Microscopic Blood Smear Segmentation and Classification using Deep Contour Aware CNN and Extreme Machine Learning

Muhammad Imran Razzak[1], $SaeedaNaz$[2]
[1]King Saud bin Abdulaziz University for Health Sciences
Riyadh, Saudi Arabia
[2]GGPGC.1, Abbottabad, Higher Education Department, KPK Pakistan
mirpak,saeedanaz292@gmail.com

Abstract

Recent advancement in genomics technologies has opened a new realm for early detection of diseases that shows potential to overcome the drawbacks of manual detection technologies. In this work, we have presented efficient contour aware segmentation approach based based on fully conventional network whereas for classification we have used extreme machine learning based on CNN features extracted from each segmented cell. We have evaluated system performance based on segmentation and classification on publicly available dataset. Experiment was conducted on 64000 blood cells and dataset is divided into 80% for training and 20% for testing. Segmentation results are compared with the manual segmentation and found that proposed approach provided with 98.12% and 98.16% for RBC and WBC respectively whereas classification accuracy is shown on publicly available dataset 94.71% and 98.68% for RBC & its abnormalities detection and WBC respectively.

Keywords: *RBC, Blood Sample Analysis, cell morphology, image analysis, ELM, KWFLICM.*

1. Introduction

Health-care image analysis is one of the most sensitive and complicated research area. Blood smear peripheral analysis is one most sensitive and challenging but very important task as it is routine work of every laboratory. In pathological treatment complete blood cell (CBC) count is highly ranked in laboratory test. Morphological analysis of White blood cells (WBC), red blood cells (RBC) and platelets plays a vital role in the diagnosis of various fatal diseases, like aids, cancer and many other. Differential cell count and deviation in cell such as cell shape, size, color or structure represent the abnormality in the cell that can be analyzed through automated analysis of blood cells. Manual blood cell counting is not a reliable as it is very lingering, tedious, time consuming as well as suffers from inter-observer variation. While, it is essential to ensure high level of consistency and capabilities of high level of expertise in haematological laboratories. Whereas modern analytical platforms which are capable of examining thousands of samples per day have prompted researchers to develop techniques for automated blood cells analysis.

The automated blood cells counting and analysis by using computer vision techniques can facilitate the blood cell medical test efficiently and accurately. Most of the automated blood cell counting and analysis systems consist of three major phases including blood cells image segmentation, feature extraction and blood cells classification [5, 13, 21, 20]. The most challenging phase in the automatic blood cell counting is segmentation of complex and varying shapes along with overlapping cells. In this work, contour aware conventional neural network is applied for segmentation of cells. For classification purpose, we have used ELM on CNN features of each individual cell. The rest of the article is organized as follows: In Section 2 , we describe the recent work for blood cell segmentation and classification. In section III, we presented contour aware segmentation and ELM based classification. In section 5, detail analysis is performed and compared the result with other approaches.

2. Related Work

In the recent decade, the interest of researchers is increasing in the development of algorithms for automated analysis of medical images and especially blood cell image. Researchers have made a lot of improvements by the integration of artificial intelligence, image processing, and pattern recognition and computer vision techniques to the blood cell analysis. There are several automated systems developed for medical diagnosis that are helpful for doctors to detect diseases particularly in red blood cells(RBC) and white blood cells(WBC), which is vital for curing the dis-

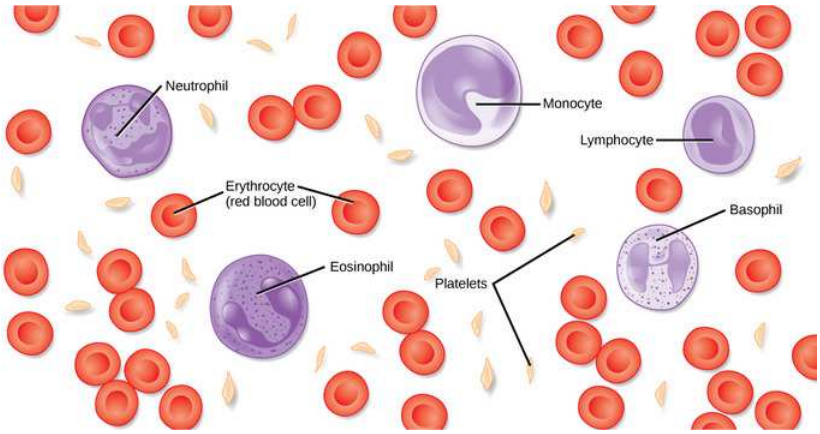

Figure 1. Blood constituents

ease.

