Arab J Sci Eng (2018) 43:7041–7058
https://doi.org/10.1007/s13369-017-2959-3

RESEARCH ARTICLE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

Leukocyte Classification using Adaptive Neuro-Fuzzy Inference System in Microscopic Blood Images

Jyoti Rawat[1] · Annapurna Singh[1] · H S Bhadauria[1] · Jitendra Virmani[2] · J S Devgun[3]

Received: 18 April 2017 / Accepted: 6 November 2017 / Published online: 21 November 2017

Abstract Microscopic pathology is still a meticulous and biased task for hematologist, which leads to the misclassification of cells and vagueness prediction of abnormal cells due to variability in the morphological structure of leukocytes. Therefore, to enhance the detection precision and diminishing the time factor, an automatic classification system for leukocytes has been proposed. In routine clinical practice, expert hematologists observed that the nucleus plays a crucial role in the identification of the blood disorders. Accordingly, in this work, the localization of leukocyte nucleus is performed by using Chan–Vase level-set method for the design of a classification framework that differentiates between four classes of the leukocytes, i.e., eosinophils, polymorphs, monocytes and lymphocytes based on the nucleus. A dataset consisting of 162 leukocyte microscopic images is used. The images in the dataset are classified on the basis of texture, shape and color features. The feature selection method based on the linguistic hedge is applied on evaluated feature space of 92. The selected features are fed to an adaptive neuro-fuzzy classifier for the classification. The proposed framework obtained an accuracy of 98.7% after applying the adaptive neuro-fuzzy classification on selected 46 informative features. The correlation of best features and data extorted from the different microscopic images may yield a dramatic increase in diagnostic consistency in clinical pathology. The results obtained by utilization of selected optimal features and adaptive neuro-fuzzy classification system indicate that it can be routinely used in clinical environment for differential diagnosis between different classes of leukocytes.

Keywords Leukocyte segmentation · Chan–Vase method · Texture features · Shape features · Color features · Leukocyte classification · Adaptive neuro-fuzzy classifier

✉ Jyoti Rawat
jyotisonirawat@gmail.com

1 G.B. Pant Engineering College, Pauri Garhwal, UK, India

2 CSIR-Central scientific instruments organization, Chandigarh, India

3 M. M. Institute of Medical Sciences & Research, Solan, HP, India

1 Introduction

New insights and technologies have the potential to optimize the use of digital pathology or cytopathology for extracting information from digital slides of leukocytes that will redefine approaches for the diagnosis of hematologic diseases. In digital microscopy, images of the blood slides can be shared with other experts easily, in the case of any complexity. These digital slides can be easily analyzed by using different algorithms that prove to be useful for automating the manual process. These algorithms can also be used to enhance the process of diagnosis and reduce uncertainty and subjectivity [1].

The present work is carried out with the purpose of analyzing the outstanding issues in the computer-aided classification (CAC) systems using microscopic images for localizing leukocyte and classifying it into four different classes, i.e., eosinophils, polymorphs, monocytes and lymphocytes. All the leukocytes work for the immune system of the human body as they kill parasites, fungi, generate cytokines and perform as scavengers to remove dead tissue [2].

For the enhancement of the diagnosis process in the field of hematology, many CAC system designs have been developed in the recent past [2]. However, automatic analysis of microscopic blood smear images is still an open research problem, the design of an effective CAC system that is capa-

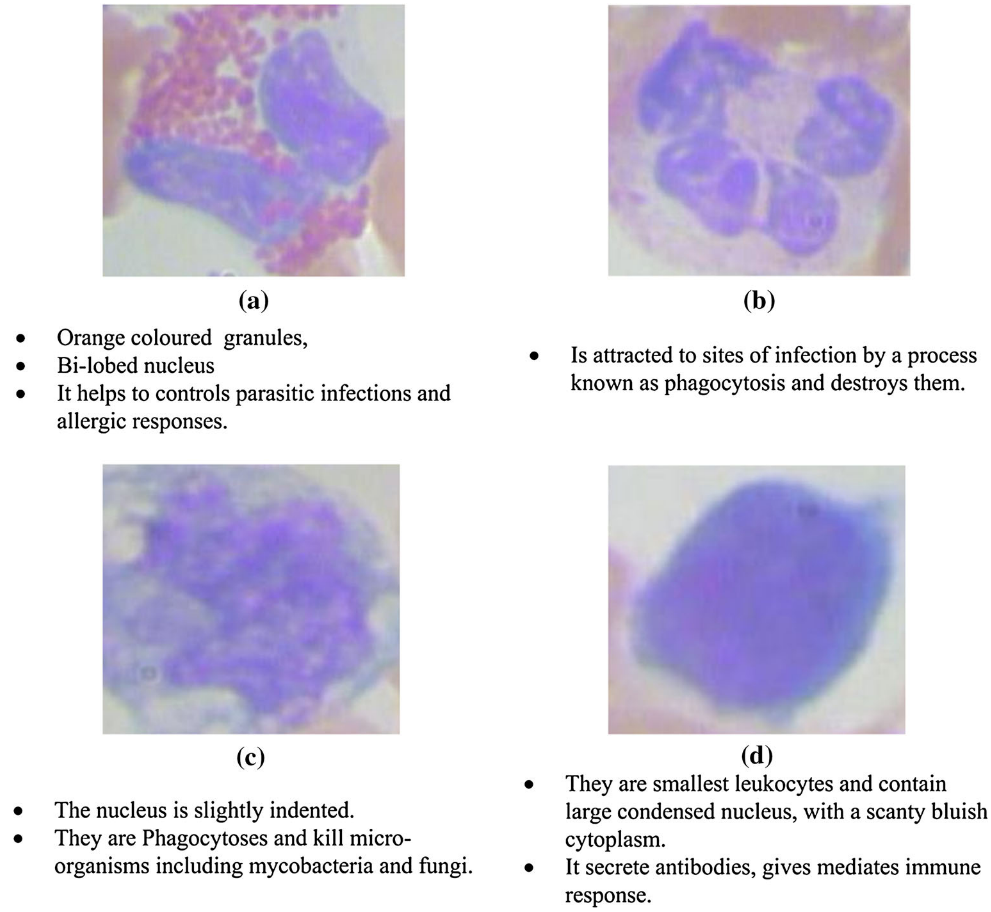


Fig. 1 Characteristics and the functions of the leukocytes present in human body (Pang et al. [8]). *Note* **a** eosinophils, **b** polymorphs, **c** monocytes and **d** lymphocytes

ble of extracting valuable information from the smear images like their morphological structure, functions and position in a bone marrow tissue is still a challenging task [3]. To reduce the gap between bone marrow tissue and its morphological structure [4,5], adaptive neuro-fuzzy inference system (ANFIS) is adopted for this work. The major benefit of ANFIS is its learning capability. On the other hand, neural network suffers from the learning capability. Thus ANFIS is prominent applications of fuzzy inference system and neural network [6].

The description of the leukocyte categories is given in Fig. 1.

A description of studies carried out on leukocyte classification is given in Table 1.

After the extensive study of the literature, it has been observed that the leukocyte classification can be performed on the basis of different classes of different morphology (i) 5-class classification problem (ii) 4-class classification problem and (iii) 3-class classification problem. It is worth mentioning that the maximum classification accuracy for 5-class classification problem is 98.8 %, for 4-class classification, it is 98.0 %, and for 3-class classification problem, it is 98.0 %. In the work [12] for 5-class leukocyte classification, 98.8 % accuracy is obtained by neural network classifier and extracting color and shape features. For 4-class leukocyte classification [25], 98.0 % accuracy is obtained by applying SVM classifier with statistical features. Similarly, 98.0 % accuracy is achieved for 3-class leukocyte classification by using decision tree classifier by extracting shape features [32].

In the present work, 4-class leukocyte classification is directly compared with [25] in which texture features are extracted and SVM is used for discriminating the four classes of leukocyte. This work proposes a leukocyte classification system that is able to localize the nucleus and classify them into leukocyte class automatically. The level-set (Chan–Vese) method has been used for the nucleus localization. Then texture feature (statistical texture measure as GLCM, NGTDM, transfer domain texture measure as 2D Gabor wavelet transform and signal processing-based texture measure as Law's 3), shape features and color features are extracted from the segmented sub-image for the discrimination between different leukocytes. Feature selection techniques have been effectively utilized as a part of numerous ranges particularly in medical diagnosis to diminish the extent of features gath-

Table 1 Studies carried out for leukocyte classification

Considered class	Authors, year	Extracted features	Classifier used	Images	Accuracy (%)
5-Class basophiles, eosinophils, polymorphs, monocytes, lymphocytes.	Pang et al. 2015 [7]	TFV	SVM	298	95.5
	Nazlibilek et al. 2015 [8]	TFV	ANN	_	95.0
	Ravikumar et al. 2015 [9]	SFV, TFV	RVM	85	91.0
	Nazlibilek et al. 2014 [10]	SFV, TFV	ANN	240	95.0
	Habibzadeh, et al. 2013 [11]	SFV, TFV, CFV	SVM	140	84.0
	Rezatofighi et al. 2012 [12]	*SFV, CFV*	*ANN*	*400*	*98.8*
	Huanga et al. 2012 [13]	TFV	SVM	_	96.4
	Ramesh et al. 2012 [14]	SFV, CFV	LDA	1983	93.9
	Lina et al. 2012 [15]	TFV, CFV	_	500	75.9
	Rezatofighi et al. 2010 [16]	TFV	SVM	90	93.0
	Xie et al. 2010 [17]	SFV	ANN	230	89.6
	Ghosh et al. 2010 [18]	SFV	Naive bayes	150	83.2
	Rodrigues et al. 2008 [19]	SFV, TFV	SVM	241	85.4
	Yampri et al. 2006 [20]	SFV	_	50	92.0
	Piuri et al. 2004 [21]	SFV	ANN	34	_
	Bikhet, et al. 2000 [22]	SFV, CFV	_	71	91.0
	Bacus et al. 1972 [23]	SFV, TFV, CFV	MGC	523	93.0
	Young, et al. 1972 [24]	SFV, CFV	DT	74	92.4
4-Class eosinophils, polymorphs, monocytes lymphocytes.	*Malek et al. 2005* [25]	*TFV*	*SVM*	*50*	*98.0*
	Sarrafzadeh et al. 2013 [26]	SFV, TFV, CFV	SVM	149	97.7
	Tabrizi 2010 [27]	SFV, TFV, CFV	SVM	302	97.0
	Stadelmann et al. 2012 [28]	SFV, TFV, CFV	AdaBoost	461	91.3
	Suapang et al. 2015 [29]	SFV, TFV, CFV	ANN	134	88.1
	Mircic et al. 2006 [30]	SFV	ANN	200	86.0
	Ferri et al. 1994 [31]	SFV	*k*NN	45	80.0
3-Class polymorphs, monocytes, lymphocytes.	P. Hiremath et al. 2010 [32]	SFV	DT	100	98.0

Italic values show the maximum accuracy

SFV shape feature vector, *TFV* texture feature vector, *CFV* color texture feature vector, *MGC* multivariate Gaussian classifier, *DT* decision tree

ered at the time of the clinical testing and trials. Extracted features are fed to the classification module for the class decision. The procedure of selecting prominent features is imperative since it decreases the dimensionality of the input data and empowers classifier to work better by diminishment of the calculation time, and accordingly increases the exactness of the classification model [33,34].

