A deep learning AlexNet model for classification of red blood cells in sickle cell anemia

Hajara Abdelkarim Aliyu[1], Mohd Azhar Abdul Razak[2], Rubita Sudirman[3], Norhafizah Ramli[4]
[1,2,3,4]School of Electrical Engineering Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia
[1]Jigawa State Polytechnic Dutse, Jigawa State, Nigeria

Article Info

ABSTRACT

Sickle cell anemia (SCA) is a serious hematological disorder, where affected patients are frequently hospitalized throughout a lifetime and even can cause death. The manual method of detecting and classifying abnormal cells of SCA patient blood film through a microscope is time-consuming, tedious, prone to error, and require a trained hematologist. The affected patient has many cell shapes that show important biomechanical characteristics. Hence, having an effective way of classifying the abnormalities present in the SCA disease will give a better insight into managing the concerned patient's life. This work proposed algorithm in two-phase firstly, automation of red blood cells (RBCs) extraction to identify the RBC region of interest (ROI) from the patient's blood smear image. Secondly, deep learning AlexNet model is employed to classify and predict the abnormalities presence in SCA patients. The study was performed with (over 9,000 single RBC images) taken from 130 SCA patient each class having 750 cells. To develop a shape factor quantification and general multiscale shape analysis. We reveal that the proposed framework can classify 15 types of RBC shapes including normal in an automated manner with a deep AlexNet transfer learning model. The cell's name classification prediction accuracy, sensitivity, specificity, and precision of 95.92%, 77%, 98.82%, and 90% were achieved, respectively.

Corresponding Author:

Mohd Azhar Abdul Razak,
Department of Electrical and Computer Engineering,
Universiti Teknologi Malaysia,
81310 UTM Johor Bahru, Malaysia.
Email: mohdazhar@utm.my

1. INTRODUCTION

Sickle cell anemia (SCA) is a type of inherited RBC disorder connected with abnormal hemoglobin Sickle (HbS). When HbS molecules polymerize into RBCs based on lack of oxygen, they greatly affect the adhesion, shape, and size properties of the RBCs. The red blood cells become easily damaged with a wide proportion of diverse shapes in the cell population [1], which makes this problem a perfect prospect for the examination of morphological heterogeneity. RBCs are a biconcave shape. All the sides of the cell surface curve inward like the interior of a sphere. This shape gives RBCs the ability to move through tiny blood vessels to deliver oxygen to the organs and tissues. RBCs are also a key factor in determining human blood type. Blood type is decisive by the presence or absence of certain identifiers on the surface of RBCs. The diameter of red blood cells varies from 7 to 8 microns [2]. Each red blood cell has around 280 million of hemoglobin molecules, and also the lifetime of each healthy person RBCs cell is 100 to 120 days. Then RBCs get destroyed in the spleen while for the SCA patient last for 10 to 20 days [3]. Hence, SCA is

affected by several risks of life-threatening complications such as organ damage over time, stroke, and high mortality rate.

Based on the study conducted in 2013, about 3.2 million people are SCA patients, while 43 million are the carrier of SCA, and the mortality rate rises to 176,000 in 2013, mostly African origin. The study estimates that Nigeria and the Democratic Republic of the Congo will have to increase the birth of SCA in 2015 to 140,800 and 44,700, respectively [4]. The main key challenge of SCA is the variability in its clinical severity level from one patient to another. Available methods for treating SCA are mostly supportive and focus on symptom control but lack disease prediction in different clinical stages [5]. Recent progress in machine learning and computer image processing techniques could provide an effective method in monitoring the status of SCA patients.

