Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images

Q1 Roopa B. Hegde [a,b], Keerthana Prasad [a,*], Harishchandra Hebbar [a], Brij Mohan Kumar Singh [c]

Q2 [a] School of Information Sciences, MAHE, Manipal 576104, India
[b] Dept. of ECE, NMAM Institute of Technology, Nitte, Karkala 574110, India
[c] Dept. of Immunohematology and Blood Transfusion, KMC, MAHE, Manipal, Karnataka 574106, India

ARTICLE INFO

Article history:
Received 14 October 2018
Received in revised form
28 January 2019
Accepted 30 January 2019
Available online xxx

Keywords:
Peripheral blood smear analysis
White blood cells
Classification
Machine learning
Deep learning

ABSTRACT

Automated classification and morphological analysis of white blood cells has been addressed since last four decades, but there is no optimal method which can be used as decision support system in laboratories due to biologically complex nature of the cells. Automated blood cell analysis facilitates quick and objective results and can also handle massive amount of data without compromising with efficiency. In the present study, we demonstrate classification of white blood cells into six types namely lymphocytes, monocytes, neutrophils, eosinophils, basophils and abnormal cells. We provide the comparison of traditional image processing approach and deep learning methods for classification of white blood cells. We evaluated neural network classifier results for hand-crafted features and obtained the average accuracy of 99.8%. We also used full training and transfer learning approaches of convolutional neural network for the classification. An accuracy around 99% was obtained for full training CNN.

1. Introduction

Peripheral blood smear analysis is used for diagnosis of many diseases such as anemia, leukemia, malaria, etc. It consists of red blood cells (RBCs), white blood cells (WBCs) and platelets. There are five types of WBCs namely lymphocyte, monocyte, neutrophil, eosinophil and basophil as shown in Fig. 1(a–e). Increase or decrease in number of WBCs in peripheral blood indicates disorders. Also morphological variations such as shape, size and color variations help in diagnosis. Thus detection and classification of WBCs play a vital role in peripheral blood smear analysis [1]. Manual evaluation of WBCs includes total WBC count, differential WBC count and morphological analysis. Manual assessment of WBCs is laborious and prone to inter-observer variations [2,3].

* *Corresponding author at*: School of Information Sciences, MAHE, Manipal 576104, India.
E-mail address: keerthana.prasad@manipal.edu (K. Prasad).
https://doi.org/10.1016/j.bbe.2019.01.005

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

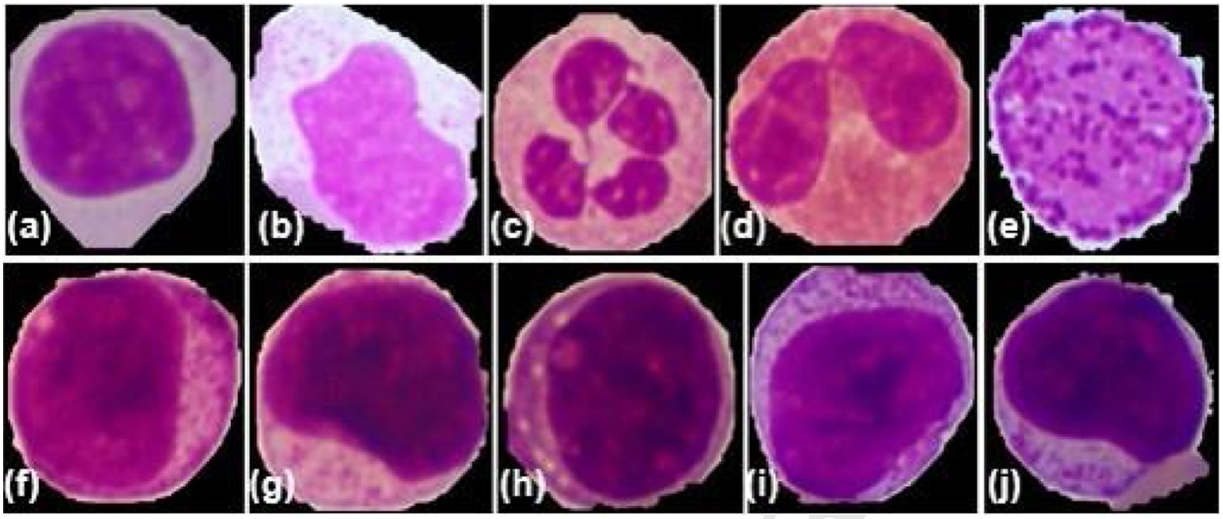

Fig. 1 – WBC types: (a) lymphocyte, (b) monocyte, (c) neutrophil, (d) eosinophil, (e) basophil, and (f–j) abnormal.

Computer aided WBC analysis is gaining popularity to reduce the workload of hematologists [4]. Automated blood cell analysis facilitates quick and objective results [5,6]. It can also handle massive amount of data efficiently [7]. There are two ways to achieve automated classification namely traditional image processing approach and deep learning approach.

Traditional image processing approach involves interconnected steps such as segmentation, feature extraction and classification. Segmentation of WBCs using image processing approach is not a difficult task due to presence of dark nucleus, but accurate detection of border of WBCs and separation of overlapped cells are challenging tasks in automated detection and classification of WBCs [8]. Many methods have been proposed for classification of WBCs using neural networks. It is a supervised machine learning algorithm which consists of input layer, hidden layer and output layer. Hence the prediction depends on given features in this method.

Deep learning method operates directly on raw pixels and learns features by itself. Hence deep learning approaches such as stacked autoencoders and CNN have also been adapted for medical image analysis which is robust to noise and illumination variations.

Many studies reported the use of stacked autoencoders for medical image segmentation and classification [9–12]. Autoencoder belongs to deep learning category of machine learning which functions similar to neural network. It is a learning method which transforms inputs to outputs with minimum possible error. Input to autoencoder can be a feature set or images. It consists of encoders, decoders and a loss function. Encoder is a neural network which produces output y depending on size of hidden layer from given input x. Decoder is also a neural network which produces output x from given input y. The output of an encoder is given as input to the decoder. It uses output functions such as sigmoid, softmax, etc ., for classification.

CNN is commonly used in deep learning for image classification. It can be employed for medical image classification in three different ways namely full-training CNN, transfer learning and CNN as feature generator. In case of full training, network is designed from scratch. This requires large number of labeled data which is hard to get in medical domain, whereas in transfer learning approach, network is pre-trained with large number of non-medical data which can be fine-tuned for a specific application. In the third approach, features are extracted from certain layers of the pre-trained network. These features can be later used to train external classifiers.