The main motivation of the present work is to develop a second opinion tool for practising hematologists so that they can diagnose hematological disorders with minimum human error.

2 Material and Methods

2.1 Database Description

In the research area of microscopic pathology, the lack of publicly available database is the prevalent occurred problem. The same problem is also discussed in [8,10,17,21,23]. There are a few sample images used in many proposed classification systems which are not accessible publicly that may result in limited innovation.

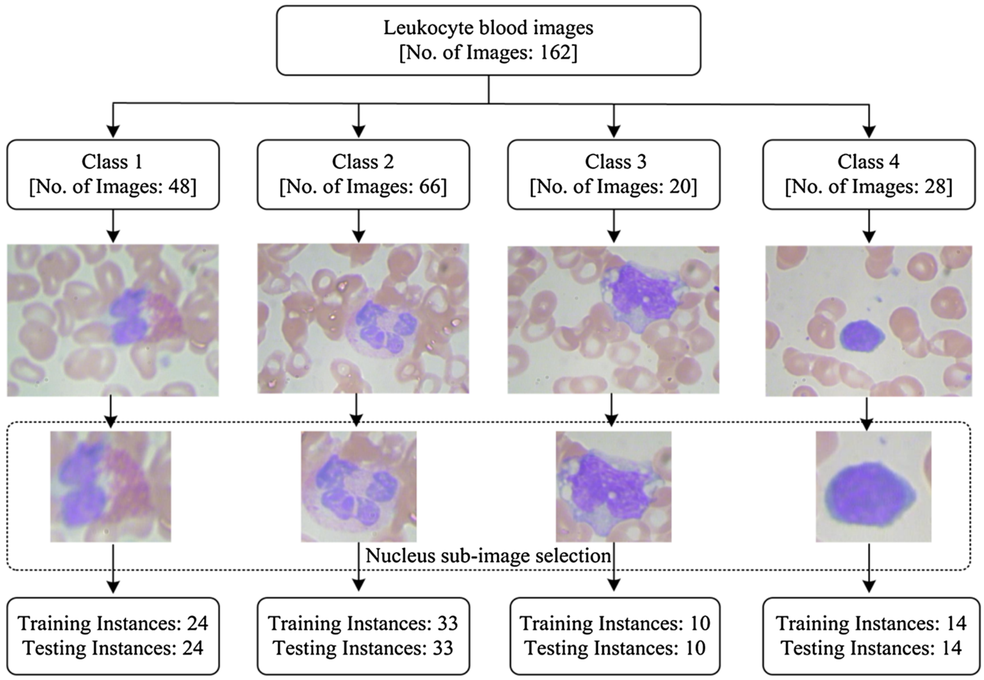


Fig. 2 Dataset description

In order to develop efficient classification framework, it is necessary to train the classifiers with a comprehensive image database with representative images from each subclass. In the present work, 162 Gismo-right stained peripheral white blood cell images are taken from online available image repository at MATLAB file exchange [35] which provides a broad range of good quality web-based image library for white blood cell images that are labeled by an expert pathologist. Further, these cells are labeled by the participating pathologist who has more than 20 years of experience in hematology department and classified the ground truth for the complete dataset into four classes, i.e., eosinophils, polymorphs, monocytes and lymphocytes. Thus the used dataset consists of 48 instances of eosinophils, 66 instances of polymorphs, 20 instances of monocytes and 28 instances of lymphocytes. A total of 162 sub-images are extracted from 162 microscopic images for removing unwanted cells and outliers. A light microscope using 100× objective Lens is used and the VGA image resolution of captured images is 640 × 480 pixels. The brief description of the database and its further bifurcation into training and testing set is shown in Fig. 2.

The following protocols were followed for dataset preparation:

1. The judgment regarding the diagnostic quality (free from artifacts) and representativeness of image class, i.e., eosinophils, polymorphs, monocytes and lymphocytes was made by four domain experts (one co-author of this paper) with minimum 10 and maximum 26 years of experience in field of pathology .
2. Selection of sub-image is discussed in Sect. 2.2.2 of the same document.
3. Experienced participating pathologists confirmed the type of leukocyte by assessment criteria. Interpretation by pathologists is done on the basis of visual inspection of microscopic features (morphological, textural and chromatic appearance of cell image) according to their expertise.

2.2 Proposed CAC System

The proposed system comprises of segmentation module, feature extraction module, feature selection and classification module. The flow chart for analysis of microscopic blood image for classification is shown in Fig. 3. The Segmentation module is used for the extraction and identification of nucleus from leukocyte. To find the morphological differences between the different type of leukocytes, different characteristic of their shape, color and texture are extracted by using feature extraction methods. These extracted features are processed by feature selection method using linguistic hedge then the machine learning module known as a classification module is used to classify the images into one of the four classes of leukocyte. The brief explanation of each module is given in next section of this document.

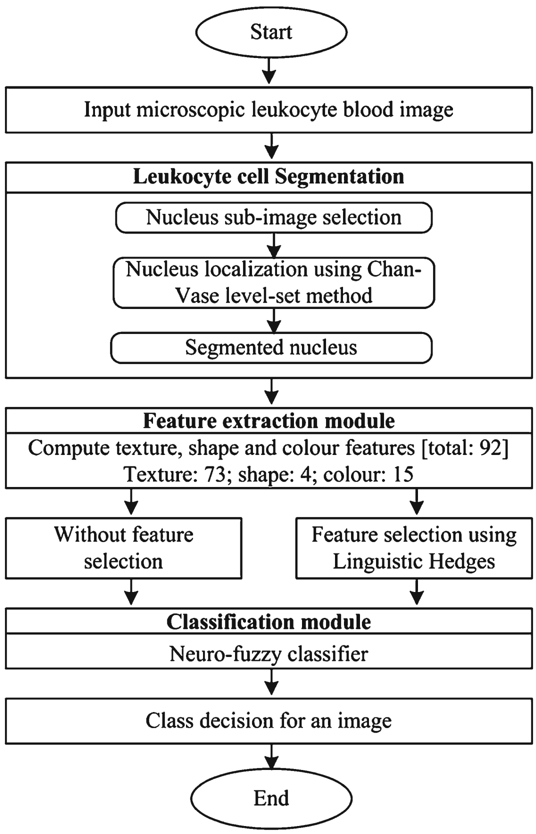


Fig. 3 Classification framework for proposed system

2.2.1 Leukocyte Segmentation

The leukocytes are differentiated from each other on the basis of their nuclei, as shown in many medical studies, especially for hematological diseases caused by a specific type of leukocyte [36]. The step involved in segmentation is shown in Fig. 4.

The following steps are used in proposed leukocyte segmentation method: (i) Cell nucleus is located by sub-image selection. (ii) For segmentation of nucleus, the initial seed point is placed inside the nucleus. (iii) The Chan–Vese algorithm is applied to the sub-images and repeats step (ii) until the entire nucleus is localized.

2.2.2 Nucleus Sub-image Selection

The presence of adjacent cells and leukocyte agglomerates is an important problem for testing of microscopic leukocyte images, so it is important to remove all the leukocytes located on the boundary of the image non-leukocytes elements, which results in erroneous measurements at a later stage of the testing procedure [37–39]. Leukocytes are identified by using the process of sub-image selection that can be simplified by obtaining the sub-image. Input image is automatically sub-imaged by cropping the smallest rectangle that entirely holds a connected part of nucleus and cytoplasm (cell membrane) on the basis of pixel values. Separating a particular leukocyte cell membrane from the entire background is done by sub-imaging as shown in Fig. 5.

2.2.3 Nuclei Segmentation

Image segmentation is carried out with a purpose to segregate the image into consequent parts that are strongly correlated with the objects contained in the image [40–45]. In the present work, Chan–Vese level-set method is used that relies on intensity, area and other global properties, rather than local properties, such as gradients. It is an illustration of an active contour model based on an energy minimization

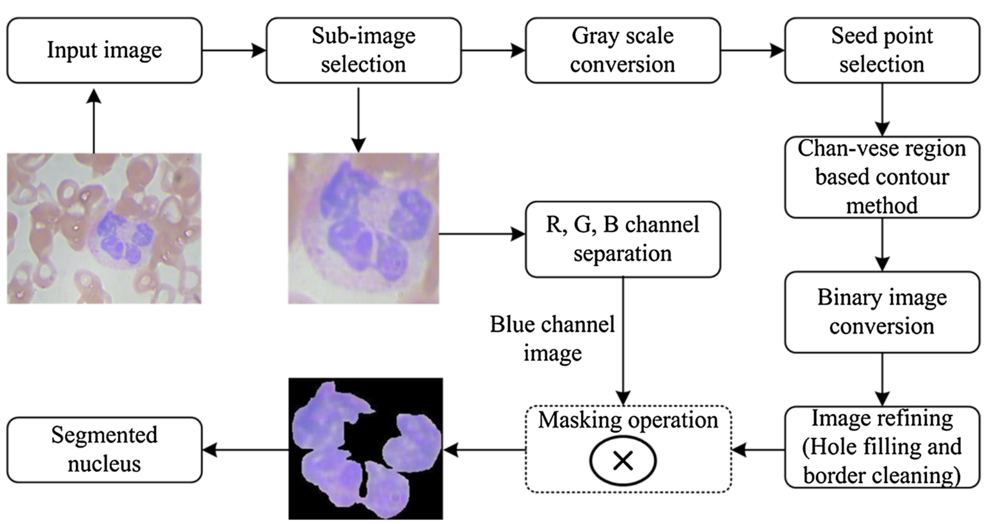


Fig. 4 Segmentation of nucleus

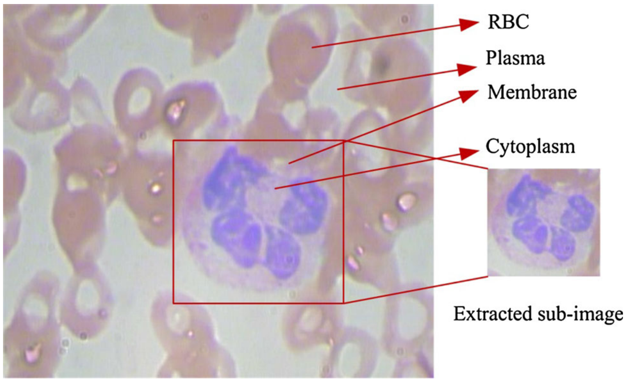


Fig. 5 Nucleus sub-image selection of microscopic leukocyte blood image

problem. It pulls off good results even when the image is blurred and noisy.