Many researchers work on SCA today to detect the disease but not considering the named classification RBCs shapes present in SCA patient. Das *et al.*[6] Proposed a method that used machine learning techniques to characterized RBCs in anemia based on blood smear images, and 25 sets of features were used for identification but observed that using a small set of a feature will present higher accuracy than they achieved. Xu *et al.*[7] used deep convolution neural network to achieved shape classification framework, which achieved high accuracy but the work drawback is that not all abnormalities in SCA patients are captured based on less number of the patient. Alom *et al.*[8] used a deep learning approach known as the Inception Recurrent Residual Convolutional Neural Network (IRRCNN) model for RBC and WBC classification to classify the white blood cells into five classes and red blood cell into nine class of abnormality with a single normal class. Syed *et al.* segment and classify RBCs based on fuzzy C-mean segmentation and classify the cells using the Extreme Learning Machine (ELM) and Support Vector Machine (SVM) approach using different texture and geometric features extracted from the cells [9]. Aliyu *et al.* compare SVM and deep learning AlexNet models to classify a normal and four abnormal cells which observed SVM performs better than AlexNet due to lack of RBCs sample and recommends data collection for RBCs disease detection for the straight forward AlexNet model to detect the abnormalities cell of SCA patients [10]. Aliyu *et al.*[11] used form factor and perimeter features as descriptors to detect normal and four abnormal cells, but SCA patients have abnormalities that are more than the ones predicted in the study. Tajkia *et al.*[12] compare K- Nearest Neighbour (KNN), SVM, and ELM classifiers for sickle cell detection, which shows that ELM performs better in terms of accuracy, but considering the learning rate, SVM and KNN outperform ELM classifier. Alzubaidi *et al.*[13] Proposed a method to classify sickle cell anemia using convolution neural network and error-correcting output as a classifier to classify the cells into normal, abnormals, and miscellaneous cells and the method achieved 92.06% accuracy. In this paper, we automatically crop normal and 14 abnormal RBCs present in 130 SCA patients and classify them using deep learning AlexNet model.

2. RESEARCH METHOD

This paper presents a framework that includes image acquisition, pre-processing image phase, automatic RBC blotch extraction, and the trained AlexNet model.

2.1. SCA blood cells image acquisition

The blood specimens were obtained from 130 different patients in Murtala Muhammad Specialist Hospital (MMSH) and Aminu Kano Teaching Hospital (AKTH) Nigeria with approved ethic numbers of MOH/off/797/T.1/849 and NHREC/21/08/2008/AKTH/EC/2307 respectively. For the preparation of a stained blood smear, a drop of blood was dripped on a glass slide at an angle of 25°. The blood smear was air-dried and stained with Giemsa. The stained slides of RBCs were then captured using a microscope (Hundwetzler H600) that connected with a 3M pixel Amscope camera (FMA050) using a ×40 magnification objective lens. Three images at a different location were captured on each patient blood smear slide and was automatically cropped to obtain 750 single RBCs images (over 9000 single RBCs).

A computer with a specification of Intel core i7, NVIDIA Geforce MX1502GB, 1TB 7200rpm SATA, 128 GB SSD, 8GB DDR4 was used. The study used MATLAB R2017b software to run the image processing and machine learning task.

2.2. Image pre-processing

The pre-processing process is needed to remove unwanted noise for further processing. The steps of pre-processing phase are given as follows:
1. Image Greyscale, and Thresholding

An input image resolution of the cells 1024×768 pixels is converted into a greyscale that has different grey shades. Thresholding is a global segmentation method applied to separate objects from the

background. Otsu method is the well-known global image threshold [14]. The threshold image contains small unwanted particles like debris and noise due to the nature of the dataset [15-17]. A binary image depends mainly on the greyscale image to extract the foreground and background. Each pixel of an image is replaced with value zero and one if intensity value is lower and higher than the threshold, respectively. The threshold value is set to 0.4 based on the debris present in the dataset. Then, the border-image is eliminated to clear the incomplete cells. The incomplete cells that are located at the border were removed. The image border clearance used morphological operations for the reconstruction of the image using dilation and erosion process [18]. Equation (1) is used for thresholding.

$$t^* = \arg max_t \delta_b^2(\text{t})$$ (1)

t^* is the optimal threshold that maximizes the functions δ_b^2 which is based on second-order statistics.

