Original Article

The Best Texture Features for Leukocytes Recognition

Omid Sarrafzadeh[1,4], Alireza M. Dehnavi[2,4], Hossein Y. Banaem[4], Ardeshir Talebi[3], Arshin Gharibi[4,5]

[1]Department of Biomedical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, [2]Department of Biomedical Engineering, School of Advanced Technologies in Medicine, [3]Department of Pathology, School of Medicine, [4]Medical Image and Signal Processing Research Center, [5]Department of Biomedical Engineering, PhD student at UNM, University of New Mexico, Albuquerque

Abstract

Background: Differential counting of white blood cells (WBCs or leukocytes) is a common task to diagnose many diseases such as leukemia, and infections. An accurate process for recognizing leukocytes is to evaluate a blood smear under a microscope by an expert. Since, this procedure is manual, time-consuming and tedious, making the procedure automatic would overcome these problems. In an automated CAD (Computer-Aided-Design) system for this purpose, a crucial module is leukocytes recognition. In this paper, we are looking for the best features in order to recognize five types of leukocytes (Monocyte, Lymphocyte, Neutrophil, Eosinophil and Basophil) from microscopic images of blood smear in an automated cell counting system. **Methods:** In this work, we focus on the texture features and seven categories: GLCM features, Haralick features, Spectral texture features, Wavelet-based features, Gabor-based features, CoALBP and RICLBP are analyzed to find the best features for leukocytes detection. The best features of each category are selected using stepwise regression and finally three well-known classifiers called K-NN, LDA and NB are utilized for classification. **Results:** The proposed system is tested on a self-provided dataset composed of 200 cell images. In our experiments, to evaluate the process, the accuracy of each leukocyte type and the mean accuracy are computed. RICLBP features achieved the best mean accuracy (85.53%) for LDA classifier. **Conclusions:** In our experiments, although the maximum mean accuracy (85.53%) went with RICLBP features, but the accuracies of all five leukocyte types weren't maximized for RICLBP features. This result directs us to design and develop a system based on multiple features and multiple classifiers to maximize the accuracies even for each individual cell type in our future work.

Keywords: *Automatic leukocytes recognition, best texture features, blood smear, computer-aided design (CAD) system, microscopic images*

Introduction

Differential counting of white blood cells (WBCs or leukocytes) is the first step and a routine task in laboratories to diagnose many diseases, such as leukemia, anemia, and numerous infections. As thousands of cells must be assessed by the pathologist to have a precise diagnosis, and this procedure is manual, time-consuming and tedious, making the procedure automatic would overcome these problems.[1,2] In an automated CAD (computer-aided design) system in this regard, a vital module is leukocytes recognition. Many efforts have been done by researchers to meet different demands of a CAD system. Some method focus more on automatic segmentation of cells,[3,4] some on classification part,[5,6] and others covers the entire process.[7-10] A recent survey on the image analysis of microscopic images is done by Saraswat and Arya[1] for leukocytes identification. Also a new review on classification techniques of WBCs in microscopic blood images is presented by Rawat *et al.*[11] A system is proposed by Huang *et al.*[12] for leukocyte recognition. In their method, the nuclei of WBCs are first segmented, and geometry and Gray-Level Co-occurrence Matrix (GLCM)-based texture features are then extracted from nuclei and finally, a *K*-means clustering approach is used for classification. Since the cytoplasms characteristics differ through WBCs' types, their method utilized just the information of nuclei not the entire cell. An automatic method is proposed by Rezatofighi and Soltanian-Zadeh[13] for the detection and recognition of leukocytes. They utilized Gram–Schmidt-based algorithm as well as a snake model to segment the cells' nuclei and cytopasms. Then various geometry and texture features are extracted from the cells, and next, features are reduced using a Sequential

How to cite this article: Sarrafzadeh O, Dehnavi AM, Banaem HY, Talebi A, Gharibi A. The Best Texture Features for Leukocytes Recognition. J Med Sign Sens 2017;7:220-7.

