort to solve erythrocyte problems, Sheikh *et al.* used neural network to identify abnormality in Red blood cell, White blood cells and platelets in anemic condition (Sheikh, Zhu, & Micheli-Tzanakou, 1996). Abnormal red blood cell was characterised based on parametric deformable but high number of images was not considered template (Bronkorsta, Reinders, Hendriks, Grimbergen, Heethaar, & Brakenhoff, 2000). MLP (Multilayer perception) was used to classify normal RBCs and sickle cells by Poomcokrak but other abnormalities are not considered (Poomcokrak & Neatpisarnvanit, 2008). A method using Hough transform and neural network to detect healthy and unhealthy RBCs (Elsalamony, 2016). Work was done using machine learning techniques to characterized red blood cell in anemia based on blood smear images and the 25 sets of features are used for identification but observed that using small set of feature will present higher accuracy than they achieved (Das, Chakraborty, Mitra, Maiti, & Ray, 2013). Deep learning was used to classify plant disease but worked on high number of samples and does not consider less dataset (Mohanty, Hughes, & Salathé, 2016). A work was proposed based on depth map for RBCs classification on the surface type in which the work was successful in red blood cell classification but is mainly based on counting the number of red blood cell the blood smear image, and it does not establish link between the surface type and red blood cell disease detection (Vromen & McCane, 2006). A work was done by measuring angular distribution of lights scattered from a collection of red blood cells. Then by using a method based on fast Fourier transform (FFT), they calculated the size and refractive index of cells but the diseases were not detected (Ghosh, Buddhivant, Uppal, Majumder, Patel, & Gupta, 2006). The main advantage of the ANN is its response time, which means how fast is the prediction of the future values was made. Then when the learning process was completed, is the most slowest of the step in use of ANN, the neural network is ready for use, obtaining results very quickly compared to other more complex prediction models as ARIMA (Cortez, Rio, Rocha, & Sousa, 2012). These methods need special apparatus and setup and are computationally expensive. In this paper we compare SVM and Deep learning classification to automatically detect normal, achantocyte, elliptocyte, sickle cell and teardrop red blood cell (RBC) in the blood smear image to classify the disease fast and effective for emergency purpose without following the process of preprocessing stage. The rest of the paper is organized as follows: Section 2 describes the proposed method, Results and discussions are listed in section 3, Conclusion in section 4.