Chan–Vese level-set method The Chan–Vese algorithm evolves this contour via a level-set method that is used widely in the biomedical imaging, particularly for the localization [46]. Consider X is a set of bounded area of Δ^2 where ∂X represents its boundary. Let f_0 is an image and $f_0 : \overline{X} \rightarrow \Delta$, and $\delta(s)$is a piecewise $\delta^1[0, 1]$ parameterized curve. Now, denote the inside region δas ω, and the region outside δ as$\overline{X}/\omega$. Moreover, $\delta_1$will denote the average pixel's intensity insideδ and will represent the average intensity outside $\delta(i.e., \delta_1 = \delta_2(\delta), \delta_2 = \delta_2(\delta))$.

The objective function of Chan–Vese algorithm is to minimize the energy functional $F(c_1 c_2, \delta)$, defined in equation 1.

$$
\begin{aligned}
F(c_1, c_2, \delta) = {} & \mu \cdot \text{Length}(\delta) + v \cdot \text{Area}(\text{inside}(\delta)) \\
& + \lambda_1 \int\limits_{\text{inside}(C)} |f_0(x, y) - c_1|^2 \, \mathrm{d}x\mathrm{d}y \\
& + \lambda_2 \int\limits_{\text{outside}(C)} |f_0(x, y) - c_2|^2 \, \mathrm{d}x\mathrm{d}y
\end{aligned} \tag{1}
$$

where $\mu \geq 0$, $v \geq 0$, $\lambda_{1,}\lambda_2 > 0$are fixed by the user. Initially, the preferred settings are $v = 0$, $\lambda_1, \lambda_2 = 1$.

The obtained segmented nucleus of an input cell is shown in Fig. 6.

Figure 6a–d represents original images from a dataset, i.e., eosinophils, polymorphs, monocytes and lymphocytes, (e–h) sub-image selection, (i–l) shows the beginning curve drawing with seed point and initial level set, (m–p) represents segmented image by Chan–Vese approach to get regions with edges, (q–t) binary conversion of segmented image, (u–x) segmented images after masking of output image of Chan–Vese level-set method with blue band image of original input image.

2.3 Feature Extraction and Selection Module

In this module, three different types of features have been extracted namely shape, texture and color features. Shape descriptors (IFV1), namely convex area, perimeter, diameter, extent have been extracted from the nucleus [47]. In the present work, GLCM features (IFV2) [48–50], NGTDM features (IFV3) [51], Laws' features of kernel width 3 (IFV4) [52,53] and Gabor wavelet transform features (IFV5) for 3 magnitudes and 5 orientations are computed [34,54]. The frequently used color descriptors are color histograms and color moments; therefore, mean, standard deviation and second-, third-, fourth-order moments are computed for red, green and blue channels [55] as IFV6. The extracted input feature vectors (IFVs) are reported in Table 2.

2.4 Feature Selection Using Linguistic Hedge

This work implements a fuzzy feature selection method based on the linguistic hedges strategy in view of for leukocyte classification. The proposed strategy is utilized to accomplish a quick, straightforward and productive computer-aided classification framework. Linguistic hedges are applied to the fuzzy sets of rules, and are tailored by scaled conjugate gradient algorithm. By along these lines, some prominent features are underlined by using power values, and some unessential features are desponded with power values [33].

The extracted feature space for the classification of leukocyte cells consists a total of ninety-two features (i.e., 73 texture, 4 shape and 15 color features). Among these features, it is not necessary that all features are relevant for the classification task. Therefore feature selection is applied to the extracted feature space. In the present work, feature selection is performed using a wrapper-based approach for the adaptive neuro-fuzzy classifier (ANFC) using linguistic hedges (LH) [56,57].

Let A_1 and A_2 be fuzzy sets on X_1 and X_2 feature, and Y be the output, respectively. The P_1 and P_2 represent the LH values of those fuzzy sets. The function F_1 for variable A_1 is expressed as $A_1\{F_1 = f_1(A_1, A_2) = A_1\}$. It shows that the F_1 depends only on the variable A_1, irrespective of the value of A_2. For the function, F_1 FS can be defined by the product and power operators as shown in Table 3.

In this case, a general fuzzy classification rule is defined instead of the Boolean function as:

Rule1: *IF* X_1*is* A_1 *with* P_1 *hedge AND* X_2 *is* A_2 *with* P_2 *hedge THEN Y is* C_1;

where C1 represents the output class label.

According to this fuzzy rule, the functions F_1 and F_2 are redefined in fuzzy logic with a similar meaning:

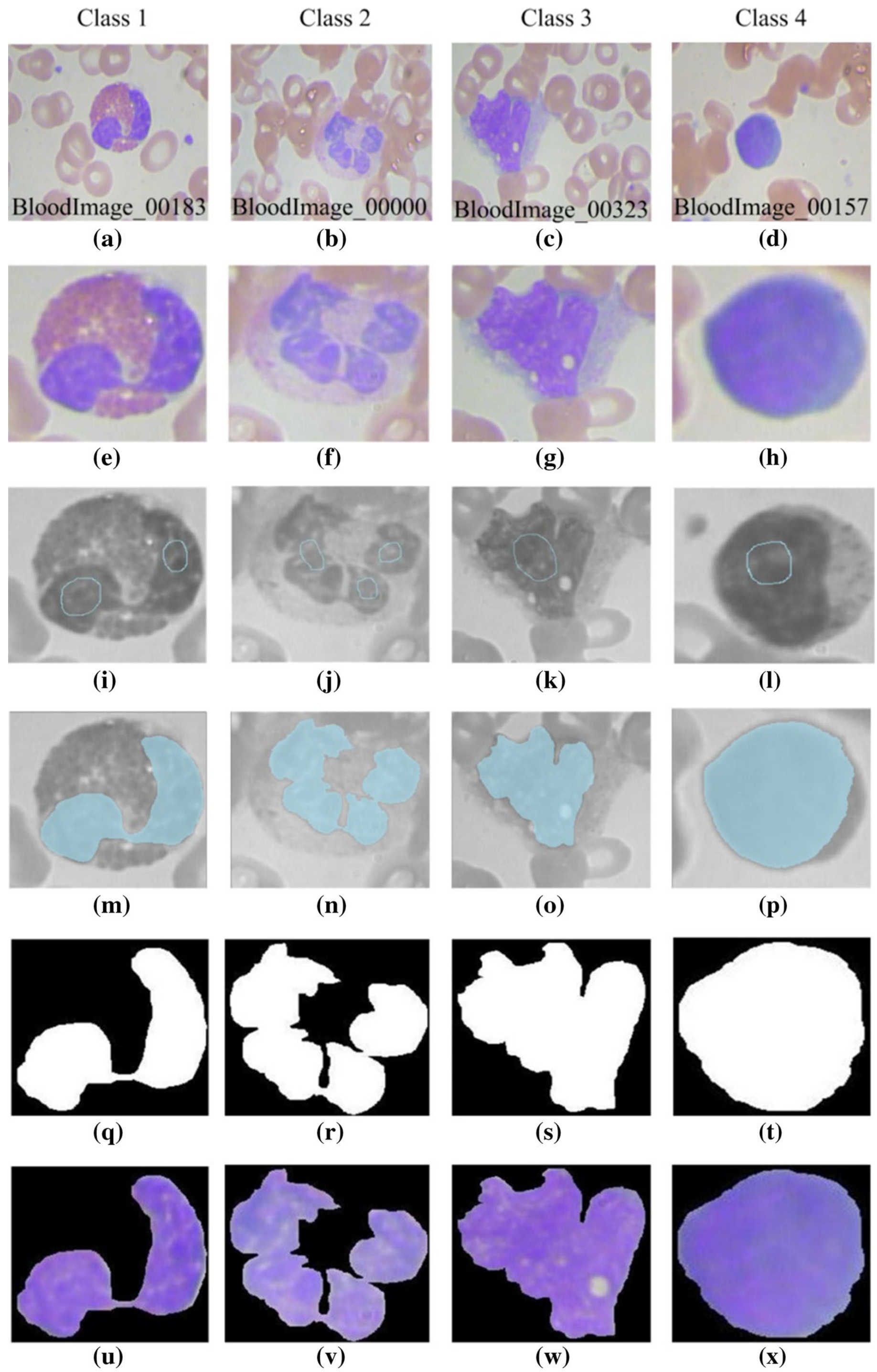


Fig. 6 Result of *Nuclei segmentation*

R_1 : *IF* X_1 *is* A_1 *with* $P_1 = 1$ *hedge AND* X_2 *is* A_2 *with* $P_2 = 0$ *hedge THEN Y is* F_1.
R_2 : *IF* X_1 *is* A_1 *with* $P_1 = 0$ *hedge AND* X_2 *is* A_2 *with* $P_2 = 1$ *hedge THEN Y is* F_2.

These rules can be reduced to the following rules:

R_1 : *IF* X_1 *is* A_1 *with* $P_1 = 1$ *hedge THEN Y is* F_1.
R_2 : *IF* X_2 *is* A_2 *with* $P_2 = 1$ *hedge THEN Y is* F_2.

The reduced rules contain only the selected features. If the LH value of a fuzzy set of any feature for any class equals to one, this feature will be important for that class [58].

In the present work, the length of feature space is ninety-two and LH value is 0.354. After applying the FS based on LH feature space is reduced to forty-six on the basis

Table 2 Various texture and shape and color features extracted for leukocyte classification

Models	IFVs	Extracted features	l
Shape features	IFV1	Convex area (number of pixels in convex polygon), perimeter (total no of the pixel representing the boundary of nucleus), diameter and extent	4
Texture features	IFV2	Variance, correlation, inverse difference moment (homogeneity), sum variance, contrast, sum average, information measures of correlation-1 and information measures of correlation-2	8
	IFV3	Contrast, coarseness, business, complexity and strength	5
	IFV4	Mean, standard deviation, kurtosis, skewness and entropy for six rotational invariant images	30
	IFV 5	Mean, standard deviation for 15 Gabor output images resultant of 3 scale value and 5 orientations	30
Color features	IFV6	MV_red, MV_green, MV_blue, STDV_red, STDV_green, STDV_blue, MOMENTS3_red (skewness), MOMENTS3_green, MOMENTS3_blue, MOMENTS4_red (kurtosis), MOMENTS4_blue, MOMENTS4_green, MOMENTS5 (high order moments)_red, MOMENTS5_green and MOMENTS5_blue	15
Total			92

IFV input feature vectors, l length of IFV

Table 3 LH and power value for A_1 and A_2

A_1	A_2	P_1	P_2	$F_1 = A_1^{P^1} \wedge A_2^{P^2}$	P_1	P_2	$F_2 = A_1^{P^1} \wedge A_2^{P^2}$
0	0	1	0	0	0	1	0
0	1	1	0	0	0	1	1
1	0	1	0	1	0	1	0
1	1	1	0	1	0	1	1

