revista principia
Vol. 60, N. 4
REVISTA DE DIVULGAÇÃO CIENTÍFICA E TECNOLÓGICA DO INSTITUTO FEDERAL DE CIÊNCIA, EDUCAÇÃO E TECNOLOGIA DA PARAÍBA

DOI: http://dx.doi.org/10.18265/1517-0306a2021id6456
ORIGINAL ARTICLE

SUBMITTED November 5, 2021
APPROVED April 25, 2022
PUBLISHED ONLINE May 18, 2022
PUBLISHED October 10, 2023
ASSOCIATE EDITOR
Pedro Bertemes Filho

Detection and classification of abnormal red blood cells with computational intelligence techniques: a review

Fernanda D'Amico Silva [1]*
Evair Borges Severo [2]
Victor Henrique Alves Ribeiro [3]
Gilberto Reynoso-Meza [4]

[1] fernanda.damico@pucpr.edu.br
[4] g.reynosomeza@pucpr.br
Pontifícia Universidade Católica do Paraná (PUCPR), Brazil

[2] evair.severo@hitechnologies.com.br
Artificial Intelligence Team, Hilab, Brazil

[3] vhrique@gmail.com
Silex, Brazil

*Corresponding author.

ABSTRACT: Red blood cells (RBCs), or erythrocytes, are usually disc-shaped. However, pathological conditions can change their shape. The complete blood count (CBC) is a test that can detect abnormal RBCs. Yet it is a manual test susceptible to errors, so there are efforts to automate the detection and classification of abnormal cells. A total of 31 papers were reviewed, and all the selected studies focus on the detection or classification of abnormal RBCs. Different approaches were applied to tackle this issue, including image processing techniques, classification using machine learning, and convolutional neural networks for detection, among other methods. Furthermore, machine learning techniques are recently presenting promising results for abnormal RBCs detection and classification. This review also brings a discussion on the computational intelligence methodologies applied.

Keywords: abnormal cells classification; computational intelligence; machine learning; red blood cells.

Detecção e classificação de hemácias anormais com técnicas de inteligência computacional: uma revisão

RESUMO: *Células vermelhas ou eritrócitos, possuem normalmente um formato de disco. Entretanto, condições patológicas podem alterar seu formato. O hemograma completo é um exame que consegue detectar células anômalas, entretanto é um teste manual e suscetível a erros, assim existem esforços para automatizar a detecção e classificação de células anômalas. Um total de 31 artigos foram revisados, todos os estudos selecionados focaram na detecção ou classificação de eritrócitos anômalos. Diferentes abordagens foram aplicadas para atacar o problema, incluindo: técnicas de processamento de imagens, classificação utilizando aprendizado de máquina, redes neurais convolucionais para detecção, entre outros métodos. Ademais, técnicas com aprendizado de máquina vêm apresentando resultados promissores para detecção e*

CC BY NC SA
ISSN (on-line): 2447-9187

classificação de células vermelhas anormais. A revisão também traz uma discussão acerca das metodologias de inteligência computacional utilizadas.

Palavras-chave: *aprendizado de máquina; classificação de células anômalas; eritrócitos; inteligência computacional.*

1 Introduction

Red blood cells (RBCs), or erythrocytes, are responsible for the transport of oxygen and nutrients throughout the blood. All healthy RBCs are disc-shaped; however, pathological conditions can cause abnormalities in cells. These changes in the surface of the cell can compromise, or even obstruct, blood circulation (PARK *et al*., 2010).

A variety of pathological conditions cause a change in the shape of cells, such as hereditary disorders or even parasitic infections. The most common is anemia, a condition that affects a quarter of the global population (BALARAJAN *et al*., 2011).

With a complete blood count (CBC) much information can be extracted from blood, including the number, shape, and size of cells. Nowadays most blood counts are performed manually by specialists with a blood smear preparation using a microscope. It is a cheap method that requires attention, and it is susceptible to errors (ACHARJEE *et al*., 2016; LOU *et al*., 2016).

To improve the reliability of blood smear tests, there have been several studies in the literature to automate the detection and analysis of RBCs and white blood cells (WBCs) (HABIBZADEH; KRZYZAK; FEVENS, 2013; SARRAFZADEH *et al*., 2015; ZHAO *et al*., 2017).

This review will focus on RBCs abnormalities in shape and size and on the efforts in the literature to provide automation solutions to identify these cells.

Several methods were applied in RBCs images to detect and classify abnormal cells. Image processing was applied for feature extraction, and machine learning techniques were used to classify extracted features (ALAM; ISLAM, 2019; PARAB; MEHENDALE, 2021; SAFCA *et al*., 2018).

Most recently, convolutional neural networks (CNNs) were applied to automatically extract features from images, and artificial neural networks (ANNs) were used to classify manually extracted features (KAJÁNEK; CIMRÁK, 2019; MOLINA-CABELLO *et al*., 2018).

Machine learning (ML) performs automatic pattern recognition and inferences with data. With it, a computer program can learn from experience while performing tasks and evaluating performance. As the program executes more tasks and acquires more experience, the performance improves (MITCHELL, 1997).

Deep learning, which encompasses CNNs, is a subfield of ML that can learn complex assumptions from massive and complex datasets (WALTER *et al*., 2021).

Deep learning is solving problems in several applications such as science, business, and government with image recognition, speech recognition, and other activities. And still, it has a lot of potential due to its need for little engineering performed by its developers, so it can take advantage of great amounts of data (LECUN; BENGIO; HINTON, 2015).

The objective of this review is to understand the approaches used in the literature to tackle the automatic classification and detection of abnormal RBCs and to map gaps in the reviewed approaches.

This paper is organized as follows: in section 2 there is the problem description, the methods used for the literature review, and the results from 31 studies; then in section 3, there is a discussion of the results that were found during the review process; finally, in section 4 there is a conclusion.

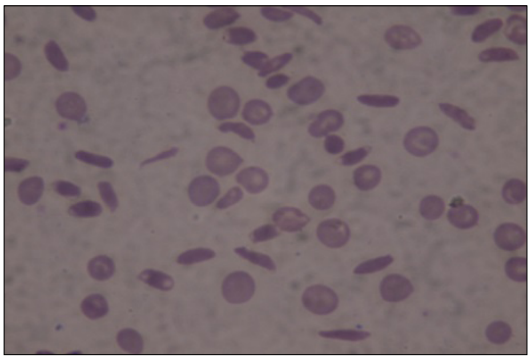
2 Literature review

In this section, the problem tackled in this review will be explained. Then, the methods used to select the studies reviewed will be presented. Finally, in section 2.3, the 31 analyzed studies will be presented.

2.1 Problem description

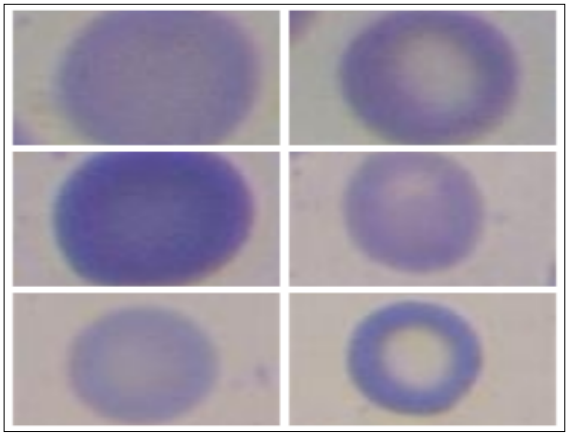
The abnormal RBCs identification approaches addressed in the literature can be divided into two procedures, detection, and classification. Detection takes place in images with several cells, as in Figure 1. In this situation, it is necessary to separate each cell in the image to classify it later.

Figure 1 ▶
Image from a blood smear test with several blood cells.
Source: González-Hidalgo et al. (2014)



The classification occurs when the dataset used in the study contains RBCs already separated and cropped as in Figure 2.

Figure 2 ▶
Individual RBCs.
Source: González-Hidalgo et al. (2014)



The majority of the studies examined reside on multiclass problems. Each of the abnormalities found in the CBC is considered a different class. An example of different types of abnormalities found in RBCs is in Figure 3.

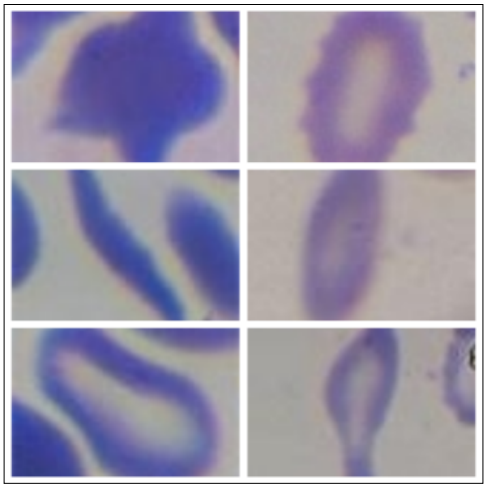


Figure 3 ▶ Examples of abnormalities found in RBCs. *Source: González-Hidalgo et al. (2014)*

2.2 Methods for literature review

Several databases were used in the papers, but only seven studies used public datasets. Among the public databases, there are the erythrocytes IDB[1] and ALL-IDB[2]. Twenty-two examined papers used private datasets and two studies did not disclose the datasets used.

[1] Available in: http://erythrocytesidb.uib.es/

[2] Available in: https://homes.di.unimi.it/scotti/all/

The criteria for choosing the papers analyzed in this review consist of papers related to the detection and classification of abnormalities in RBC. The keywords used to find the appropriate papers were “abnormal red blood cells” combined with “detection abnormal RBC” and “abnormal red blood cells” combined with “anemia detection neural network”. Also “abnormal erythrocyte” combined with “abnormal detection”. Originally 36 papers were selected, however, studies that focused on detecting and classifying normal cells were not considered. Papers that concentrated on WBCs were also disregarded, as for papers that focused on malaria and cell count. Other criteria used to discard articles were works that did not include measurable results. The papers examined are from the period comprehended between 2013 and 2021, to deal with the most up-to-date proposed solutions. In total, 31 articles met the established criteria.