2. Hole Filling and Filtration
The hole filling process considers debris or noise that is less than 150 pixels to be discarded from the blood smear image and leave only the RBCs. Equation (2) suppresses the pixels that do not belong to RBCs and will lead to cropped the cells automatically into a folder to justify correct pixel labeling.

$$ID = \text{Area-}\mu > \text{B}\sigma$$ (2)

Where *ID* refers to a filter, μ is a mean, σ is a standard deviation (SD) of the pixel value that belongs to RBCs, and B is a constant threshold value. The *ID* is the filtered location of a single blotch image that computed in this study to capture the exact location of each cell on a blood smear image. The value of 0.4 is the determined threshold value based on the dataset images.

2.3. RBCs blotch extraction and trained model

After the preprocessing stage, the image was automatically cropped to obtain 750 single RBC images. This was done using the RBCs blotch extraction technique based on features such as eccentricity, area, and bounding box. The images were then used in training deep learning AlexNet model.

1. Eccentricity (E) - The eccentricity of an image is the ratio of the minor axis to the major axis of the detected cell [19]. The value of the eccentricity value is between 0 and 1. In this study, a cell is referred to as a circular, if E ≤ 0.6, and if E > 0.6, the cell is referred as other shapes or abnormal cells. Equation (3) defines the eccentricity.

$$E = \frac{\text{major axis} - \text{minor axis}}{\text{major axis}}$$ (3)

2. Area (A) - This is the number of pixels enclosed inside the cell boundary, Equation (4-5) are used to calculate the area of the circular and not circular cells, respectively.

$$A_{circle} = \pi r^2$$ (4)

where r is the radius of the circle cell.

$$A_{not\ circle} = \pi r_1 r_2$$ (5)

where r_1, r_2 referred to the radius of the non- circle cell for the minor and major axis, respectively.

3. Bounding Box – This is a rectangular box containing the Region of Interest (RoI) of the single RBC image. The bounding box kept the coordinates of the rectangular border that enclosed the cell image[20].

4. AlexNet Transfer Deep Learning Model - The AlexNet comprises of five convolution layers and three fully connected layers with a SoftMax layer as the final layer. Those same convolution layers 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) [21]. Equation (6) determines the ReLU. Figure 1 is the sample dataset for the 15 cell classes after the automatic RBCs blotch image achieved. The single images are sub-divided into training folders of normal and abnormal cells with 1 and 14 categories, respectively. The abnormal cells are the cells that cause anemia diseases [22].
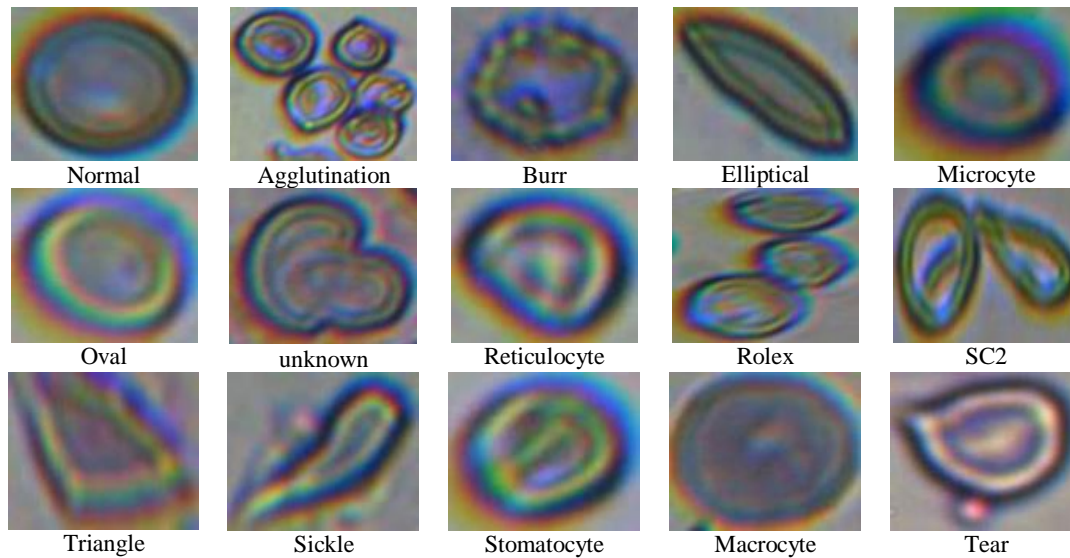
$$(z_i) = \max(0, z_i)$$ (6)

