

FAST AND RELIABLE DISEASE DIAGNOSIS THROUGH DIGITISED IMAGES OF BLOOD

M. Arunachalam

Department of Computer Science and Engineering, K.L.N College of Information Technology, India
E-mail: er.arun80@gmail.com

Abstract

Principal effectors in the blood i.e. Erythrocyte, Leukocyte and Platelets of the blood play a crucial role in supplying blood to the various parts of the body, resisting foreign particles and in clotting of blood respectively. The effectors are identified and skipped for our observation. In microscopic images of the diseased blood cells, the diagnosis is based on the evaluation of some general features of the blood cells such as color, shape, and border, and the presence and the aspect of characteristic structures. Perception of these structures depends both on magnification and image resolution. The introduction of vision system associated with image processing techniques enables the numerical and objective description of some pattern cells features. This paper is the implementation of a simple, fast and reliable method for automatically diagnosing diseases through digitized images of blood. A new algorithm is proposed which will efficiently reduce the overall running time of object recognition.

Keywords:

Medical Imaging, BPN, Components of Blood, Running Time

1. INTRODUCTION

Hematopathology reveals that there is an intrusion of the disease cell in the blood having identical characteristics for each disease. Principal effectors in the blood that are Erythrocyte, Leukocyte and Platelets of the blood play a crucial role in supplying blood to the various parts of the body, resisting foreign particles and in clotting of blood respectively. The effectors are identified and skipped for our observation. In microscopic images of the diseased blood cells, the diagnosis is based on the evaluation of some general features of the blood cells such as color, shape, and border, and the presence and the aspect of characteristic structures. Perception of these structures depends both on magnification and image resolution. The introduction of vision system associated with image processing techniques enables the numerical and objective description of some pattern cells features[1]-[4]. This paper implements a simple, fast and reliable method for automatically diagnosing diseases through digitized images of blood. A new algorithm is introduced which will efficiently reduce the overall running time of object recognition. Concept of Back Propagation network is used to set appropriate error values during object recognition. Experimental results using proposed technique gives a way for new research area in Bio mechanics.

2. EXPERIMENTAL SETUP AND IMPLEMENTATION

2.1 EXPERIMENTAL SETUP

The technology used in solid state imaging sensors is based principally on charged coupled devices (CCD), which can

convert the images from analog to digital format. Simple strategies can be adapted to grab the images from the Electron Microscope. The experiment apparatus consists of structured light source, Electron Microscope, CCD camera and vision system[5]-[6]. The relative cells of the blood are first measured. Before the actual diagnosis is made, the real coordinates and world coordinates must be determined in the camera coordinate system by a calibration process. The camera acquisition requires the knowledge of calibration analysis. The camera can be viewed simply as a device for mapping a 3D image into a 2D image. A detailed experimental study is carried out in diseased cell structure with the normal one. The real time experiment is to illustrate the actual implementation of the estimation method developed by spatial encoding technique. The innovative hardware and vision-processing algorithm with menu driven user interface was used to analyze the condition of the blood cells in the diseased cell structure. The experiment is planned taking into consideration of the above requirements and detection of diseases through microscopic images of the diseased cell structure. The software facilities capturing of high resolution images in color or in black and white, comparison of built up model (Template) with that of subjected model and the images can be sized and stored for future reference. The last stage of the experimental setup in which only one of the diseases would be activated from the output layer of the back propagation algorithm denoting the disease that is to be determined. The Setup is shown in Fig.1.