2.3 Results

Distinct techniques have been used in the literature to tackle the issue of detecting and classifying anomalous blood cells, most of them consisting of image processing and machine learning applications.

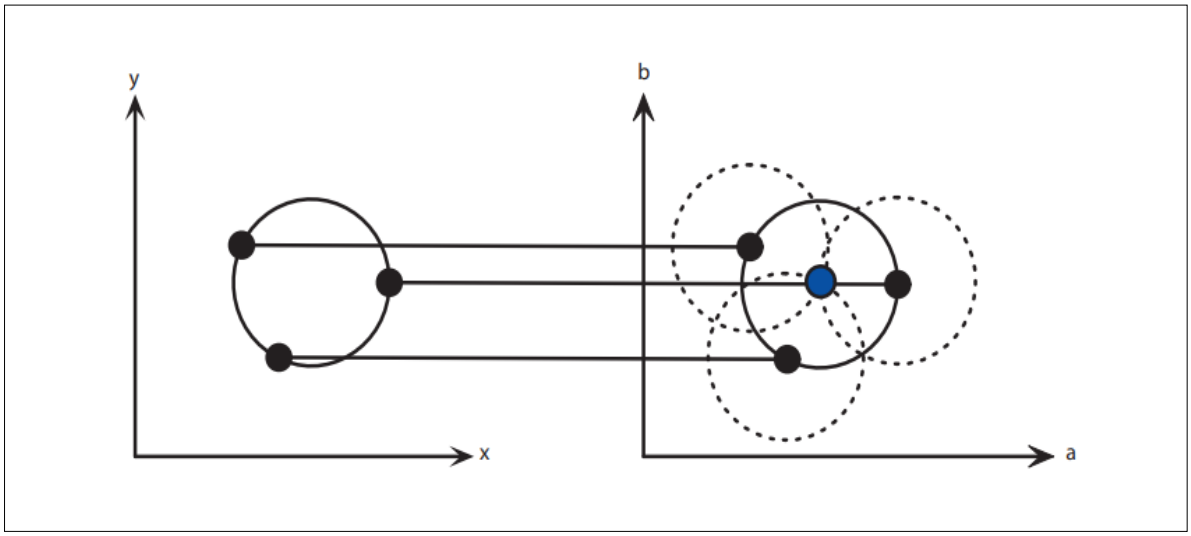
In Alzubaidi *et al*. (2020) the main objective was to classify RBCs in microscopy images to diagnose sickle cell anemia (SCA). Same domain transfer learning was used to tackle the lack of training data, transfer learning is a technique where deep learning models are trained on large datasets and then tuned to be applied in the smaller datasets. In the paper, transfer learning was applied to three different datasets

combined to perfect the CNN model. The authors addressed the binary classification, which is very common in SCA papers, by including a third class called 'other blood content' to prevent faulty classification. Also, data augmentation was applied to minimize the overfitting effect. The data was extracted from several public datasets, including the erythrocytes database from González-Hidalgo *et al*. (2014), the Wadsworth Center dataset, and several other images found across the internet. The model proposed by the authors obtained an accuracy of 99.54% and 99.98% with the model plus a multi-class support vector machine (SVM) classifier. An accuracy of 98.87% was obtained in the erythrocytes database.

Chy and Rahaman (2019) proposed a comparative analysis of different methods to classify sickle cells. The authors applied image pre-processing techniques to identify the cells using a gray image, noise filtering, enhancement of image, and Fuzzy C means clustering for segmentation. For feature extraction, the geometrical and statistical features were used. Finally, three techniques were applied and compared for classification, K-nearest neighbors (KNN), SVM, and Extreme Learning Machine (ELM). The database used is private and consists of 80 images. From the dataset, the authors used 50 images for training and 30 for testing. The authors obtained accuracies of 73.33% using KNN, 83.33% with SVM, and 87.73% using ELM. The F1 score results for KNN, SVM, and ELM were 81.81%, 88.89%, and 91.30%, respectively.

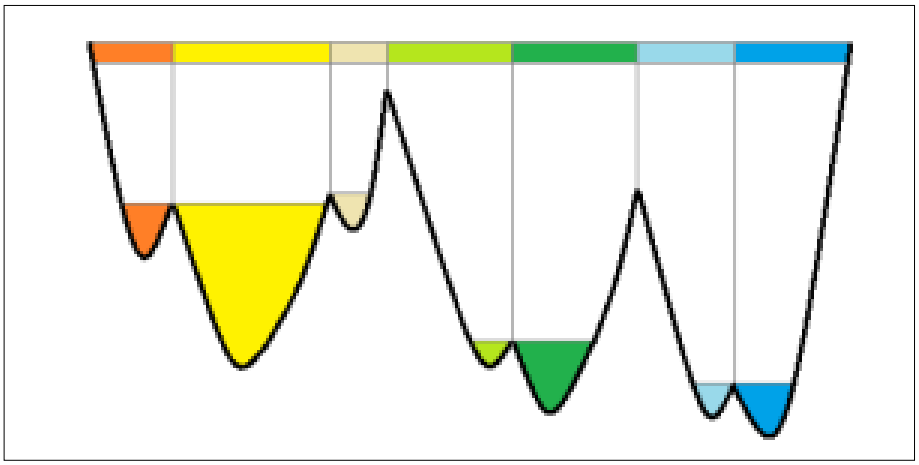
In Fadhel, Humaidi and Oleiwi (2017) the main objective was detecting and counting normal and sickle cells. Circle Hough transform (CHT) and Watershed Transform (WT) were used and compared. CHT is a technique used to detect circles by applying an accumulator array computation then, after calculating the peaks in the array computed, there is a center estimation and a radius estimation, the representation of CHT is represented in Figure 4.

Figure 4 ▼
CHT detects the edges (left) and computes the center and radius estimation (right).
Source: Pedersen (2007)



WT is often used when objects are overlapping, it consists of a method that detects ridges between adjacent local minima and considers the whole gray image as a surface, then, the dark pixels are considered low and the light pixels are high so the “watershed ridge lines” are found and the objects are separated, as demonstrated in Figure 5. The authors concluded that CHT had a better counting efficacy and was more robust than WT.

Figure 5 ▶
Application of watershed transform in 1-D.
Source: Bai and Urtasun (2017, p. 5223)



González-Hidalgo *et al*. (2014) proposed a method for detecting elongated erythrocytes and regular cells. The method used in the paper applies ellipse adjustments to detect cells efficiently. Another method for efficient detection of convex and concave points of interest in a contour was also used. The images were not preprocessed. The dataset used is open to the scientific community, it is called erythrocytesIDB. The results obtained in the classification were 100% for normal cells and 98% for sickle cells.

Xu *et al*. (2017) suggested the use of CNNs for the classification of RBCs in SCA. The pipeline used in the research consists of patch extraction of each cell in the image, then the patch was normalized and labeled. Finally, a CNN model was applied. The cells were classified into five categories, discocytes, granular, elongated, oval, or sickle. The detection of RBCs was performed by splitting the image into blocks, calculating information entropy for each block, creating a region of interest (ROI) mask, then detecting ROI and identifying the boundaries of ROI, finally, the cells were split into single patches.

The applied CNN architecture consists of 10 layers including 3 convolutional, 3 dropout layers, and a fully connected layer. Rectified Linear Unit (ReLU) was applied as an activation function. Finally, logistic regression combining the softmax function with a cross-entropy loss function was implemented to obtain the learning probability and to predict labels. In addition to the classification, the authors also perform a shape factor analysis for each cell, the shape parameters were extracted from the contour analysis. The database used in the research is private and, it was obtained from 434 microscopic images of 8 different patients; to increase the dataset, data augmentation was applied in each of the patches, rotating in 90º, 180º, 270º, and flipping the image horizontally and vertically. 5-fold cross-validation was carried out to evaluate the performance of the CNN and the mean evaluation accuracy obtained was 87.50%.

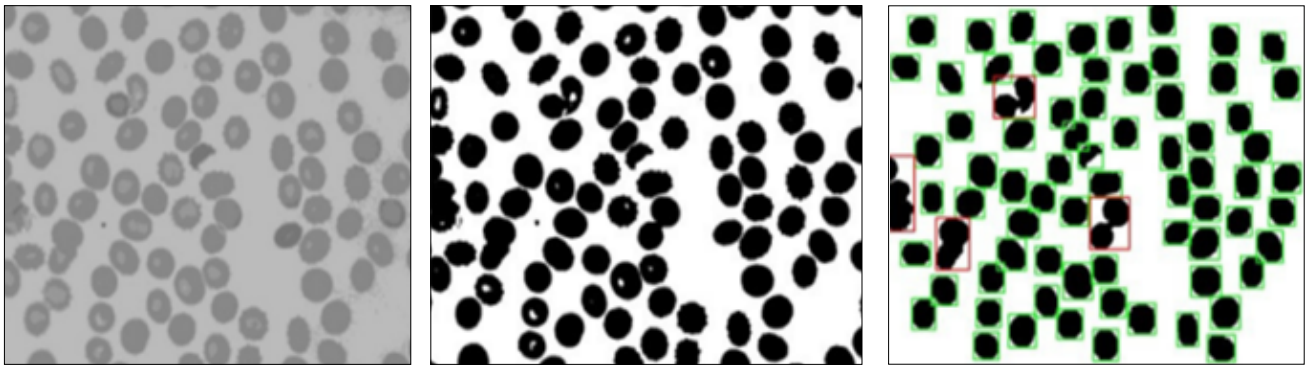
The main objective of Acharya and Kumar (2017) was to identify sickle cells in a CBC. The dataset used in the study is from blood smear images collected from Kasturba Medical College, Manipal, Karnataka, and Atlas of Hematology. The experiment was performed with 1000 images. Firstly, the authors addressed the issue of separating the RBCs from the WBCs. The K-Medoids algorithm was used to segment the image. The image was segmented into four colors where the intensity of cells was used to separate WBCs from RBCs among morphological operations and watershed transform to separate touching objects. Then, the image was converted to grayscale and converted to binary using Otsu thresholding. Erosion was applied to eliminate platelets. The authors extracted 8 features to classify the cells, being: area, perimeter, diameter, shape geometric features, area proportion, deviation, central pallor, and form factor. Finally,

the cells were classified based on their features. The accuracy obtained with the proposed method was 98%.