Figure 1. Dataset of Normal and Abnormal of RBC for Classification

5. Accuracy, Sensitivity, Specificity, and Precision Percentage

The accuracy, sensitivity, specificity, and precision for sickle cell name classification prediction are determined using (7-10), respectively. The proportion of true-false, positive-negative, was carried out. The accuracy of the named cell detection system is measured in the percentage of true-positive (TP), false positive (FP), true negative (TN), and false-negative (FN) [23-24].

where,

TP - Number of named cells correctly identified as an abnormal cell on the blood smear image
TN - Number of named cells correctly identified as a normal cell on the blood smear image
FP - Number of named cells incorrectly identified as an abnormal cell on the blood smear image
FN - Number of named cells incorrectly identified as a normal cell on the blood smear image

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{8}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{9}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{10}$$

3. RESULTS ANALYSIS AND DISCUSSION

1. Image sample and conversion - Image samples in Figure 2 are the acquired image from SCA patients. These input images are then undergoing resizing, and greyscale image conversion process and the outcomes are shown in Figure 3.
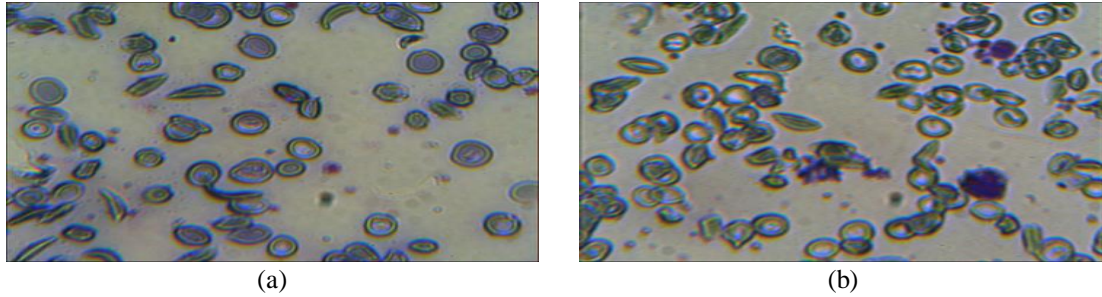
(a)   (b)

Figure 2. Samples of acquired sickle cell blood smear images
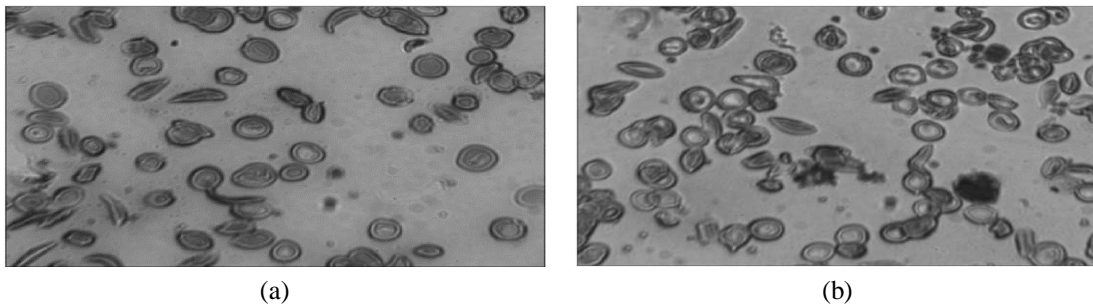

(a)   (b)