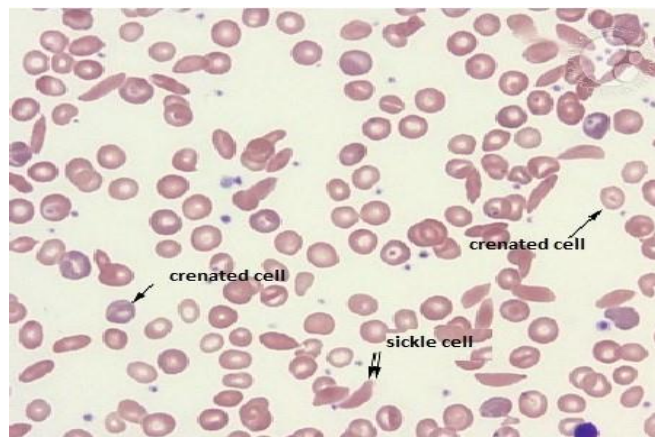
## 2. Methodology

### 2.1 Image Acquisition

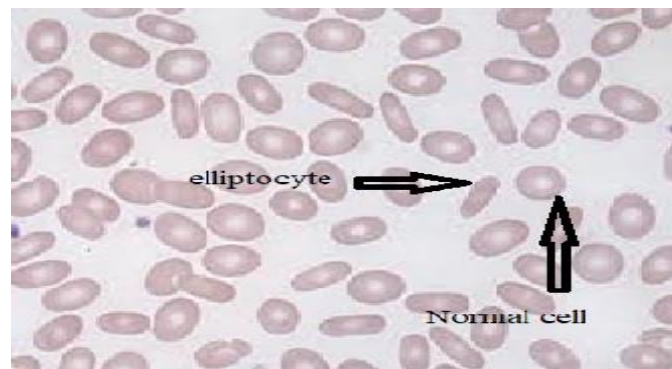
The input images are acquired from Pinterest online open source haematology database. We analyse 105 and 250 images of Red blood cells respectively, which have a spread of 5 class labels normal, achantocyte, sickle cell, teardrop and elliptocyte cell assigned to them for first analysis in Table 1, 21 images each class and for Table 2, 50 images each class are used. Each class label is a blood disease with a normal cell, and we make an attempt to predict the RBCs disease pair given just the image of the blood smear image. Fig 2 shows one example each from every RBCs disease from online haematology images. In all the approaches described in this paper, the images were resized to  $256 \times 256$  pixels, and both the model prediction performance on these downscaled images. Across the experiments, we use the images in colour version as for the SVM and Deep learning classification.



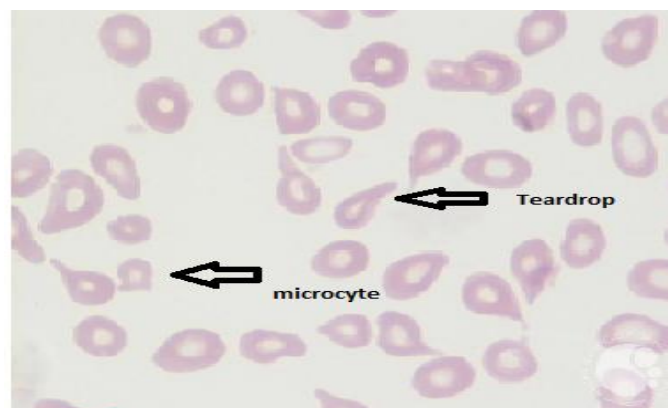
(a)



(b)



(c)



(d)

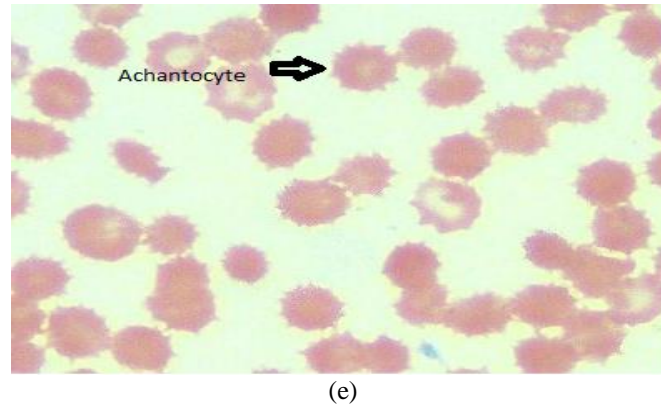


Fig 2- example of acquired mages on blood smear image (a) normal cell; (b) sickle cell; (c) elliptocyte cell; (d) teardrop cell; (e) achantocyte.

### 2.2 Cropped Images

The images in Fig. 3 are cropped from the blood smear to generate the dataset of normal, achantocyte, sickle cell, teardrop and elliptocyte cell in order to train the SVM model and deep neural network for the classification of the diseases using colored images in order to predict the classifier performance.

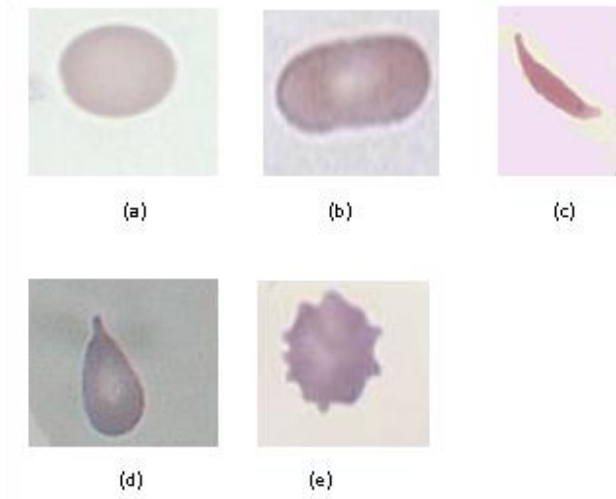


Fig. 3- (a) normal cell; (b) elliptocyte cell; (c) sickle cell; (d) teardrop cell; (e) achantocyte.

