Research Article

Classification and Segregation of Abnormal Lymphocytes through Image Mining for Diagnosing Rheumatoid Arthritis Using Min-max Algorithm

[1]S.P. Chokkalingam and [2]K. Komathy
[1]Department of Information Technology, Saveetha School of Engineering, Saveetha University,
[2]Department of Computer Science and Engineering, Easwari Engineering College,
Anna University, Chennai, India

Abstract: Advances in the acquisition of complex medical images and storing it for further analysis through image mining have significantly helped to identify the root causes for various diseases. Mining of medical image data set such as scanned images or blood cell images require extraction of implicit knowledge from the data set through hierarchical image processing techniques and identifying the relationships and patterns that are not explicitly stored in a single image. Rheumatoid Arthritis (RA) is an autoimmune disease and it cause chronic inflammation of the joints. Causes of the RA is unknown due to that need to find out in the early stage is required. Diagnosis of RA based on blood cell types and shapes requires computational analysis. An assistive technology for the doctor to detect and investigate rheumatoid arthritis is therefore required. The objective of the proposed work is to analyze the shapes of lymphocytes, a key component of blood cells that causes RA complications, to automate the process of identifying abnormal lymphocytes by estimating the centroids of lymphocytes using AIT centroid technique and thereby finding a differential count. The process involves cropping nucleus from the blood cell image, segmenting it and to investigate further whether the shapes of the lymphocytes are irregular and dissimilar. Features are extracted from each cell components for comparison and the abnormal lymphocytes are segregated from the normal. To enhance the segregation process, neural network based perceptron classifier tool is used.

Keywords: Centroid analysis, image mining, lymphocytes, perceptron classifier, Rheumatoid Arthritis (RA), segmentation, white blood cells

INTRODUCTION

Image mining is an incredible technology to produce all the patterns without specifying any information of the image content. Establishment of an image mining system has been frequently tortuous processed because it implies joining diverse techniques ranging from image retrieval and indexing schemes up to data mining and pattern recognition. The fundamental challenging in image mining is to determine how low-level pixel representation contained in a raw image or image sequence can be efficiently and effectively processed to identify the relationships.

Rheumatoid Arthritis is a chronic disease, which attacks the synovial tissue that lubricates the joints of the human skeleton. It is the disease that affects the musculoskeletal system including bones, muscles, joints and tendons that Contribute to loss of function and range of movement and difficulties in performing activities of daily living. Rheumatoid arthritis typically occurs in joints on both sides of the body (such as hands, knees, or wrists).

The process of life is maintained by blood, which is a specialized body fluid consisting of plasma and blood cells. To perform a proper diagnosis of the disease, identification of the blood cells and their relative quantity in the blood samples must be known. A blood film or peripheral blood smear is a thin layer of blood smeared on a microscope slide and then stained in such a way to allow the various blood cells to be examined microscopically. Due to the development in technology, this traditional blood examination is digitized. By connecting a high resolution digital camera to the microscope, the blood cell images are captured by adjusting the microscope magnification to obtain a fairly good resolution image. For identifying different types of blood cells and for counting their quantity in blood smear, image processing is applied. This study attempts to analyze the shapes of lymphocytes in the blood cell images resulting in a better investigation of Rheumatoid Arthritis (RA) through an automated process. Differential count and shapes of lymphocytes in white blood cells provide valuable information that aids in diagnosis of RA.

Corresponding Author: S.P. Chokkalingam, Department of Information Technology, Saveetha School of Engineering, Saveetha University, Chennai, India

MATERIALS AND METHODS

Analysis of White Blood Cells (WBCs) requires a segmentation algorithm, which separates the Region of Interest (ROI) from the blood cell image. The accuracy of the segmentation algorithm has a great impact on the final results of the analysis. Various approaches are available for segmenting ROI from the image. Fast segmentation approach for automatic differential counting of White Blood cells (Anoraganinrum, 1999). This involves simple localization of WBCs using some prior information about the blood smear images and with recursively applying thresholding; the various components of WBCs are separated. Then morphological operations such as erosion and dilation are applied to smoothen the segmented image. In order to classify the WBCs, only the information about the nucleus is adequate (Sharma and Sahni, 2011). By applying the localization of white blood cell method (Mamatha *et al*., 2012), the problem of cells that touch each other can be eliminated. The segmentation of nucleus alone from the smear image is easier than segmenting the entire cell (Arifin and Asano, 2006). This can be achieved by using mathematical morphology of WBCs. Differential counting can be automated using an automated approach, which extract features from the segmented image with the help of 'Eigen faces' a widely used method in face recognition (Yampri *et al*., 2006) A combination of Principal Component Analysis (PCA) and parametric feature detection can be used for this purpose (Brunelli and Mich 2000). The above system uses a set of known features that are projected into a feature space that holds significant variations known as Eigen cells.