Address for correspondence:
Dr. Omid Sarrafzadeh, Department of Biomedical Engineering, Mashhad Branch, Islamic Azad University, Ostad Yosefi Street, Shahrake Gharb, Mashhad, Khorasan Razavi, Iran.
E-mail: o.sarrafzade@gmail.com

Website: www.jmss.mui.ac.ir

Forward Selection method. Finally, the cells are classified using ANN and SVM classifiers. Habibzadeh *et al.*[14] proposed a convolutional neural network (CNN) framework for differential counting of WBCs, which is useful for low-resolution images, and compared their system with SVM-based classifiers. A texture approach is proposed by Sabino *et al.*[15] to identify WBCs. They separated nucleus and cytoplasm of a cell, and then, extract statistical features from each segment based on GLCM matrix. In another work performed by Sabino *et al.*,[16] color, morphology, and texture features are combined to recognize leukocytes. They utilized a Naive Bayesian (NB) classifier to classify the cells. In a recent work done by Zhao *et al.*,[17] granularity features and SVM classifier are used to separate basophil and eosinophil from other leukocytes. Then, a CNN model is utilized to extract features from other leukocytes in a high level and a random forest is carried out to recognize residual WBCs. In another recent study by Li and Cao,[18] a new island clustering texture is proposed to automatically recognize WBCs. They combined the texture features with the geometry features of WBCs to improve the performance of their system. In most previous works, a combination of different color, geometry, and texture features are utilized to recognize WBCs. In addition, in some study, nucleus and cytoplasm of cells are first separated and then features are extracted individually from each part. In the routine process of blood smear preparation in clinics or laboratories, we often don't face a typical form of cells, and leukocytes are deformed due to staining and the pressure of adjacent cells such as red blood cells. So the geometry features aren't useful features in these cases. In addition, the color of WBCs varies among different samples because of unstable staining, technician's skills, lightening, and so on. Therefore, color features are not a proper feature in most occasions. As a result, any changes in geometry and color of the cells can significantly affect the geometry and color features, respectively, and this is undesirable. On the other hand, geometry and color changes of cells have the least effect on texture features. In addition, there are more powerful tools for texture analysis, which are invariant to image rotation and robust to illumination changes. That's why we focus on the texture features in this work that is not covered previously by researchers. In this study, we focus on seven categories of texture features: GLCM features, Haralick features, Spectral texture features, Wavelet-based features, Gabor-based features, Co-occurrence of Adjacent LBP (CoALBP), and Rotation Invariant Co-occurrence of LBP (RICLBP), and try to find the best features and best category for the recognition of leukocytes. The paper is organized as follows. Section two introduces the methodology of our work. The dataset used in this work is explained in this section. Section three presents the results of our study and finally the work is concluded in Section four.

Materials and Methods

In this paper, we aim to find the best texture features for the classification of WBCs. The details properties about different types of WBCs are well gathered in[19] and here, we just summarize leukocytes' characteristics. Leukocytes are generally divided into two groups. First group is named granulocytes, in which cytoplasms contain granules and nuclei are segmented. The other group is known as agranulocytes whither cytoplasms have no granules and nuclei are mononuclear. The properties of different types of WBCs are explained in Table 1. As it is seen from Table 1, several features such as geometry, texture, and color features are utilized by experts to differentiate various types of leukocytes. However, in clinical experiments, during the preparation of blood smears, the appearance of cells may get away from typical forms due to artifacts, irregular staining, lightening, and so on. Figure 1 shows some examples of these variations. It is seen from Figure 1 that geometry and color properties of cells might be extremely varies in routine slide preparation. Although texture properties might be changed, but there are strong descriptors to capture texture characteristics. In addition, the geometry changes of cells due to the pressure of adjacent cells highly affect the geometry descriptors belong to the same cell's

Table 1: The properties of different types of white blood cells (or leukocytes)

Leukocytes		Characteristics	Typical cells
Granulocyte	Neutrophil	~60% in blood, nucleus (2–5 lobes, stains dark purple), cytoplasm (transparent, faintly pink-purple granules), size (12–16 μm)	
	Eosinophil	~3% in blood, nucleus (2 lobes connected by a thin strand, stains blue), cytoplasm (full of pink-orange granules), size (14–16 μm)	
	Basophil	~1% in blood, nucleus (2–3 lobes, stains purple), cytoplasm (rich in dark purple granules hiding nucleus), size (14–16 μm)	
Agranulocyte	Monocyte	~6% in blood, nucleus (kidney or horseshoe shaped, stains pale bluish-violet, fine chromatin), cytoplasm (stains blue-gray with tiny granules), size (14–20 μm)	
	Lymphocyte	~30% in blood, nucleus (large round or oval shaped, dark staining), cytoplasm (tiny size, stains pale blue), size (8–15 μm)	