Lee and Chen (2014) suggested a hybrid neural network architecture to classify single-cell images. The classifier used features like shape and texture to cluster the images into two distinct groups, in addition, separated input layers were used to process the features separately in the classifier. A hybrid neural network was proposed to detect multiple types of RBCs abnormalities instead of other neural networks that can only identify singular blood cell diseases. The authors used a private dataset with 200 abnormal cell images. The proposed network obtained an 88.25% accuracy in cell abnormality classification and 91% for cell disease classification.

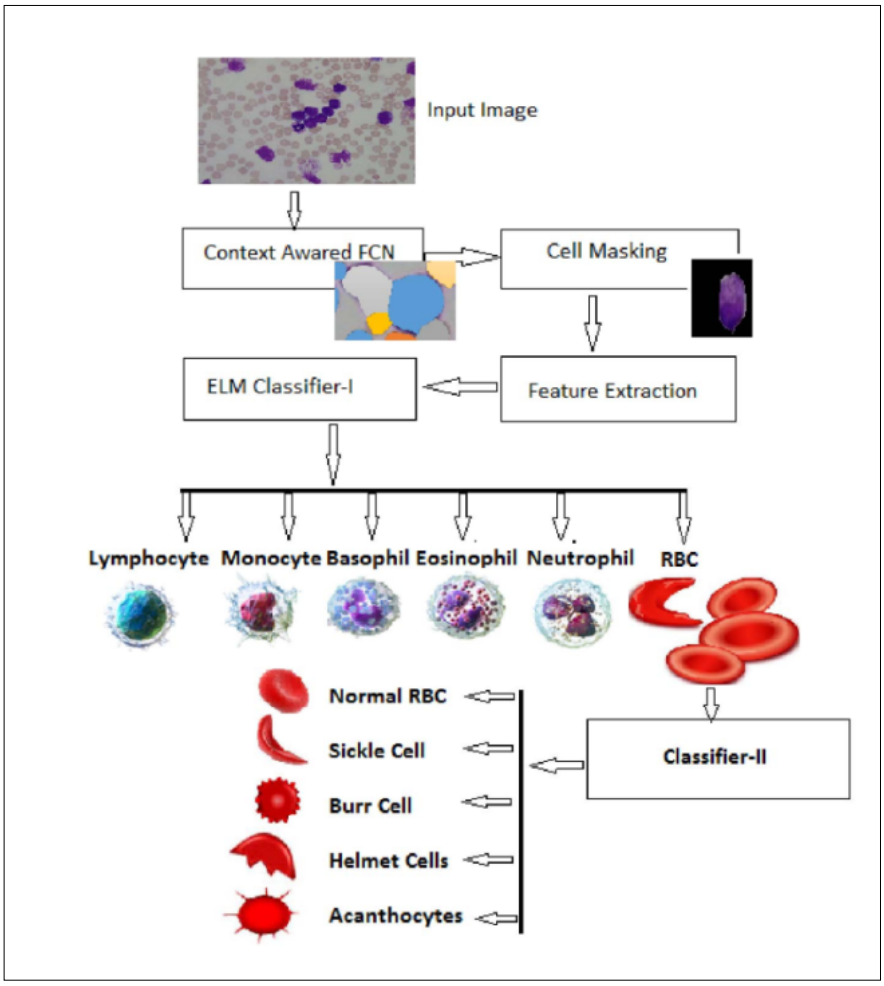
Tomari *et al*. (2014) proposed a computer-aided system for RBCs detection and classification in blood smear images. The images were collected for the research and are private. The authors segmented the image, and then proceed to feature extraction and finally classification. Firstly, the authors separated the RGB image into a single-channel color representation to determine the optimal color channel able to distinguish the object from the background. Then, segmentation using Otsu is applied, followed by post-processing to eliminate noise, holes, and cells around the borders. The process of segmentation can be visualized in Figure 6. Then, the feature extraction was performed using features such as compactness and moment invariant values. The moment method was used to analyze and recognize the object, and seven HU moment features were used to represent the RBCs' shape. The classification module was performed by an artificial neural network. The results after the classification were 83% accuracy, 82% precision, and 76% recall.

Figure 6 ▼
(a) Image in a single channel; (b) Segmentation applying Otsu; (c) Result after post-processing in Tomari *et al*. (2014).
Source: Tomari et al. (2014, p. 209)


(a) (b) (c)

Razzak and Naz (2017) proposed a microscopic blood smear segmentation and classification using deep contour-aware CNN and ELM. In the dataset used in the research, the RBCs and WBCs are combined so the objective is to identify the five types of WBCs: monocytes, lymphocytes, basophil, eosinophil, and neutrophil; the platelets and also, normal RBCs and abnormal RBCs. The dataset is public, called ALL-IDB and it contains 108 JPG images. The authors used contour aware approach based on a fully connected conventional network for segmentation. For classification, a CNN-based ELM was used, the methodology applied can be visualized in Figure 7. An accuracy of 94.71% was obtained for RBCs and anomaly detection, and 98.68% for WBCs classification.

Figure 7 ▶
Methodology proposed by Razzak and Naz (2017).
Source: Razzak and Naz (2017, p. 804)



Elsalamony (2017) suggested anemia cell detection based on shape signature using neural networks. The objective is to detect normal RBCs, elliptocytosis, and sickle cells. Firstly, for normal cells, the author used circular Hough transform, watershed segmentation, and morphological functions. For anomalous cells, parameters such as perimeter, eccentricity, convex area, area, solidity, ratio, and absolute deviations were used to train and test the neural network. Then, two neural networks with one hidden layer and ten neurons were applied. The accuracy obtained was 100%.

In Elsalamony (2016), the author proposed a similar methodology applied in Elsalamony (2017), the difference lies in the addition of microcytes that were included in the consideration. The accuracy after the parameters were inputted in the convolutional neural network was 97.8%.

Abdulkarim, Sudirman and Razak (2019) suggested a method to detect normal and abnormal RBCs using image processing. The method consisted in converting RGB to grayscale then, segmentation using Otsu thresholding, noise removal, hole filling, and cropping images followed by feature extraction, and finally, SVM was used to classify the images into normal and abnormal. The abnormal cells were downloaded from a public bank called Ash Bank and the normal cells are private and were acquired from a hospital. The dataset consists of 65 normal blood cells and 100 abnormal cells. The features that were extracted are perimeter, area, and form factor. After the use of SVM, the results obtained in the study were an accuracy of 94.12%, a specificity of 89.47%, and a sensitivity of 100%.

Hortinela *et al*. (2019) proposed a process to detect abnormal cells using image processing and SVM. The authors proposed the implementation of a device used to analyze the blood samples. The device built uses Raspberry Pi, a Raspberry Pi camera, a microscope, and a power supply. The images used in the research were captured from blood smear samples from the Department of Laboratories at the Philippine General Hospital, 120 images were used. The pipeline of image processing used consists of Sobel Edge Detection followed by watershed segmentation and, feature extraction. The features that were extracted and later used for classification were diameter, shape geometric factor, central pallor, and target flag. Seven different types of labels were selected to classify the blood cells: Normal RBC, echinocytes, elliptocytes, dacrocytes, spherocytes, target cells, and stomatocytes. The classification using SVM obtained an accuracy of 93.33%.

Batitis *et al*. (2020) suggested image classification of abnormal RBCs using a decision tree algorithm. The system proposed by the authors classified the RBCs into 10 classes: spherocytes, codocytes, stomatocytes, ovalocytes, elliptocytes, degmacytes, drepanocytes, dacrocytes, acanthocytes, and echinocytes. 40 images composed of 600 sample cells were used in the study, they were collected from books and journals. After the image acquisition, there was a converting process to grayscale and edge detection applying Canny edge detection then, the images were binarized, and finally, they went through a contour finding stage. Then, the decision tree was applied to classify the different abnormalities in cells. The attributes used in the tree were corners, central pallor, the elongation of the central pallor, the axis of symmetry, elongation, and even the distribution of spikes. Each of the 10 different classes of cells was classified using the attributes. The average reliability obtained was 89.31% and an average error rate of 10.69%.

Rayappan and Karthik (2021) proposed a deep neural network to detect normal and abnormal blood cells. The authors compared the performance of the deep neural network (DNN) with an SVM. The images used in the paper were obtained from online sources and consisted of more than 200 images. The image processing pipeline applied consists of color conversion, noise removal, and resizing the image. The feature extracted was the form factor, which consists of the ratio between the cell area and the square of its perimeter, compactness, and eccentricity. Then, SVM with radial basis function (RBF) kernel was employed to classify the cells. Moreover, DNN was applied after the pre-processing stage, the authors used the AlexNet architecture. Finally, SVM was applied after the feature extraction using the DNN to compare the results. The usage of SVM obtained an average accuracy of 92.4% and DNN with SVM achieved an average accuracy of 93%.

Dalvi and Vernekar (2016) suggested an application to detect and classify abnormal RBCs. The proposed classification separated the cells into four types of abnormalities: echinocytes, elliptocytes, teardrop cells, and macrocytes. The methodology applied in the image processing stage included a conversion to grayscale, Otsu thresholding, and then the image was complemented, followed by a noise removal stage, and feature extraction. Thirteen features were extracted: area, major axis, minor axis, perimeter, form factor, diameter, shape geometry, compactness, eccentricity, solidity, bounding box ratio, equidiameter, and extent. Finally, two classifiers were applied, an ANN, and a decision tree. The dataset used in the classification process was balanced, composed of 250 instances, 50 instances of each abnormality, and 50 normal cells. The ANN applied used the thirteen features as 13 input nodes, 10 hidden nodes, and 5 output nodes (4 for the abnormal cells and 1 for the normal cells). For the decision tree, the parameters used include a maximal depth set to 20, a confidence level of 0.25, a minimal gain of 0.1, and a minimal leaf size of 2. To evaluate the performance, 10-fold cross-validation was applied. The average accuracy obtained for the decision tree was 89.6% and for the ANN was 90.54%.