A segmentation method, which uses the benefit of active contour, can be utilized for segmenting WBCs (Yampri *et al*., 2006). The active contour method involves converting a color image into a binary image using a threshold and placing the initial circular shape (snake) inside the roughly identified position of WBCs. By using gradient flow vector force the initial circular shape is allowed to grow until it fits the exact shape of the nucleus. Then the individual WBC is separated from the smear image using the extracted contour Subrajeet *et al*. (2010).

Poisson equation based approach can be used to extract the number of segments of the nucleus. The inner distances can be used to represent the shape features of the nucleus segments (Theerapattanakul *et al*., 2004). In the above technique, two different shape features are used for training the neural network. The concept of the random walk is used as a solution to the Poisson equation and is combined with an inner distance to extract these features Singh and Singh (2012).

A two stage color segmentation strategy along with fuzzy clustering is used for separating WBCs from other blood components (Bharanidharan and Ghosh, 2012). After extracting the Leukocytes, Lymphocytes are identified from the segmented image as a subclass. Various features such as fractal features, shape features and other texture features are extracted from the sub class of Lymphocytes Sumathi and Ravindran (2011). In addition to all the above features, two new features Hausdr off Dimension and contour signature are employed for further classification Padmavathi *et al*. (2010).

To measure the degree of cell clumping in terms of area and the number of cells it contains, the approach proposed in Sharma and Sahni (2011) has been recalled. In order to classify a large set of blood smear images into good, sparse and clumped, an integrated approach using Shannon entropy is used Kavallieratou and Stamatatos (2006). To detect good working areas in an image both the feature of spatial distribution and cell clumping algorithms are employed Nosrati *et al*. (2012). Image classification is enhanced using optimal feature selection. Pre-processing, Segmentation, Feature extraction, Classification and Prediction are the steps concentrates for identifying the shapes of Lymphocytes Al-amri *et al*. (2010).

RESULTS AND DISCUSSION

De-noising of images: The initial phase of the image processing is the preprocessing of the image. It is done to enhance the quality of image obtained from various sources such that it satisfies the requirement of further processing. Image De-noising is one of the preprocessing mechanisms it is used to remove noise from acquired image. Noise is the common term to describe visual distortion. The noise will affect the quality of the image and, the main factor is that noise can coat and reduce the visibility of certain features within the image. Stray marks, marginal noise and salt-and-pepper noise are independent of size; location of the underlying content. Similarly the texture of the observed speckle pattern is independent of the underlying content (Agrawal and Doermann, 2009) Regular noise is always showing a consistent behavior in terms of these properties. On the other hand, noise such as salt-and-pepper noise that does not show a consistent behavior is classified under "irregular noise (Mamatha *et al*., 2012). In this study we have made an attempt to study the four common types of noises like Gaussian, salt and pepper, Poisson and speckle noise. Gaussian noise also called Random Variation Impulsive Noise (RVIN) or normal noise T is a type of statistical noise in which the amplitude of the noise follows that of a Gaussian distribution (Mamatha *et al*., 2012). Salt-and-pepper noise is also called as Fat-tail distributed or impulsive noise or spike noise. An image containing salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions. Statistical

Quantum Fluctuations induce a prominent noise type in the lighter parts of an image from an image sensor. This noise is called photon shot noise or Poisson noise. The noises at different pixels are independent of each other. Speckle noise is a granular noise that increases the mean grey level of a local area in an image. This type of noise makes it difficult for image recognition and interpretation.

In image processing, filters are mainly used to suppress either the high frequencies in the image, i.e., smoothing the image, or the low frequencies, i.e., enhancing or detecting edges in the image. Noise removal is easier in the spatial domain as compared to the frequency domain as the spatial domain noise removal requires very less processing time (Nichol and Vohra, 2004). Average filtering replaces each pixel value in an image with the mean value of its neighbors, including itself. The simplest procedure would be to calculate the mask for all the pixels in the image. For all the pixels in the image which fall under this mask, it will be considered as the new pixel (Patidar *et al.*, 2010). This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Average filter is also considered to be a convolution filter or a mean filter.