Automated classification of WBCs using traditional image processing approach has been addressed in the last four decades. These methods either used grayscale representation of original images or any color components. Jaroonrut et al. [13] used Naive Bayes classifier for detection of WBC types which provided an accuracy around 98%. Naive Bayesian classifier with incremental learning classifier was used by Mathur et al. [14] for classification of WBCs and reported overall accuracy around 92%. Classification of WBCs was also performed considering neural network classifier [15,16]. Sedat et al. [15] compared neural network results with and without using PCA algorithm and achieved overall accuracy around 95% with PCA and 65% without using PCA. Seyed et al. [16] evaluated the performances of SVM and NN classifier for classification of WBCs into five types, and reported classification accuracy between 90% and 96% using SVM. Also, Neelam et al. [17] compared the results of SVM and NN for classification of WBCs, thereby reporting an average accuracy around 97% using NN and 94% using SVM. Qingli et al. [18] evaluated morphological parameters for classification of WBCs and reported classification accuracy of 94%. Omid et al. [19] proposed a method for detection and classification of normal WBCs which resulted in correct classification rate of 93%. Siroic´ and Moradi et al. [20,21] used SVM for detection of lymphoblast WBCs and reported an average accuracy of 96%. Rawat et al. [22] used a combination of genetic algorithm and SVM classifier for detection of acute leukemia from peripheral blood smear images. Mohapatra et al. [23] proposed a method for detection of ALL using PBS images. Several classifiers namely NB, KNN, MLP, SVM and ensemble of classifiers were evaluated for detection of ALL. They obtained an average classifier sensitivity around 90% using ensemble of classifiers.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

WBC analysis using traditional image processing methods is a challenging task due to its complex biological appearance, staining method used and illumination variations in acquired images [24]. A few researchers used autoencoders for medical image segmentation and classification. Chang et al. [9] showed that, stacked autoencoder can be effectively used for unsupervised feature learning when obtaining labeled samples is difficult. Andrew et al. [10] proposed stain normalization method for histopathology images using sparse autoencoders. Jose et al. [11] proposed a method for segmentation of brainstem using stacked autoencoders. MRI images of brain cancer patients were used in their study and reported similarity of 90% with respect to ground truth. Bejoy et al. [12] proposed a method for diagnosis of prostate cancer from MRI images using a combination of sparse autoencoders and random forest classifiers. The authors reported an accuracy of 93.65%.

Many researchers employed CNN for classification of WBCs in recent years [25–37]. Zhang et al. [25] proposed an ensemble method which was a combination of SVM (traditional method) and CNN for identification of WBCs in peripheral blood smear images. The reported average classification accuracy was about 92.5% using SVM classifier, 89.5% using CNN and 93.5% using ensemble method. The results were reported using 10,000 WBC images and 10,000 'non-WBC' images for detection of WBCs. Zhao et al. [26] proposed an automated method for detection and classification of WBCs using CNN approach. Detection of eosinophils and basophils were reported using handcrafted features by training SVM classifier. Classification of lymphocytes, monocytes and neutrophils was carried out by extracting features from CNN to train Random forest classifier. The reported accuracy of classification was between 74% and 97% depending on the type of WBCs. Shahin et al. [27] used deep CNN for classification of WBCs. Classification was performed by designing CNN from scratch and also by using pre-trained network. Overall system accuracy of about 96.1% was reported using 2551 images. Luis et al. [28] presented a method for diagnosis of leukemia using transfer learning approach of CNN. Several pre-trained networks such as AlexNet, CaffeNet and VGGNet were used for feature extraction. Further, gain ratio method was considered for feature selection. SVM was trained using the selected features to detect leukemic WBCs and obtained accuracy around 99% for detection of leukemic WBCs. Feiwei et al. [29] proposed a method for classification of WBCs using deep residual network. The deep residual neural classifier was tested on microscopic image dataset with 40 leukocyte categories. The average accuracy on the test set was nearly 76.84%. Choi et al. [30] proposed an eight layered CNN for classification of blood cells into ten classes and reported an accuracy around 96%. Tiwari et al. [31] developed a seven layered CNN for classification of WBCs into the five types. Results of the network were compared with that of SVM and Naive Baye's classifiers. The best accuracy of 95% was reported for the CNN. Rehman et al. [32] developed a CNN for classification of subtypes of ALL. Experimental results were compared with that of Naïve Bayee's, KNN, and SVM classifiers. The authors reported that, the best accuracy of 97.78% was achieved for the CNN. Shafique and Tehsin [33] employed AlexNet for classification of sub-types of ALL and obtained accuracy around 96%. CNN was efficiently used for HEp-2 cell classification [34], malaria diagnosis [35,36] and for automatic detection of tuberculosis [37] using microscopic images.

In this paper, we present automated classification of WBCs based on morphological features such as shape, color, size and texture using traditional image processing method. This study also attempts to explore application of deep learning method for classification of WBCs. A comparative study based the results obtained from both the approaches is provided.

2. Materials and methods

This section provides the details of data collection, feature extraction and classification. Features such as shape, color and texture are extracted in the present study. The details of the segmentation and extracted features are given in Section 2.2. We considered classification of WBCs into five subtypes and abnormal as shown in Fig. 1. We experimented on the classification by considering three approaches namely NN, autoencoders and CNN. The details of the classifiers are provided in Section 2.3.

2.1. Data collection

We collected 1159 PBS images consisting of 1418 cropped images which include 170 lymphocytes, 109 monocytes, 297 neutrophils, 154 eosinophils, 81 basophils and 607 abnormal WBCs. Abnormal WBCs include reactive lymphocytes, degenerated cells, myelocytes and leukemic leukocytes which are generally called as blasts. These images contained RBCs, WBCs and platelets. The original images were cropped to obtain cropped images using bounding box method which consist of single leukocyte. Hence, the pre-processing step in the proposed method involves cropping the original images to obtain single leukocyte images with variable image size based on the size of leukocytes. In CNN approach, input images with equal number of rows and columns are expected. Also, in order to compare the performance of the proposed CNN with that of AlexNet, the cropped images were resized to a size of 227×227.

In order to obtain the size of dataset to be used in traditional image processing approach, sample size calculation was performed as mentioned in [38]. It was found to be 162 with significance level of 0.5 and error rate of 12%. However, we considered 280 cropped images consisting of 26 lymphocytes, 28 monocytes, 66 neutrophils, 31 eosinophils, 32 basophils and 137 abnormal WBCs so as to balance the cells of each type of the WBCs. This dataset used in traditional image processing consists of 320 cropped images.

The acquired dataset consists of 1418 cropped images. Training a deep network using CNN requires large dataset. Hence, data augmentation was considered to increase the number of images in the dataset. Data augmentation methods used in the present study are color-balancing method and non-uniform brightness variations. Color and brightness variations can be commonly observed in microscopic images. Hence in the data augmentation, color and brightness variations are introduced onto the original images.