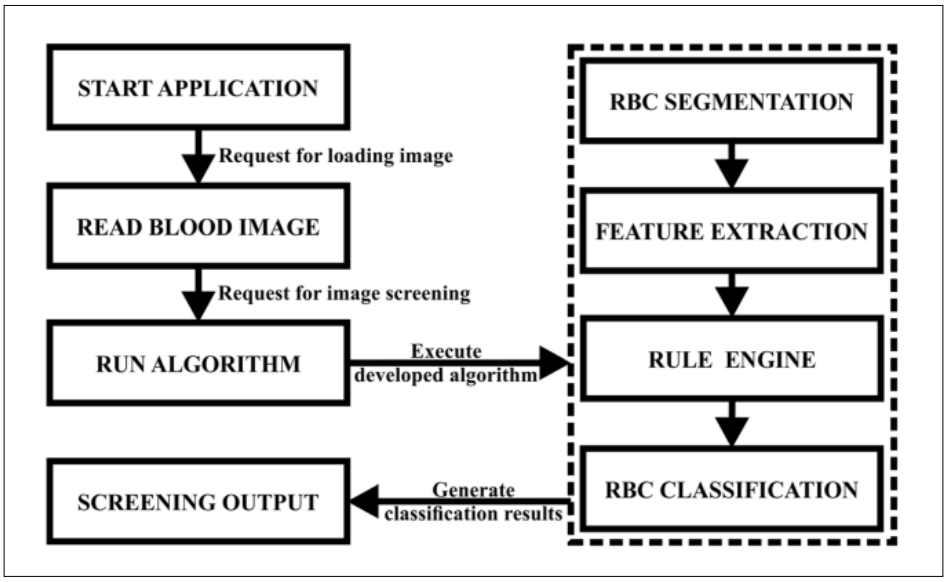
Yuningsih and Mustikasari (2020) proposed a method to classify anemia based on abnormal RBCs morphology by applying a convolutional neural network. The study aimed to classify 5 kinds of abnormal cells (helmet cells, megaloblastic cells, normocytes cells, stomatocytes cells, and teardrop cells) into 2 types of anemia: iron deficiency anemia and megaloblastic anemia. The data was obtained from BCCD (Blood Cell Count and Detection), an MIT-licensed dataset available to the public. The dataset contains 410 images of blood cells. The data was split into 70% for training and 30% for testing. The method applied a pre-processing stage that included transforming RGB images to grayscale images, and image resizing. Then, feature extraction and classification were performed by CNN. For the feature extraction process, the activation function used was ReLu then, the parameters passed through a hidden layer, and finally, the output of the hidden layer was entered as input in the classification CNN. In this network, 2 layers represent the two types of anemia.

In this method by Yuningsih and Mustikasari (2020), different hyperparameters were tested in the classification stage, including the batch size, the number of epochs, the learning rate, and momentum. The best accuracy found in the study was 87.25%, using 200 epochs, a batch size of 256, a learning rate of 0.1, and a momentum value of 0.95. The precision value obtained was 90.74% and the recall of 94.06%.

Durant *et al*. (2017) proposed a CNN to perform a morphological classification of RBC. The dataset used in the study is private and it had 3737 cell images, the data split was 80:20, with 80% for training and 20% for testing. The training dataset was also split into 80% for training and 20% for validation. The 10 morphological classes evaluated in the study include schistocytes, dacrocytes (teardrop cells), acanthocytes, elliptocytes, stomatocytes, spherocytes, codocytes (target cells), echinocytes, overlap, and normal. Data augmentation was performed in the training stage rotating and mirroring the images. Then, DenseNet was applied, and the architecture of this network contained 39 convolutional layers. The network was trained for 300 epochs. Three replicates of DenseNet were trained using unique random seed initializers to calculate the ensemble. After the training of each replicate, the probability for each of the 10 classes was evaluated, and the class with the greater probability was selected. The average precision among the 10 classes was 0.953, the average recall was 0.924 and the average *f1* score was 0.936.

Maity *et al*. (2017) suggested an ensemble of decision trees to classify RBCs based on morphology. For the study, blood smear samples were collected from 100 patients between the age groups of 25 to 50 years. The samples were collected from Medinipur Medical College and Hospital in West Bengal, India, and the cells were separated into 9 groups: microcytic, macrocytic, teardrop, sickle, elliptocyte, echinocyte, acanthocytosis, keratocyte, and normal. The pipeline proposed by the authors is depicted in Figure 8 and it consists of an application that requests the upload of a blood sample image then, the algorithm performs RBCs segmentation, feature extraction, rule engine, and RBCs classification. Finally, the output returns to the user. In the study, the contrast was enhanced using CLAHE (contrast limited adaptive histogram equalization), the image was partitioned into desired regions of interest (ROI) and then, the histogram of each ROI was equalized. Then, the authors performed a background correction, noise reduction, entropy thresholding, and erythrocyte segmentation. In total, 41 shape features were extracted, however, 35 were used to generate the 10 decision trees. Then, the 10 decision trees were used for the ensemble. The authors compared a general C4.5 algorithm, a single rule engine, and the proposed decision tree ensemble, where the suggested method obtained better results, a sensitivity of 97.33%, a specificity of 99.71%, and a precision of 98%, and an accuracy of 97.81%.

Figure 8 ▶
Pipeline proposed by Maity *et al*. (2017). *Source: Maity et al. (2017, p. 56)*



Rahman *et al*. (2021) proposed an automatic detection of abnormal RBCs using color and morphology variation and central pallor. Two datasets were selected for the study, both private from local hospitals, one with 300 images and the other with 250 images. The proposed method divided cells into 15 classes: normocytes, microcytes, macrocytes, echinocytes, acanthocytes, elliptocytes, sickle cells, teardrop, elliptic spur, clump, stomatocyte, target, hypochromic, normochromic, and hyperchromic. In the preprocessing stage, the images were enhanced and quantized, then, the segmentation was performed for the morphology of the RBC. The cells and their central pallors were analyzed geometrically. The method proposed by the authors, Cross-validation Accuracy Weighted Probabilistic Ensemble (CAWPE) consists of a heterogeneous ensemble built on weighting the probability approximations of classifiers with an estimation of the accuracy designed through cross-validation on the train data. CAWPE was compared to MLP, AdaBoost, Random Forest, and SVM and, it obtained the best result, with an average accuracy of 97.1% for central pallor base classification and 96.9% for shape base classification.

Chy and Rahaman (2018) proposed a method to detect and classify sickle cells as normal or sickle cells. The method consists of a pre-processing phase after the image input. In this stage, there is a grayscale image conversion, image enhancement, and noise filtering. Then, in the processing stage, threshold segmentation using Otsu is applied followed by morphological operations. The extracted features were geometrical as metric value, aspect ratio, entropy, mean, standard deviation, and variance. Finally, the images were used to train the SVM algorithm. The data was split in half, with 40 images to train and 40 images to test. The results obtained were 95% accuracy and 96.55% sensitivity.

Akrimi *et al*. (2014) suggested an approach to detect two classes of cells, normal and abnormal. The authors used an imbalanced dataset with 1000 images obtained from patients in the Department of Hematology Serdang Hospital in Kuala Lumpur, Malaysia. The data consisted of 100 normal and 900 abnormal cells. For the segmentation, algorithms such as k-means, seed filling, and bilinear algorithms were applied. Then, edge smoothing was performed with a mean filter. The smoothing process was performed three times to increase the smoothness. For the feature extraction process, the authors extracted 271 different features for each RBCs image, the features include geometric, texture, and color. Finally, an SVM classifier was applied, for the training set 30 normal images and 300 abnormal images were used. Cross-validation was used to validate

the algorithm. The results obtained consist of 99.9% accuracy, 100% sensitivity, and 99.8% specificity.

Lotfi *et al*. (2015) proposed a method to detect three types of abnormalities in RBC, dacrocytes, schistocytes, and elliptocytes. The dataset was obtained from Sayad-Al-Shohada Hospital in Isfahan, Iran, and the Department of Medical Sciences of the University of Isfahan and it consists of 30 blood smear samples. The dataset was balanced, with 100 cells of each of the 4 classes. The pre-processing stage consisted, in contrast, of enhancement, using Histogram Equalization. Then, segmentation was performed by dividing the image into nine equal parts and applying the Otsu algorithm. The authors removed the marginal cells in the image and proceed to detection. To perform feature extraction, boundary, region descriptors, and Hu moments were used to extract 33 features. To classify the extracted features three different classifiers were compared, KNN, SVM, and a Neural Network. For the neural network, there were 33 features as input, 4 outputs (classes), and 10 neurons in the hidden layer. 70% was used to train the network, 15% for validation, and 15% for the test set. The result obtained was 85% accuracy. For SVM and KNN, the authors used a One Against All approach. So, the algorithm was trained three different times considering the 3 abnormalities for each classifier, in KNN, k=1 was used. Then, the authors combined the classifiers to obtain a better result. After maximum voting, the result obtained was a precision of 95% and 99% accuracy for dacrocytes. A precision of 97.5% and accuracy of 97% for elliptocytes, and, for schistocytes, a 100% precision, and 100% accuracy.

Tyas *et al*. (2017) proposed a methodology for classifying abnormal cells, which are seen in thalassemia, a genetic disorder that impedes the production of RBCs and the distribution of oxygen to the body. The blood for the dataset was collected from the clinical pathology laboratory of medical faculty in Indonesia. Based on the images obtained, the region of interest (ROI) was selected manually. Then, the image was converted from RGB to grayscale, histogram equalization was applied, followed by image binarization, morphology operation, and hole-filling operation. For the feature extraction stage, 43 features were extracted using a variety of techniques, such as invariant moments, shape features, texture extraction, and color features. Then, two different classifiers were applied, an artificial neural network (ANN) and a convolutional neural network (CNN). The architecture of the artificial neural network with the best result obtained consists of the following hyper-parameters, 10000 epochs, momentum value = 0.05, learning rate = 0.1, and 6 neurons in the hidden layer. The result obtained was 93.24%. For the CNN the input was the RGB images, resized to 32x32 pixels. The architecture applied was LeNet-5 with a few modifications, it consisted of two convolution layers and one fully connected layer. The highest accuracy was obtained with 5000 epochs and a learning rate = 0.01, the accuracy was 92.55%.