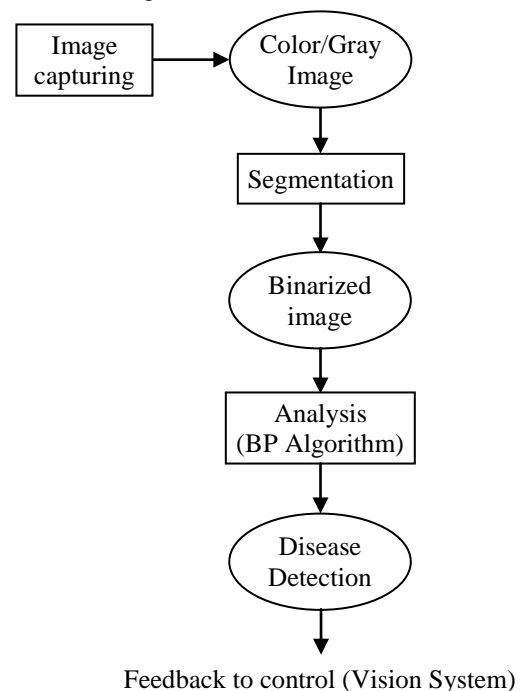


Fig.1. Experimental Setup

Stage 1: Image capturing

The blood strip taken from the patient's body is placed in the slide. This is viewed through the electron microscope. A frame is retrieved using the frame grabber. From the frame the region of interest is extracted according to the height and width in number of pixels as required. Thus the image is made to fix into the standard dimensions as shown below.

For example:

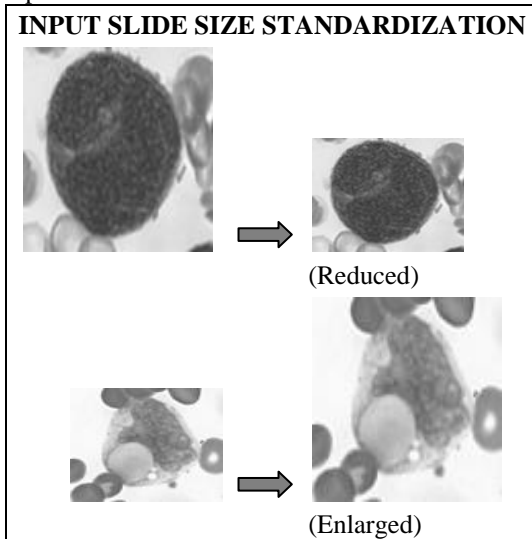


Fig.2. Input slide size standardization

Stage 2: Color/Gray image

The image is then passed to the above component from which the color image is converted to gray scale using the necessary formula. Weights of the Red, Blue and Green are mixed in proper proportions and then we calculate a single value as required in gray scale.

Stage 3: Segmentation

The image is then segmented according to the required values of height and width. When there is some exceeding part of the block they are filled with appropriate values of padding.

Stage 4: Binarized Image

The Image is then binarized into 0 and 1 using necessary approximations. For example terms with values 0.65 are approximated to 1 and terms with 0.35 are considered as 0.

Stage 5: Back propagation algorithm

Back propagation (explained in section 2.4.1) possesses certain formula so that an image that almost resembles a particular image of pattern is considered to be that pattern. For example if the difference between the two images is negligible or fixed to a certain error rate is considered to have the same pattern. Comparisons are made repeatedly so that the inputted slide is found to either pattern matched or not pattern matched.

Stage 6: Disease detection

The last stage of the experimental setup in which only one of the diseases would be activated from the output layer of the back propagation algorithm denoting the disease that is to be determined.

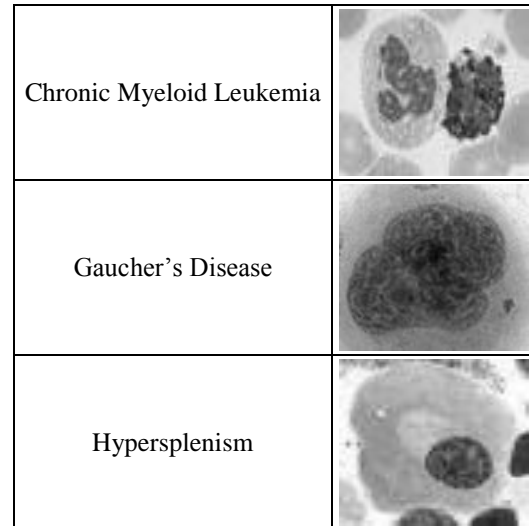


Fig.3. Slides containing blood cells

2.2 BLOOD: THE VITAL FLUID

Blood is a red fluid pumped by the heart to various parts of the body through the circulatory system. No other artificial fluid can replace the function of blood. Blood contains hemoglobin which makes the blood appear red in color. Blood contains 45% of solid content and the remaining liquid. Blood contains plasma that is 12% consisting of the solid particles and the remaining liquid.