In the color-balancing method, arithmetic operations of mean values of R, G and B components and mean value of

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

grayscale representation of the original image were performed. The steps of color-balancing method are as follows:

(a) Read original image I_{RGB}
(b) Extract each component of I_{RGB} to obtain I_R, I_G and I_B
(c) Compute mean of each of the components to obtain $Mean_R$, $Mean_G$ and $Mean_B$
(d) Convert I_{RGB} to grayscale representation to obtain I_{gray}
(e) Calculate mean value of I_{gray} to obtain $Mean_{gray}$
(f) Use the relation given in Eq. (1) to make all the components with equal mean

$$Im_{channel} * \frac{Mean_{gray}}{Mean_{channel}} \quad (1)$$

where 'channel' corresponds to R, G and B components which results in Im_R, Im_G and Im_B (g). Combine Im_R, Im_G and Im_B to obtain color-balanced image I_{RGB1}. In this method, R, G and B components were multiplied with mean value of grayscale representation of original images and the resultant images were divided by mean values of R, G and B components respectively as given in Eq. (1). A few resultant images are shown in Fig. 2. In the figure, images from A1–A4 are original images and B1–B4 are the resultant images after the application of 'color-balancing' method. Color variation can be observed between original images and the augmented images. Though there is a color variation, the leukocytes did not lose the original information such as texture and color of nuclei. The method was applied on the 'training-set' which consists of 1196 cropped images and also on 'test-set' of size 387.

The images were subjected to non-uniform brightness variation. This was achieved by using a scaling profile $S(k)$ which was multiplied to each row of the original image as given in Eq. (2).

$$Im_{non\text{-}uniform} = Im_{original} * S(k) \quad (2)$$

where $S(k)$ is a scaling profile with linearly spaced values. The algorithm steps to generate $S(k)$ are as follows:

(a) Obtain the original image I_{RGB}
(b) Obtain the size of the original image as 'm' and 'n'
(c) For horizontal variation, consider limit = 'm', else limit = 'n'
(d) Define scaling factor $a = 0.25$
(e) Define start = $1 - a$ and stop = $1 + a$
(f) For index = 1: limit, compute S(index) as given in Eq. (3)

$$S(index) = start + (index - 1) * \frac{(end - start)}{(limit - 1)} \quad (3)$$

The scaling function thus generated consists of 'limit' number of samples which are linearly varying between 0.75 and 1.25. To introduce non-uniform illumination variation horizontally across an image, the scaling function thus obtained was multiplied to each row of the image. Similarly, to introduce non-uniform variation from top to bottom of an image in a vertical manner, the scaling function was multiplied to each column of the image. The images thus generated are shown in Fig. 3(C1–C4). Increase in brightness from left to right can be observed in Fig. 3(C1, C2) and vertical brightness variation can be observed in Fig. 3(C3, C4). Images of abnormal class and images of neutrophils were excluded for non-uniform brightness variation to balance the number of images in each class.

The augmented dataset consists of 2697 cropped images in 'training-set' and 816 cropped images in 'test-set'. A care was taken so that images in the 'test-set' were not included in the 'training-set'. The 'training-set' was used for designing the CNN from scratch. The 'test-set' was used to evaluate the performance of the designed CNN.

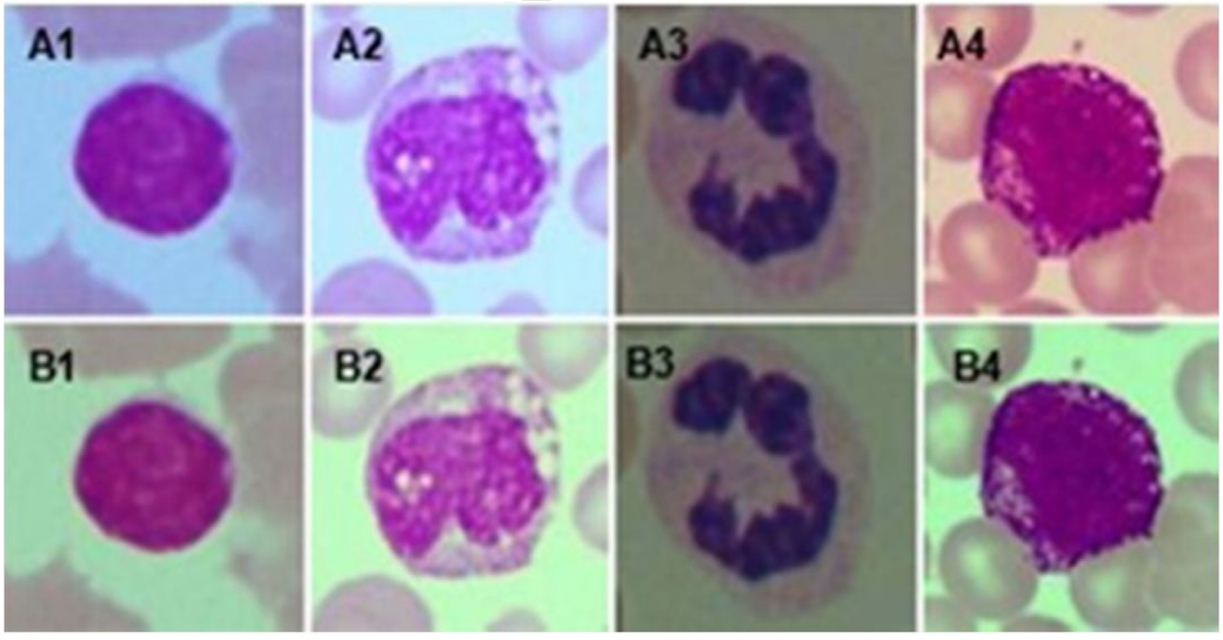

Fig. 2 – Results of 'color-balancing' method; A1–A4: original images, B1–B4: augmented images.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