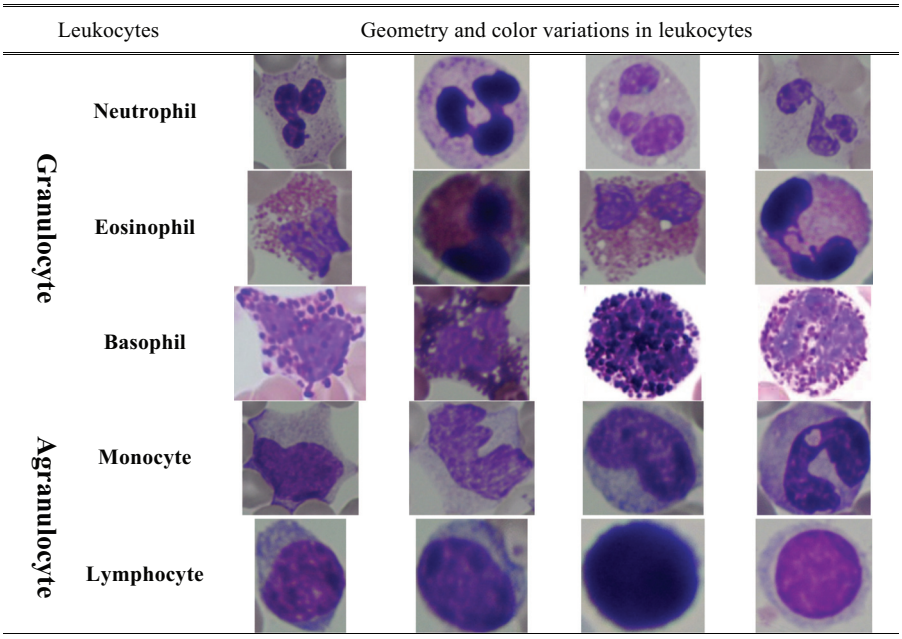

Figure 1: Geometry and color variations in different types of leukocytes

type which is not desirable, while will slightly affect the texture descriptors. Likewise, color variations due to staining greatly influence the color descriptors but insignificantly alter texture features. Furthermore, there are powerful texture features that are robust to illumination variations. Therefore, in this study, we've focused on texture descriptors to find the best ones for the recognition of different types of leukocytes. The details description of the proposed system is described in the following.

Dataset

To test and evaluate our system, a dataset has been provided with the cooperation of Medical Image & Signal Processing Research Center and Department of Pathology at Isfahan University of Medical Sciences. Blood smear of 10 patients are prepared with Giemsa staining and images from slides are captured with a Nikon V1 camera mounted on a Nikon ECLIPSE 50i microscope with the magnification of 100×, and finally, images are saved with JPEG format and 2448 × 3246 resolution. Forty cells of each type and in total 200 cells are gathered from images. It is noted that as our goal is to assess the features and find the best ones, the cells are segmented manually by an expert to overcome the difficulties of automatic segmentation. As Basophils are the lowermost component in the blood (below 1%), and we didn't have enough Basophil, to increase the number of Basophils to 40 and make the distribution of cells in each group uniform, we utilized the microscopic images of Basophils from the Web to complete our dataset. In the following, the steps of our proposed method are explained.

Preprocessing

As the illumination, brightness, and size of images vary in our dataset, some preprocessing steps are applied to each image to normalize it. First, the image in RGB color space is converted to grayscale. Next, grayscale image is resized to 512 × 512 pixels to be normalized in dimension. Resized image is then filtered with a median filter of 5 × 5 window size to decrease the noise while keeping the edges of the image. Next, the filtered image is enhanced using Contrast Limited Adaptive Histogram Equalization algorithm.[20] Then the corresponding mask is imposed on the enhanced image and the intensities are changed so that 1% of data is saturated at the lowest and highest intensity of the enhanced image, respectively. This resultant image is called $\boldsymbol{I_n}$ referring to final normalized image and is utilized for next processing.

Feature extraction

As discussed previously, we are focused on texture features to find the best ones for leukocytes classification. Seven texture features are studied in our work: GLCM features, Haralick features, Spectral texture features, Wavelet-based features, Gabor-based features, CoALBP, and finally RICLBP features which are introduced in the following.