Figure 3. Greyscale Image

2.   Otsu Thresholding and Border Clearance - Otsu threshold was applied to the greyscale image to obtain the binary image. The results are shown in Figure 4.
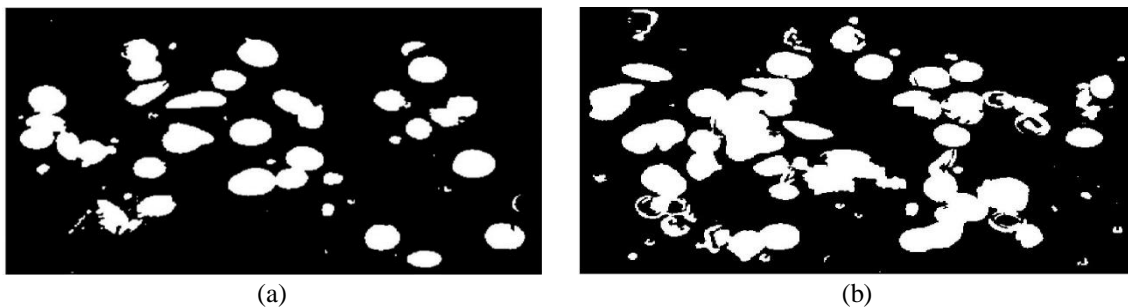

(a)   (b)

Figure 4. Otsu thresholding and border clearance

3.   Hole Filling and Filtration- Holes inside the cells were filled to complete the cell shape. The images captured contains many debris and noise. Hence filter is needed to remove them. Figures 5(a) and 5(b) are the images that undergo hole filling and filtration processes, respectively.

After the completion of the pre-processing stage on the images, the analysis was continued with differentiate the normal and abnormal cells based on the eccentricity value the abnormal cells were manually grouped into 14 classes. The deep learning AlexNet model was trained using the RBCs classes with some changes in the mini-batch size of 64, 0.9 monumentum, and weight decay of 0.0001. The training process observed that a small amount of weight decay improves the model training performance as it reduces the training error rate. The model was trained using 650 cells for each class and the detected process were tested on 100 images. The cell classification prediction presented in Figure 6.

<center>(a)   (b)</center>
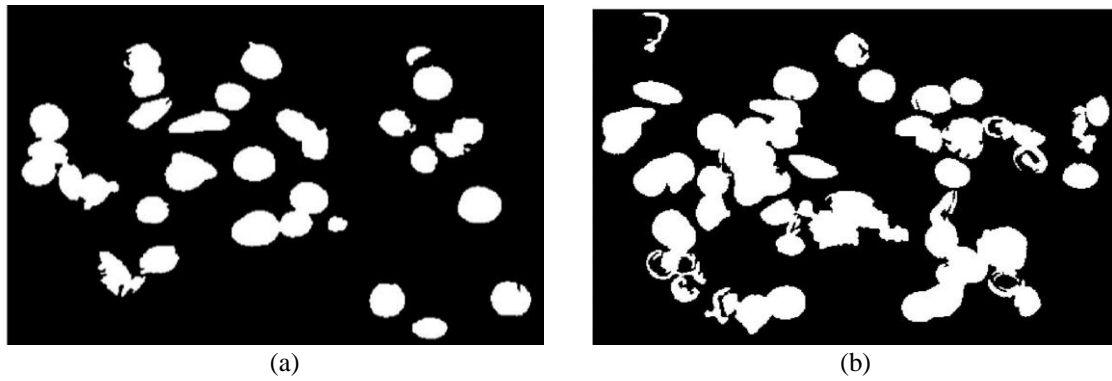
Figure 5. Hole filling and filtration

Based on the cell detected in the blood smear image of the patients in Figure 6, the accuracy, sensitivity, specificity, and precision were calculated. Table 1 illustrates the result of patients analyzed by the system. The system name classification prediction performance on the accuracy, sensitivity, specificity, and precision were 95.92%, 77%, 98.82%, and 90.1%, respectively.