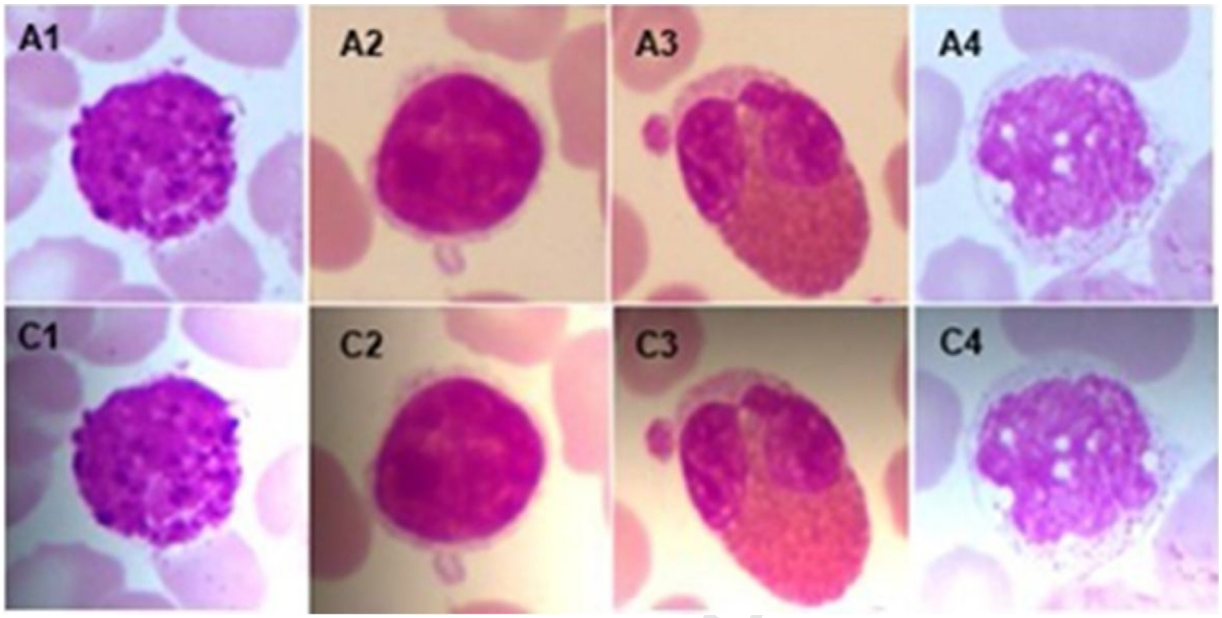

Fig. 3 – Results of non-uniform brightness variation A1–A4: original images C1–C4: augmented images.

2.2. *Pre-processing and feature extraction in traditional image processing approach*

In order to detect WBCs, a semi-automated method was considered. Detection of nucleus from the PBS image was considered as a first step using an automated method as described in Ref. [39]. Location of nucleus was used to crop the original images using bounding box method. This resulted in single leukocyte images with variable image size based on the size of leukocytes. In order to detect WBCs, a suitable threshold value was applied on the cropped image. The threshold value was varied between 0.4 and 0.75 manually based on the color and brightness of the images. Area filter, morphological opening and morphological closing operations were applied to obtain the region of WBCs. Morphological opening with disk shaped structuring element of radius 15 was applied and all the areas below 10,000 pixels were removed using area filter. In order to obtain a smooth boundary of the WBCs, morphological closing of disk shaped structuring element of radius 10 was applied. Region of cytoplasm was obtained by subtracting the region of nucleus from the WBC. Results of detection of nucleus, WBC and cytoplasm are shown in Fig. 4.

2.2.1. *Feature extraction*

Features extracted from ROI region play significant role in classification of WBCs in traditional image processing approach. Shape features were extracted from nuclei, color and texture features were extracted from both nuclei and cytoplasm.

2.2.2. *Shape features*

Shape features such as area, perimeter, circularity, convexity and solidity were extracted from the nuclei of WBCs. We also computed nucleus to cell ratio (NC ratio) as given in Table 1. All these shape features were extracted from binary representation of the nuclei.

2.2.3. *Color features*

Cytoplasm of WBCs vary in color depending on the type as shown in Fig. 1(a–e). Also nuclei of abnormal WBCs show color variation as shown in Fig. 1(f–j). Hence, mean and variance of R, G and B components of RGB image, H, S and V components of HSV color space and L, A and B components of CIE LAB color space were extracted from the nuclei and cytoplasm of WBCs.

2.2.4. *Texture features*

Significant texture variations can be observed in cytoplasm of WBCs as shown in Fig. 1. We extracted texture features such as mean, variance, skewness, kurtosis, Spatial Gray Level Dependence Matrix (SGLDM) features and Laws texture features from the grayscale representation of original images and Local Binary Pattern (LBP) representation of WBCs. The feature images are shown in Fig. 5, images in first row are grayscale representation of original images and images in second row are LBP representation of WBCs.

LBP is a unified approach for statistical and structural texture analysis and is a powerful tool to describe local textures [40]. We used circular LBP of radius 1. The decimal value of resulting LBP given a pixel at (x, y) can be computed as follows:

$$\mathrm{LBP}_{(P,R)}(x,y) = \sum_{p=0}^{p-1} S(i_p - i_c)2^p \tag{4}$$

where P is neighborhood, R is radius, i_c central pixel value, i_p neighboring pixel value. Function $S(x)$ is defined as follows:

$$S(x) = \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ if } x < 0 \end{cases} \tag{5}$$

The texture features were computed from these two feature images. SGLDM computes 13 features in four directions for

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

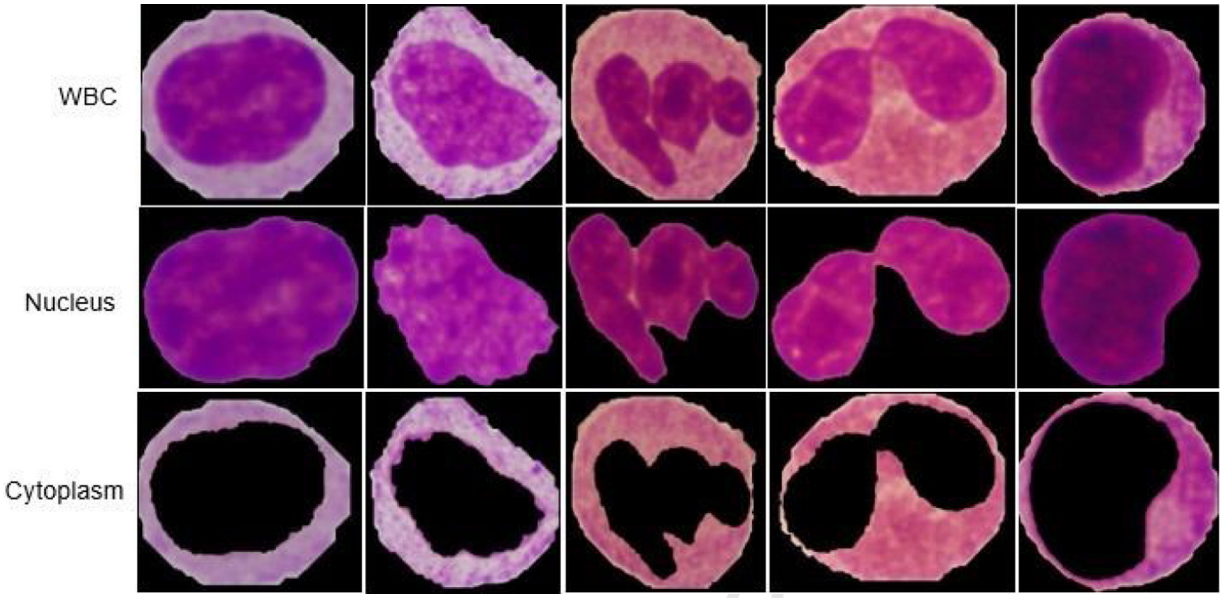

Fig. 4 – Results of segmentation.

Table 1 – Details of the shape features.

Feature	Relation
Area	Total number of pixels
Perimeter	Path length that surrounds area
Circularity	Perimeter^2/4*pi*area
Convexity	Perimeter of convexhull/perimeter of nucleus
Solidity	Area of nucleus/area of conxhull
NC ratio	Area of nucleus/area of WBC

angles 0°, 45°, 90° and 135°. The mean and the range of these features over the four angles are taken as feature set. The extracted statistical measures include angular second moment, contrast, correlation, variance, inverse difference moment, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation, information measure of correlation. The mean and range of these features in all four directions results in 26 texture features.