### 2.3 Support Vector Machine

The application of the multiclass Support Vector Machine using feature selection is an interesting procedure on red blood cells (Russakovsky, Deng, Krause, Satheesh, & Berg, 2015). The classifier allows the feature selection using shape feature in which the process is shown in Fig.4.

Matlab 2017b software has used for developing this classification method using Support Vector Machine (SVM) with Radial Basis Function (RBF) default setting. The study on the behavior of the features on the cells has been identified in classifying the RBCs cells types in the work. The feature ranking was done on almost majority of the features and may also be done for all the features that are working together. The method is based on the idea, that the absolute values of the weights of a linear classifier trained on the whole set of features produces a feature ranking (Ciresan, Meier, Masci, Maria Gambardella, & Schmidhuber, 2011).The features that are associated with the larger weight are more important than those associated with smaller one.

The training samples  $\{y_i x_i\}_{i=1}^N$ , have a label of  $y_i \in (-1, +1)$  shows a class that the feature vector is  $x_i \in R^d$ . The SVM has separated hyper plane with maximum-margin in higher feature space is induced by RBF kernel function  $k(x, z)$ . The SVM classifier in normal form is defined as follows:

$$F(x) = \sum_{i=1}^n \alpha_i y_i k(x, v_i) + b \tag{1}$$

Where  $\{v_i\}_{i=1}^n$  are the support vectors which make the classifier to used little set of training data close to the separating hyper plane.  $x$  is the feature vector and is normalized to unit length before the training and the classification.

The performance of multilayer support vector machine was tested on the RBCs images using the training and testing images 80% to 20% respectively. The basic result such as the overall accuracy was predicted.

### 2.3.1 Feature Extraction

The healthy red blood cell looks like a thin circular disk. In order to analyze the shape of red blood cell quantitatively, the following shape features are extracted: Form factor, Compactness, Eccentricity, Minor axis and Major axis.

- Form factor: This refers to the ratio between the cell area and the square of its perimeter which shows the circularity measurement of the cell

$$\text{Form factor} = \frac{4 \pi A}{p^2} \tag{2}$$

Where A is the total area of the cell and p is perimeter.

- Compactness = 
$$\frac{\sqrt{\frac{4 A}{\pi}}}{\text{Maximum Diameter}} \tag{3}$$

- Eccentricity = 
$$\sqrt{1 - \frac{L_{minor}^2}{L_{major}^2}} \tag{4}$$

Where  $L_{minor}$  is the length of minor principle axis while  $L_{major}$  is the length of major principle axis.

- Major Axis: The number of pixels of the longest diameter of red blood cells passing through its centre.
- Minor Axis: The number of pixels of the smallest diameter of the red blood cell that pass through the centre of the red blood cell and it's perpendicular to the major axis.