<center>(a)</center>
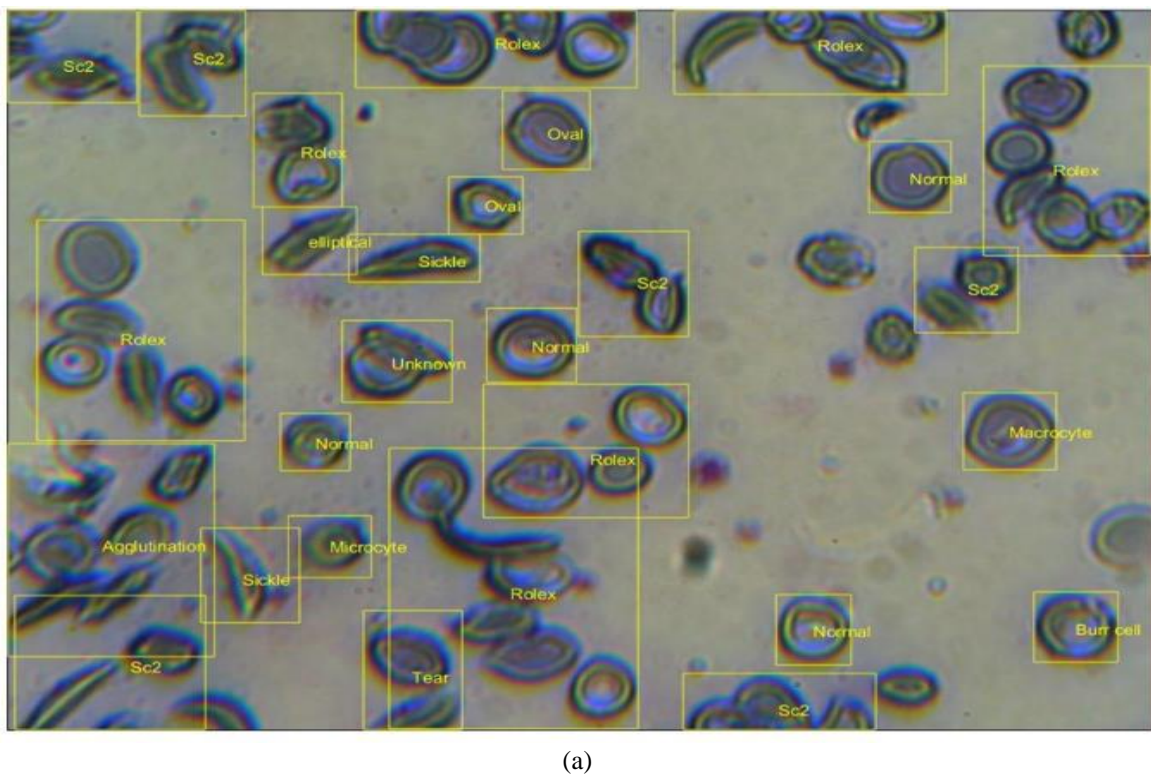
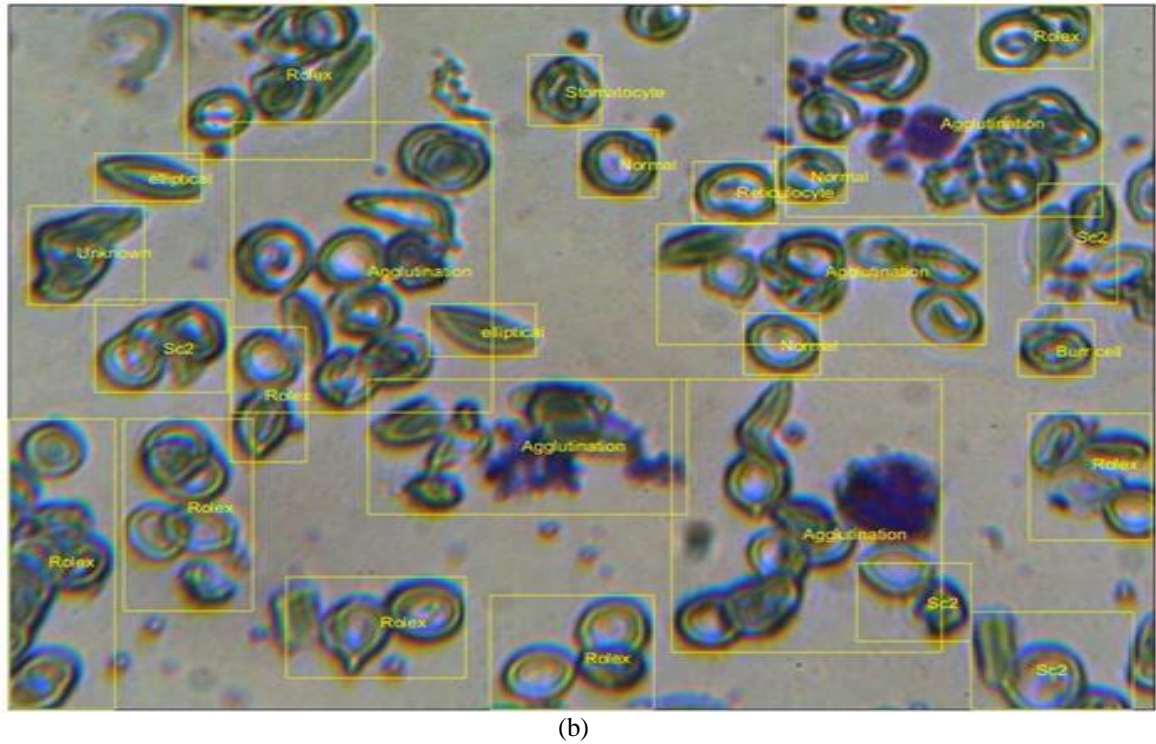Figure 6. Deep AlexNet transfer learning detection result *(continue)*