Syahputra, Sari and Rahmat (2017) suggested the use of a radial basis function network to classify abnormal RBCs based on their shape. The radial basis function network (RBFN) consists of a neural network that combines guided and non-guided training, also known as hybrid training, the network contains an input layer unit, a hidden layer unit, and an output layer unit. The images used are 50x50 pixels, and 55 images for each category were used. The categories studied were normal cells, acanthocytosis, and cigar-shaped cells. The dataset was divided into 90% for training and 10% for testing. The pipeline applied by the authors consisted of conversion to grayscale, conversion to binary, and Canny edge detection to find the edges on the images. Finally, the data were classified using RBFN. The accuracy obtained was 83.3%.

Sen *et al*. (2021) proposed a methodology for detecting different types of abnormalities in cells, the cells were separated into circular, elongated, and others. The dataset used

was erythrocytes IDB1. Firstly, the images were converted to grayscale, the median filter was used to remove noise, and small objects were removed. For segmentation, Otsu thresholding was applied, as watershed segmentation, and morphological operations. Then, geometrical, statistical, and texture features were extracted, such as circularity, aspect ratio, eccentricity, variance, entropy, and standard deviation. Finally, several classifiers were applied and compared. Random forest, Naive Bayes, SVM, and logistic regression were used. The results can be visualized in Table 1.

Table 1 ▶
Results obtained by each classifier in Sen *et al*. (2021).
Source: Sen et al. (2021)

Table of random forest			
Parameters	**Circular**	**Elongated**	**Others**
Precision	95%	92%	91%
Recall	95%	95%	89%
F1 score	95%	93%	90%
Accuracy	92%	92%	92%
Table of Naïve Bayes			
Parameters	**Circular**	**Elongated**	**Others**
Precision	84%	88%	90%
Recall	96%	98%	93%
F1 score	90%	70%	80%
Accuracy	88%	88%	88%
Table of SVM			
Parameters	**Circular**	**Elongated**	**Others**
Precision	97%	86%	89%
Recall	92%	95%	84%
F1 score	94%	90%	86%
Accuracy	90%	90%	90%
Table of logistic regression			
Parameters	**Circular**	**Elongated**	**Others**
Precision	98%	86%	86%
Recall	97%	86%	87%
F1 score	86%	98%	87%
Accuracy	90%	90%	90%

Sharma, Rathore and Vyas (2016) proposed the detection of abnormal cells caused by sickle cell anemia and thalassemia in RBCs using image processing. The cells were separated into different classes, sickle-shaped, dacrocytes, ovalocytes, and normal cells. 100 images were used in the dataset. The process consisted of a pre-processing stage, applying a median filter, then, image segmentation using watershed, followed by morphological operations such as RGB to binary, complementing image, hole filling, and labeling. Features such as the variance of radial signature, metric value, and aspect ratio were extracted, and finally, the classification was performed by applying KNN. The accuracy obtained was 80.6% and the sensitivity of 87.2%.

Kannadaguli (2020) suggested the application of principal component analysis (PCA) and SVM to detect abnormalities in RBC. The author used a balanced dataset with 54000 blood smear images. The images were divided into normal, echinocytes, elliptocytes, and sickle cells, each of the classes with 13500 examples. The pipeline used consisted of a pre-processing stage where the images were converted to grayscale, then, Watershed has applied to segment the image. Also, morphology operations were applied. And, as some images were from videos, a median filter and histogram matching were applied frame by frame. For feature extraction, K-means clustering and PCA were used. Principal component analysis (PCA) consists of a technique used for dimensionality reduction. In total, 500 features were extracted and saved as a bag of features (BoF). Finally, SVM was applied to classify the features, the recall obtained was 0.94 and the precision was 0.97.

Rakshit and Bhowmik (2013) proposed image processing techniques to detect abnormal RBCs and, diagnose SCA. The methodology applied consists of pre-processing using a Weiner filter followed by Sobel edge detection to find the cells. The authors calculated a metric, *M*, for each cell, and any deviation was detected to diagnose SCA. M correlates the area to the perimeter of the cell. The equation has a range between $0 < M < 1$, for a normal cell the value would be between 0.84 and 0.92. Abnormal cells, such as sickle cells had a range between $M = 0.42$ and $M = 0.59$. The solution presented an overall accuracy equal to 95.8%.

Tyagi, Saini and Dahyia (2016) suggested the application of an artificial neural network to detect abnormal RBC, and therefore, anemia. The dataset used was collected from Haematologycal Department AIIMS in New Delhi, India and it consisted of 100 images of different blood samples. Firstly, the images were pre-processed converting to grayscale and the contrast was increased to differentiate cells from the background. Secondly, segmentation was performed by applying Canny edge detection. Then, in the feature extraction stage, the features were extracted by region descriptors and GLCM. Three types of region descriptors were extracted: geometric properties, texture, and moment invariants. In total, 34 features were used in the classification performed by an ANN with 20 hidden layers, and 5 output layers. The output layers are the 5 different classes of cells, discocyte (normal cell), elliptocyte, dacrocyte, degmacyte, and schistocyte. The dataset was balanced, with 50 cells of each class. The split between the data consisted of 70% for training, 15% for validation, and 15% for testing. The results obtained are presented in Table 2. The average accuracy obtained was 77.75% and the average precision of 54.75%.

Table 2 ►
Results obtained for each type of cell in Tyagi, Saini and Dahyia (2016).
Source: Tyagi, Saini and Dahyia (2016)

Cell	TP	FP	TN	FN	Precision	Accuracy
EL[a]	37	13	112	21	74%	81%
DA[b]	26	24	123	20	52%	77%
DE[c]	25	24	124	17	51%	78%
SH[d]	21	29	128	20	42%	75%

a – Number of Elliptocytes; b – Number of Dacrocytes; c – Number of Degmacytes; d – Number of Schistocytes

Qiu *et al*. (2020) proposed a multi-label detection and classification of RBC. The authors implemented a Faster-RCNN on the RBCs' microscopic images. The images were collected from the University of Pittsburgh Medical Center and Massachusetts General Hospital and it contained 313 images. Six types of RBCs were classified: discocytes and oval, elongated and sickle, reticulocytes, granular, echinocytes, and stomatocytes.

The algorithm was trained on whole microscopic images with touching and overlapping cells and, also on patches with individual cells. 1080 single-cell patches were processed and 1389 patches with touching and overlapping cells. The methodology applied uses a Faster-RCNN with the structure of a ResNet-50 and pre-trained on ImageNet to extract the features. The learning rate used was 10^{-3}, decay 10^{-6}, the momentum of 0.9, batch size of 10, and 1000 epochs. Finally, with the output from the neural network, a Gradient Boosting Classifier was applied, and the classifier obtained an accuracy of 85.1% in a 5-fold cross-validation.

3 Discussion of the results

As mentioned previously, there are two types of problems in the literature, detection, and classification. In the 31 papers analyzed during the review, Alzubaidi *et al*. (2020), Durant *et al*. (2017), and Syahputra, Sari and Rahmat (2017) tackled classification problems while the others worked on a detection approach.

For detection, the majority of papers applied image processing techniques to segment or detect the cells, and five studies applied machine learning techniques to detect RBC. The list of each paper and their respective approach applied can be visualized in Table 3. The column 'Results' presents the results with the same metrics that were presented in the studies.

Table 3 ▼
Comparative of papers.
Source: research data

Paper	Problem	Dataset	Classes	Cell detection/ segmentation		Cell classification		Techniques		Results
				Image Processing	ML	Expert System	ML	Feature Extraction	Classification	
Alzubaidi *et al*. (2020)	Classification	Balanced	Normal, sickle cell and other				X	CNN	CNN	Accuracy CNN: 99.54% Accuracy (CNN + SVM): 99.98%
Chy and Rahaman (2019)	Detection	Imbalanced	Normal, sickle cell	X			X	Fuzzy C means	KNN, SVM, and ELM	Accuracy: KNN = 73.33%, SVM = 83.33% e ELM = 87.73%. FI Score: KNN = 81.81%, SVM = 88.89% e ELM = 91.3%.
Fadhel, Humaidi and Oleiwi (2017)	Detection	Imbalanced	Normal and abnormal	X		X		CHT and WT	Expert system	CHT detected 138 normal cells WT detected 123 normal cells
González-Hidalgo *et al*. (2014)	Detection	Imbalanced	Normal and sickle-shaped	X		X		Ellipse fitting	Expert system	Accuracy for normal cells: 100% Accuracy for sickle cells: 98%
Xu *et al*. (2017)	Detection	Imbalanced	Discocyte, echinocyte, elongated, granular, oval, reticulocyte, sickle, stomatocyte	X			X	Random walk method with seeds obtained from distance transform result	CNN	Accuracy: 87.50%, AUC: 0.94