The median filter is an effective method that can suppress isolated noise without blurring sharp edges. In Median Filtering, all the pixel values are first sorted into numerical order and then replaced with the middle pixel value (Murali *et al.*, 2012) The Gaussian filter is a non-uniform low pass filter. The kernel coefficients diminish with increasing distance from the kernel's centre. Central pixels have a higher weighting than those on the periphery.

The inverse filtering is a restoration technique for de-convolution, i.e., when the image is blurred by a known low pass filter, it is possible to recover the image by inverse filtering or generalized inverse filtering. However, inverse filtering is very sensitive to additive noise.

Initial phase of the image processing is the preprocessing of the image. It is done to enhance the quality of image obtained from various sources such that it satisfies the requirement of further processing. Image De-noising is one of the preprocessing mechanisms it is used to remove noise from acquired image. Noise is the common term to describe visual distortion.

In image processing, filters are mainly used to suppress either the high frequencies in the image, i.e., smoothing the image, or the low frequencies, i.e., enhancing or detecting edges in the image. Noise removal is easier in the spatial domain as compared to the frequency domain as the spatial domain noise removal requires very less processing time (Nichol and Vohra, 2004) Average filtering replaces each pixel value in an image with the mean value of its neighbors,

including itself. The simplest procedure would be to calculate the mask for all the pixels in the image. For all the pixels in the image which fall under this mask, it will be considered as the new pixel (Patidar *et al.*, 2010). This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Average filter is also considered to be a convolution filter or a mean filter.

The median filter is an effective method that can suppress isolated noise without blurring sharp edges. In Median Filtering, all the pixel values are first sorted into numerical order and then replaced with the middle pixel value (Murali *et al.*, 2012). The Gaussian filter is a non-uniform low pass filter. The kernel coefficients diminish with increasing distance from the kernel's centre. Central pixels have a higher weighting than those on the periphery.

The inverse filtering is a restoration technique for de-convolution, i.e., when the image is blurred by a known low pass filter, it is possible to recover the image by inverse filtering or generalized inverse filtering. However, inverse filtering is very sensitive to additive noise.

Preprocessing image setup and results: A set of 108 images had been taken for analysis. Each image is subjected to different types of noise mentioned above. Each image with added noise is subjected to different types of filters. The filtered image is compared against the original image using the following image quality measures.

Peak Signal-to-Noise Ratio (PSNR): The Peak Signal to Noise Ratio is calculated by:

$$PSNR = 20log_{10}\left(\frac{N}{MSE}\right) dB \tag{1}$$

For the image quality measures, if the value of the PSNR is very high for an image with a particular noise type then it is the best quality image. Table 1 shows the aggregated PSNR value of each image subjected to different type of filters. According to PSNR, it is clear that the median filter gives best result over salt and pepper noise and Wiener is more suitable for Gaussian noise, Poisson noise and Speckle noise.

Mean Square Error (MSE): Mean square error is given by:

Table 1: PSNR value for images subjected to various filters over different types of noise

Filter/noise	Salt and pepper noise	Gaussian noise	Poisson noise	Speckle noise
Mean filter	26.0075	26.2839	28.1714	26.0728
Median filter	32.6813	26.0729	29.5074	24.4857
Weiner filter	21.2104	26.6245	30.4699	26.5855
Gaussian filter	23.4399	24.1645	29.0728	23.8969

Table 2: MSE value for images subjected to various filters over different types of noise

Filter/noise	Salt and pepper noise	Gaussian noise	Poisson noise	Speckle noise
Mean filter	163.0537	152.9985	99.0697	160.6209
Median filter	35.0713	160.6164	72.8347	231.4779
Weiner filter	492.0871	141.4579	58.3566	142.7363
Gaussian filter	294.5029	249.2496	80.5016	265.0888

Table 3: NK value for images subjected to various filters over different types of noise

Filter/noise	Salt and pepper noise	Gaussian noise	Poisson noise	Speckle noise
Mean filter	0.9849	0.9893	0.9937	0.9839
Median filter	0.9976	0.9966	0.9963	0.9918
Weiner filter	0.9865	0.9911	0.9962	0.9861
Gaussian filter	0.9878	0.9922	0.9967	0.9868