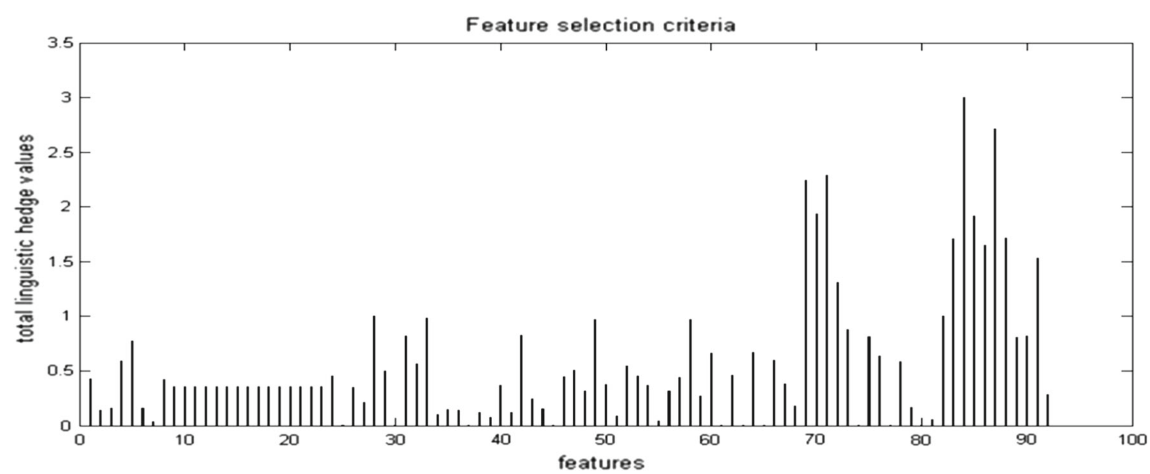


Fig. 7 Input features and their respectively calculated power of LH value

of the power of features to discriminate between different leukocyte cells. The relationship between input features and their, respectively, calculated power of LH value is shown in Fig. 7.

2.5 Classification Module

Image classification is a machine learning process used to predict the class membership of unknown data instance based on the training set of data whose class membership is known.

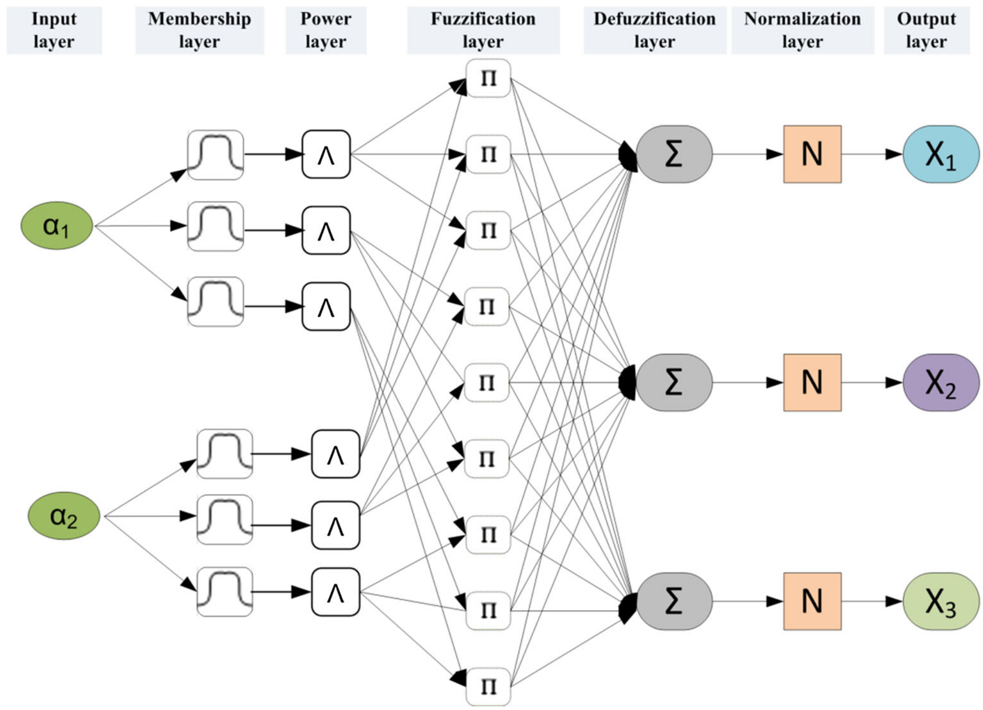


Fig. 8 Architecture of ANFC

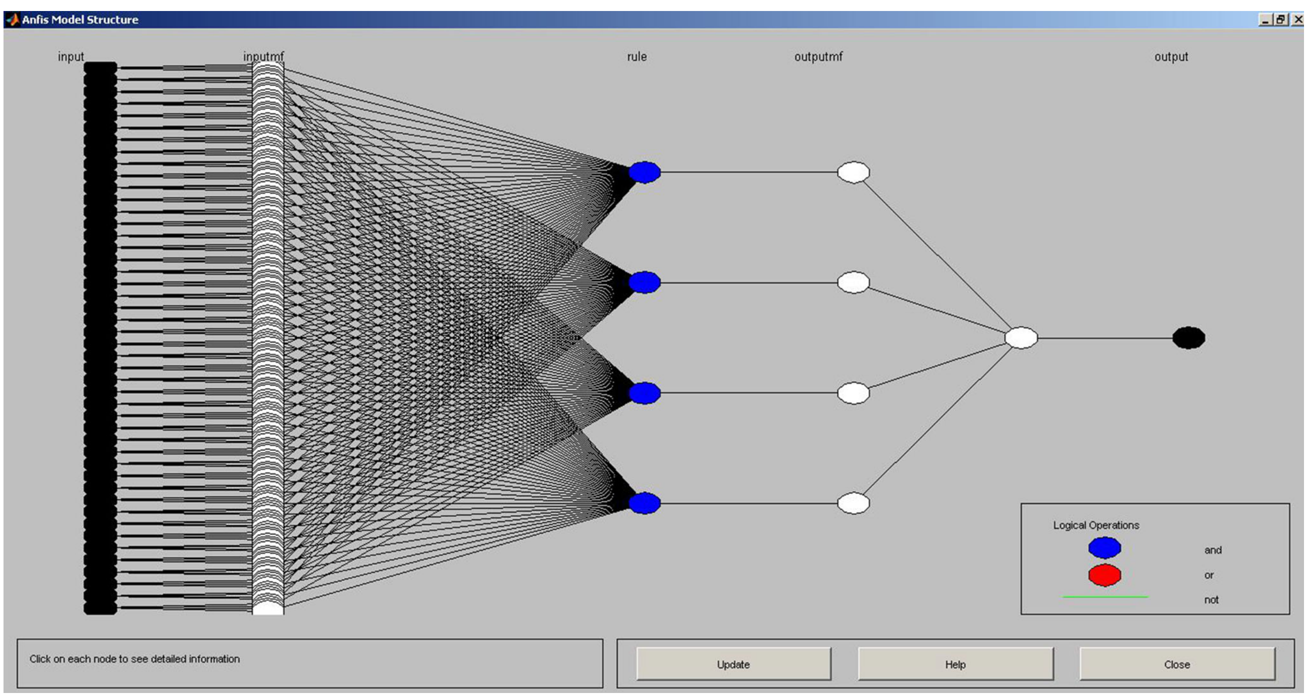


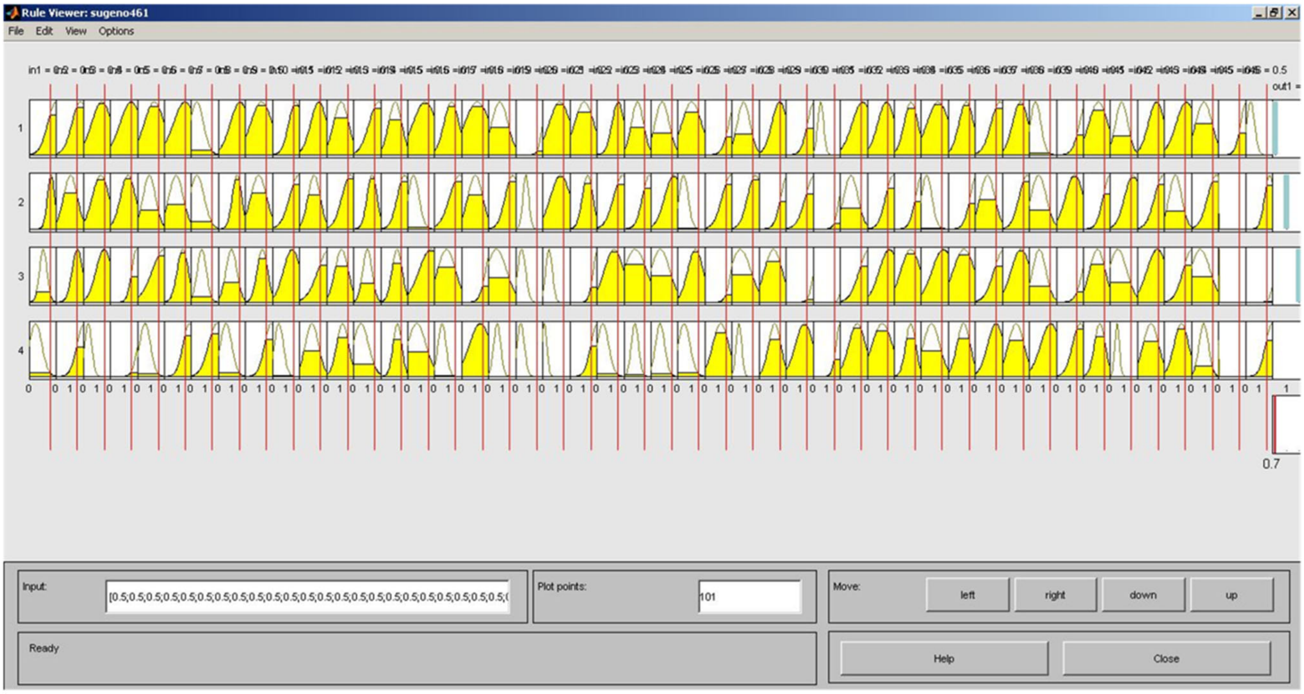
Fig. 9 ANFIS model for WBC classification

In the present work, adaptive neuro-fuzzy inference system (ANFIS) is used for the classification of leukocyte cells.