Cont. on next page

Table 3 continued

Acharya and Kumar (2017)	Detection	Imbalanced	Normocyte, normocyte w/o central pallor, macrocyte w/o central pallor, spherocyte, normochromic macrocyte hypochromic microcytes, hypochromic macrocytes, hypochromic, normocyte, normochromic microcyte, elongated cell, sickle cell	X		X		K-Medoids, WT, Otsu	Expert system	Accuracy: 98%
Lee and Chen (2014)	Detection	Balanced	Burr cell, sickle cell, horn cell, elliptocyte cell	X			X	Hybrid Neural Network	Hybrid Neural Network	Accuracy for cell classification: 88.25%
Tomari *et al*. (2014)	Detection	Imbalanced	Normal, abnormal	X			X	Otsu, HU moments	Neural Network	Accuracy: 83% Precision: 82% Recall: 76%
Razzak and Naz (2017)	Detection	Imbalanced	Normal, sickle cell, burr cell, helmet cell, acanthocyte	X	X		X	DCAN	Convolutional ELM	Accuracy: 94.71%
Elsalamony (2017)	Detection	Imbalanced	Normal, sickle, elliptocytosis	X			X	CHT and WT	Neural Network	Accuracy: 100%
Elsalamony (2016)	Detection	Imbalanced	Normal, elliptocytosis and sickle cells	X			X	CHT and WT	Neural Network	Accuracy: 97.8%
Abdulkarim, Sudirman and Razak (2019)	Detection	Imbalanced	Normal and abnormal	X			X	Otsu	SVM	Accuracy: 94.12% Specificity: 89.47% Sensitivity: 100%.
Hortinela *et al*. (2019)	Detection	Imbalanced	Normal, echinocytes, elliptocytes, dacrocytes, spherocytes, target cells, stomatocytes, and unknown	X			X	Sobel Edge Detection and WT	SVM	Accuracy: 93.33%
Batitis *et al*. (2020)	Detection	Imbalanced	Spherocytes, codocytes, stomatocytes, ovalocytes, elliptocytes, degmacytes, drepanocytes, dacrocytes, acanthocytes, echinocytes	X			X	Canny Edge Detection	Decision Tree	Reliability rate: 89.31% Average error rate: 10.69%
Rayappan and Karthik (2021)	Detection	Imbalanced	Achantocyte, elliptocyte, sickle cell, teardrop, normal		X		X	SVM/DNN	SVM	SVM: Average accuracy: 92.4% DNN + SVM: Average accuracy: 93%
Dalvi and Vernekar(2016)	Detection	Balanced	Elliptocyte, echinocyte, teardrop, macrocyte, and normal	X			X	Otsu	Decision Tree/ NN	Decision Tree: Accuracy = 89.6% ANN: Accuracy = 90.54%

Cont. on next page

Table 3 continued

Yuningsih and Mustikasari (2020)	Detection	Imbalanced	Helmet, megaloblastic, normocytes, stomatocytes, teardrop		X		X	CNN	CNN	Accuracy = 87.25% Precision = 90.74% Recall = 94.06%
Durant *et al.* (2017)	Classification	Imbalanced	Schistocytes, dacrocytes acanthocytes, elliptocytes, stomatocytes, spherocytes, codocytes, echinocytes, overlap, and normal	X			X	DenseNet	Ensemble	Average F1 Score = 93.6% Average Precision = 95.3% Average Recall = 92.4%
Maity *et al.* (2017)	Detection	Imbalanced	Normal, microcytic, macrocytic, teardrop, sickle, elliptocyte, echinocyte, acanthocytosis, keratocyte	X			X	WT	Decision Tree/ Ensemble	Sensitivity: 97.33% Specificity: 99.71% Precision: 98% Accuracy: 97.81%.
Rahman *et al.* (2021)	Detection	Imbalanced	Normal, microcyte, macrocyte, echinocyte, acanthocyte, elliptocyte, sickle cell, teardrop, elliptic spur, clump, stomatocyte, target, hypochromic, normochromic, hyperchromic	X			X	Expert system on a binary image	Cross-validation Accuracy Weighted Probabilistic Ensemble (CAWPE)	Accuracy based on morphology: 96.9% Accuracy based on central pallor: 97.1%
Chy and Rahaman (2018)	Detection	Not disclosure	Normal and sickle-shaped	X			X	Otsu	SVM	Accuracy: 95% Sensitivity: 96.55%
Akrimi *et al.* (2014)	Detection	Imbalanced	Normal and abnormal	X			X	K-Means	SVM	Accuracy: 99.9% Sensitivity: 100% Specificity: 99.8%
Lofti *et al.* (2015)	Detection	Balanced	Normal, dacrocyte, elliptocyte, and schistocyte	X			X	Otsu	KNN, SVM, and neural network	Dacrocytes: accuracy = 99% precision = 95% Elliptocyte: accuracy = 97% precision = 97% Schitocyte: accuracy = 100% precision = 100%
Tyas *et al.* (2017)	Detection	Not disclosure	Teardrop, target cell, cigar cell acantocyte, normal	X			X	Invariant Moments, GLCM, colour moments	ANN and CNN with LeNet architecture	Accuracy ANN = 93.24% Accuracy CNN (LeNet) = 92.55%
Syahputra, Sari and Rahmat (2017)	Classification	Balanced	Normal, acanthocytosis, and cigar-shaped cells	X			X	Canny Edge Detection	Radial Basis Function Network	Accuracy = 83.3%

Cont. on next page

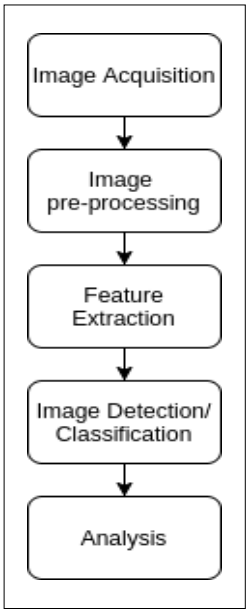
Table 3 continued

Sen *et al.* (2021)	Detection	Imbalanced	Circular, sickle-shaped, and other	X			X	Otsu Thresholding, Watershed segmentation	Random forest, Naive Bayes, SVM, and Logistic Regression	RF: F1 Score: 92.67% NB: F1 Score: 80% SVM: F1 Score: 90% LR: F1 Score: 90,3%
Sharma, Rathore and Vyas (2016)	Detection	Not disclosure	Sickle-shaped, dacrocytes, ovalocytes, and normal cells	X			X	Watershed segmentation	KNN	Accuracy: 80.6% Sensitivity: 87.2%
Kannadaguli (2020)	Detection	Balanced	Normal, echinocytes, elliptocytes, sickle-cell	X			X	K-means clustering, principal component analysis (PCA)	SVM	Recall: 94% Precision: 97%
Rakshit and Bhowmik (2013)	Detection	Not disclosure	Normal and sickle-shaped	X		X		Sobel edge detection	Expert system	Accuracy: 95.8%
Qiu *et al.* (2020)	Detection	Not disclosure	Discocytes and oval, elongated and sickle, reticulocytes, granular, echinocytes, stomatocytes		X		X	Faster-RCNN with ResNet-50 structure pre-trained on ImageNet	Gradient boosting classifier	Accuracy: 85.1%
Tyagi, Saini and Dahyia (2016)	Detection	Balanced	Normal, elliptocyte, dacrocyte, degmacyte, schistocyte	X			X	Canny edge detection, region descriptors, and GLCM	ANN with 20 hidden layers	Accuracy: 77.75%

The pipeline depicted in Figure 9 is a very common approach for cell detection and classification, and it was applied in 28 studies.

Yuningsih and Mustikasari (2020) and Qiu *et al.* (2020) proposed detecting the cells by applying a neural network. Further studies applying neural networks are necessary to determine their efficacy in detecting abnormal RBC.

Figure 9 ▶
Pipeline used in 28 of the analyzed studies.
Source: elaborated by authors (2022)



To classify the cells after the feature extraction, Fadhel, Humaidi and Oleiwi (2017), González-Hidalgo *et al*. (2014), Acharya and Kumar (2017) and Rakshit and Bhowmik (2013) proposed an expert system approach. An expert system behaves as a human expert in the field, and the decision-making process is treated as a selection of a particular alternative from a set of alternatives (BOHANEC; RAJKOVIC, 1990). However, according to Table 3, it is possible to infer that more recent papers apply machine learning techniques to classify cells instead of focusing on expert systems.

An issue found in most studies is the application of accuracy to measure the proposed method. According to Tohka and Van Gils (2021), accuracy can be useless when the classes are imbalanced since it consists of the number of correct classifications over the total number of samples to classify, and it is represented in Equation 1. Using an example, if there were 1000 test cases, with 990 healthy people and 10 people with diseases and the classifier classified all the cases as healthy, the accuracy would be 99.0%, however, it would be ineffective since it could not detect the disease.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

where *TP* refers to true positive tests when the classifier correctly classifies a disease. *TN* is true negative when the classifier correctly classifies a negative test. *FP*, or false positive, occurs when the classifier incorrectly classifies a healthy person for a person with a disease. Finally, *FN*, a false negative, occurs when a person with a disease is classified as healthy.

For imbalanced datasets, it is interesting to observe measures as precision, defined by Equation 2, recall, in Equation 3, and *F1*-score in Equation 4 (RIBEIRO; REYNOSO-MEZA, 2020).

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = \frac{2\,Precision \times Recall}{Precision + Recall} \tag{4}$$

Another interesting aspect found during the review is the lack of data regarding abnormal RBCs. Most datasets contain few images and are private. However, with an imbalanced dataset, some considerations must be taken into account.

An additional issue present in the 31 papers analyzed is that 19 contain imbalanced datasets which is due to the fact abnormal cells are rare. As mentioned previously, it is fundamental to use the appropriate measure for error. Also, while applying

cross-validation (CV) it is necessary to use a stratified CV, which occurs when the folds contain the same proportion of classes for training and testing. It is also important to be careful while applying oversampling because some techniques do not increase the number of training data according to Tohka and Van Gils (2021).

4 Conclusions

This review presented the latest efforts in the literature to detect and classify abnormal RBCs. In total, 31 works were reviewed, and several techniques were applied. The application of a robust and reliable methodology for detecting and classifying RBC can reduce the time spent in validating these tests and increase their reliability. It is possible to perceive that in the last few years, more solutions are applying convolutional neural networks to extract features and neural networks to detect cells, due to the advances in machine learning in the past 5 years. However, there is still a lack of datasets available to the scientific community to experiment with new solutions. It is also critical to be careful while dealing with imbalanced datasets and applying the appropriate metrics to measure the reliability of the proposed solutions.

The application of CNNs to detect abnormal RBCs, using precise metrics to measure the solutions obtained, requires further studies.

Funding

This paper was partially funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação Araucária (FAPPR) – Brazil. Projects PQ-2/310079/2019-5 and Universal/408164/2021-2.