Laws texture measure is performed using convolution operation and non-linear windowing operation on images. The 2D filters for convolution operation are generated from 1D filters namely average gray level (L), edges (E), spots (S) waves (W) and ripples (R) of length 5. The 1D convolution kernels are as given L = [1 4 6 4 1], E = [1 2 0 2 1], S = [1 0 2 0 1], W = [1 2 0 2 1], R = [1 4 6 4 1]. It computes energy at pixel by summing absolute value of convolved output. This produces nine energy maps whose average values are used as texture measures [41].

2.3. *Classification*

The goal of classification is to detect subtypes of WBCs and abnormal WBCs using the extracted features. In the present

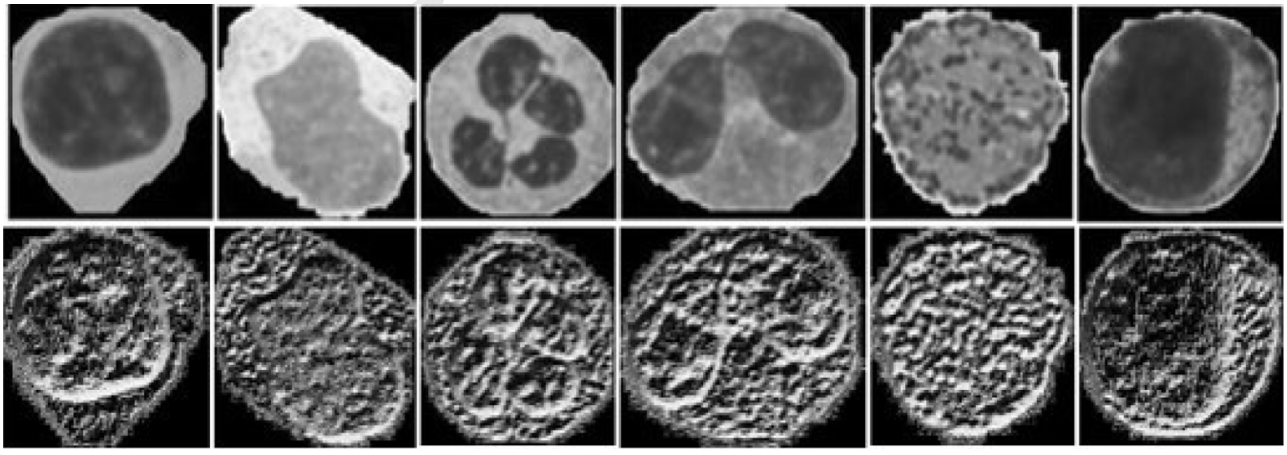

Fig. 5 – Feature images (first row) gray scale representation of original images (second row) LBP images.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

study, classification of WBCs using NN, autoencoders and CNN is proposed.

2.3.1. *Neural network*

Neural network is a supervised machine learning algorithm which is inspired by the biological nervous system. It consists of input layer, hidden layer and output layer which are linked by weighted connections. These weights are numerical values; the classifier performance can be approximated to the desired result by altering these values. The input layer takes in the inputs as features namely shape, color and texture. The output layer produces classified output namely normal and abnormal WBCs based on weights, number of neurons in the hidden layer and error function. We used cross-entropy error function with a hidden layer consisting of 11 neurons.

2.3.2. *Autoencoders*

In the present study, the autoencoder network was considered for two different types of inputs namely the extracted features as inputs and original images as inputs. In latter approach, we used 320 cropped images as in the case of traditional image processing approach.

A stacked autoencoder with three hidden layers and softmax function in the output layer was considered for classification of WBCs using the extracted features. We used total of 132 features as input to the network. Three hidden layers of size 130, 90 and 50 were used in the present application. We used 1000 epochs to train each layer and softmax layer with cross-entropy loss function.

We also considered WBC images as input to train stacked autoencoder network. Input image of size $30 \times 30 \times 3$ with 80% of the images for training and remaining 20% for testing was considered. We considered three layered stacked autoencoder network with hidden layer sizes of 1500, 500 and 50. Each layer was trained for 500 epochs and the output softmax layer was trained for 1000 epochs.

2.3.3. *Convolution neural network*

We used both full-training from scratch and transfer learning method in the present application. In full-training, we trained CNN by adding layers each time to evaluate the performance. This was considered to investigate the impact of the number of layers on classification of WBCs. The designed CNN is as listed in Table 2. We initially considered layer 1 and layer 8 for training and testing the performance. We added layer 7 and layer 6 which resulted in network with one convolutional layer and three fully connected layers. We considered layers 1, 2 and 8 for the classification and later added layer 7 and layer 6 followed by layer 3, layer 4 and layer 5. The results of each layer combination are provided in the result section. We used stochastic gradient descent method (SGDM) for training with initial learn rate of 0.001, mini batch size of 20 and maximum epoch of 1000. The input layer consists of WBC images of size $227 \times 227 \times 3$.

Table 2 – Architecture details of full training CNN.

Layer	Details
Layer 1	Convolution layer [11,11], 96, stride (4,4), padding (0,0) ReLU Max pooling layer (3, stride 2)
Layer 2	Convolution layer [5,5], 256, stride (1,1), padding (2,2) ReLU Max pooling layer (3, stride 2)
Layer 3	Convolution layer [3,3], 384, stride (1,1), padding (1,1) ReLU Max pooling layer (3, stride 2)
Layer 4	Convolution layer [3,3], 384, stride (1,1), padding (1,1) ReLU
Layer 5	Convolution layer [3,3], 256, stride (1,1), padding (1,1) ReLU Max pooling layer (3, stride 2)
Layer 6	Fully connected layer (500)
Layer 7	Fully connected layer (256)
Layer 8	Fully connected layer (6)

AlexNet [42] which is a pre-trained network was also employed for classification of WBCs. It is an eight layered architecture with five convolution layers and three fully connected layers as shown in Fig. 6. ReLU and max pooling layers are also included after each convolution layer.

The cropped WBC images were resized to $227 \times 227 \times 3$ as per the requirement of input layer of the network. We considered changing the number of nodes in the last fully connected layer to 6 for the present application. We used SGDM for training the network with initial learning rate of 0.001, mini batch size of 20 and maximum epoch of 100. We used 80% of the data for training and remaining 20% for testing in our study.

3. Results

In this section we provide the performances of NN, autoencoders and CNN for classification of WBCs into six types namely lymphocyte, monocyte, neutrophil, eosinophil, basophil and abnormal WBCs. We also provide the comparison of classification results of NN with deep learning algorithms namely stacked autoencoders and CNN for classification of WBCs.