GLCM features

These features are extracted from GLCM[21] of the image. GLCM shows the distribution of cooccurring intensities in an image at given specific distance (D) and angle (θ). Some parameters are needed to create GLCM that are named: D, θ, and NL. D and θ determine the offset and angle between two adjacent pixels, and NL denotes the number of gray-levels in the image. The values of 4, 6, and 8 are considered for NL, and for each value, the intensities of the image are first scaled to NL. Four angles {0°,45°,90°,135°} are chosen for θ to take into account the rotations of the image. For each angle, a GLCM matrix is produced. For an image with NL intensity values, a $NL \times NL$ matrix is

generated.These four created GLCM matrices are averaged to have a GLCM matrix with $NL \times NL$ dimensions. This matrix is finally converted to a vector to create GLCM feature vector. In our study, distance parameter, D, is swept from 1 to 30.

Haralick features

Haralick features are first proposed by Haralick *et al.*[21] These features are statistical descriptors depicting overall texture of an image that are computed using GLCM matrix. In our work, GLCM matrix is computed the same as introduced beforehand. Here, the values of NL are also chosen the same, meaning 4, 6, and 8. In addition, the parameter D is changed from 1 to 30. For given parameters D and NL, a GLCM matrix is created the same as before and first 13 Haralick features[21] are computed. So our proposed Haralick feature vector for a given cell has 13×1 dimensions regardless of the value of NL.

Spectral features

Spectral measurements introduced in[22] are a way to capture the texture of an image based on the Fourier spectrum. The spectrum is considered in polar coordinates to produce function $S(r, \theta)$. For each direction θ, and for each frequency r, $S(r, \theta)$ will be a one-dimensional function $S_\theta(r)$ and $S_r(\theta)$, respectively. Global spectral textures are then computed using by $S(r) = \sum_{\theta=0}^{\pi} S_\theta(r)$ and $S(\theta) = \sum_{r=1}^{Ro} S_r(\theta)$, where R_0 is the radius of a circle centered at the origin. For computing, $S(r)$, is considered between 0 and 180° in increments of 1°, and R_0 is chose by min (row, col) where *row* and *col* are the row and column of the image, respectively. Functions $S_r(\theta)$ and $S_\theta(r)$ are normalized by dividing to the maximum value and the following criteria are calculated: mean, median, mean–median, variance, energy, entropy, and the location of the highest value. Also all above criteria except the last one are computed from the spectrum. So the final spectral feature vector will be of dimension 20.

Wavelet features

Wavelet transform provides a multiresolution analysis that is widely used for texture studies.[23,24] To extract wavelet features, different multilevel wavelet decompositions are considered. The energy of approximation and details subbands are computed using different wavelet filters and different levels. Numerous wavelet filters are considered such as Haar, Daubechies (db2, . . . , db10), Discrete Meyer, Symlets (sym2, . . . , sym10), Coiflets (coif1, . . . , coif5), Biorthogonal (bior1.1, . . . , bior6.8), and Reverse Biorthogonal (rbio1.1, . . . , rbio6.8). In addition, wavelet decompositions with different levels (1, 2, 3, and 4) are calculated. Since for the level L_i, $(3 \times L_i + 1)$ subbands exist, final wavelet feature vector will be of size $3 \times \sum_{i=1}^{4} L_i + 4 = 34$.

Gabor features

Gabor filters are directional filters that are greatly exploited to extract features invariant to rotation, translation, scale, and illumination. A recent framework to extract Gabor-based features is proposed by Haghighat *et al.*[25] and is also considered in this study to extract Gabor features. Gabor filters are designed with 1, 2, . . . , and 5 scales, 6, 8, and 10 orientations with the sizes of 39×39. For the number of scales, s and the number of orientations, d, the feature vector proposed by Haghighat *et al.*[25] have the dimension of $\text{row} \times \text{col} \times s \times d$. They utilized down-sampling at first stage to reduce the dimension. In our study, the columns and rows are down-sampled with the factors of $2^i, i (= \{0, 1, \ldots \cdots, 8\}$ and in each stage, the mean, variance, energy, and entropy of down-sampled Gabor vectors are computed as final Gabor features. So the final Gabor feature vector will be of size 36.