In literature the researchers have shown increasing interest in the development of various algorithms for blood cell analysis. The quantification of red blood cells (RBC) is performed to find the size and shape of the cells. The RBC count is a perfect indicator of diseases such as anaemia, cancer, leukaemia, aids, renal tumour and many other diseases. Generally, microscopic based evaluation technique is used as the standard method to perform blood cells segmentation and analysis [20] [25] [16] [15] shirazi2015accurate [5]. Tomari et al presented computer aided system for the automated analysis of RBC. They have extracted RBC regions from the background by applying thresholding on the green channel of RGB image [24] . They have performed image enhancement by morphological filters. In automated blood cell analysis image segmentation plays key role. The image segmentation directly affect the next processing steps of feature extraction and classification. In case of human blood cell image segmentation numerous algorithms were proposed to obtain better results but still efforts are required to improve the accuracy of segmentation [25][19]. Huge amount of the diagnostic information can be extracted by the detection of size and shape of RBC. The count of red blood cells in a particular area in blood sample and its ratio the whole area can also be used for the detection of disease [8]. In medical field, the analysis of RBC yields information relevant to pathological diseases. The deformability of shape of RBC has a connection with the relevant diseases like Myalgic Encephalomyelities (ME) and many other diseases. That is why the accurate diagnosis is vital in identifying the correct treatment [25] [26]. The traditional blood cells counting in laboratory is called hemocytometer or haematology analyser. These procedures fully rely on the expertise skill to count the cells by reviewing the sample under microscope. In automated blood cells counting the segmentation of overlapped cells is a challenging task. In literature most of the previously proposed segmentation techniques of blood cells are suffering from errors and inconsistencies. Segmentation of RBC from cytoplasm region of WBC is more difficult due to the complex nature of the blood cells and colour similarity. For this purpose statistical model based approaches are the most successful and practical for the boundaries detection and texture distribution [4]. Due to complex background it is difficult to clearly detect boundaries [1]. Different approaches for RBC feature extraction are used in literature. Morphological operators like erosion and dilation were used to enhance the boundary of objects [3]. Morphological operators are useful to preserve the size and shape features. Ross et al. have exploited local thresholding of levels to enhance the early detection. Their focus is on classification problem to distinguish among parasite species like malaria, vivax or oval, and therefore; they have used two phase machine learning strategy for this purpose [18]. For classification they have used colour and texture features in the first phase. In the second phase they have used shape and relative size features along with number of parasites per cell and relative position of the parasite within the cell. Various methods are used for the classification of blood cells in literature. Most commonly found methods were Bayesian classifier which were used by Sinha et.al [22] and Ghosh et al.[9] . They have extracted different features like area ratio, number of nucleus lobs, average value of colour components, and have achieved 82% accuracy in classifying leukocytes. K-Nearest neighbour was used by Piuri eta al. and achieved with accuracy of 70.6% [14]. Tomari have used ANN for the classification or RBC and have achieved 82% accuracy [24].A new classification technique extreme learning machine is introduced in literature which is suitable for single layer classification. Several authors have used ELM as a classifier for the classification of different images processing task and provided promising results [23] [12] [27] [11].

3. Blood Smear Analysis

Automated screening of prepared blood films consist of two main phases: segmentation of cells into individual and classification of segmented cells in to appropriate class. In order to meet the challenges of segmentation, in this work, we have applied contour aware fully conventional neural network chen for efficient segmentation of cells. Deep contour aware notwork. In the second section, we introduce CNN features for WBC and plus several other features for RBC and forward to ELM for classification.

3.1. Contour Aware Segmentation

Blood cells exhibit wide variations of cell morphology and size that make them difficult to be segmented accurately especially due to overlapping. To address this issue, several methods have been proposed as discussed in section II. However, these methods are not efficient due to large variations, overlapping of cells. Contour aware conventional network [7] generate multi-scale feature representation to deal with large variations and overlapped cells. Recently DCAN (contour aware segmentation) showed promising result and is the winner

Segmentation of touching and overlapped cells could be enhanced by using contextual features and CNN with auxiliary supervision. As segmentation accuracy is affected due to complexity of cells in traditional CNN as network with single receptive field size cannot satisfactorily deal with the large variations of blood smear images. Thus, leveraging contextual information could improve the segmentation performance by recognizing its structure. In this work, we have used multilevel contextual feature representation that include contextual information. The classification scores from FCN are established based on the intensity information from the given receptive field. By harnessing the multilevel contextual features with auxiliary supervision, the network can produce good probability maps of gland objects. However, its still quite hard to separate the touching glands by leveraging only on the likelihood of gland objects due to the essential ambiguity in touching regions. This is rooted in the down-sampling path causing spatial information loss along with feature abstraction. The boundary information that make cure somehow provides good complementary cues for splitting objects