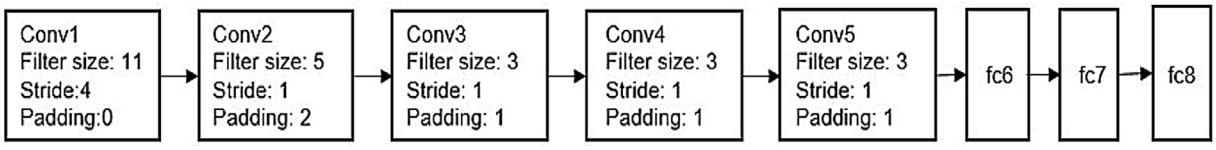

Fig. 6 – Architecture of AlexNet.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

Table 3 – Performance of NN classifier.

Performance measure (%)	Training	Testing
Accuracy	99.6	98.4
Sensitivity	99.3	97.2
Specificity	99.5	99.3

3.1. *Performance of NN*

Neural network performance was evaluated for the various features extracted from nuclei, cytoplasm and WBCs. The average values of accuracy, sensitivity and specificity for the classifier are listed in Table 3. The classifier was evaluated by considering 80% of the data for training and the remaining 20% for testing. It is evident from the table that, the testing accuracy is close to that of training accuracy.

The bar graph of overall performance of the classifier for the types of WBCs are shown in Fig. 7. The accuracy of lymphocyte, basophil and abnormal WBC detection is 100%. The performance of the classifier is less for detection of neutrophils and eosinophils compared to the other types of WBCs.

Performance of SVM and ensemble classifiers were also evaluated using the extracted features for the purpose of comparison. Average accuracies of 92.5% and 97.2% were obtained using multi-class SVM with quadratic kernel and the ensemble classifier respectively. However, NN was found to be better with an average accuracy of 99.8%.

Performance of the NN classifier was also evaluated using 5-fold cross validation. An overall accuracy of 99.6% was obtained with overall sensitivity of 98.9%.

3.2. *Performance of autoencoder*

The extracted features were used to train the autoencoder with 80% of the data for training and the remaining 20% for testing. We obtained an average accuracy of 96.72%. The average accuracy was computed for 100 iterations.

We used WBC images as input to the autoencoder network to evaluate the performance. The confusion matrix for WBC classification is shown in Fig. 8. The average accuracy of 72.4% was obtained as shown in the figure which is less compared to the accuracy obtained using the extracted features.

3.3. *Performance of CNN*

We considered designing CNN from scratch and also transfer learning approach was used for classification of WBCs. The 'training-set' consisting of 2697 cropped images were used to design and train the CNN from scratch and the 'test-set' consisting of 816 cropped images were used to test the designed CNN.

The details of the full training approach are as listed in Table 2. The layer wise accuracy of full training CNN is shown in Fig. 9. It can be observed from the figure that, an overall accuracy around 99% is obtained for the four convolutional layers (layer 1 to layer 4) and two fully connected layers (layer 7 and layer 8).

Training and testing performances in terms of accuracy, sensitivity and specificity are given in Table 4. It can be observed from the table that, performance values of the training and testing are almost similar.

AlexNet was used in transfer learning approach for the classification. The 'test-set' consisting of 816 cropped images were used in this approach. An average accuracy of 95% was obtained for the classification of WBCs which is less compared to that of the designed CNN from scratch (Table 5).

4. Discussion

In this paper, classification WBCs using traditional image processing approach and deep learning approach were considered. The motivation for the present study is, deep learning methods usually operate on images with small size

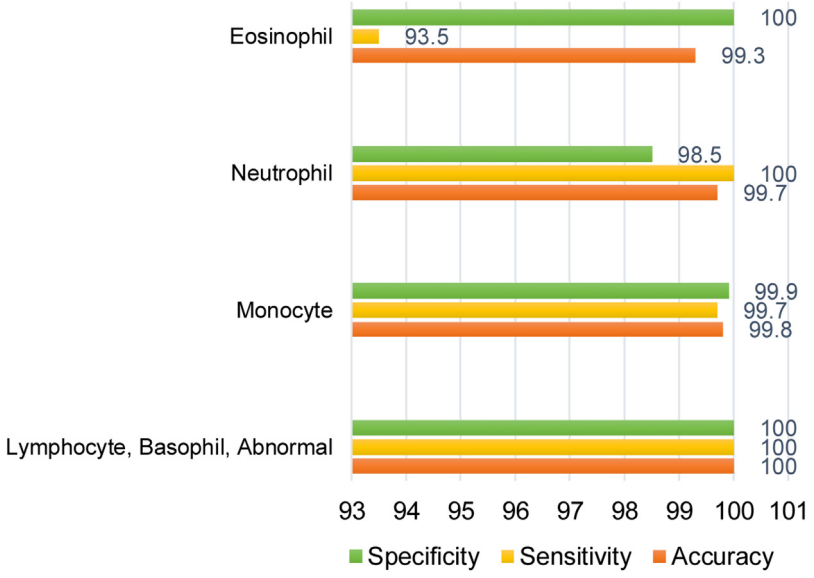

Fig. 7 – Plot for performance of NN classifier.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

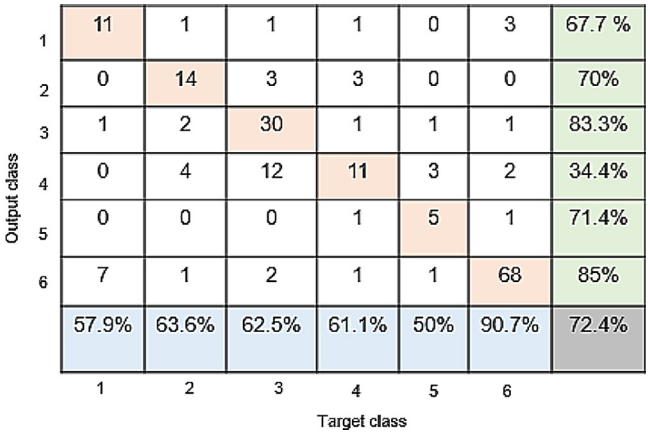

Fig. 8 – Confusion matrix for WBC classification using autoencoder.

which is suitable for cell images due to smaller image size. Also deep learning techniques require large number of training samples which is not difficult to obtain in case of peripheral blood smear images. The main attraction of deep learning method is that, it operates directly on raw pixels and processes in hierarchical manner. Feature extraction and classification steps are taken care by the network.

We demonstrated the results for classification of WBCs using NN, autoencoders and CNN. The results obtained for NN using traditional image processing approach is around 99%. We considered cropped WBCs in the present study. We focused on feature extraction and classification. The obtained accuracy is around 100% due to use of LBP representation to extract texture features along with the other features. LBP provides the local texture patterns, which is suitable to capture the texture patterns of types of WBCs. Hence these features can be considered for classification of WBCs. The result of segmentation and features used greatly affect the classification results in traditional image processing approach. Also, such methods require images acquired under controlled environment such as uniform lighting conditions, consistent staining method, etc. Our dataset contains images with brightness and color variations. Hence, we cropped the images to the size of WBCs and used different threshold values for segmentation of WBCs based on its brightness level. The drawback of our approach using traditional method is, we did not focus on segmentation of WBCs from the peripheral blood smear images. Classification of WBCs depends mainly on segmentation of WBCs from peripheral blood smear images and extraction of features. Hence, accurate detection of WBCs plays an important role in automated detection and classification of WBCs using traditional approach.