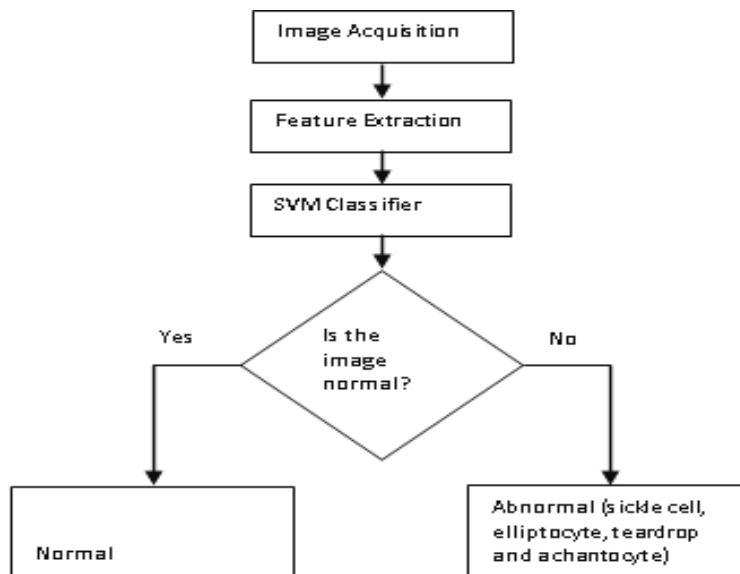


Fig.4- svm workflow

### 2.4 Deep Neural Network Architecture

Deep learning is a machine learning in which the model learns to perform its classification tasks directly from the images, sound, or text. Deep learning usually is implemented by using the neural network architecture. The term “deep” is the number of layers in the network. The more layers present, the deeper the network (Guyon & Elisseeff, 2003). Traditional neural networks contain only 2 or 3 layers, while deep neural networks can have hundreds. The machine learning can be

categorized into Supervised learning, Unsupervised learning and Re-enforced learning. But in this study supervised learning is considered Fig 5 shows the flow chart for deep neural network classification.

The deep learning model was trained using Matlab 2017b AlexNet architecture default setting was used for the training. The AlexNet architecture is not that different from another well-known LeNet-5 architecture in which the variants or local receptive fields are normally a couple of arranged convolution layers succeeded by one or more fully connected layers (Mohanty, Hughes, & Salathé, 2016). Those same convolution layers can also have normalization and pooling layers right after them, and traditionally all the layers are initiated or activated using the Rectifying linear unit or function (ReLU). The ReLU applies a transformation to the output of each neuron, and then maps the output to the highest possible value or to zero if negative. It adaptively learns the parameters of rectifiers thereby improving accuracy at negligible extra computational cost. ReLU is given as:

$$f(z_i) = \max(0, z_i) \tag{5}$$

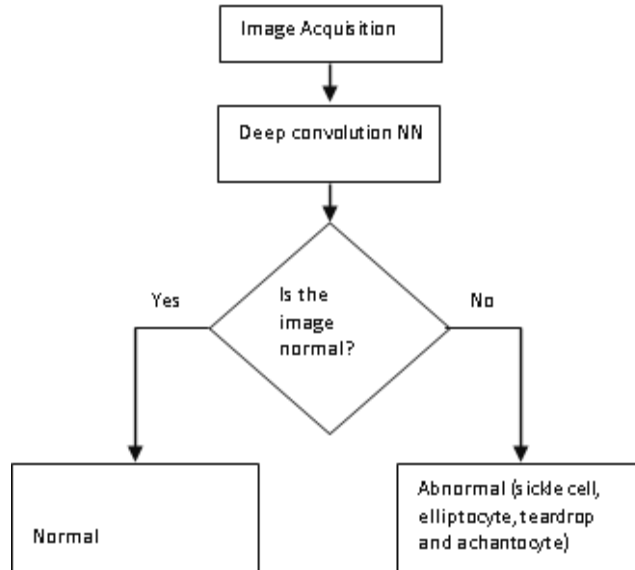
Where  $z_i$  represents the input of nonlinear activation function  $f$  on the  $i^{th}$  channel.

The pooling layer further transform the output of the activation step by reducing the dimensionality of the features map considering the output of the small region of neurons into a single output. The AlexNet comprises of 5 convolution layers and 3 fully connected layers with a softMax layer as the final layer. Each individual convolution layer  $C_{Li}$  has maps of the same size for the two directions of  $x$  &  $y$  of the image given as  $M_{ix}$  and  $M_{iy}$  with kernel sizes  $k_{ix}$  and  $k_{iy}$  respectively. Then given the number of pixels to skip during the traverse at both directions denoted as  $S_{ix}$  and  $S_{iy}$ , the final output map size could be given as (Russakovsk, Deng, Krause, Satheesh, & Berg, 2015).

$$M_{ix}^L = \frac{M_{ix}^{L-1} - k_{ix}^L}{S_{ix}^L + 1} + 1 \tag{6}$$

$$M_{iy}^L = \frac{M_{iy}^{L-1} - k_{iy}^L}{S_{iy}^L + 1} + 1 \tag{7}$$

Where  $L$  denotes the layer, thereby the next layer of  $L^n$  is connected to through most maps in the previous layer. The performance of deep neural network using transfer learning. Alexnet architecture was tested on the same set of RBCs images that was used on the SVM classifier.