The three major components of blood [7]-[9] are:

- Erythrocytes
- Leukocytes
- Platelets

as shown in the below Fig.4.

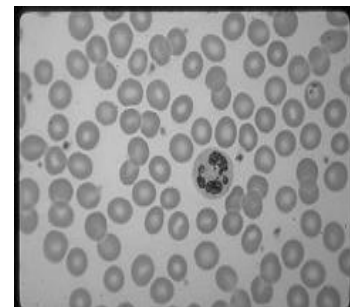


Fig.4. Components of Blood

Erythrocytes

Erythrocytes are also termed as Red blood cells. The function of Erythrocytes is to carry oxygen to various organs of the human body. The size of red blood cell is 7.5 microns in diameter. It is spherical in shape. In a human body weighing 75 kg containing 5 litre of blood can have about 1000 red blood cells in it. Red blood cells are almost same in diameter.

Leukocytes

Leukocytes are also termed as White blood cells. The function of Leukocytes is to fight against the foreign organisms or the cells getting into the blood due to the various diseases. When compared to Erythrocytes, Leukocytes are large in

number. The main way in differencing Erythrocytes and Leukocytes is the nucleus that is only possessed by the Leukocytes. The size of Leukocytes is also bigger in diameter compared to that of Erythrocytes.

Platelets

Platelets play a vital role in performing the clotting of blood. Platelets do not have a definite structure.

2.3 DIGITAL IMAGE PROCESSING

Digital image processing is the processing of the image data for storage, transmission and representation for autonomous machine perception. A digital image is composed of a finite number of elements each of which has a particular location and value. These elements are referred to as picture elements, image elements and pixels. "Pixels" is the term most widely used to denote the elements of the digital image. Vision is the most advanced of our senses, so it is not surprising that images play the single most important role in human perception. Thus digital image processing encompasses a wide and varied field of applications. There are fields such as computer vision whose ultimate goal is to use computers to emulate human vision, including learning and being able to make inferences and take actions based on visual inputs. This area is itself a branch of artificial intelligence whose objective is to emulate human intelligence. The area of image analysis also called as "Image understanding" is in between image processing and computer vision. Lets us now discuss about Medical Imaging that is part of Digital image processing.

2.3.1 Medical Imaging:

The parts of the digital image processing that are exclusively used in medical related fields define Medical Imaging[10]. There are numerous applications of the Digital image processing concepts used in medical fields like that of Gamma ray and X-ray imaging.

2.4 NEURAL NETWORK CONCEPTS INVOLVED

The architecture of the human brain is significantly different from the architecture of a conventional computer. The conventional computers are not suited to perform pattern recognitions problems. We therefore borrow features from physiology of the brain as the basis for our new process models. Hence the technology has come to be known as Artificial Neural Systems technology (ANN's) or simple Neural Networks. ANN's are composed of interconnecting artificial neurons which may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. Thus we are going to a concept involved in the Neural Networks termed as Back Propagation networks that plays a vital role in our project in the pattern recognition problem.

2.4.1 Back propagation network:

Back propagation network can be used to solve complex pattern matching problems. The network learns a predefined sets of input/output example pairs by using a two phase propagate adapt cycle. After an input pattern has been applied as a stimulus to the first layer Xi of the network units, it is propagated through

hidden layer Hj until an output is generated in output layer Y. This output pattern in then compared with the desired output and an error signal is computed for each output unit.