CoALBP features

CoALBP features proposed by Nosaka *et al.*[26] are texture descriptors for an image based on spatial co-occurrence of local binary patterns (LBPs). LBP takes into account only the differences intensities between a central pixel and its neighbors, and is robust to illumination variations. However, in the process of producing LBP histogram, the spatial relations between LBPs are not considered. CoALBPs attains the co-occurrence of LBPs using four autocorrelation matrices and each matrix has $N_p \times N_p$ dimensions. $N_p$is the number of all probable LBPs, $N_p = 2^N$, where N is the number of neighbor pixels in LBP. Four autocorrelation matrices are finally merged to create a feature vector with $4N_p^2$ dimension. To have less computational time, two configurations are considered for LBP, LBP(+) describing four horizontal and vertical neighbor pixels, and LBP $(\times)$ depicting four diagonal neighbor pixels. So by these definitions, for each LBP(+) and LBP $(\times)$ we have: $N = 4$ and $N_p = 16$, and finally the CoALBP feature vector have $4N_p^2 = 1024$ dimension. We also considered LBP(■) which indicates the average of LBP(+) and LBP $(\times)$ features, so LBP(■) is also a feature vector of 1024 dimension. Two parameters, s and r, are needed to produce CoALBP features, where s is the scale of LBP and r is the displacement between a LBP pair. We chose the range of 1 to 10 for s and range of 1 to 30 for r.

RICLBP features

RICLBP[27] is a LBP-based feature and is a developed version of CoALBP introduced previously. As by rotating a target object, CoALBP features cannot ably change, CoALBP is extended to RICLBPs by utilizing the idea of rotation equivalence class. RICLBP is invariant to image rotations and produces a final feature vector with the dimension of $N_p(N_p + 1)/2$, where N_p has the same definition as explained for CoALBP.[27] Four neighbor pixels are chosen, so $N_p = 2^N = 16$, and then the final feature vector will have 136 dimension. The same as CoALBP, two parameters, s and r, are required to generate RICLBP features. We also chose the range of 1 to 10 for s and range of 1 to 30 for r.

Feature selection

To select the best features, stepwise regression (SWR)[28] is utilized. SWR is an analytical technique that considers a multilinear regression model and then adds and removes features from the model in a stepwise manner dependent upon their statistical significance. The algorithm starts with an initial model and in a systematic process, adds the most important feature or removes the less meaningful feature throughout each step. The *P* value of a *F*-statistic is used to test the model. In each step, if any feature not in the model has *P* values lower than *P-enter*, the feature with the minimum *P* value is added to the model. Contrariwise, if any feature in the model has *P* values higher than *P-remove*, the feature with the maximum *P* value is removed from the model. The steps proceeds and the process stop while no further improvement achieves in next step. In our study, *P-enter* and *P-remove* are chosen as 0.05 and 0.1, respectively. Stepwise methods might be locally optimal due to initial model and the order of entering/removing of features. To improve the performance of the method and converge, the model to a global optimum, we utilized two-fold cross-validation during stepwise analysis, and the process is repeated 200 times and selected feature set is voted in each iteration. Finally, the feature set with maximum score is considered as the best feature set.

Classification

To classify the leukocytes, the following three well-known classifiers are chosen: K-Nearest Neighbors (K-NN),[29] Linear Discriminant Analysis (LDA),[30] and Naive Bayesian (NB).[29] For K-NN classifier, five nearest neighbors and Euclidean distance are used in the classification. For LDA classifier, the prior probabilities of all groups are assumed equal. For NB classifier, normal distribution is considered for modeling the data. For validation, eight-fold cross-validation is performed and also the classifiers are iterated 100 times and the confusion matrices are added. For evaluation, the accuracy of each class (ACC_i), the number of correctly classified cells in class *i*/total number of cells in class *i*, and the mean accuracy (MACC), the total number of cells that are correctly classified/total number of cells.

Results and Discussion

As dataset introduced in the "Dataset" section, there are 200 cells equally distributed in five types: Neutrophil (**N**), Eosinophil (**E**), Basophil (**B**), Monocyte (**M**), and Lymphocyte (**L**). The results of classification for GLCM features are shown in Table 2. The best MACC for GLCM features is 81.49% achieved by K-NN classifier. As the classification results for Haralick features shows in Table 3, maximum MACC is 74.12% for K-NN classifier which is lower than that of GLCM features. Spectral features yielded the maximum MACC of 64.69% with LDA classifier [Table 4] that is lower than Haralick and GLCM features. For Wavelet features, the highest MACC is 76.4% for LDA classifier [Table 5], which is higher than Haralick and