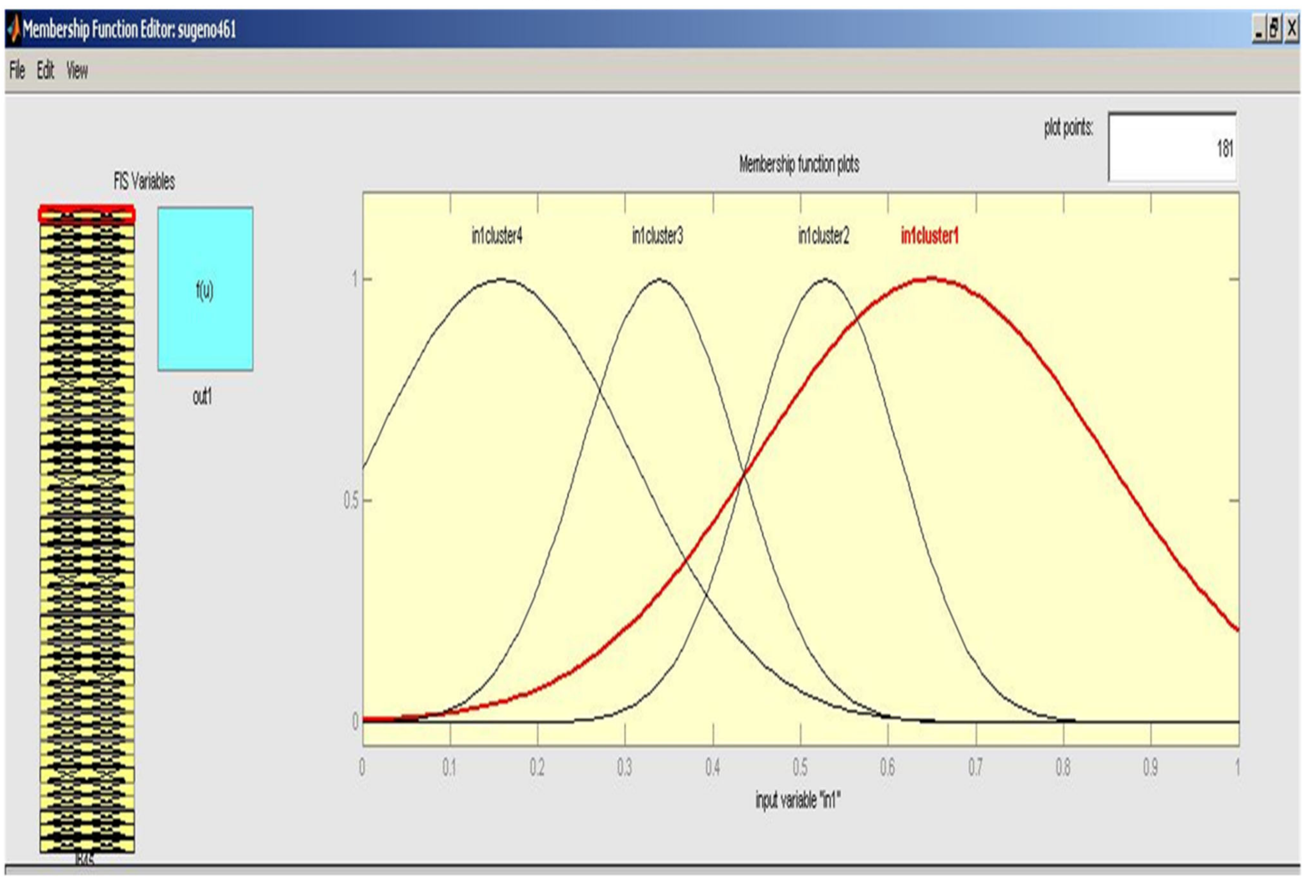
2.5.1 Adaptive Neuro-fuzzy Classifier (ANFC)

An ANFC is a multilayer feed-forward network consisting of the input layer, membership layer, power layer, fuzzification layer, defuzzification layer, normalization layer and output layer [58,59]. The ANFC works on ANFIS which is a tool that unites the IFVs, input membership function, output membership function, defined rules and the output class. The architecture of ANFC is shown in Fig. 8.

Figure 8 depicts an ANFC for three classes $\{C_1, C_2, C_3\}$ described by using two features $\{\alpha_1, \alpha_2\}$, and every input



(a) Sugeno rule-base viewer



(b) membership function for input 1

Fig. 10 The Sugeno rule-base viewer and membership function for input 1 of designed ANFC for the 4-class classification of leukocyte cells

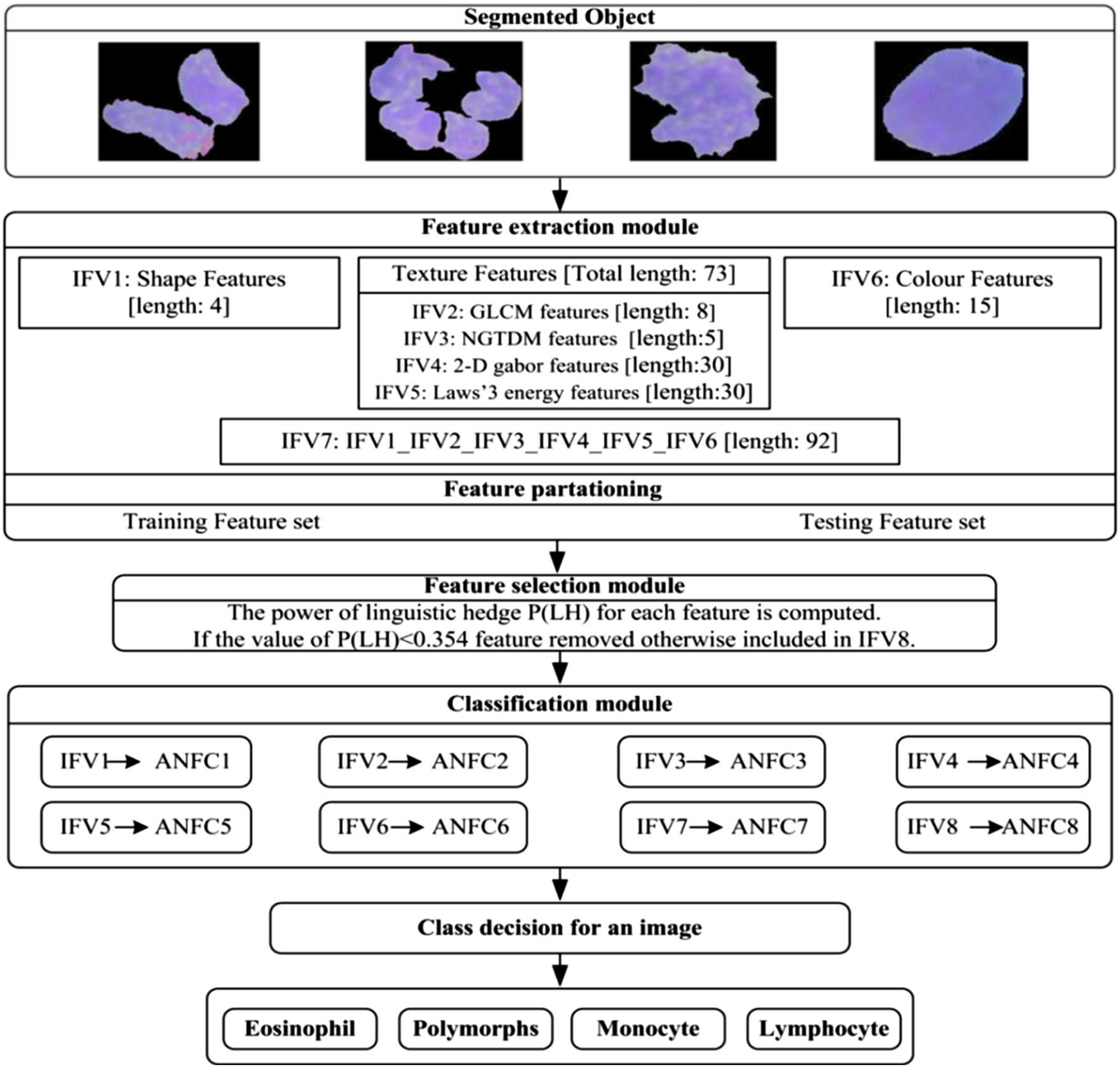


Fig. 11 Experimental work flow diagram for the 4-class classification of leukocytes

is defined with three linguistic variables; thus, there are nine fuzzy rules [58]. The NFC works on the adaptive neuro-fuzzy inference system (ANFIS). ANFIS is a tool which combines the input dataset (feature vectors), input membership function (inputmf), output membership function (outputmf), rule base and the output class [59]. The details about the ANFIS learning system are found in study [60]. An ANFIS structure used by the ANFC for the present work of WBC classification is shown in Fig. 9.

2.5.2 *Performance Measurement*

The performance of the proposed system is measured in terms of individual class accuracy and overall classification accuracy using given expression.

$$\mathrm{OCA} = \frac{\mathrm{CC}}{\mathrm{TI}} \times 100 \tag{2}$$

where OCA stands for overall classification accuracy, CC stands for a total number of correctly classified instances and TI stands for total number of testing instances.

$$\mathrm{ICA}_{\mathrm{class}} = \frac{\mathrm{ICC}}{\mathrm{TIC}} \times 100 \tag{3}$$

where $\mathrm{ICA}_{\mathrm{class}}$ stands for individual class accuracy, ICC stands for a total number of correctly classified instances belonging to a particular class and TIC stands for total number of testing instances belonging to a particular class.

The performance of the ANFC classifier is also measured in terms of root-mean-square-error (RMSE) value. RMSE is calculated by equation (4).

Table 4 Classification performance of ANFC (ANFC1 to ANFC6) classifier for 4-class leukocyte classification without any feature selection for the IFV1 to IFV6

IFV	*l*	ANFC	CM					Accuracy (%)	
IFV1	4	ANFC1		ESO	POL	MON	LYM	ICA	OCA
			ESO	*19*	4	0	1	79.1 (19/24)	72.8 (59/81)
			POL	3	*22*	0	8	66.6 (22/33)	
			MON	1	0	*9*	0	90.0 (9/10)	
			LYM	5	0	0	*9*	64.2 (9/14)	
IFV2	8	ANFC2		ESO	POL	MON	LYM		83.9 (68/81)
			ESO	*20*	3	1	0	83.3 (20/24)	
			POL	0	*30*	0	3	*90.9 (30/33)*	
			MON	4	0	*6*	0	60.0 (6/10)	
			LYM	0	1	1	*12*	85.7 (12/14)	
IFV3	5	ANFC3		ESO	POL	MON	LYM		86.4 (70/81)
			ESO	*17*	2	5	0	70.8 (17/24)	
			POL	2	*30*	0	1	90.9 (30/33)	
			MON	0	0	*10*	0	100 (10/10)	
			LYM	0	0	1	*13*	92.8 (13/14)	
IFV4	30	ANFC4		ESO	POL	MON	LYM		88.8 (72/81)
			ESO	*24*	0	0	0	*100 (24/24)*	
			POL	0	*29*	4	0	87.7 (29/33)	
			MON	0	0	*10*	0	*100 (10/10)*	
			LYM	0	0	5	*9*	64.2 (9/14)	
IFV5	30	ANFC5		ESO	POL	MON	LYM		86.4 (70/81)
			ESO	*20*	0	3	1	83.3 (20/24)	
			POL	0	*29*	0	4	87.7 (29/33)	
			MON	1	0	*9*	0	90.0 (9/10)	
			LYM	1	0	1	*12*	85.7 (12/14)	
IFV6	15	ANFC6		ESO	POL	MON	LYM		88.8 (72/81)
			ESO	*23*	1	0	0	95.8 (23/24)	
			POL	7	*26*	0	0	78.7 (26/33)	
			MON	0	0	*10*	0	100 (10/10)	
			LYM	1	0	0	*13*	*92.8 (13/14)*	

Italic values show the maximum accuracy

IFV input feature vector, *l* length of IFV, *ANFC* adaptive neuro-fuzzy classifier, *ESO* eosinophils, *POL* polymorphs, *MON* monocytes, *LYM* lymphocyte, *ICA* individual class accuracy, *OCA* overall classification accuracy

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} (\hat{y}_i - y_i)} \tag{4}$$

where $\hat{y}_i$ is predicted value and y_i is the observed value. For the ideal system, RMSE value tends to zero.