Table 4: NAE value for images subjected to various filters over different types of noise

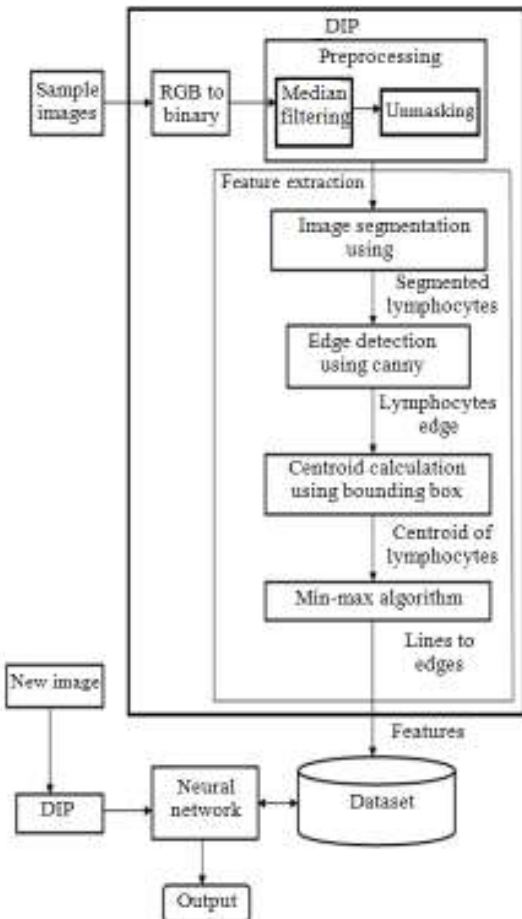| Filter/noise	Salt and pepper noise	Gaussian noise	Poisson noise	Speckle noise
Mean filter	0.0439	0.0515	0.0374	0.0519
Median filter	0.0170	0.0548	0.0351	0.0661
Weiner filter	0.0406	0.0508	0.0316	0.0508
Gaussian filter	0.0291	0.0701	0.0389	0.0734

Fig. 1: Block diagram of feature extraction process

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[g(i,j) - f(i,j)]^2 \qquad (2)$$

where, M and N are the total number of pixels in the horizontal and the vertical dimensions of image, g denotes The Noise image and *f* denotes the filtered image. The lowest mean square error represents the best quality image.

Table 2 shows that the MSE value of an image subjected to median filter has the lowest value of salt and pepper noise it shows that median filter is the best choice for Salt and pepper noise. Image subjected to Weiner filter gives the lowest MSE value and Wiener is more suitable for Gaussian noise, Poisson noise and Speckle noise.

Normalized correlation (NK): The closeness between two digital images can also be quantified in terms of the correlation function. These measures measure the similarity between two images, hence in this sense they are complementary to the difference based measures. All the correlation based measures tend to 1, as the difference between two images tend to zero. It is calculated using the formula:

$$NK = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j).f'(i,j)]}{\sum_{i=1}^{M}\sum_{j=1}^{N}(f(i,j))^2} \qquad (3)$$

For image-processing applications in which the brightness of the image and the template can vary due to lighting and exposure conditions, the images can be first normalized.

This is typically done at every step by subtracting the mean and dividing by the standard deviation. If the normalized cross correlation tends to 1, then the image quality is deemed to be better.

Table 3 shows that the NK value of an image subjected to Median filter over all types on noise is near to 1, which is followed by Gaussian filter and Weiner filter and finally the Mean filter shows the poor correlation value over all other types of filters.

Normalized Absolute Error (NAE): Normalized Absolute Error should be the minimum in order to minimize the difference between original and obtained image. It is calculated using the formula:

$$NAE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|[f(i,j).f'(i,j)]|}{\sum_{i=1}^{M}\sum_{j=1}^{N}|f(i,j)|} \qquad (4)$$

The normalized absolute error indicates how different both the de-noised image and the original image are with the value of zero being the perfect fit. A large value of NAE represents the poor quality of the image.

Table 4 shows that NAE value of an image subjected to Median filter over salt and pepper noise have the lowest value; for all other types of noise. Image subjected to Weiner filter gives the lowest NAE value. The mean filter shows an average performance over all type noise then followed by Gaussian filter.