3.2. Morphological Analysis

Differential cell count and deviation in cell such as cell shape, size, color or structure represent the abnormality in the cell are the main objective of automated analysis of blood cells. In the first phase,we have used contour aware CNN for segmentation of individual cells. Each segmented cell is then extracted to remove the background cells. At this stage, color descriptors are enough to segregate cells into RBC and WBC, thus based on the color intensity features, cells are segregated into two types of blood constituents. In order to classify WBC into five sub types, we have extracted minimum enclosing rectangle of segmented cell as consider as region. For each cropped region, we have extracted CNN features computed at early stage and forward to ELM for classification into five sub types. In order to classify RBC into normal and identify abnormalities, we have extracted several features i.e. area, circularity, parameter, centroid, medial axis ratio, cell deform ratio, roughness, regularity, uniformity and coarseness etc. Extracted features are forwared to ELM for classification.

To train the ELM we only need to adjust the hidden neurons. We have chosen sigmoid function for the activation function. We have determined number of neurons through cross validation. In case of WBC segmentation, we forwarded CNN features to ELM for classification. CNN based features are extracted based on minimum enclosing rectangle of cell. Whereas for RBCs, extracted features are forwarded to ELM classifier for identification to classify into normal and further abnormalities.

4. Result and Discussion

In this work, automatic system used to help the diagnosis of some important blood diseases is developed. To evaluate the performance of proposed approach, experiment is conducted on blood cell images dataset ALL-IDB. Constituents of blood smear images are RBCs, WBCs and platelets. Differential cell count is time consuming and prone to error. This process is automated by segmenting the RBCs from other blood constituent and analyzing each segmented RBC.

Figure 2 describe the blood cell segmentation and classification framework. Results shows that deep contour aware network is capable of segmenting the touching as well as overlapped cells efficiently as compare to FCN based segmentation as shown in figure 3. After the segmentation of individual, the next task is the classification into blood consistent.

Each identified cell is segregated from other neighbouring cells and based on minimum area covered by object is segmented using minimum enclosing rectangle. As for WBC classification into sub types, we have use CNN based extreme machine leaning. To optimize the algorithm performance, we have extracted the CNN features from already computed CNN based on region. In our case, region is the MER of each identified cell. For RBC classification, we have used texture and shape defining featuure base on the fact that texture feature and shape features are very important for the classification and differential RBCs into normal cells and other abnormalities. Healthy RBC shape is biconcave with central pale area, thus any deviation in cell shape