Conflict of interest

The authors declare no conflict of interest.

References

ABDULKARIM, H. A.; SUDIRMAN, R.; RAZAK, M. A. A. Normal and abnormal red blood cell recognition using image processing. **Indonesian Journal of Electrical Engineering and Computer Science**, v. 14, n. 1, p. 100-104, 2019. DOI: https://dx.doi.org/10.11591/ijeecs.v14.i1.pp96-100.

ACHARJEE, S.; CHAKRABARTTY, S.; ALAM, M. I.; DEY, N.; SANTHI, V.; ASHOUR, A. S. A semiautomated approach using GUI for the detection of red blood cells. *In*: 2016 INTERNATIONAL CONFERENCE ON ELECTRICAL, ELECTRONICS, AND OPTIMIZATION TECHNIQUES (ICEEOT), 2016, Chennai. **Proceedings** […]. Chennai: IEEE, 2016. p. 525-529. DOI: https://dx.doi.org/10.1109/ICEEOT.2016.7755669.

ACHARYA, V.; KUMAR, P. Identification and red blood cell classification using computer aided system to diagnose blood disorders. *In*: 2017 INTERNATIONAL

CONFERENCE ON ADVANCES IN COMPUTING, COMMUNICATIONS AND INFORMATICS (ICACCI), 2017, Udupi. **Proceedings** […]. Udupi: IEEE, 2017. p. 2098-2104. DOI: https://dx.doi.org/10.1109/ICACCI.2017.8126155.

AKRIMI, J. A.; SULIMAN, A.; GEORGE, L. E.; AHMAD, A. R. Classification of red blood cells using support vector machine. *In*: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY AND MULTIMEDIA, 6., 2014, Putrajaya. **Proceedings** […]. Putrajaya: IEEE, 2014. p. 265-269. DOI: https://dx.doi.org/10.1109/ICIMU.2014.7066642.

ALAM, M. M.; ISLAM, M. T. Machine learning approach of automatic identification and counting of blood cells. **Healthcare Technology Letters**, v. 6, n. 4, p. 103-108, 2019. DOI: https://dx.doi.org/10.1049/htl.2018.5098.

ALZUBAIDI, L.; FADHEL, M. A.; AL-SHAMMA, O.; ZHANG, J.; DUAN, Y. Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis. **Electronics**, v. 9, n. 3, p. 427, 2020. DOI: https://doi.org/10.3390/electronics9030427.

BAI, M.; URTASUN, R. Deep watershed transform for instance segmentation. *In*: 2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2017, Honolulu. **Proceedings** […]. Honolulu: IEEE, 2017. p. 5221-5229. DOI: https://doi.org/10.1109/CVPR.2017.305.

BALARAJAN, Y.; RAMAKRISHNAN, U.; ÖZALTIN, E.; SHANKAR, A. H.; SUBRAMANIAN, S. V. Anaemia in low-income and middle-income countries. **The Lancet**, v. 378, n. 9809, p. 2123-2135, 2011. DOI: https://doi.org/10.1016/S0140-6736(10)62304-5.

BATITIS, V. M. E.; CABALLES, M. J. G.; CIUDAD, A. A.; DIAZ, M. D.; FLORES, R. D.; TOLENTINO, R. E.; Image classification of abnormal red blood cells using decision tree algorithm. *In*: INTERNATIONAL CONFERENCE ON COMPUTING METHODOLOGIES AND COMMUNICATION (ICCMC), 4., 2020, Erode. **Proceedings** […]. Erode: IEEE, 2020. p. 498-504. DOI: https://dx.doi.org/10.1109/ICCMC48092.2020.ICCMC-00093.

BOHANEC, M.; RAJKOVIC, V. DEX: an expert system for decision support. **Sistemica**, v. 1, n. 1, p. 145-157, 1990. Available at: http://kt.ijs.si/MarkoBohanec/pub/Sistemica90.pdf. Accessed on: 4 May 2022.

CHY, T. S.; RAHAMAN, M. A. Automatic sickle cell anemia detection using image processing technique. *In*: 2018 INTERNATIONAL CONFERENCE ON ADVANCEMENT IN ELECTRICAL AND ELECTRONIC ENGINEERING (ICAEEE), 2018, Gazipur. **Proceedings** […]. Gazipur: IEEE, 2018. p. 1-4. DOI: https://dx.doi.org/10.1109/ICAEEE.2018.8642984.

CHY, T. S.; RAHAMAN, M. A. A comparative analysis by KNN, SVM & ELM classification to detect sickle cell anemia. *In*: 2019 INTERNATIONAL CONFERENCE ON ROBOTICS, ELECTRICAL AND SIGNAL PROCESSING TECHNIQUES (ICREST), 2019, Dhaka. **Proceedings** […]. Dhaka: IEEE, 2019. p. 455-459. DOI: https://dx.doi.org/10.1109/ICREST.2019.8644410.

DALVI, P. T.; VERNEKAR, N. Computer aided detection of abnormal red blood cells. *In*: 2016 IEEE INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ELECTRONICS, INFORMATION & COMMUNICATION TECHNOLOGY (RTEICT), 2016, Bangalore. **Proceedings** […]. Bangalore: IEEE, 2016. p. 1741-1746. DOI: https://dx.doi.org/10.1109/RTEICT.2016.7808132.

DURANT, T. J. S.; OLSON, E. M.; SCHULZ, W. L.; TORRES, R. Very deep convolutional neural networks for morphologic classification of erythrocytes. **Clinical Chemistry**, v. 63, n. 12, p. 1847-1855, 2017. DOI: https://doi.org/10.1373/clinchem.2017.276345.

ELSALAMONY, H. A. Anaemia cells detection based on shape signature using neural networks. **Measurement**, v. 104, p. 50-59, 2017. DOI: https://doi.org/10.1016/j.measurement.2017.03.012.

ELSALAMONY, H. A. Healthy and unhealthy red blood cell detection in human blood smears using neural networks. **Micron**, v. 83, p. 32-41, 2016. DOI: https://doi.org/10.1016/j.micron.2016.01.008.

FADHEL, M. A.; HUMAIDI, A. J.; OLEIWI, S. R. Image processing-based diagnosis of sickle cell anemia in erythrocytes. *In*: 2017 ANNUAL CONFERENCE ON NEW TRENDS IN INFORMATION & COMMUNICATIONS TECHNOLOGY APPLICATIONS (NTICT), 2017, Baghdad. **Proceedings** […]. Baghdad: IEEE, 2017. p. 203-207. DOI: https://dx.doi.org/10.1109/NTICT.2017.7976124.

GONZÁLEZ-HIDALGO, M.; GUERRERO-PENA, F. A.; HEROLD-GARCÍA, S.; JAUME-I-CAPÓ, A.; MARRERO-FERNÁNDEZ, P. D. Red blood cell cluster separation from digital images for use in sickle cell disease. **IEEE Journal of Biomedical and Health Informatics**, v. 19, n. 4, p. 1514-1525, 2014. DOI: https://dx.doi.org/10.1109/JBHI.2014.2356402.

HABIBZADEH, M.; KRZYZAK, A.; FEVENS, T. White blood cell differential counts using convolutional neural networks for low resolution images. *In*: RUTKOWSKI, L.; KORYTKOWSKI, M.; SCHERER, R.; TADEUSIEWICZ, R.; ZADEH, L. A.; ZURADA, J. M. (ed.). **Artificial Intelligence and Soft Computing (ICAISC 2013)**. Berlin: Springer, 2013. p. 263-274. (Lecture Notes in Computer Science, v. 7895). DOI: https://doi.org/10.1007/978-3-642-38610-7_25.

HORTINELA, C. C.; BALBIN, J. R.; FAUSTO, J. C.; DIVINA, P. D. C.; FELICES, J. T. F. Identification of abnormal red blood cells and diagnosing specific types of anemia using image processing and support vector machine. *In*: IEEE INTERNATIONAL CONFERENCE ON HUMANOID, NANOTECHNOLOGY, INFORMATION TECHNOLOGY, COMMUNICATION AND CONTROL, ENVIRONMENT, AND MANAGEMENT (HNICEM), 11., 2019, Laoag. **Proceedings** […]. Laoag: IEEE, 2019. p. 1-6. DOI: https://dx.doi.org/10.1109/HNICEM48295.2019.9072904.

KAJÁNEK, F.; CIMRÁK, I. Evaluation of detection of red blood cells using convolutional neural networks. *In*: 2019 INTERNATIONAL CONFERENCE ON INFORMATION AND DIGITAL TECHNOLOGIES (IDT), 2019, Zilina. **Proceedings** […]. Zilina: IEEE, p. 198-202. 2019. DOI: https://dx.doi.org/10.1109/DT.2019.8813664.

KANNADAGULI, P. Microscopic blood smear RBC classification using PCA and SVM based Machine Learning. *In*: INTERNATIONAL CONFERENCE

ON MULTIMIDIA PROCESSING, COMMUNICATION & INFORMATION TECHNOLOGY (MPCIT), 3., 2020, Shivamogga. **Proceedings** [...]. Shivamogga: IEEE, 2020. p. 82-86. DOI: https://dx.doi.org/10.1109/MPCIT51588.2020.9350389.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436-444, 2015. DOI: https://doi.org/10.1038/nature14539.

LEE, H.; CHEN, Y.-P. P. Cell morphology based classification for red cells in blood smear images. **Pattern Recognition Letters**, v. 49, p. 155-161, 2014. DOI: https://doi.org/10.1016/j.patrec.2014.06.010.

LOU, J.; ZHOU, M.; LI, Q.; YUAN, C.; LIU, H. An automatic red blood cell counting method based on spectral images. *In*: INTERNATIONAL CONGRESS ON IMAGE AND SIGNAL PROCESSING, BIOMEDICAL ENGINEERING AND INFORMATICS (CISP-BMEI), 9., 2016, Datong. **Proceedings** [...]. Datong: IEEE, p. 1391-1396. 2016. DOI: https://doi.org/10.1109/CISP-BMEI.2016.7852934.