Compares the performance of four spatial domain filters mean filter, Median Filter, Gaussian filter and Weiner Filter to de-noise the images subjected to four different types of noise Salt-and-Pepper, Gaussian, Poisson and Speckle noise which can be accumulated during image acquisition phase of Microscopic image processing. From the results shown above, it is clear that the median filter shows its best performance over salt-and-pepper noise and Weiner filter shows good performance over Gaussian, Poisson and Speckle noise. It also shows an optimum performance over salt-and-pepper noise and hence it is concluded that Weiner filter is an optimum filter that can be applied to microscopic images (Fig. 1).

Image segmentation: Image segmentation is used to identify the regions of interest in an image or to annotate the data. White blood cell image contains neutrophil, basophil, eosinophil, lymphocyte, monocyte and platelets. Figure 2 contains lymphocytes with other types. Using threshold mechanism RPG of white blood cell image is converted into a binary image. If the image is clumsy then the sub imaging is performed using adaptive contour method (Zang and Chen, 2001). After microscopic images are obtained, it needs to be preprocessed for more accuracy.

ROI retrieval: As Rheumatoid Arthritis (RA) needs to consider only the lymphocytes as the Regions of Interest (ROI) and the essential features are extracted from this ROI. Segmentation isolates lymphocytes from white blood cells shown in Fig. 2.

Preparation of image set: Samples of stained blood slides are collected from various laboratories. The images are captured with a digital microscope under 100x oil immersed setting and with an effective magnification of 1000. Due to excessive staining, noise may accumulate in images. Weiner filter is an optimum filter shows better performance for microscopic images so Weiner filter is used in preprocessing module and adaptive threshold into noise detection process led more reliable and more efficient detection of noise. Figure 3 shows the output image obtained from preprocessing.

Image segmentation using threshold: In our proposed thresholding method values are assigned to the slider controls and images are segmented based on the slider values (Fig. 4).

The RGB images are converted into gray scale image. Sizes of images are in the matrix forms and it can be assigned into row and column. It is required to process up to the *n* number of rows and columns and each time it will check whether row and column values are greater than slider values and if yes, it can be converted into a white pixel otherwise it can be taken as black.

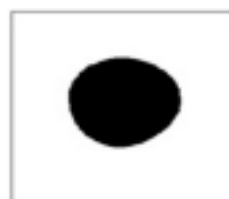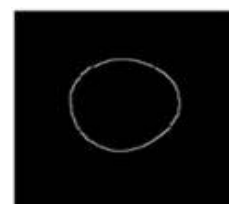Fig. 2: Lymphocytes with platelets

Fig. 3: Segmented lymphocytes

Fig. 4: Edge detection for the segmented lymphocytes using canny edge method

Algorithm 1: To separate lymphocytes from WBC
Input: White Blood Cell Image contains neutrophil, basophil, eosinophil, lymphocyte, monocyte and platelets
Output: Image containing Lymphocyte
Procedure Threshold (Image, Slider_value)
Image is the input and segmented from the whole image.
Slider values are taken as a threshold value and applied to the Image for Segmentation.
begin
      if Image is in RGB then
    image = RGB-to-Grey (Image)
  end
   if slider_value gets changed then
    size (row, col) = size (Image)
    for each row of Image do
     for each col of Image do
     if pixel_val (row, col) >slider_value then
      pixel_val (row, col) = 255
     else
       pixel_val (row, col) = 0
    end
    end
   end
  end
  return Image
end