(b)

Figure 6. Deep AlexNet transfer learning detection result

Table 1. Computation of TP, FP, TN, and FN

Image No	Total number of Predicted cells on blood smear images	Cells correctly identified as a normal name on the blood smear image (TP)	Cells incorrectly identified as a normal name on the blood smear image (FP)	Cells correctly identified as an abnormal name on the blood smear image (TN)	Cells incorrectly identified as an abnormal name on the blood smear image (FN)
(a)	28	3	1	24	0
(b)	26	2	0	24	0
(c)	23	5	0	18	0
(d)	21	0	0	18	3
Total	98	10	1	84	3

4. CONCLUSION

The framework predicts and classifies the normal and 14 abnormalities of RBCs based on the longitudinal study of the abnormalities present in a sickle cell patient. The RBC of the patients shows several abnormalities in which the framework that used the AlexNet model is robust and can even detect other type of cells such as thallasaemia and hereditary elliptocytosis. The AlexNet model suits the framework because it can classify as many as 1000 class of abnormalities [25]. The system achieved considerable high accuracy and precision. The specificity has lower percentage due to less number of normal cells present in the blood smear images of the SCA patients. The work concludes that hematologist could use this framework to investigate sickle cell anemia patients that offer a fast and affordable process of patient health management.

REFERENCES

[1] R. M. Fasano *et al.*, "Red blood cell alloimmunization is influenced by recipient inflammatory state at time of transfusion in patients with sickle cell disease," *Br. J. Hematol.*, no. 7, pp. 291–300, 2015.
[2] V. Acharya and P. Kumar, "Identification and red blood cell automated counting from blood smear images using

computer-aided system," *Springer Med Biol Eng Comput*, vol. 56, pp. 483–489, 2018.

[3] H. A. Elsalamony, "Healthy and unhealthy red blood cell detection in human blood smears using neural networks," *Micron*, vol. 83, no. 4, pp. 32–41, 2016.

[4] S. I. Hay, S. Gupta, D. J. Weatherall, and T. N. Williams, "Global Burden of Sickle Cell Anaemia in Children under Five , 2010 – 2050 : Modelling Based on Demographics Excess Mortality , and Interventions," *PLosmedicine*, vol. 10, no. 7, 2013.

[5] J. N. Milton, V. R. Gordeuk, J. G. Vi, T. Mark, M. H. Steinberg, and P. Sebastiani, "Prediction of Fetal Haemoglobin in Sickle Cell Anemia Using an Ensemble of Genetic Risk Prediction Models," *Circ Cardiovasc Genet.*, vol. 7, no. 2, pp. 110–115, 2015.

[6] D. K. Das, C. Chakraborty, B. Mitra, A. K. Maiti, and A. K. Ray, "Quantitative microscopy approach for shape-based erythrocytes characterization in anaemia," *J. Microsc.*, vol. 249, no. 2, pp. 136–149, 2013.

[7] M. Xu, D. P. Papageorgiou, S. Z. Abidi, M. Dao, H. Zhao, and G. E. Karniadakis, "A deep convolutional neural network for classification of red blood cells in sickle cell anemia," *Plos Comput. Biol.*, pp. 1–27, 2017.