3 Experiments and Results

In the present work, a framework for 4 class leukocyte classification is proposed using the adaptive neuro-fuzzy classifier. Firstly, leukocytes are extracted using the Chan–Vase method and 4 shape features, 73 texture features and 15 color features (a total of ninety-two features) have been extracted. Subsequently, each feature vector group is bifurcated into training and testing instances. Finally, testing instances are fed to the classifiers to classify them into respectively class. The experimental work flow diagram for the 4-class classification of leukocytes is shown in Fig. 11.

The Sugeno rule-base viewer and membership functions for input 1 of designed ANFC for the 4-class classification of leukocyte cells is shown in Fig. 10.

The experiments conducted in the present work are given here.

Experiment 1 4-class leukocyte classification on IFV1 to IFV6 using ANFC (ANFC1 to ANFC6) without feature selection.

Experiment 2 4-class leukocyte classification on the concatenation of IFV1 to IFV6 using ANFC7 without feature selection.

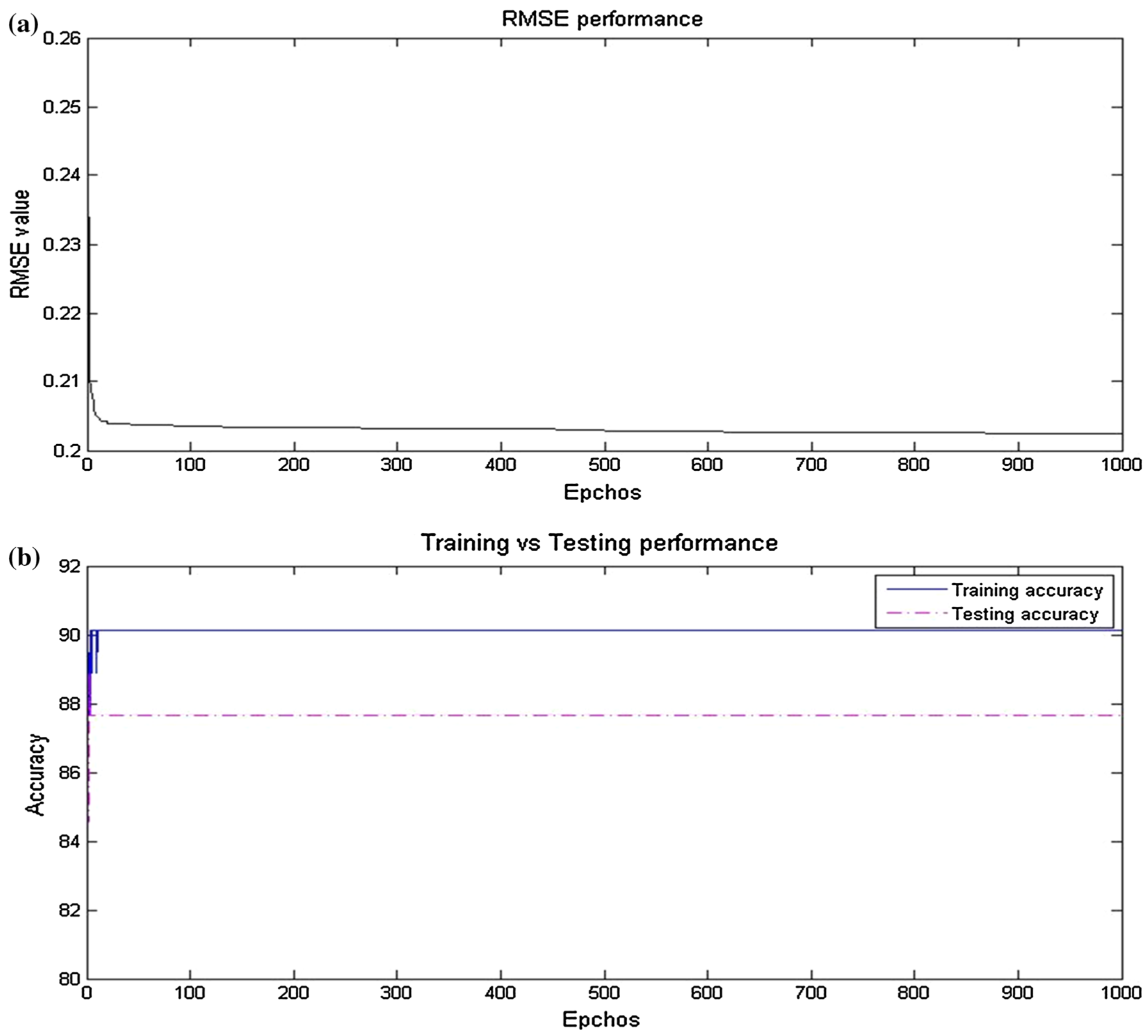


Fig. 12 **a** RMSE performance, **b** training versus testing performance for ANFC6

Table 5 Classification performance of ANFC7 classifier for 4-class leukocyte classification using NFC without feature selection on concatenation of the IFV1 to IFV6 without any feature selection

l	CM					Accuracy (%)	
92		ESO	POL	MON	LYM	ICA	OCA
	ESO	*24*	0	0	0	100 (24/24)	93.8 (76/81)
	POL	0	*32*	1	0	96.9 (32/33)	
	MON	0	0	*10*	0	100 (10/10)	
	LYM	0	0	4	*10*	71.4 (14/14)	

Italic values show the maximum accuracy
l length of IFV, *ESO* eosinophils, *POL* polymorphs, *MON* monocytes, *LYM* lymphocyte, *ICA* individual class accuracy, *OCA* overall classification accuracy

Experiment 3 4-class leukocyte classification on selected feature set of Experiment 2 using ANFC8.

3.1 Experiment 1: 4-class Leukocyte Classification Using ANFC Without Feature Selection

The classification performance of ANFC classifier for 4-class leukocyte classification without any feature selection for the IFV1 to IFV6 is reported in Table 4.

From Table 4, it can be observed that the maximum OCA value of 88.8 % is achieved from IFV4 and IFV6. The ICA values of 100, 87.7, 100 and 64.2% are obtained for ESO, POL, MON and LYM classes, respectively, from ANFC4. It is also observed that ICA values of 95.8, 78.7, 100 and 92.8% are obtained for ESO, POL, MON and LYM classes, respectively, from ANFC6.

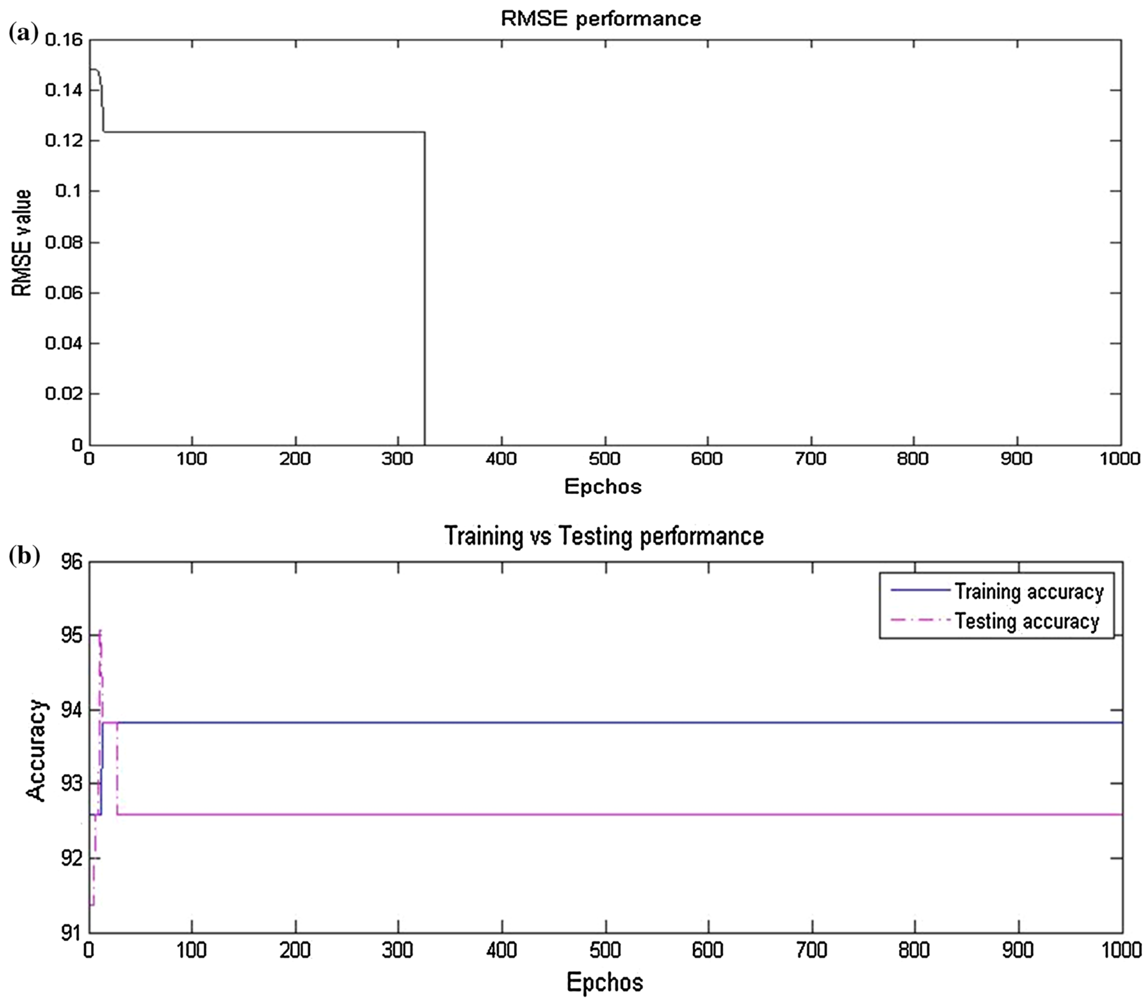


Fig. 13 **a** RMSE performance, **b** training versus testing performance for ANFC7

Table 6 Classification performance for 4-class leukocyte classification using ANFC8 with feature selection on experiment 2 feature space

L	CM					Accuracy (%)	
46		ESO	POL	MON	LYM	ICA	OCA
	ESO	*24*	0	0	0	100 (24/24)	98.7 (80/81)
	POL	0	*32*	1	0	96.9 (32/33)	
	MON	0	0	*10*	0	100 (10/10)	
	LYM	0	0	0	*14*	100 (14/14)	

Italic values show the maximum accuracy
l length of IFV, *ESO* eosinophils, *POL* polymorphs, *MON* monocytes, *LYM* lymphocyte, *ICA* individual class accuracy, *OCA* overall classification accuracy

The RMSE performance graph and training vs. testing performance graph of ANFC6 for 1000 epochs are shown in Fig. 12.