**Fig. 5- deep neural network workflow**

In this section, the performance of proposed RBC classification system is evaluated for recognizing the RBC cell type by using 105 and 250 images respectively by directly using colored cropped cells from the blood smear images in Fig. 3. This shows that using SVM shows high performance learning improved compared with deep neural network on the same RBCs images while considering the 250 images the deep neural improved.

### 2.5 Accuracy Calculation

The execution of the two classification model has been assessed by considering the accuracy of each model. The benchmarks of them include false positive (FP), false negative (FN), true positive (TP) and true negative (TN). A false positive choice occurs when the model detect healthy cell (positive expectation) as abnormal cell and the chosen of false negative occurs when the model marks a negative (abnormal) cell. And a true positive choice occurs when the positive expectation of the classifier correspond correctly with positive forecast and true negative choice also occurs when both the classifier and the division proposes the non attendance of a positive expectation. The accuracy have been characterized as proportion of effectively ordered cells and is equivalent to all TP and TN that were isolated by the aggregate number of RBC's (Elsalamony, 2016).

$$\text{Accuracy} = \frac{TP+TN}{N} \tag{8}$$

### 3. Results and Discussion

Feature extraction was done on SVM classifier using shape and size features, each image produced the features of its geometrical shape with size and for the training and testing the model used 80% to 20% respectively in which the performance of the model is high but comparing to deep learning architecture the absence of feature selection lead to the deformity of the model in order to classify the RBCs data. Our approach on deep learning is based on AlexNet architecture and coloured images were used directly but shows that the performance of the approach is less to the performance of SVM that the model can withhold any amount of data with less pixels. But deep learning need much and clear data for good performance. The challenge on RBCs dataset is they are not in abundance for deep learning approach in order to work with recent state of art classification method. Furthermore the deep neural network learns to perform classification directly from the images while the SVM model trained by specifying the features needed for the process.

Considering achantocyte, teardrop and normal cells from Table 1 is observed that it achieved maximum accuracy using SVM model with colored images this shows that the trained features in the model behaved excellently in the classification while the same images achieved 0% accuracy using deep learning model. This result shows many challenges using deep learning architectures for RBCs image:

- Blood cell images requires certified medical expert. Thus, generating training data on the scale of ImageNet and AlexNet architecture need to be worked on to make the dataset available for research purpose because Small training sets are inevitable in recent state of art (Stoyanov, 2018).
- Deep learning architecture need much generated RBCs dataset in order perform efficiently on blood smear images (Mohanty, Hughes, & Salathé, 2016).

Furthermore considering Table 2 shows that achantocyte, teardrop and normal improved in accuracy which the work suggested an excellent output will be achieved using deep learning if worked on the data collection of RBCs.

Moreover elliptocyte and sickle cell anemia achieved 67% and 83% respectively on Table 1 which shows even the SVM model accuracy improved on the disease accuracy when compared with the increased of dataset on Table 2, while deep learning achieved 25% on elliptocyte and 25% on sickle cell and improved to 33% each respectively.

Table 1 and 2 respectively shows the accuracy difference between SVM and Deep learning models.