We demonstrated the classification of WBCs using deep networks namely autoencoders and CNN. We obtained poor accuracy using stacked autoencoders compared to NN and CNN. This could be due to the size of the WBC images considered in the study. The size $30 \times 30 \times 3$ may lead to loss of information. Increasing the size of the images may improve the result. We could not experiment on this by increasing the image size due to on board memory limitations which is only 2 GB. Also the number of images used might be insufficient to train the network. Hence by increasing the number of images and size of the images the performance of stacked autoencoder could be improved.

We also demonstrated the classification using full training and transfer learning approach of CNN. We obtained accuracy around 99% for full training CNN. We obtained the highest accuracy for the network with four convolutional layers and two fully connected layers (6 layered network). The classification accuracy for other layers as shown in Fig. 7 is not as good as 6 layered network. Changing filter size, number of filters, training samples, batch size, etc., may affect the performance. Hence deciding optimal network in full-training is purely based on experimentation which is time consuming. An accuracy of 95% was obtained using transfer learning approach. The obtained result using AlexNet is for changing only the fc8 layer from 1000 to 6 in the present study. It can also be considered with fine-tuning each layer for even better

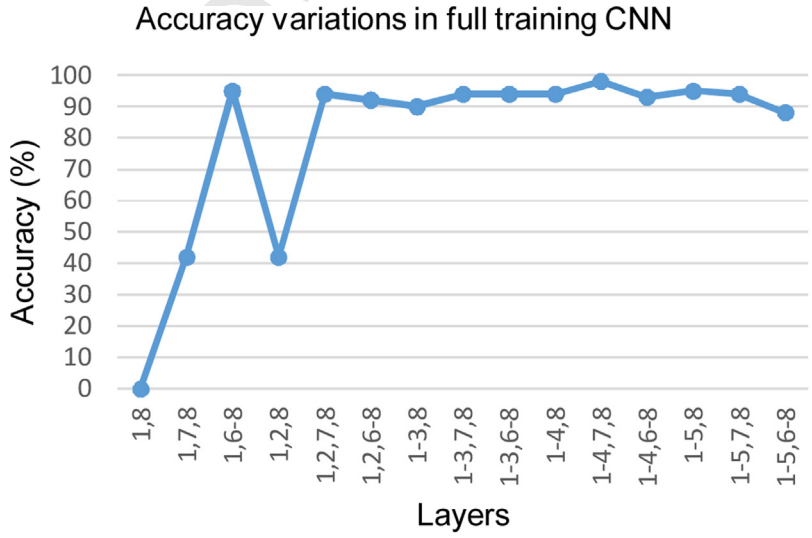

Fig. 9 – Accuracy for full training CNN for different layers.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

Table 4 – Performance of CNN from scratch.

Performance measure	Accuracy (%)	Sensitivity (%)	Specificity (%)
Training	99.5	100	99.8
Testing	98.7	99	99.4

Table 5 – Performance of AlexNet.

Performance (%)	AlexNet
Accuracy	95
Sensitivity	90
Specificity	98.6

performance. Performance of the designed CNN from scratch is found to be better compared to that of AlexNet. This is due to the fact that, AlexNet is trained with non-medical images whereas the designed CNN is trained with leukocyte images. Hence deep learning approach can be used for classification of WBCs but the resource requirement is high compared to NN. Another issue in deep learning is, requirement of large number of images for training the network. We used data augmentation in the present application to increase the size of dataset. We obtained comparatively good result for the present dataset. In order to design a deep network which could be used for classification of WBCs, the present application needs to be experimented by collecting more peripheral blood smear images. Since CNN operates directly on images, it is suitable for smaller size images rather than resizing larger images to smaller size. There is a possibility of loss of information by reducing the size of images. This implies that CNN approach is suitable for classification of WBC images. Hence, we cropped peripheral blood smear images to obtain images of smaller size.

In this study, we compared the increasing complexity of the networks used and demonstrated the performance of these networks. Computation resource increases from traditional approach to deep learning approach. Also, the number of images to be used for training increases. However, classification of WBCs can be considered using both the approaches depending on the availability of the resources. Designing CNN from scratch can be considered due to availability of data. Also, due to small image size.

5. Conclusion

In the present study, we demonstrated the classification of WBCs using traditional image processing approach and deep learning approach. Both the methods performed equally well with overall accuracy and sensitivity of 99%. In traditional image processing approach, accuracy of classification depends on the accuracy of segmentation and feature extraction. Use of deep learning technique eliminates this. It learns the feature by itself irrespective of the image variations, but it requires large number of labeled data and good infrastructure facilities. CNN can be used for classification of WBCs due to availability of data and smaller size of cell images as per the network requirement. Future scope of the present study would be working on robust segmentation method to handle the variations of peripheral blood smear images. We have obtained equally good result using traditional image processing approach which is simpler and involves less computational resources.

Funding

No funding was used for the proposed work.

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions

Roopa B Hegde: Designed and performed experiments, analyzed data and co-wrote the paper.

Keerthana Prasad: Planned and performed experiments, and co-wrote the paper.

Harishchandra Hebbar: Reviewed the paper and helped in drafting it.

Brij Mohan Kumar Singh: Provided the required data and also provided the ground truth.

REFERENCES

[1] Jianwei Z, Minshu Z, Zhenghua Z, Jianjun C, Feilong C. Automatic detection and classification of leukocytes using convolutional neural network. Med Biol Eng Comput 2017;55(8):1287.
[2] Fuyong X, Lin Y. Robust nucleus/cell detection and segmentation in digital pathology and microscoy images: a comprehensive review. IEEE Rev Biomed Eng 2016;9:234.
[3] Moshavash Z, Danyali H, Helfroush MS. An automatic and robust decision support system for accurate acute leukemia diagnosis from blood microscopic images. J Digital Imag 2018;1–16.
[4] Humayun I, Antoine V, Ludovic R, Daniel R. Methods for nuclei detection, segmentation and classification in digital histopathology: a review current status and future potential. IEEE Rev Biomed Eng 2014;7:97.
[5] Cao F, Cai M, Chu J, Zhao J, Zhou Z. A novel segmentation algorithm for nucleus in white blood cells based on low-rank representation. Neural Comput Appl 2017;28(1):503.
[6] Pan C, Park DS, Yang Y, Yoo HM. Leukocyte image segmentation by visual attention and extreme learning machine. Neural Comput Appl 2012;21(6):1217.
[7] Muhammad R, Saeeda Imran N. Microscopic blood smear segmentation and classification using deep contour aware CNN and extreme machine learning. Proc IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017. pp. 49–55. 10.1109/CVPRW.2017.111.
[8] Andrew J, Anant M. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 2016;7(29):18.
[9] Shin HC, Orton MR, Collins DJ, Oran SJ, Leach MO. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. IEEE Trans Pattern Anal Mach Intell 2013;35(8):1930–43.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005