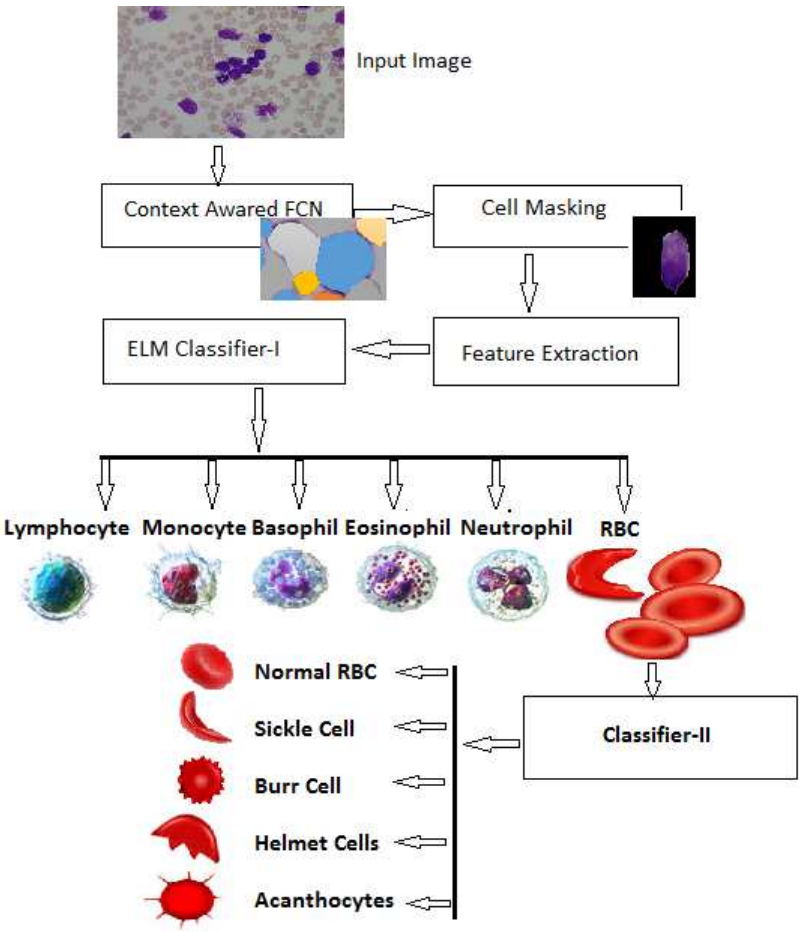

Figure 2. Proposed Methodology

(size, volume etc.) or texture deviation (color etc.) would result in abnormalities in cells, that will further classified into abnormalities types as shown in figure below.

For implementation purpose we have used Caffe: a open source library. In our work, network randomly corp 450 X 450 region form input blood smear image as input image and give prediction mask of each individual object as output. In order to label the contours, we have used [17] and [21] to segment the cells and then counter check by pathologist to validate. After review of each segmented cell, we have used disk filter to enlarge the contour.

4.1. Dataset

To evaluate the proposed classification method, we have used ALL-IDB blood cell dataset [31]. ALL-IDB is public and freely available dataset of microscopic blood images. The dataset is especially developed for the evaluation and comparison of blood cell segmentation and classification. It is composed of 108 JPG format images with 24 bit color depth, resolution 2592 x 1944. Dataset consist of about 39000 blood elements that are captured captured with an optical laboratory microscope coupled with a Canon PowerShot G5 camera. As this detest is collected for mainly for Leukemia, thus in order to include RBC abnormalities, we have included several images publicly available on web such as DPDx, ASH image bank and other images available on Google . Statistics of collected dataset are described in table 2. We have used keywords ”sickle cells microscopic”, ”eliptocytes cells microscopic image” , ”burr cells image” etc. to search images and extracted 32, 29, 23, 25, 23 and 23 for sickle, eliptocytes and ovalocytes, helmet and acathocytes and burr cell. After careful observation, we have selected the images that meet criteria and discarded rest of the images as discussed 2. We have divided the dataset into training (approximately 80%) and testing (approximately 20%). The extracted images from web contribute of 46.2% and consist of approximately 25000 blood cells. Thus our total dataset consist of 64000 blood cell.

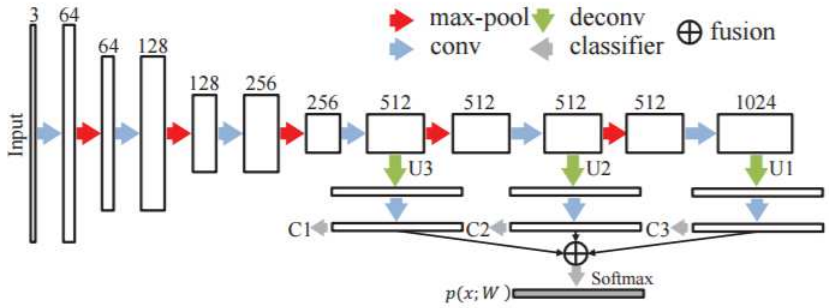

Figure 3. Contour Aware CNN [7]

Table 1. Overall Segmentation result

Target	TP	FP	TN	FN	Recal.	Prec.	Acc.
RBC.	4219	21	724	42	99.17	98.36	98.12
WBC	522	9	811	23	95.93	98.30	98.6

Table 2. Dataset Statistics

Dataset	Images	Training	Testing
ALL_IDB	108	88	20
Sickle Cell	21	16	5
Elliptocytes and Ovalocytes	22	17	5
Helmet	17	12	5
Acanthocytes	18	12	6
Burr cells	15	11	4
Total	201	156	45

Table 3. Overall RBC classification result

Target	TP	FP	TN	FN	Prec.	Acc.
Normal	350	36	541	24	90.67	93.69
Sickle	99	20	368	35	83.19	89.46
Elli.&Ova	100	29	387	25	77.51	90
Acantho.	75	24	338	38	75.75	86.94
Burr	80	28	346	33	74	87.47
Helmet	85	23	371	30	78.7	89.58
Overall	**789**	**160**	**2351**	**185**	**83.14**	**90.10**