Centroid estimation: After detecting the edges using canny edge method, bounding box is drawn around the irregular lymphocyte to crop out the nucleus. Then the size of the cropped nucleus image is computed and the nucleus values are assigned to respective rows and a column. Mean-x and Mean-y can be calculated based on the area of the image (Fig. 5).
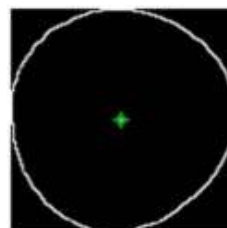
Area calculation: The entire image is scanned from left to right and total size of the image is stored and will take one row and check the entire column are read and parallel the pixel values are check with 0 and if yes, counting of area is increased (Fig. 6).
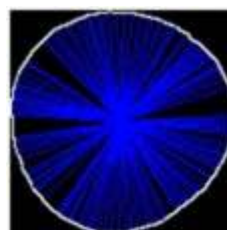
Algorithm 2: To estimate the centroid for the lymphocyte image
Input: ROI (Lymphocytes) Image
Output: Centroid of ROI (Lymphocytes) Image
Procedure Centroid for Irregular Shapes (LymphoImage)
   Image is the threshold image containing only lymphocytes
begin
   Image = canny_edge_detector (Image) /* Trace the boundaries of each closed regions*/
   Image = bwlabel (Image) /* Label each bounded region*/
   Image = Bounding Box (Image) /* Draw a rectangle around each labeled region*/
   Image = Crop (Image) /*crop the nucleus of lymphocyte)
   Size (rows, cols) = size (Image)
   x = ones (rows, 1)*[1: cols]
   y = [1: rows]'*ones (1, cols)
   area = sum (sum (Imagem))
   meanx = sum (sum (double (Image).*x)) /area
   meany = sum (sum (double (Image).*y)) /area
   return (meanx, meany)
end

Length of the line from the centroid to edges: In Image cropped Lymphocyte image centroid of X-Coordinates and Y- Coordinates are assigned to meanx and meany variables. Initially calculate the minimum and maximum distance from centroid to edges and divide it ten intervals. Images are scanned from left to right.

Algorithm 3: Area calculation
Input: Nucleus of Lymphocyte Image
Output: Area of the Image
Procedure area (Image)
   Type of the image is Segmented Lymphocytes
begin
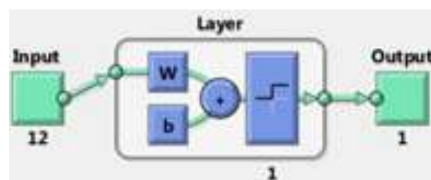   size (row, col) = size (Image)
   area = 0;

Fig. 5: Centroid estimation for lymphocytes

Fig. 6: Line drawing from centroid to edges

Fig. 7: Neuron process model

   for each row of Image do
     for each col of Image do
      if pixel_val (row, col) == 0 then
       area = area + 1
     end
   end
   end
   return area
end

Classification using neural network: The perceptron learning rule is a method for finding the weights in a network. Single Layer perceptron is used as classifier. This method is used to solve the problem of supervised learning for classification (Fig. 7):

- Initialize the weights (either to zero or to a small random value)
- Pick a learning rate μ (this is a number between 0 and 1)
- Until stopping condition is satisfied (e.g., weights don't change)

For each training pattern (x, t):

- Compute output activation
  $y = f (w x)$
- If y = t, don't change weights

Table 5: Features for normal lymphocyte image (not infected)

Image	Area	Perimeter	<33-35	35-37	37-39	39-41	41-43	43-45	45-47	47-49	49-51	51-and >53
			Regions									
Image 1	5141	305	0	21	66	96	69	51	2	0	0	0
Image 2	5681	324	0	17	49	49	71	48	47	28	15	0
Image 3	5665	333	0	3	41	41	57	167	24	0	0	0
Image 4	4882	315	0	19	71	207	18	0	0	0	0	0
Image 5	5467	329	7	46	34	30	50	87	75	0	0	0
Image 6	4801	298	0	41	84	144	29	0	0	0	0	0
Image 7	4948	310	5	37	81	93	51	43	0	0	0	0
Image 8	5531	331	0	0	2	83	164	80	2	0	0	0
Image 9	5021	315	6	76	25	63	83	56	6	0	0	0
Image 10	5181	308	0	41	59	60	78	45	21	4	0	0

Table 6: Features of abnormal lymphocyte image (infected)