[10] Andrew J, Ajay B, Anant M. Stain normalization using sparse autoencoders (StaNoSA): application to digital pathology. Comput Med Imaging Graph 2017;57:50.
[11] Dolz J, Betrouni N, Quidet M, Kharroubi D, Leroy HA, Reyns N. Stacking de-noising autoencoders in deep network to segment the brainstem on MRI brain cancer patients: a clinical study. Comput Med Imaging Graph 2016;52:8.
[12] Bejoy A, Madhu SN. Computer-aided diagnosis of clinically significant prostate cancer from MRI images using sparse autoencoder and random forest classifier. Biocybern Biomed Eng 2018;38(3):733.
[13] Jaroonrut P, Charnchai P. Segmentation of white blood cells and comparison of cell morphology by linear and nave bayes classifiers. Biomed Eng Online 2015;14(63). http://dx.doi.org/10.1186/s12938-015-0037-1
[14] Mathur A, Tripathi S, Kuse M. Scalable system for classification of white blood cells from Leishman stained blood stain images. J Pathol Inform 2013;4(15):6.
[15] Sedat N, Deniz K, Tuncay E, Murat S, Husnu, Osman K, et al. Automatic segmentation, counting, size determination and classification of white blood cells. J Measure 2014;55:58.
[16] Hamid Seyed R, Soltanian Hamid Z. Automatic recognition of five types of white blood cells in peripheral blood. Comp Med Imaging Graph 2011;35:333.
[17] Neelam S, Ramakrishnan GA. Automation of differential blood count. Proc IEEE Conference on Convergent Technologies for the Asia-Pacific Region, vol. 2. 2003. p. 547.
[18] Qingli L, yiting W, Hongying L, Xinofu H, Dongrong X, Jianbiao W, et al. Leukocytes cells identification and quantitative morphometry based on molecular hyperspectral imaging technology. Comp Med Imaging Graph 2014;38:171.
[19] Omid S, Hossein R, Ardeshir T, Hossein B, Yousefi. Selection of the best features for leukocytes classification in blood smear microscopic images. Proc. SPIE – Progress in Biomedical Optics and Imaging, vol. 9041. 2014. p. 8.
[20] Siroic´ R, Osowski S, Markiewicz T, Siwek K. Application of support vector machine and genetic algorithm for improved blood cell recognition. IEEE Trans Instr Measure 2009;58(7):2159.
[21] Morteza A, Moradi, Saeed K, Ardeshir T, Mostafa O, Ghelich. Recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and support vector machine classifier. J Med Signals Sensors 2015;5(1):49.
[22] Jyoti R, Annapurna S, Bhadauria H, Jitendra V, Singh JD. Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia. Biocybern Biomed Eng 2017;37(4):637.
[23] Mohapatra S, Patra D, Satpathy S. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. Neural Comput Appl 2014;24 (7):1887.
[24] Madhloom T, Kareem HA, Ariffin SH, Zaidan A, Alanazi AO, Zaidan HB. An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold. J Appl Sci 2010;10(11):959.
[25] Zhang J, Zhong YH, Wang X, Ni G, Du X, Liu J, et al. Computerized detection of leukocytes in microscopic leukorrhea images. Med Phys 2017;44(9):4620–9.
[26] Zhao J, Zhang M, Zhou Z, Chu J, Cao F. Automatic detection and classification of leukocytes using convolutional neural networks. Med Biol Eng Comput 2017;55(8):1287–301.
[27] Shahin AI, Yanhui Guo, Aminc KM, Amr Sharawi A. White blood cells identification system based on convolutional deep neural learning networks. Comp Methods Programs Biomed 2017. http://dx.doi.org/10.1016/j.cmpb.2017.11.015
[28] Vogado LHS, Veras RMS, Araujo FHD, Silva RRV, Aires KRT. Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. Eng Appl Artif Intell 2018;72:415–22.
[29] Qin F, Gao N, Peng Y, Wu Z, Shen S, Grudtsin A. Fine-grained leukocyte classification with deep residual learning for microscopic images. Comp Methods Prog Biomed 2018;162:243–52.
[30] Choi JW, Ku Y, Yoo BW, Kim Jung-A, Lee DS, Chai YJ, et al. White blood cell differential count of maturation stages in bone marrow smear using dual-stage convolutional neural networks. PLoS One 2017;12:e0189259.
[31] Tiwari P, Qian J, Li Q, Wang B, Gupta D, Khanna A, et al. Detection of subtype blood cells using deep learning. Cogn Syst Res 2018;52:1036–44.
[32] Rehman A, Abbas N, Saba T, Ijaz ur Rahman S, Mehmood Z, Kolivand H. Classification of acute lymphoblastic leukemia using deep learning. Microsc Res Techn 2018. http://dx.doi.org/10.1002/jemt.23139
[33] Sarmad S, Samabia S. Acute lymphoblastic leukemia detection and classification of its subtypes using pre-trained deep convolutional neural networks. Technol Cancer Res Treat 2018;17:1–7.
[34] Zhimin G, Lei W, Luping Z, Jianjia Z. HEp-2 cell image classification with deep convolutional neural networks. IEEE J Biomed Health Inform 2017;21(2):416.
[35] Zhaohui L, Andrew P, Ilker E, Mahdieh P, Kamolrat S, Kannappan P, et al. CNN based image analysis for malaria diagnosis. Proc IEEE International Conference on Bioinformatics and Biomedicine. 2016. pp. 493–6.
[36] Yuhang D, Zhuocheng J, Hongda S, David W, Pan A, Lance aVRVB, et al. Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. Proc IEEE International Conference on Biomedical and Health Informatics. 2017. pp. 101–4.
[37] Rani P, Oomman, Kaushik SK, Jeny R, Sabu M. Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybern Biomed Eng 2018;38(3):691.
[38] Charan J, Biswas T. How to calculate sample size for different study designs in medical research? Indian J Psychol Med 2013;35(2):121–6.
[39] Roopa BH, Keerthana P, Harishchandra H, Mohan BKS. Development of a robust algorithm for detection of nuclei and classification of white blood cells in peripheral blood smear images. J Med Systems 2018;42:1.
[40] Di H, Caifeng S, Mohsen A, Yunhong W, Liming C. Local binary patterns and its application to facial image analysis: a survey. IEEE Trans Syst Man Cyber C (Appl Rev) 2011;41 (6):765781.
[41] Kennith I. Laws, Rapid texture identification. Proc SPIE Image Processing for Missile Guidance, vol. 238. 1980. p. 376.
[42] Alex K, Ilya S, Geoffrey E. Hinton, ImageNet classification with deep convolutional neural networks. Proc 25th International Conference on Neural Information Processing Systems, vol. 1. 2012. p. 1097.

Please cite this article in press as: Hegde RB, et al. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.01.005