3.2 Experiment 2: 4-class Leukocyte Classification Using ANFC Without Feature Selection on the Concatenation of the IFV1 to IFV6.

The performance of ANFC classifier for 4-class leukocyte classification using ANFC7 without feature selection on the concatenation of the IFV1 to IFV6 without any feature selection is reported in Table 5.

From Table 5, it can be observed that the maximum OCA value of 93.8 % is obtained for 4-class leukocyte classification. The ICA values of 100, 96.9, 100 and 71.4% are achieved for ESO, POL, MON and LYM classes respectively.

Further, it is also observed that the OCA value is increased by 5% i.e. (93.8–88.8)% for 4-class leukocyte classification. It is also observed that the ICA value is gained by 6% in case of POL class and remain the same for ESO and MON

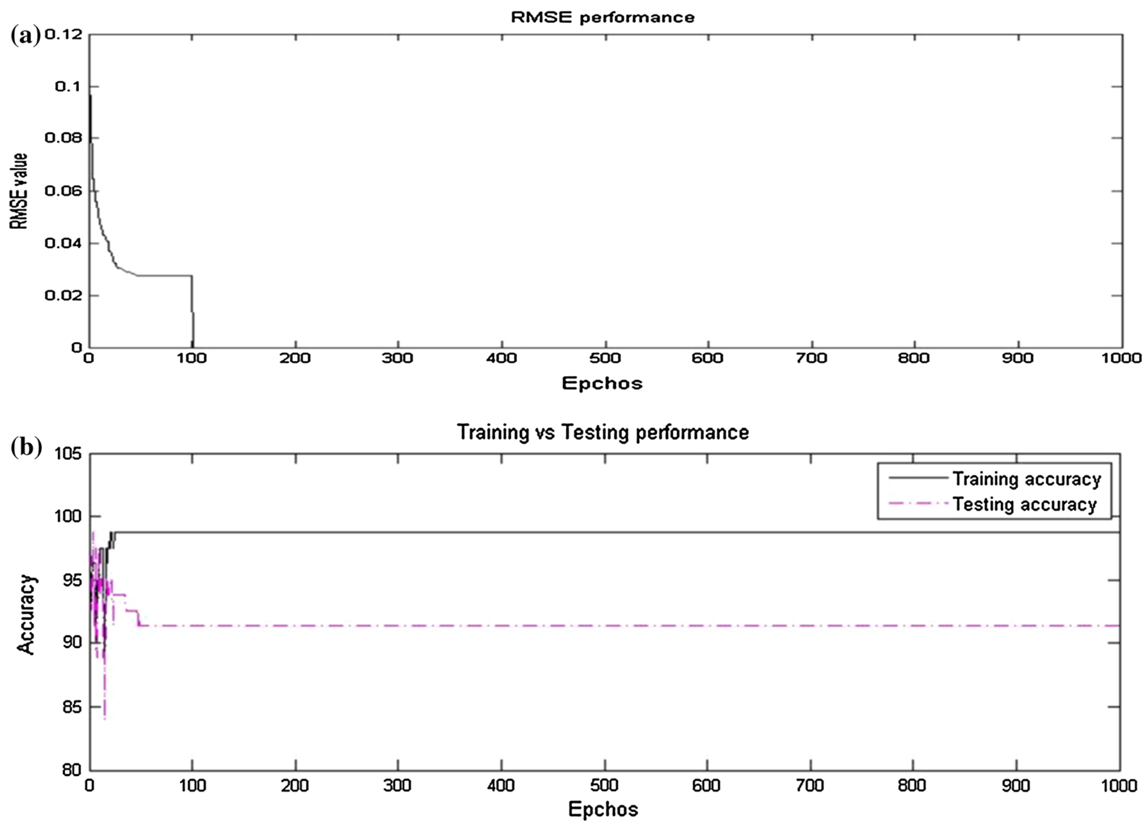


Fig. 14 **a** RMSE performance, **b** training versus testing performance for ANFC8

Table 7 Summary of misclassified instances

Exp. no.	*l*	TMI	MI_ESO	MI_POL	MI_MON	MI_LYM
1	30	9	0	5	0	4
2	92	5	0	1	0	4
3	46	1	0	1	0	0

Exp. no. experiment number, *l* length of IFV, *IFV* input feature vector, *TMI* total misclassified instance, *MI_ESO* misclassified instances of eosinophils, *MI_POL* misclassified instances of polymorphs, *MI_MON* misclassified instances of monocytes, *MI_LYM* misclassified instances of lymphocyte

classes. It is worth mentioning that ICA value is decreased 21.4% for LYM class.

The RMSE performance graph and training vs. testing performance graph of ANFC7 for 1000 epochs are shown in Fig. 13.

3.3 Experiment 3: 4-class Leukocyte Classification Using ANFC with Feature Selection on Experiment 2 Feature Space.

The performance of ANFC classifier for 4-class leukocyte classification using ANFC8 with feature selection on experiment 2 feature space is reported in Table 6.

From Table 6, the maximum OCA value of 98.7 % (80/81) is obtained for 4-class leukocyte classification. The ICA values of 100 % (24/24), 96.9 % (32/33), 100 % (10/10) and 100 % (14/14) are achieved for ESO, POL, MON and LYM classes, respectively. Further, it is also observed that the OCA value is increased by 4.9 %, i.e., (98.7–93.8) % for 4-class leukocyte classification. It is also observed that the ICA value is gained by 28.6 % in case of LYM class and remains the same for ESO, POL and MON classes.

The RMSE performance graph and training vs. testing performance graph of ANFC8 for 1000 epochs are shown in Fig. 14.

3.4 Misclassification Analysis

The total number of misclassified instances of each class from each experiment is summarized in Table 7.

After the exhaustive experimentation carried out for the 4-class classification of leukocytes, only one instance of polymorphs is misclassified. The maximum number of the misclassified instance is 9 from the *Experiment 1*. The total misclassified instances of *Experiment 1* consist of 5 instances of polymorphs and 4 instances of lymphocyte. After the *Experiment 2*, 4 instances of polymorphs are correctly classified and 4 instances of lymphocyte remain the misclassified. Thus the TMI of *Experiment 2* is 5, consisting of 1 instance of polymorphs and 4 instances of lymphocyte. After TMI of *Experiment 3* is reduced to 1 which is belongs to polymorphs class. At the end of *Experiment 3*, the achieved accuracy of the proposed system for 4-class classification system is 98.7 %, which shall be recommended for the microscopic pathology.

In this study, the feature selection based on linguistic hedges neural fuzzy classifier figured out how to dispose the redundant and irrelevant features in white blood cell dataset. Experimental results demonstrated that when the linguistic hedge value of the fuzzy classification set in any feature is near to 1, this feature is significant for that class, else it may be inappropriate. The outcomes demonstrate that utilizing of linguistic hedges in adaptive neural fuzzy classifier enhances the achievement of the classifier. The results unequivocally recommended that the proposed strategy not just able to diminish the dimensionality of large datasets but also simplify the classification task. It can assist in the diagnosis of hematological diseases by identifying leukocyte cell correctly and can be very helpful to the hematologist for their diagnostic process.

4 Conclusion

Leukocyte classification and localization is an extensively used process by a pathologist for the analysis of blood cells that is a time to consume subjective task. The present work proposed a method to conquer the lack of Otsu's thresholding to segment white cell nucleus using a Chan–Vase method that relies on global properties that give robustness for the noise by detecting object boundaries and to isolate the individual component from leukocyte image. This is essential because partitioning of the nucleus is very much simpler than the partitioning of the total cell, particularly in the bone marrow where the leukocyte density is really high. This work exhibits an efficient leukocyte classification method using the shape, color and different textural features from the nucleus and a neuro-fuzzy classification system in which, 92 features are extracted to recognize the types of leukocytes. In order to remove the redundant features and improve the classification performance of the proposed system, linguistic hedge feature selection technique is used.

In addition, experimental results indicate that neuro-fuzzy classifier is unbiased toward the classes and only one instance is misclassified, and the achieved accuracy of the proposed system for 4-class classification system is 98.7 %, which shows better classification performances. In future, the same work shall be extended for the five class classification of leukocyte with the same set of the feature set. The results obtained by utilization of selected optimal features and adaptive neuro-fuzzy classification system indicate that it can be routinely used in clinical environment for differential diagnosis between different classes of leukocytes.

In spite of the great outcomes acquired with present work, further expansions can be made to the proposed scheme. Specifically it can enhance strength and exactness of the classification task, which is a critical issue, particularly for microscopic images of blood cells. Moreover, there is an intension to apply the proposed algorithm on a bigger database.