Image	Area	Perimeter	<33-35	35-37	37-39	39-41	41-43	43-45	45-47	47-49	49-51	51-and >53
			Regions									
Image 1	8069	385	0	0	0	0	0	28	48	41	53	215
Image 2	6671	343	0	0	0	0	5	79	127	131	1	0
Image 3	6721	369	54	11	16	13	16	16	47	30	29	137
Image 4	5076	307	57	18	23	15	49	55	44	46	0	0
Image 5	8262	400	0	0	0	0	3	47	37	64	68	181
Image 6	11466	579	100	10	8	8	6	8	8	5	8	418
Image 7	9370	456	0	0	15	14	7	7	10	8	7	384
Image 8	5998	339	76	9	8	7	9	22	28	29	39	112
Image 9	7824	381	0	1	35	16	17	9	15	83	43	162
Image 10	5748	330	60	12	5	12	17	38	62	78	29	17
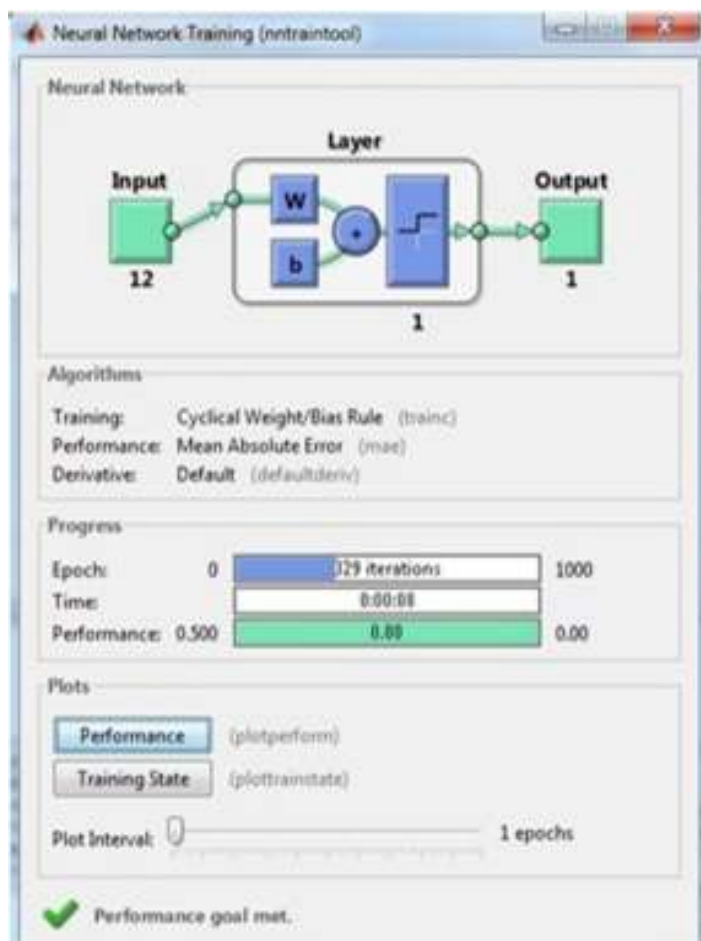
Fig. 8: Neural network training model

- If y! = t, update the weights: w (new) = w (old) + 2 μtx or
- w (new) = w (old) +μ (t - y) x, for all t

Algorithm 4: To form segmented regions from the nucleus of Lymphocyte image by MIN-MAX Algorithm.
Input: Nucleus of Lymphocyte Image
Output: Lymphocyte Image which Connected Line from Centroid to all the Edges
Procedure draw_line (Image, meanx, meany)
Cropped image having single nucleus
meanx is the x-coordinate of centroid
meany is the y-coordinate of centroid
begin
  Find min and max distance from centroid to edge pixel (Image)
  Divide the range between min and max into ten equal intervals
  size (rows, cols) = size (Image)
  for each row of Image do
    for each col of Image do
      if pixel_val (row, col) >0 then
        draw a line from current pixel to [meanx, meany]
        L = Length of the line drawn
        Increment the count value in the appropriate interval according to L
      end
    end
  end
end

The area and perimeter of normal lymphocytes of different images are given Table 5. The radial lines are drawn from the center of the Lymphocyte to its edge-contour to derive its perimeter. The area of the Lymphocyte has been divided into ten equal intervals. The radial distance from the center to its edge varies from one segment to the other segment ranging from 33 to 53. The pixel count in each region is counted and plotted into the respective column of the table shown. It is assumed to have all the pixels be located between 33 and 53 to conform normalcy else we can expect there is an abnormality in the Lymphocytes. The details in Table 6 illustrate the abnormality of Lymphocytes based on the prediction technique adopted in Table 5 and 6. Prediction technique is a method which is used to analyze the future things after presenting the training instances to the learner. Line which is started from the Centre into edges pixels counts are calculated and presented into the appropriate places. In every time a new image is processed and matches with our dataset and identified whether it is infected or not (Fig. 8).