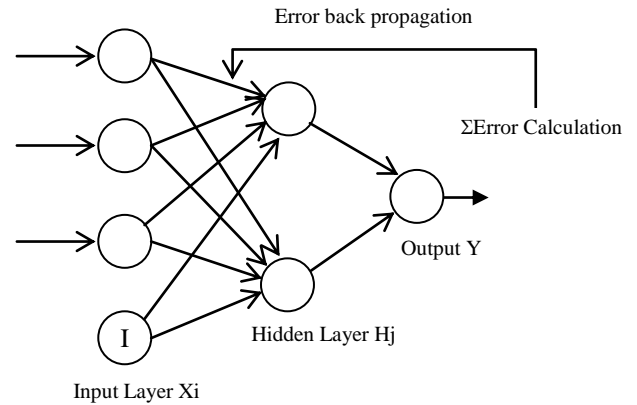


Fig.5. Back propagation Network

In our paper the input as shown in Fig 2.6 contains the input slides and the output is the table contains the diseased cell pictures.

We have considered discrete wavelet transform (DWT) for better texture modeling. A total of 12 texture images of size 512 × 512 have been used. For each image, 10 sub images of size 256 × 256 with 256 gray levels depth has been randomly selected for training and testing. The significance of the process is that as the network trains the nodes in the intermediate layers, they organize themselves such that different nodes learn to recognize different features of the total input space. After training when presented with the arbitrary input pattern that is noisy or incomplete, the units in the hidden layer of the network which will respond with the active output with the new input which contains the pattern that resembles the feature, the individual units learn to recognize during training,

$$E_p = (1/2) \sum_{k=1}^M \delta_{pk}^2$$

where, δ_{pk}^2 is the error term for the output units.

The above error value E_p is set so that the difference between the obtained and the actual image is fixed to a certain quantity.

Result analysis of Error differences by BPN		
Actual image	Obtained image	Result
		0
		1

Fig.6. Result analysis of error differences by BPN

Thus, only if the obtained image has an error difference within the allowed error term that is E_p the output is activated indicated by 1. Else if the error difference is exceeding E_p , then the output is not activated and is indicated by 0. As we know in the Back Propagation network we fix the error value. Using the error value or the threshold we could determine whether the inputted image matched the already trained image. If the difference is below the error or the threshold value then it is obvious that the inputted image is one among the trained sample of the blood. From the Fig.6, we could find the following,

- 1) When the obtained image is heavily blurred exceeding the error value and thus the result is 0.
- 2) When the obtained image is less blurred exceeding the error value and thus the result is 1.

2.4.1.1 Constructing BPN architecture:

The BPN architecture is diagrammatically represented as below.

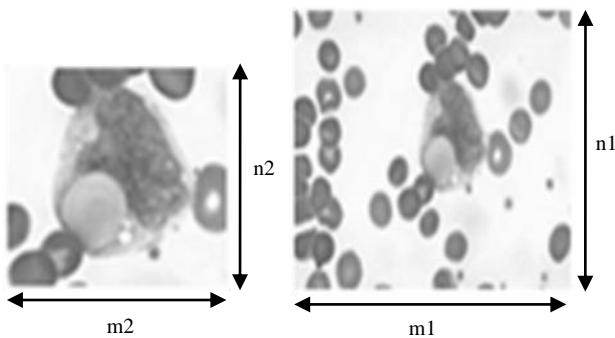


Fig.7. BPN Architecture

2.5 PATTERN MATCHING

Let's consider the following case in which the first matrix in the top to be the slide containing only the diseased cell. Let the second matrix be considered as the slide retrieved from the frame grabber containing the diseased cell.

As shown below we can observe that the first matrix is present in the second containing the values 53, 44, 67, 55. In ordinary or classic methods of pattern matching we would obviously think of solving the problem by sliding the first matrix with the second.

$$\begin{bmatrix} 53 & 44 \\ 67 & 55 \end{bmatrix}$$

$$\begin{bmatrix} 12 & 11 & 13 & 14 \\ 22 & 33 & 53 & 44 \\ 54 & 76 & 67 & 55 \\ 22 & 44 & 88 & 55 \end{bmatrix}$$

2.5.1 Basic principle of pattern matching:

So the step involved in the normal case is as below:

```

For i = 1 to m2
For j = 1 to n2
For k = 1 to m1
For l = 1 to n1
{
Processing statements

```

}
Thus the running time is as below:

$$Runtime = O(m2*n2*m1*n1)$$

where, 'O' denotes the BIG O notation representing the lower bound of any algorithm.