References

1. Putzu, L.: Computer aided diagnosis algorithms for digital microscopy. Doctoral dissertation, Universita'degli Studi di Cagliari.
2. Putzu, L.; Caocci, G.; Di Ruberto, C.: Leucocyte classification for leukemia detection using image processing techniques. Artif. Intell. Med. **62**(3), 179–91 (2014)
3. Bhattacharjee, S.; Mukherjee, J.; Nag, S.; Maitra, I.K.; Bandyopadhyay, S.K.: Review on histopathological slide analysis using digital microscopy. Int. J. Adv. Sci. Technol. **62**, 65–96 (2014)
4. Yu, W.; Lü, J.; Yu, X.; Chen, G.: Distributed adaptive control for synchronization in directed complex networks. SIAM J. Control Optim. **53**(5), 2980–3005 (2015)
5. Liu, K.; Wu, L.; Lü, J.; Zhu, H.: Finite-time adaptive consensus of a class of multi-agent systems. Sci. China Technol. Sci. **59**(1), 22–32 (2016)
6. Kumar, I.; Bhadauria, H.S.; Virmani, J.; Thakur, S.: A hybrid hierarchical framework for classification of breast density using digitized film screen mammograms. Multimed. Tools Appl. **76**, 18789–18813 (2017)
7. Pang, G.; Zhuang, Y.; Zhou, P.: Automatic leukocytes classification by distance transform, moment invariant, morphological features, gray level co-occurrence matrices and SVM. In: International Conference on Information Sciences, Machinery, Materials and Energy (ICISMME 2015), pp. 1090–1095
8. Nazlibilek, S.; Karacor, D.; Ertürk, K.L.; Sengul, G.; Ercan, T.; Aliew, F.: White blood cells classifications by SURF image matching. PCA and dendrogram. Biomed. Res. **26**(4), 633–640 (2015)
9. Ravikumar, S.; Shanmugam, A.: WBC image segmentation and classification using RVM. Appl. Math. Sci. **8**(45), 2227–37 (2014)
10. Nazlibilek, S.; Karacor, D.; Ercan, T.; Sazli, M.H.; Kalender, O.; Ege, Y.: Automatic segmentation, counting, size determination and classification of white blood cells. Measurement **30**(55), 58–65 (2014)
11. Habibzadeh, M.; Krzyżak, A.; Fevens, T.: Comparative study of shape, intensity and texture features and support vector machine for white blood cell classification. J. Theor. Appl. Comput. Sci. **7**(1), 20–35 (2013)
12. Rezatofighi, S.H.; Soltanian-Zadeh, H.: Automatic recognition of five types of white blood cells in peripheral blood. Comput. Med. Imaging Graph. **35**(4), 333–43 (2011)

13. 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. **85**(9), 2104–18 (2012)
14. Ramesh, N.; Dangott, B.; Salama, M.E.; Tasdizen, T.: Isolation and two-step classification of normal white blood cells in peripheral blood smears. J. Pathol. Inform. **3**(1), 13 (2012)
15. Na, L.; Chris, A.; Mulyawan, B.: A combination of feature selection and co-occurrence matrix methods for leukocyte recognition system. J. Softw. Eng. Appl. **5**(12), 101 (2013)
16. Rezatofighi, S.H.; Soltanian-Zadeh, H.: Automatic recognition of five types of white blood cells in peripheral blood. Comput. Med. Imaging Graph. **35**(4), 333–343 (2011)
17. Xie, E.; McGinnity, T.M.; Wu, Q.: Automatic extraction of shape features for classification of leukocytes. In: 2010 International Conference on Artificial Intelligence and Computational Intelligence (AICI), 2010 Oct 23, vol. 2, pp. 220–224. IEEE (2010)
18. Ghosh, M.; Das, D.; Mandal, S.; Chakraborty, C.; Pala, M., Maity, A.K.; Pal, S.K.; Ray, A.K.: Statistical pattern analysis of white blood cell nuclei morphometry. In: Students' Technology Symposium (TechSym), 2010 IEEE, pp. 59–66. IEEE (2010)
19. Rodrigues, P.; Ferreira, M.; Monteiro, J.: Segmentation and classification of leukocytes using neural networks: a generalization direction. In: Prasad, B., Prasanna, S.R.M. (eds.) Speech, Audio, Image and Biomedical Signal Processing Using Neural Networks, vol. 83, pp. 373–396. Springer, Berlin (2008)
20. Yampri, P.; Pintavirooj, C.; Daochai, S.; Teartulakarn, S.: White blood cell classification based on the combination of eigen cell and parametric feature detection. In: 2006 1ST IEEE Conference on Industrial Electronics and Applications, 2006 May 24, pp. 1–4. IEEE (2006)
21. Piuri, V.; Scotti, F.: Morphological classification of blood leucocytes by microscope images. In: 2004 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2004. CIMSA. 2004 Jul 14, pp. 103–108. IEEE (2004)
22. Bikhet, S.F.; Darwish, A.M.; Tolba, H.A.; Shaheen, S.I.: Segmentation and classification of white blood cells. In: 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings, vol. 4, pp. 2259–2261. IEEE (2000)
23. Bacusmber, J.W.; Gose, E.E.: Leukocyte pattern recognition. IEEE Trans. Syst. Man Cybern. **2**(4), 513–26 (1972)
24. Young, I.T.: The classification of white blood cells. IEEE Trans. Biomed. Eng. **4**, 291–8 (1972)
25. Adjouadi, M.; Zong, N.; Ayala, M.: Multidimensional pattern recognition and classification of white blood cells using support vector machines. Part. Part. Syst. Charact. **22**(2), 107–18 (2005)
26. Sarrafzadeh, O.; Rabbani, H.; Talebi, A.; Banaem, H.U.: Selection of the best features for leukocytes classification in blood smear microscopic images. In: SPIE Medical Imaging 2014 Mar 20, pp. 90410P–90410P. International Society for Optics and Photonics (2014)
27. Tabrizi, P.R.; Rezatofighi, S.H.; Yazdanpanah, M.J.: Using PCA and LVQ neural network for automatic recognition of five types of white blood cells. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2010 Aug 31, pp. 5593–5596. IEEE (2010)
28. Stadelmann, J.V.; Spiridonov, I.N.: Automated classification of leukocytes in blood smear images. Biomed. Eng. **1**, 1–5 (2012)
29. Suapang, P.; Chivaprecha, S.: Automatic leukocyte classification. Int. J. Appl. **8**(1), 39–46 (2015)
30. Mirčić, S.; Jorgovanović, N.: Automatic classification of leukocytes. J. Autom. Control **16**(1), 29–32 (2006)
31. Ferri, M.; Lombardini, S.; Pallotti, C.: Leukocyte classifications by size functions. In: Proceedings of the Second IEEE Workshop on Applications of Computer Vision, 1994, pp. 223–229. IEEE (1994)
32. Hiremath, P.S.; Bannigidad, P.; Geeta, S.: Automated identification and classification of white blood cells (leukocytes) in digital microscopic images. In: IJCA special issue on "recent trends in image processing and pattern recognition" RTIPPR, pp. 59–63 (2010)
33. Azar, A.T.; El-Said, S.A.; Balas, V.E.; Olariu, T.: Linguistic hedges fuzzy feature selection for differential diagnosis of Erythemato-Squamous diseases. Soft Comput. Appl. **195**, 487–500 (2013)
34. Rawat, J.; Singh, A.; Bhadauria, H.S.; Virmani, J.; Devgun, J.S.: Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia. Biocybern. Biomed. Eng. **37**, 637–654 (2017)
35. Mohamed, M.: Image dataset with ground truth images and code. Retrieved on 1-May-2015; MatlabFile exchange from: http://www.mathworks.com/matlabcentral/fileexchange/36634-an-efficient-technique-for-white-blood-cellsnuclei (2012)
36. Rawat, J.; Bhadauria, H.S.; Singh, A.; Virmani, J.: Review of leukocyte classification techniques for microscopic blood images. In: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1948–1954. IEEE (2015)
37. Putzu, L.; Di Ruberto, C.: White blood cells identification and classification from leukemic blood image. In: Proceedings of the IWBBIO International Work-Conference on Bioinformatics and Biomedical Engineering, pp. 99–106 (2013)
38. Theerapattanakul, J.; Plodpai, J.; Pintavirooj, C.: An efficient method for segmentation step of automated white blood cell classifications. In: 2004 IEEE Region 10 Conference TENCON 2004, pp. 191–194. IEEE (2004)
39. Otsu, N.: A threshold selection method from gray-level histograms. Automatica **11**(285–296), 23–7 (1975)
40. Gonzalez, R.C.; Woods, R.E.; Eddins, S.L.: Digital image processing using MATLAB, 3rd edn. Prentice Hall, Upper Saddle River, New Jersey (2004)
41. Rawat, J.; Singh, A.; Bhadauria, H.S.; Kumar, I.: Comparative analysis of segmentation algorithms for leukocyte extraction in the acute Lymphoblastic Leukemia images. In: 2014 International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 245–250. IEEE (2014)
42. Rawat, J.; Singh, A.; Bhadauria, H.S.: An approach for leukocytes nuclei segmentation based on image fusion. In: 2014 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 000456–000461. IEEE (2014)
43. Agaian, S.; Madhukar, M.; Chronopoulos, A.T.: Automated screening system for acute myelogenous leukemia detection in blood microscopic images. IEEE Syst. J. **8**(3), 995–1004 (2014)
44. Madhukar, M.; Agaian, S.; Chronopoulos, A.T.: Deterministic model for acute myelogenous leukemia classification. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 433–438. IEEE (2012)
45. Baker, Q.B.; Balhaf, K.: Exploiting GPUs to accelerate white blood cells segmentation in microscopic blood images. In: 2017 8th International Conference on Information and Communication Systems (ICICS), pp. 136–140 (2017)
46. Chan, T.F.; Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–77 (2001)
47. Yang, M.; Kpalma, K.; Ronsin, J.: A survey of shape feature extraction techniques. In: Pattern Recognition, pp. 43-90. InTech (2008)
48. Haralick, R.M.; Shanmugam, K.; Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
49. Rawat, J.; Singh, A.; Bhadauria, H.S.; Virmani, J.; Devgun, J.S.: Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers. Multimedia Tools Appl. **76**(18), 19057–19085 (2017)
50. Rawat, J.; Singh, A.; Bhadauria, H.S.; Virmani, J.: Computer aided diagnostic system for detection of leukemia using microscopic images. Procedia Comput. Sci. **1**(70), 748–56 (2015)

51. Abenius, T.: Classification of cell images using MPEG-7-influenced descriptors and support vector machines in cell morphology. arXiv preprint arXiv:0812.2309. (2008 Dec 12).
52. Kriti,; Virmani, J.: Breast density classification using Laws' mask texture features. Int. J. Biomed. Eng. Technol. **19**(3), 279–302 (2015)
53. Laws, K.I.: Rapid texture identification. In: 24th Annual Technical Symposium, pp. 376–381. International Society for Optics and Photonics (1980 Dec 23)
54. Lee, C.C.; Chen, S.H.: Gabor wavelets and SVM classifier for liver diseases classification from CT images. In: SMC'06. IEEE International Conference on Systems, Man and Cybernetics, 2006, vol. 1, pp. 548–552. IEEE (2006 Oct 8)
55. Han, Z.Y.; Gu, D.H.; Wu, Q.E.: Feature extraction for color images. In: Hussain, A. (ed.) Electronics, Communications and Networks V, vol. 382, pp. 215–221. Springer, Singapore (2016)
56. Cetisli, B.: Development of an adaptive neuro-fuzzy classifier using linguistic hedges: part 1. Expert Syst. Appl. **37**(8), 6093–101 (2010)
57. Cetisli, B.: The effect of linguistic hedges on feature selection: part 2. Expert Syst. Appl. **37**(8), 6102–8 (2010)
58. Kher, R.; Pawar, T.; Thakar, V.; Shah, H.: Physical activities recognition from ambulatory ECG signals using neuro-fuzzy classifiers and support vector machines. J. Med. Eng. Technol. **39**(2), 138–52 (2015)
59. Do, Q.H.; Chen, J.F.: A neuro-fuzzy approach in the classification of students' academic performance. Comput. Intell. Neurosci. **1**(2013), 6 (2013)
60. Jang, J.S.: ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. **23**(3), 665–85 (1993)