[8] M. Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, "Microscopic Blood Cell Classification Using Inception Recurrent Residual Convolutional Neural Networks," *Proc. IEEE Natl. Aerosp. Electron. Conf. NAECON*, vol. 2018-July, pp. 222–227, 2018.

[9] S. H. Shirazi, A. I. Umar, N. U. Haq, S. Naz, M. I. Razzak, and A. Zaib, "Extreme learning machine based microscopic red blood cells classification," *Cluster Computing*. pp. 1–11, 2017.

[10] H. A. Aliyu, R. Sudirman, M. Azhar, A. Razak, M. Amin, and A. Wahab, "Red Blood Cells Abnormality Classification : Deep Learning Architecture versus Support Vector Machine," *Int. Journal of Integr. Eng.*, vol. 7, pp. 34–42, 2018.

[11] Abdulkarim Aliyu Hajara, Azhar Mohd, and Sudirman Rubita, "Normal and abnormal red blood cell recognition using image processing," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 14, no. 1, pp. 96–100, 2019.

[12] T. S. Chy, "A C omparative A nalysis by KNN , SVM & ELM C lassification to D etect S ickle C ell A nemia," *Int. Conf. Robot. Signal Process. Tech.*, pp. 455–459, 2019.

[13] L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, and J. Zhang, *Classification of red blood cells in sickle cell anemia using deep convolutional neural network*, vol. 940. Springer International Publishing, 2020.

[14] H. J. Vala and P. A. Baxi, "A Review on Otsu Image Segmentation Algorihm," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 2, no. 2, pp. 387–389, 2013.

[15] H. A. Aliyu, M. A. A. Razak, and R. Sudirman, "Segmentation and detection of sickle cell red blood image," *AIP Conf. Proc.*, vol. 2173, no. November, 2019.

[16] I. Nigam, S. Agrawal, R. Singh, and S. Member, "Revisiting HEp-2 Cell Image Classification," *IEEE Access*, vol. 3, pp. 3102–3113, 2016.

[17] M. Maity, T. Mungle, D. Dhane, A. K. Maiti, and C. Chakraborty, "An Ensemble Rule Learning Approach for Automated Morphological Classification of Erythrocytes," *J. Med. Syst.*, vol. 41, no. 4, 2017.

[18] J. M. Sharif, M. F. Miswan, M. Ngadi, S. Hj, and M. Mahadi, "Red Blood Cell Segmentation Using Masking and Watershed Algorithm : A Preliminary Study," *Int. Conf. Biomed. Eng.*, no.2, pp. 27–28, 2012.

[19] C. Y. Kit *et al.*, "Mobile based Automated Complete Blood Count (Auto-CBC) Analysis System from Blood Smeared Image," *Int. J. Electr. Comput. Eng.*, vol. 7, no. 6, p. 3020, 2017.

[20] I. H. Witten, E. Frank, and M. A. Hall, *Practical Machine Learning Fourth Edition*,. 2017.

[21] X. Han, Y. Zhong, L. Cao, and L. Zhang, "Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification," *Remote Sens.*, vol. 9, no. 8, 2017.

[22] P. Maji, A. Mandal, M. Ganguly, and S. Saha, "An automated method for counting and characterizing red blood cells using mathematical morphology," *Eight Int. Conf. Adv. Pattern Recognit.*, pp. 1–6, 2015.

[23] P. T. Mohammed Khalafa, Abir Jaafar Hussaina, Robert Keighta, Dhiya Al-Jumeilya, Paul Fergusa,and Russell Keenanb, "Machine learning approaches to the application of disease modifying therapy for sickle cell using classification models.pdf," *Neurocomputing*, pp. 154–164, 2016.

[24] C. L. Chen, A. Mahjoubfar, L. Tai, I. K. Blaby, and A. Huang, "Deep Learning in Label-free Cell Classification," *Nat. Publ. Gr.*, no. January, pp. 1–16, 2016.

[25] A. K. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *J. Geotech. Geoenvironmental Eng.*, vol. 12, pp. 18 - 26 , 2012.