SUMMARY AND CONCLUDING REMARKS

Medical image mining is an interdisciplinary domain, which focuses on similarity and retrieval of patters in domain specific applications to solve challenging problem in the medical field. An automatic identification process is proposed in this study that detects the shape of Lymphocytes through image segmentation and centroid analysis. The algorithm uses threshold and classification for approximation. Weiner filter shows good performance over Gaussian, Poisson and Speckle noise. It also shows an optimum performance over salt-and-pepper noise and hence it is concluded that Weiner filter is an optimum filter that can be applied to microscopic images. The dataset is trained through a perceptron neural network classifier for accurate analysis. Through these classification and segregation process it is found that it is possible to automate the process of diagnosing RA based on the blood cell images.

REFERENCES

Agrawal, M. and D. Doermann, 2009. Clutter noise removal in binary document images. Proceeding of 10th International Conference on Document Analysis and Recognition (ICDAR '09), pp: 556-560.

Al-amri, S.S., N.V. Kalyankar and S.D. Khamitkar, 2010. A comparative study of removal noise from remote sensing image. IJCSI Int. J. Comput. Sci., 7(1).

Anoraganinrum, D., 1999. Cell segmentation with median filter and mathematical morphology operation. Proceeding of the International Conference on Image Analysis and Processing, pp: 1043-1046.

Arifin, A.Z. and A. Asano, 2006. Image segmentation by histogram thresholding using hierarchical cluster analysis. Pattern Recogn. Lett., 27(2006): 1515-1521.

Bharanidharan, T. and D.K. Ghosh, 2012. A two dimensional image classification neural network for medical images. Eur. J. Sci. Res., 74(2): 286-291.

Brunelli, R. and O. Mich, 2000. Image retrieval by examples. IEEE T. Multimedia, 2: 164-71.

Kavallieratou, E. and E. Stamatatos, 2006. Improving the quality of degraded document images. Proceedings of the 2nd International Conference on Document Image Analysis for Libraries (DIAL'06), pp: 10-349.

Mamatha, H.R., S. Madireddi and K.S. Murthy, 2012. Performance analysis of various filters for De-noising of handwritten Kannada documents. Int. J. Comput. Appl., 48(12): 0975-888.

Murali, Y., M. Babu, M.V. Subramanyam and M.N. Giri Prasad, 2012. PCA based image denoising. SIPIJ, Vol. 2.

Nichol, J.E. and V. Vohra, 2004. Noise over water surfaces in Landsat TM images. Int. J. Remote Sens., 25(11): 2087-2093.

Nosrati, M., R. Karimi and M. Hariri, 2012. Detecting circular shapes from areal images using median filter and CHT. Global J. Comput. Sci. Technol., Vol. 12.

Padmavathi, G., P. Subashini, M. Muthu Kumar and S.K. Thakur, 2010. Comparison of filters used for underwater image-preprocessing. IJCSNS Int. J. Comput. Sci. Network Secur., 10(1).

Patidar, P., M. Gupta, S. Srivastava and A.K. Nagawat, 2010. Image de-noising by various filters for different noise. Int. J. comput. Appl., 9(4).

Sharma, A.K. and S. Sahni, 2011. A comparative study of classification algorithms for spam email data analysis. Int. J. Comput. Sci. Eng. (IJCSE), 3(5).

Singh, P. and H. Singh, 2012. A comparison of high pass spatial filters using measurements and automation. Int. J. Eng. Res. Technol. (IJERT), 1(3), ISSN: 2278-0181.

Subrajeet, M., D. Patra and S. Satpathi, 2010. Image analysis of blood microscopic images for acute leukemia detection. Proceeding of International Conference on Industrial Electronics Control and Robotics.

Sumathi, P. and G. Ravindran, 2011. The performance of fractal image compression on different imaging modalities using objective quality measures. Int. J. Eng. Sci. Technol. (IJEST), 3(1): 525-530.

Theerapattanakul, J., J. Plodpai, S. Mooyen and C. Pintavirooj, 2004. Classification of White Blood Cell using Adaptive Active Contour. ICCAS, Bangkok, Thailand, pp: 1892-1894.

Yampri, P., C. Pintavirooj, S. Daochai and S. Teartulakarn, 2006. White blood cell classification based on the combination of Eigen cell and parametric feature detection. Proceeding of 1st IEEE Conference on Industrial Electronics and Applications, pp: 1-4.

Zang, P. and L. Chen, 2001. Document filters using morphological and geometrical features of characters. Image Vision Comput., 19: 847-855.