LOTFI, M.; NAZARI, B.; SADRI, S.; SICHANI, N. K. The detection of dacrocyte, schistocyte and elliptocyte cells in iron deficiency anemia. *In*: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION AND IMAGE ANALYSIS (IPRIA), 2., 2015, Rasht. **Proceedings** [...]. Rasht: IEEE, 2015. p. 1-5. DOI: https://dx.doi.org/10.1109/PRIA.2015.7161628.

MAITY, M.; MUNGLE, T.; DHANE, D.; MAITI, A. K.; CHAKRABORTY, C. An ensemble rule learning approach for automated morphological classification of erythrocytes. **Journal of Medical Systems**, v. 41, n. 4, 56, 2017. DOI: https://doi.org/10.1007/s10916-017-0691-x.

MITCHELL, T. M. **Machine learning**. New York: McGraw-Hill, 1997.

MOLINA-CABELLO, M. A.; LÓPEZ-RUBIO, E.; LUQUE-BAENA, R. M.; RODRÍGUEZ-ESPINOSA, M. J.; THURNHOFER-HEMSI, K. Blood cell classification using the hough transform and convolutional neural networks. *In*: ROCHA, A.; ADELI, H.; REIS, L. P.; COSTANZO, S. (ed.). **Trends and Advances in Information Systems and Technologies**: WorldCIST'18 2018. Cham: Springer, 2018. v. 2, p. 669-678. (Advances in Intelligent Systems and Computing, v. 746). DOI: https://doi.org/10.1007/978-3-319-77712-2_62.

PARAB, M. A.; MEHENDALE, N. D. Red blood cell classification using image processing and CNN. **SN Computer Science**, v. 2, 70, 2021. DOI: https://doi.org/10.1007/s42979-021-00458-2.

PARK, Y.; BEST, C. A.; BADIZADEGAN, K.; DASARI, R. R.; FELD, M. S.; KURIABOVA, T.; HENLE, M. L.; LEVINE, A. J.; POPESCU, G. Measurement of red blood cell mechanics during morphological changes. **PNAS**, v. 107, n. 15, p. 6731-6736, 2010. DOI: https://doi.org/10.1073/pnas.0909533107.

PEDERSEN, S. J. K. **Circular hough transform**. Aalborg University, 2007. Available at: https://cdn.manesht.ir/9961___Simon_Pedersen_CircularHoughTransform.pdf. Accessed on: 4 May 2022.

QIU, W.; GUO, J.; LI, X.; XU, M.; ZHANG, M.; GUO, N.; LI, Q. Multi-label detection and classification of red blood cells in microscopic images. *In*: 2020 IEEE INTERNATIONAL CONFERENCE ON BIG DATA (BIG DATA), 2020, Atlanta. **Proceedings** […]. Atlanta: IEEE, 2020. p. 4257-4263. DOI: https://dx.doi.org/10.1109/BigData50022.2020.9377782.

RAHMAN, S.; AZAM, B.; KHAN, S. U.; AWAIS, M.; ALI, I.; HUSSEN KHAN, R. J. Automatic identification of abnormal blood smear images using color and morphology variation of RBCS and central pallor. **Computerized Medical Imaging and Graphics**, v. 87, 101813, 2021. DOI: https://doi.org/10.1016/j.compmedimag.2020.101813.

RAKSHIT, P.; BHOWMIK, K. Detection of abnormal findings in human RBC in diagnosing sickle cell anaemia using image processing. **Procedia Technology**, v. 10, p. 28-36, 2013. DOI: https://doi.org/10.1016/j.protcy.2013.12.333.

RAYAPPAN, A. L.; KARTHIK, R. J. Deep learning approach to the normal and abnormal blood cells in human blood smear. **Annals of the Romanian Society for Cell Biology**, v. 25, n. 2, p. 3187-3196, 2021. Available at: https://www.annalsofrscb.ro/index.php/journal/article/view/1298. Accessed on: 4 May 2022.

RAZZAK, M. I.; NAZ, S. Microscopic blood smear segmentation and classification using deep contour aware CNN and extreme machine learning. *In*: 2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), 2017, Honolulu. **Proceedings** […]. Honolulu: IEEE, 2017. p. 49-55. DOI: https://doi.org/10.1109/CVPRW.2017.111.

RIBEIRO, V. H. A.; REYNOSO-MEZA, G. Ensemble learning by means of a multi-objective optimization design approach for dealing with imbalanced data sets. **Expert Systems with Applications**, v. 147, 113232, 2020. DOI: https://doi.org/10.1016/j.eswa.2020.113232.

SAFCA, N.; POPESCU, D.; ICHIM, L.; ELKHATIB, H.; CHENARU, O. Image processing techniques to identify red blood cells. *In*: INTERNATIONAL CONFERENCE ON SYSTEM THEORY, CONTROL AND COMPUTING (ICSTCC), 22., 2018, Sinaia. **Proceedings** […]. Sinaia: IEEE, 2018. p. 93-98. DOI: https://dx.doi.org/10.1109/ICSTCC.2018.8540708.

SARRAFZADEH, O.; DEHNAVI, A. M.; RABBANI, H.; GHANE, N.; TALEBI, A. Circlet based framework for red blood cells segmentation and counting. *In*: 2015 IEEE WORKSHOP ON SIGNAL PROCESSING SYSTEMS (SiPS), 2015, Hangzhou. **Proceedings** […]. Hangzhou: IEEE, 2015. p. 1-6. DOI: https://dx.doi.org/10.1109/SiPS.2015.7344979.

SEN, B.; GANESH, A.; BHAN, A.; DIXIT, S.; GOYAL, A. Machine learning based diagnosis and classification of sickle cell anemia in human RBC. *In*: INTERNATIONAL CONFERENCE ON INTELLIGENT COMMUNICATION TECHNOLOGIES AND VIRTUAL MOBILE NETWORKS (ICICV), 3., 2021, Tirunelveli. **Proceedings** […]. Tirunelveli: IEEE, 2021. p. 753-758. DOI: https://dx.doi.org/10.1109/ICICV50876.2021.9388610.

SHARMA, V.; RATHORE, A.; VYAS, G. Detection of sickle cell anaemia and thalassaemia causing abnormalities in thin smear of human blood sample

using image processing. *In*: 2016 INTERNATIONAL CONFERENCE ON INVENTIVE COMPUTATION TECHNOLOGIES (ICICT), 2016, Coimbatore. **Proceedings** […]. Coimbatore: IEEE, 2016. p. 1-5. DOI: https://dx.doi.org/10.1109/INVENTIVE.2016.7830136.

SYAHPUTRA, M. F.; SARI, A. R.; RAHMAT, R. F. Abnormality classification on the shape of red blood cells using radial basis function network. *In*: INTERNATIONAL CONFERENCE ON COMPUTER APPLICATION AND INFORMATION PROCESSING TECHNOLOGY (CAIPT), 4., 2017, Kuta Bali. **Proceedings** […]. Kuta Bali: IEEE, 2017. p. 1-5. DOI: https://dx.doi.org/10.1109/CAIPT.2017.8320739.

TOHKA, J.; VAN GILS, M. Evaluation of machine learning algorithms for health and wellness applications: a tutorial. **Computers in Biology and Medicine**, v. 132, 104324, 2021. DOI: https://doi.org/10.1016/j.compbiomed.2021.104324.

TOMARI, R.; ZAKARIA, W. N. W.; JAMIL, M. M. A.; NOR, F. M.; FUAD, N. F. N. Computer aided system for red blood cell classification in blood smear image. **Procedia Computer Science**, v. 42, p. 206-213, 2014. DOI: https://doi.org/10.1016/j.procs.2014.11.053.

TYAGI, M.; SAINI, L. M.; DAHYIA, N. Detection of poikilocyte cells in iron deficiency anaemia using artificial neural network. *In*: 2016 INTERNATIONAL CONFERENCE ON COMPUTATION OF POWER, ENERGY INFORMATION AND COMMUNICATION (ICCPEIC), 2016, Melmaruvathur. **Proceedings** […]. Melmaruvathur: IEEE, 2016. p. 108-112. DOI: https://dx.doi.org/10.1109/ICCPEIC.2016.7557233.

TYAS, D. A.; RATNANINGSIH, T.; HARJOKO, A.; HARTATI, S. The classification of abnormal red blood cell on the minor thalassemia case using artificial neural network and convolutional neural network. *In*: INTERNATIONAL CONFERENCE ON VIDEO AND IMAGE PROCESSING (ICVIP 2017), 2017, Singapore. **Proceedings** […]. New York: Association for Computing Machinery, 2017. p. 228-233. DOI: https://doi.org/10.1145/3177404.3177438.

WALTER, W.; HAFERLACH, C.; NADARAJAH, N.; SCHMIDTS, I.; KÜHN, C.; KERN, W.; HAFERLACH, T. How artificial intelligence might disrupt diagnostics in hematology in the near future. **Oncogene**, v. 40, p. 4271-4280, 2021. DOI: https://doi.org/10.1038/s41388-021-01861-y.

XU, M.; PAPAGEORGIOU, D. P.; ABIDI, S. Z.; DAO, M.; ZHAO, H.; KARNIADAKIS, G. E. A deep convolutional neural network for classification of red blood cells in sickle cell anemia. **PLoS Computational Biology**, v. 13, n. 10, e1005746, 2017. DOI: https://doi.org/10.1371/journal.pcbi.1005746.

YUNINGSIH, N.; MUSTIKASARI, M. Anemia classification based on abnormal red blood cell morphology using convolutional neural network. **IOSR Journal of Computer Engineering (IOSR-JCE)**, v. 22, n. 1, p. 22-29, 2020. DOI: https://dx.doi.org/10.9790/0661-2201042229.

ZHAO, J.; ZHANG, M.; ZHOU, Z.; CHU, J.; CAO, F. Automatic detection and classification of leukocytes using convolutional neural networks. **Medical & Biological Engineering & Computing**, v. 55, p. 1287-1301, 2017. DOI: https://doi.org/10.1007/s11517-016-1590-x.