2.5.2 Enhanced pattern matching:

We now find a method from which we could obviously reduce the running time of the pattern matching involved. The steps involved are described as below,

- Step 1:** Number of pixels of a sample and the slide for height and width is collected for every sample.
- Step 2:** The first element of the sample is compared with each and every element of the slide. Comparison is done along with the error value of the BPN already prefixed.
- Step 3:** If at all the first element is equal (considering the error value) to an element then the remaining $m2$ elements in that row are found for match.
- Step 4:** If a row is matched a counter is set, comparison is done simply by skipping $m1$ value thus saving runtime.
- Step 5:** Pattern is matched if the counter value equals to $n2$. Depending upon the matched values that exceeds the threshold value gives a true result.

Thus the Enhanced pattern matching analysis is done as below.

*Running time = $O(m1*m2)$ comparatively less than $O(m2*n2*m1*n1)$*

*Thus time saved = $O(m2*n2*m1*n1) - O(m1*m2)$*

In order to still reduce the running time of the algorithm we can select some points may be in some fashion and find out appropriate match. This is represented in the following diagram,

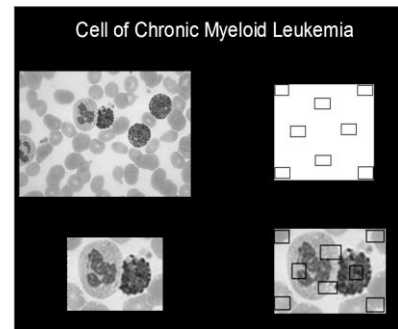


Fig.8. Block Extraction

This is represented as: Considering Slide dimensions as below the total no of comparisons required are $100 \times 100 = 10000$.

Table.1. Analysis of Improved algorithm

Block dimensions	No. of blocks required in minimum	Total no of comparisons	Percentage of running time saved
20 x 20	8	3200	68%
10 x 10	16	1600	84%
5 x 5	32	800	92%

The above is graphically represented below:

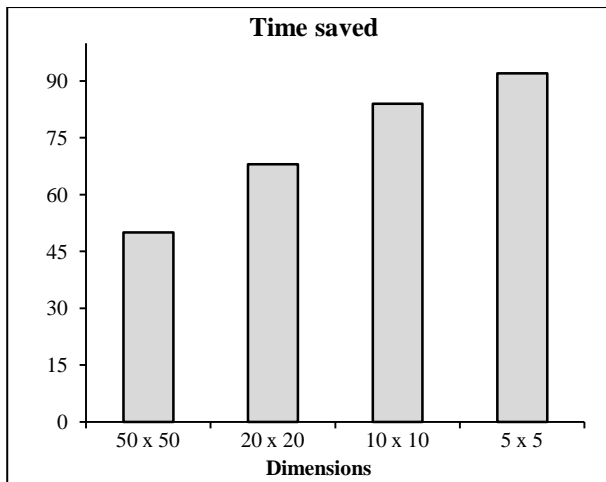


Fig.9. Graphical representation of algorithm analysis

Thus improved running time considering dimensions,

$$10 \times 10 = 16\% [O(m1*m2)]$$

Thus total time saved (Improved algorithm),

$O(m2*n2*m1*n1) - 16\%[O(m1*m2)]$ that is considerably more to that of $O(m2*n2*m1*n1) - O(m1*m2)$

3. RESULTS AND DISCUSSION

Blood cell recognition, disease diagnosis and some of the morphological methods are briefly reviewed in this section. This includes experimental investigations and image processing algorithms that have been developed to diagnosis diseases. To overcome the problems associated with blood cells, disease constituents and effectors during the diagnosis process, the image processing technique have been formulated with vision system to provide closed loop control. Several experiments were carried out to characterize the disease of the main modeling steps, i.e., establishing corresponding cells from a set of learning shapes and deriving a statistical model.