Table 2: The accuracies of leukocytes classification for GLCM features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	71.83	85.1	84.5	74.1	91.93	**81.49**
LDA	58.43	85.25	83.18	79.33	88.93	79.02
NB	59.05	92.5	86.98	67.93	87.45	78.78
Average	63.1	87.62	84.89	73.79	89.44	79.76

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Table 3: The accuracies of leukocytes classification for Haralick features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	52.15	82.5	73.9	66.63	95.43	**74.12**
LDA	59	82.25	88.45	64.28	60.63	70.92
NB	35.55	84.75	72.38	77.68	82.33	70.54
Average	48.9	83.17	78.24	69.53	79.46	71.86

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Table 4: The accuracies of leukocytes classification for Spectral features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	35.75	59.55	81.45	41.15	77.83	59.15
LDA	30.63	68.75	90.35	51.9	81.8	**64.69**
NB	5.9	62.15	85.68	40.15	78.78	54.53
Average	24.09	63.48	85.83	44.4	79.47	59.46

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Table 5: The accuracies of leukocytes classification for Wavelet features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	56.85	86.1	85.5	41.35	88.6	71.68
LDA	59.3	77.85	76.68	70.93	97.25	**76.4**
NB	64.2	74.58	90.8	50.28	84.83	72.94
Average	60.12	79.51	84.33	54.19	90.23	73.67

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Spectral features but lower than GLCM features. Maximum MACC for Gabor features is 56.83% for NB classifier [Table 6] which is the least MACC among all studied features. For CoALBP features, the best MACC is 81.74% for LDA classifier [Table 7], which is slightly higher than GLCM features. Finally, the best MACC for RICLBP is 85.53% for LDA classifier [Table 8], which is greater than of CoALBP features. It is seen from Tables 2 to 8 that the maximum and minimum averages of MACC of classifiers are achieved for RICLBP (82.33%) and Gabor (56.83%) features, respectively. The best results of each texture feature are compared in Table 9. Maximum MACC is achieved by RICLBP (85.53%) and minimum by Gabor (56.83%) features. In addition, RICLBP features have the best accuracies for monocyte (85.43%) and lymphocyte (89.43%). Wavelet features have the best accuracy for basophil (97.25%), Spectral features for neutrophil (90.35%), and CoALBP features for eosinophil (78.33%), respectively. By assessing results in Tables 2 to 8, we see that although the best MACC is achieved for RICLBP features, the accuracies of all five leukocyte types aren't maximized for RICLBP features. This issue encourages one to design a multifeature-classifier system to maximize even the accuracies for all types, which is out of scope of this work. The details information about the dimensions and

Table 6: The accuracies of leukocytes classification for Gabor features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	23.38	60.63	65.18	37.98	79.5	53.33
LDA	18.77	67.47	66.78	35.83	75.05	52.78
NB	2.95	73.7	77.18	46.35	83.98	**56.83**
Average	15.03	67.27	69.71	40.05	79.51	54.31

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Table 8: The accuracies of leukocytes classification for RICLBP features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	64.95	96.85	89.78	62.38	84.68	79.72
LDA	85.43	89.43	90.18	73.43	89.18	**85.53**
NB	64.58	91.3	88.95	76.4	87.43	81.73
Average	71.65	92.53	89.64	70.74	87.1	**82.33**

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Table 7: The accuracies of leukocytes classification for CoALBP features

Classifier	Accuracies (%)					MACC
	M	L	N	E	B	
K-NN	65.28	89.35	68.6	68.9	91.9	76.81
LDA	75.97	83.98	80.83	78.33	89.58	**81.74**
NB	80.22	85	81.18	53.98	86.95	77.47
Average	73.82	86.11	76.87	67.07	89.48	78.67

B = basophil, E = eosinophil, K-NN = K-nearest neighbors, L = lymphocyte, LDA = linear discriminant analysis, M = monocyte, MACC = mean accuracy, N = neutrophil, NB = naive Bayesian.