Table 4. Overall WBC classification result

Target	TP	FP	TN	FN	Prec.	Acc.
Neutro.	181	3	195	4	98.36	98.17
Eosino.	47	1	331	3	97.91	98.6
Basophil	29	1	348	1	99.4	90
Lympho.	67	3	329	2	95.75	98.8
Monocyte	46	2	328	3	95.8	98.68

4.2. Evaluation Criteria

To evaluate the proposed method, commonly used evaluation criteria for quantitative measurement such as classification False Positive, True Negative, False Negative, True Negative, sensitivity and specificity were analyzed. On the basis of these parameters we have calculated precision, recall and over all accuracy shown in equation below.

$$Precision = TP/(Tp + Fp) \quad (1)$$

$$Recall = TP/(Tp + FN) \quad (2)$$

where TP, FN, TN, and FP are the number of true positives, false negatives, true negatives and false positives, respectively. AUC is one of the most popular methods for evaluating the performance of the binary classifier. A perfect classifier provides an AUC of 1.

5. Conclusion

In this work, we have presented efficient contour aware segmentation approach based based on fully conventional network whereas for classification we have used CNN based extreme machine learning. We formulate this challenge segmentation and classification problem by using contour aware CNN and Faster-Region CNN. In our case, We have used Deep Contour aware CNN for region extraction. We have evaluated system perfor-

Table 5. Comparative analysis

Ref	Dataset	Prep.	Segm.	Classifi.	Acc
[20]	ALL_IDB	Morph	ad.Otsu	SVM, ANN	85%
[24]	ALL_IDB	Morph	ad.Otsu,	NN	83%
[10]	400 Images		Canny	Hybrid	88%
[19]	UTHM	Morph	Watershed	ANN	87%
[26]	100 Images	Median	Contour	SVM	
[2]	UTM	Morph	CHT	Manual	91%
[6]	All-IDB	Hist. equ	Thres.	KNN	93%
[17]	All-IDB+other	image norm.	KWFLICM,	ELM	91.10%
Proposed	All-IDB+other	image norm.	DCAN	FRCN+ELM	95.10%

mance based on segmentation and classification on publicly available dataset. Experiment was conducted on 64000 blood cells and dataset is divided into 80% for training and 20% for testing. Segmentation results are compared with the manual segmentation and found that proposed approach provided with 98.12% and 98.16% for RBC and WBC respectively whereas classification accuracy is shown on publicly available dataset 94.71% and 98.68% for RBC & its abnormalities detection and WBC respectively. System performance showed that segmentation result on both datasets are fast, robust, efficient and coherent. Meanwhile, the classification and categorization of cells into normal and abnormal cells showed high sensitivity with slightly better results.

References

[1] Robiyanti Adollah, MY Mashor, NF Mohd Nasir, H Rosline, H Mahsin, and H Adilah. Blood cell image segmentation: a review. In *4th Kuala Lumpur International Conference on Biomedical Engineering 2008*, pages 141–144. Springer, 2008.

[2] Yazan M Alomari, Siti Norul Huda Sheikh Abdullah, Raja Zaharatul Azma, and Khairuddin Omar. Automatic detection and quantification of wbcs and rbcs using iterative structured circle detection algorithm. *Computational and mathematical methods in medicine*, 2014, 2014.

[3] Jesus Angulo and Georges Flandrin. Automated detection of working area of peripheral blood smears using mathematical morphology. *Analytical cellular pathology*, 25(1):37–49, 2003.

[4] Rongtai Cai, Qingxiang Wu, Rui Zhang, Lijuan Fan, and Chengmei Ruan. Red blood cell segmentation using active appearance model. In *Signal Processing (ICSP), 2012 IEEE 11th International Conference on*, volume 3, pages 1641–1644. IEEE, 2012.

[5] Tamalika Chaira. Accurate segmentation of leukocyte in blood cell images using atanassov's intuitionistic fuzzy and interval type ii fuzzy set theory. *Micron*, 61:1–8, 2014.

[6] Niranjan Chatap and Sini Shibu. Automatic blasts counting for acute leukemia cells based on blood samples. *International Journal of Computer Science & Applications (TIJCSA)*, 3(01), 2014.

[7] Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis*, 36:135–146, 2017.

[8] Cecilia Di Ruberto, Andrew Dempster, Shahid Khan, and Bill Jarra. Analysis of infected blood cell images using morphological operators. *Image and vision computing*, 20(2):133–146, 2002.