We applied the modeling procedure to distinct the incoming image with the normal one. We test our methods on 15 images of healthy and 30 images with disease. In our approach, we established the corresponding cells prior to pose and scaling estimation during statistical modeling. This assumes that the position of the cells does not change significantly with different pose and scaling, which is justified by the high accuracy of the model.

To assess how well the statistical model generalizes to unseen objects, we measured how close a shape not part of the learning sample can be approximated with the statistical shape model. To compute the robustness of the method with respect to different parameter settings, the entire modeling was carried out for a range of parameter at a time while keeping the remaining parameters at their optimal values. The images used are presented here, which have exhibited a wide variety of lesions and confusing manifestations. On this difficult data set, our methods successfully detect the disease in most of the reliable cases.

4. CONCLUSION

In this paper, we have presented a method to automatically construct statistical blood cells from segmented microscopical

images. Standard models are used to diagnosis the disease that can also be used for haematological applications. Corresponding parameters on the surface of the shape are established automatically by the adaptation of back propagation algorithm to segmented volumetric images. To assess the robustness of the method with respect to the parameter settings, the entire statistical model was carried out for range of parameter values, varying one at a time while keeping the remaining parameters at their optimal values. MS Visual C++ and Matrix Inspector image processing software are used for the experimental work. Nevertheless, this would be an interesting future investigation, since our approach is sufficiently general to be applied to modeling tasks from other two dimensional or three dimensional application domains. Nevertheless, this would be an interesting future investigation, since our approach is sufficiently general to be applied to modeling tasks from other two dimensional or three dimensional application domains.

REFERENCES

- [1] G. Schneider, G. Artmann and G. Henning, "Automated image processing system for shape recognition of single red blood cells based on out-of-focus images", Elsevier Publications, Vol. 32, No. 2, pp. 237-238, 1995.
- [2] G.H. Landeweerd, E.S. Gelsema, J.F. Brenner, W.D. Selles and D.J. Zahniser, "Pattern recognition of nucleated cells from the peripheral blood", *Pattern Recognition*, Vol. 16, No. 2, pp. 131-140, 1983.
- [3] Daniela Mayumi Ushizima Sabino, Luciano da Fontoura Costa, Edgar Gil Rizzatti and Marco Antonio Zago, "A texture approach to leukocyte recognition", *Real-Time Imaging, Imaging in Bioinformatics: Part III*, Vol. 10, No. 4, pp. 205-216, 2004.
- [4] G.H. Landeweerd, T. Timmers and E.S. Gelsema, "Binary tree versus single level tree classification of white blood cells", *Pattern Recognition*, Vol. 16, No. 6, pp. 571-577, 1983.
- [5] G.H. Landeweerd, T. Timmers and E.S. Gelsema, "Classification of normal and abnormal samples of peripheral blood by linear mapping of the feature space", *Pattern Recognition*, Vol. 16, No. 3, pp. 319-326, 1983.
- [6] J.M. Lester, H.A. Williams, B.A. Weintraub and J.F. Brenner, "Two graph searching techniques for boundary finding in white blood cell images", *Computers in Biology and Medicine*, Vol. 8, No. 4, pp. 293-308, 1978.
- [7] K. Parthenis, Metaxaki-Kossionides and B. Dimitriadis, "An automatic computer vision system for blood analysis", *Microprocessing and Microprogramming*, Vol. 28, No. 1-5, pp. 243-246, 1990.
- [8] Philip E. Norgren, Ashok V. Kulkarni and Marshall D. Graham, "Leukocyte image analysis in the diff3 system", *Pattern Recognition*, Vol. 13, No. 4, pp. 299-314, 1981.
- [9] M. Schönfeld and R. Grebe, "Automatic shape quantification of freely suspended red blood cells by isodensity contour tracing and tangent counting", *Computer Methods and Programs in Biomedicine*, Vol. 28, No. 4, pp. 217-224, 1989.
- [10] Xubo B. Song, Yaser Abu-Mostafa, Joseph Sill, Harvey Kasdan and Misha Pavel, "Robust image recognition by fusion of contextual information", *Information Fusion*, Vol. 3, No. 4, pp. 277-287, 2002.