Table 9: Comparison of results for different texture features

Features	Accuracies (%)					MACC
	M	L	N	E	B	
GLCM	71.83	85.1	84.5	74.1	91.93	81.49
Haralick	52.15	82.5	73.9	66.63	95.43	74.12
Spectral	30.63	68.75	**90.35**	51.9	81.8	64.69
Wavelet	59.3	77.85	76.68	70.93	**97.25**	76.4
Gabor	2.95	73.7	77.18	46.35	83.98	56.83
CoALBP	75.97	83.98	80.83	**78.33**	89.58	81.74
RICLBP	**85.43**	**89.43**	90.18	73.43	89.18	**85.53**

B = basophil, CoALBP = co-occurrence of adjacent local binary pattern, E = eosinophil, GLCM = gray-level co-occurrence matrix, L = lymphocyte, M = monocyte, MACC = mean accuracy, N = neutrophil, RICLBP = rotation invariant co-occurrence of local binary pattern.

Table 10: Details information about the dimensions and final parameters of the features

Features	Initial dimension	Second dimension	Final parameters	MACC
GLCM	64	4	D=28, NL=8	81.49
Haralick	**13**	**2**	D=15, NL=8	74.12
Spectral	20	3	–	64.69
Wavelet	34	4	Wavelet filter: bior3.7	76.4
Gabor	36	3	Scales: 2, orientations: 6	56.83
CoALBP	1024	13	s=1, r=10, config:LBP(+)	81.74
RICLBP	136	5	s=2, r=14	**85.53**

CoALBP = co-occurrence of adjacent local binary pattern, GLCM = gray-level co-occurrence matrix, MACC = mean accuracy, RICLBP = rotation invariant co-occurrence of local binary pattern.

final parameters of the features are shown in Table 10. In Table 10, *initial dimension* column refers to the first extracted feature dimension discussed in the "Materials and Methods" section; *second dimension* column depicts the number of selected features by SWR for each feature set; *final parameters* column shows the parameters related to the best MACC achieved by the texture features. It is seen from Table 10 that the best MACC (85.53%) is achieved with five RICLBP features with parameters $s = 2$ and $r = 14$. These five features are 7th, 10th, 63rd, 93rd, and 129th components of RICLBP features.

Conclusion and Future Works

In this paper, we were looking for the best texture features to recognize five types of leukocytes from the microscopic images of blood smear. A dataset was prepared composed of 200 cells, 40 cells of each type. The border of the cells was manually segmented by an expert to disregard the problems in automatic segmentation. Some preprocessing procedures were applied to the images to normalize them. Then, seven texture features named GLCM features, Haralick features, Spectral texture features, Wavelet-based features, Gabor-based features, CoALBP features, and RICLBP features were extracted from the cells. Next, the best features were selected for each category using SWR. Finally, three common classifiers, called K-NN, LDA, and NB, were utilized to classify leukocytes. To evaluate the process, the accuracy of each leukocyte type and the mean accuracy were computed. RICLBP features achieved the best mean accuracy (85.53%) for LDA classifier. In our experiments, although the maximum mean accuracy went with RICLBP features, but the accuracies of all five leukocyte types weren't maximized for RICLBP features. This result directs us to design and develop a system based on multiple features and multiple classifiers to maximize the accuracies even for each individual cell type in our future work.

Financial support and sponsorship

This work was supported by Medical Research Center, Isfahan University of Medical Sciences, Isfahan, Iran, [194100].

Conflicts of interest

There are no conflicts of interest.