[9] Madhumala Ghosh, Devkumar Das, Chandan Chakraborty, and Ajoy K Ray. Automated leukocyte recognition using fuzzy divergence. *Micron*, 41(7):840–846, 2010.

[10] Sang Yun Lee, Hyun Joo Park, Catherine Best-Popescu, Seongsoo Jang, and Yong Keun Park. The effects of ethanol on the morphological and biochemical properties of individual human red blood cells. *PloS one*, 10(12):e0145327, 2015.

[11] Javier López-Fandiño, Pablo Quesada-Barriuso, Dora B Heras, and Francisco Argüello. Efficient elm-based techniques for the classification of hyperspectral remote sensing images on commodity gpus. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6):2884–2893, 2015.

[12] Abdul Adeel Mohammed, Rashid Minhas, QM Jonathan Wu, and Maher A Sid-Ahmed. Human face recognition based on multidimensional pca and extreme learning machine. *Pattern Recognition*, 44(10):2588–2597, 2011.

[13] Sedat Nazlibilek, Deniz Karacor, Tuncay Ercan, Murat Husnu Sazli, Osman Kalender, and Yavuz Ege. Automatic segmentation, counting, size determination and classification of white blood cells. *Measurement*, 55:58–65, 2014.

[14] Vincenzo Piuri and Fabio Scotti. Morphological classification of blood leucocytes by microscope images. In *Computational Intelligence for Measurement Systems and Applications, 2004. CIMSA. 2004 IEEE International Conference on*, pages 103–108. IEEE, 2004.

[15] Muhammad Imran Razzak. Automatic detection and classification of malarial parasite. *International Journal of Biometrics and Bioinformatics (IJBB) Vol*, 9:1, 2015.

[16] Muhammad Imran Razzak and Bandar Alhaqbani. Automatic detection of malarial parasite using microscopic blood images. *Journal of Medical Imaging and Health Informatics*, 5(3):591–598, 2015.

[17] Muhammad Imran Razzak and Saeeda Naz. Automated red blood cell segmentation and classification of red cell disorders using extreme learning machine and kernel weighted fuzzy local information c-means. *CVPR17, Submitted*.

[18] Nicholas E Ross, Charles J Pritchard, David M Rubin, and Adriano G Duse. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Medical and Biological Engineering and Computing*, 44(5):427–436, 2006.

[19] J MISWAN Sharif, MF Miswan, MA Ngadi, Md Sah Hj Salam, and Muhammad Mahadi bin Abdul Jamil. Red blood cell segmentation using masking and watershed algorithm:

A preliminary study. In *Biomedical Engineering (ICoBE), 2012 International Conference on*, pages 258–262. IEEE, 2012.

[20] Syed H Shirazi, Arif Iqbal Umar, Saeeda Naz, and Muhammad I Razzak. Efficient leukocyte segmentation and recognition in peripheral blood image. *Technology and Health Care*, 24(3):335–347, 2016.

[21] Syed Hamad Shirazi, Arif Iqbal Umar, Nuhman Ul Haq, Saeeda Naz, and Muhammad Imran Razzak. Accurate microscopic red blood cell image enhancement and segmentation. In *International Conference on Bioinformatics and Biomedical Engineering*, pages 183–192. Springer, 2015.

[22] Neelam Sinha and AG Ramakrishnan. Automation of differential blood count. In *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, volume 2, pages 547–551. IEEE, 2003.

[23] Jiexiong Tang, Chenwei Deng, and Guang-Bin Huang. Extreme learning machine for multilayer perceptron. *IEEE transactions on neural networks and learning systems*, 27(4):809–821, 2016.

[24] Razali Tomari, Wan Nurshazwani Wan Zakaria, Muhammad Mahadi Abdul Jamil, Faridah Mohd Nor, and Nik Farhan Nik Fuad. Computer aided system for red blood cell classification in blood smear image. *Procedia Computer Science*, 42:206–213, 2014.

[25] Joost Vromen and Brendan McCane. Red blood cell segmentation from sem images. *Image and Vision, New Zealand*, 2009.

[26] Ruihu Wang. Red blood cell surface segmentation based on shape reconstruction and statistics feature extraction. *Journal of Computers*, 8(4):1083–1089, 2013.

[27] Weiwei Zong, Hongming Zhou, Guang-Bin Huang, and Zhiping Lin. Face recognition based on kernelized extreme learning machine. In *International Conference on Autonomous and Intelligent Systems*, pages 263–272. Springer, 2011.