**Table 1 – Percentage accuracy.**

Diseases	SVM Model (%)	Deep learning model (%)
Achantocyte	100	0
Elliptocyte	67	25
Sickle cell	83	25
Teardrop	100	0
Normal	100	0



**Table 2 – Percentage accuracy.**

Diseases	SVM	Deep learning
	Model (%)	model (%)
Achantocyte	100	11
Elliptocyte	73	33
Sickle cell	90	33
Teardrop	100	10
Normal	100	10

Deep learning approach unperformed on less data set, these conclude that the dataset of red blood cell is to be generated in order to blend with the state of the art in future because of high population in the world and RBCs abnormality that becomes part of the hematologist challenges due to high demand of the examinations.

#### 4. Conclusion

The paper compared the performance of Support Vector machine classification and deep neural network classification on RBCs using 105 and 250 images in Table 1 and Table 2 respectively, for the experiment and the procedure was done on MATLAB 2017b. This shows that when considering SVM in RBCs images the model performance of these approach depends heavily on the underlying predefined features and it support less dataset while deep learning approach unperformed on less data set these conclude there is a need to generate many RBCs dataset to overcome the deficiency of deep learning on RBCs in future in order to detect RBC'S disease using the state of the art.

#### Acknowledgement

This work is supported by Research Universiti Grant of Universiti Teknologi Malaysia under vot number 14J23.

#### References

- [1] H. A. Aliyu. (2017). Detection of Accurate Segmentation in Blood Cells Count – A Review. *International Journal of Science & Engineering Development Research*, 2(8), 28–32.
- [2] Webster, J.G. and Cazzanti, S.C.(2004). *Bioinstrumentation-Hematology*. John Wiley & Sons, Inc., 170-188.
- [3] Dalvi, P. T., & Vernekar, N. (2016). Computer aided detection of abnormal red blood cells. *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 1741-1746.
- [4] Blood Cells Images, Retrieved from <https://bigpictureeducation.com/blood-cells-images/>
- [5] Sheikh, H., Zhu, B., & Micheli-Tzanakou, E. (1996). Blood cell identification using neural networks. *IEEE Bioengineering Conference, Proceedings of the 1996 IEEE Twenty-Second Annual Northeast*, 119-120.
- [6] Bronkorsta, P. J. H., Reinders, M. J., Hendriks, E. A., Grimbergen, J., Heethaar, R. M., & Brakenhoff, G. J. (2000). On-line detection of red blood cell shape using deformable templates. *Pattern Recognition Letters*, 21(5), 413-424.
- [7] Poomcokrak, J., & Neatpisarnvanit, C. (2008). Red blood cells extraction and counting. *The 3rd International Symposium on Biomedical Engineering*, 199-203.
- [8] Elsalamony, H. A. (2016). Healthy and unhealthy red blood cell detection in human blood smears using neural networks. *Micron*, 83, 32-41.
- [9] Das, D. K., Chakraborty, C., Mitra, B., Maiti, A. K., & Ray, A. K. (2013). Quantitative microscopy approach for shape - based erythrocytes characterization in anaemia. *Journal of microscopy*, 249(2), 136-149.
- [10] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.
- [11] Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.
- [12] Mathworks. (2017). *Deep Learning with Matlab*. Introduction Deep Learning with Matlab.
- [13] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211-252.
- [14] Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L., & Schmidhuber, J. (2011). Flexible, high performance convolutional neural networks for image classification. *Proceedings-International Joint Conference on Artificial Intelligence (IJCAI)*, 22(1), 1237.
- [15] Stoyanov, D. (2018). *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 10553, 178–185.
- [16] Vromen, J., & McCane, B. (2006). Red blood cell segmentation using guided contour tracing. Presented at SIRC 2006 - The 18th Annual Colloquium of the Spatial Information Research Centre University of Otago, Dunedin,



New Zealand.

- [17] Ghosh, N., Buddhiwant, P., Uppal, A., Majumder, S. K., Patel, H. S., & Gupta, P. K. (2006). Simultaneous determination of size and refractive index of red blood cells by light scattering measurements. *Applied physics letters*, 88(8), 084101.
- [18] Cortez, P., Rio, M., Rocha, M., & Sousa, P. (2012). Multi - scale Internet traffic forecasting using neural networks and time series methods. *Expert Systems*, 29(2), 143-155.