References

1. Saraswat M, Arya KV. Automated microscopic image analysis for leukocytes identification: A survey. Micron 2014;65:20-33.
2. Sarrafzadeh O, Rabbani H, Dehnavi AM, Talebi A. Detecting different sub-types of acute myelogenous leukemia using dictionary learning and sparse representation. Quebec City, QC: IEEE International Conference on Image Processing (ICIP); 2015. p. 3339-43. doi: 10.1109/ICIP.2015.7351422.
3. Sarrafzadeh O, Dehnavi AM. Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing. Adv Biomed Res 2015;4:174-9.
4. Sarrafzadeh O, Dehnavi AM, Rabbani H, Talebi A. A simple and accurate method for white blood cells segmentation using K-means algorithm. IEEE Workshop on Signal Processing Systems (SiPS): Design and Implementation, vol 2015, 2015.
5. Supardi NZ, Mashor MY, Harun NH, Bakri FA, Hassan R. Classification of blasts in acute leukemia blood samples using k-nearest neighbour. 8th IEEE International Colloquium on Signal Processing and its Applications, 2012, pp. 461-5.
6. Sarrafzadeh O, Rabbani H, Mehri Dehnavi A, Talebi A. Analyzing features by SWLDA for the classification of HEp-2 cell images using GMM. Pattern Recognit Lett 2016;82:44-55.
7. Shirazi SH, Umar AI, Naz S, Razzak MI. Efficient leukocyte segmentation and recognition in peripheral blood image. Technol Heal Care 2016;24:335-47.
8. Theera-Umpon N. White blood cell segmentation and classification in microscopic bone marrow images. Fuzzy Systems and Knowledge Discovery. Berlin, Heidelberg: Springer; 2005. p. 787-96.
9. Jung C, Kim C, Chae SW, Oh S. Unsupervised segmentation of overlapped nuclei using Bayesian classification. Biomed Eng IEEE Trans 2010;57:2825-32.
10. Sarrafzadeh O, Dehnavi AM, Rabbani H, Ghane N, Talebi A. Circlet based framework for red blood cells segmentation and counting. IEEE Workshop on Signal Processing Systems, SiPS: Design and Implementation, vol 2015, 2015.
11. Rawat J, Bhadauria HS, Singh A, Virmani J. Review of leukocyte classification techniques for microscopic blood images. 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 1948-54.
12. Huang D-C, Hung K-D, Chan Y-K. A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images. J Syst Softw 2012;85:2104-18.
13. Rezatofighi SH, Soltanian-Zadeh H. Automatic recognition of five types of white blood cells in peripheral blood. Comput Med Imaging Graph 2011;35:333-43.
14. Habibzadeh M, Krzy ak A, Fevens T. White blood cell differential counts using convolutional neural networks for low resolution images. Artificial Intelligence and Soft Computing. Berlin, Heidelberg: Springer 2013. p. 263-74.
15. Sabino DMU, Da Fontoura Costa L, Rizzatti EG, Zago MA. A texture approach to leukocyte recognition. Real-Time Imaging 2004;10:205-16.
16. Sabino DMU, Costa LF, Rizzatti EG, Zago MA. Toward leukocyte recognition using morphometry, texture and color. 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano, vol 2, 2004, pp. 121-4.
17. Zhao J, Zhang M, Zhou Z, Chu J, Cao F. Automatic detection and classification of leukocytes using convolutional neural networks. Med Biol Eng Comput 2016:1-5.
18. Li X, Cao Y. A robust automatic leukocyte recognition method based on island-clustering texture. J Innov Opt Health Sci 2016;9:1650009.
19. Sarrafzadeh O, Rabbani H, Talebi A, Banaem HU. Selection of the best features for leukocytes classification in blood smear microscopic images. Proc. SPIE 9041, Medical Imaging 2014: Digital Pathology, 90410P, 2014. doi: 10.1117/12.2043605.
20. Zuiderveld K. Contrast limited adaptive histogram equalization. In: Paul SH, editor. Graphics gems IV. San Diego, CA, USA: Academic Press Professional, Inc.; 1994, p. 474-85.
21. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3:610-21.

22. Gonzalez RC, Woods RE, Eddins SL. Digital Image Processing Using MATLAB. Upper Saddle River, NJ: Prentice Hall; Pearson Education. 2004.
23. Arivazhagan S, Ganesan L. Texture classification using wavelet transform. Pattern Recognition Letters 2003;24:1513-21.
24. Ji H, Yang X, Ling H, Xu Y. Wavelet domain multifractal analysis for static and dynamic texture classification. IEEE Trans Image Process 2013;22:286-99.
25. Haghighat M, Zonouz S, Abdel-Mottaleb M. CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. Expert Syst Appl 2015;42:7905-16.
26. Nosaka R, Ohkawa Y, Fukui K. Feature extraction based on co-occurrence of adjacent local binary patterns. 5th Pacific Rim Conference on Advances in Image and Video Technology, 2012, pp. 82-91.
27. Nosaka R, Suryanto CH, Fukui K. Rotation invariant co-occurrence among adjacent LBPs. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Berlin, Heidelberg: Springer; 2013. p. 15-25.
28. Draper NR, Smith H, Pownell E. Applied Regression Analysis. Vol 3 New York: Wiley 1966.
29. Mitchell TM. Machine Learning. New York City, New York, US: McGraw-Hill; 1997.
30. Guo Y, Hastie T, Tibshirani R. Regularized linear discriminant analysis and its application in microarrays. Biostatistics 2007;8:86-100.