Pattern Recognition Letters 49 (2014) 155–161

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Cell morphology based classification for red cells in blood smear images ☆

Howard Lee, Yi-Ping Phoebe Chen *

Department of Computer Science and Computer Engineering, La Trobe University, Melbourne, Australia

ARTICLE INFO

Article history:
Received 14 October 2013
Available online 16 July 2014

Keywords:
Red blood cell
Anemia
Classification
Visual information
Hybrid ANN model

ABSTRACT

Red blood cells are the most common type of blood cell and are responsible of delivering oxygen to the body tissues. Abnormalities in red blood cell may change the physical properties of the red cell or shorten its life spend, and may lead to stroke or anemia. In this paper, we proposed a hybrid neural network based classifier, which utilize the visual information extracted from the red blood cell images to determine whether a red cell is normal or abnormal. Based on the feature properties, we clustered the visual features into two main groups, namely shape and texture cluster groups. The input feature clusters were processed using parallel and cascade architecture with multiple input layers. Our experimental result has shown significant improvement in classification accuracy in our proposed system as compared to the single input layer classifier with recent feature selection algorithms.

1. Introduction

The circulation system is one of the most vital organizations in the human body. In addition to the importance of cardiovascular system, some tiny red blood cells have the vital duty of carrying nutrient and oxygen to different parts of the body and collecting unwanted materials from them. Otherwise, the whole body would suffer from the diseases that prevent red blood cells from doing their tasks. Much diagnostic information can be extracted from the shape and size of the red blood cells, the number of the red blood cells in a particular sample of the blood, and the ratio between the area containing oxygen and the whole area of each cell [1,8,10,14]. Analyzing the hematological images manually is tiresome and time consuming. It also suffers from inter and intra observer variability. Hence, automation of this task will be helpful in identifying diseases related to the red blood cells accurately, besides saving the precious time of the hematologists.

Detection of the red blood cells is the first and most important step, which presents several challenges. In addition, overlapping of the cells is another problem in detecting the red blood cells. Moreover, the hematological images suffer from variations in color and intensities [7,8]. Therefore, the techniques employed for this purpose should be robust to these problems.

Many literatures in blood cell classification focused on separating red blood cells and white blood cells [1,3,8,9]. In [4], the image intensities were used as features and the Principal Component Analysis (PCA) was employed for feature reduction. Using local linear map neural network as a classifier, they managed to detect the leukocytes in fluorescent images. On the other hand, in [9], Fisher Linear Discriminant (FLD) was used to improve the features extracted based on image's intensities, and Multi-Layer Perceptron (MLP) was employed for detecting the leukocytes in the fluorescent images. However, using intensities as features may not be very effective in detecting the overlapping cells. The second-order edge detection methodology for detecting the overall number of cells irrespective of their stain was proposed in [7]. Then, proliferating cells were located using PCA of the color image along with histogram thresholding. Vromen and McCane [18] used depth map to classify the red blood cell according to its surface type. Although the work has successfully classified the red blood cell into eleven different surface types (shape characteristics), the work was mainly for counting the number of red blood cells in the blood smear image, and did not establish the links between these surface types and the red cell diseases.

To improve the classification result, feature selection algorithms have been investigated to obtain the optimum feature space to improve the classification result [2,5,10,11,16,22–24]. The subset of the features are selected by evaluating the mutual information [2,12,21], redundancy [22,24], or the relevance [24,26]. These algorithms are effective to model situations where large feature space is available. However, with the limited features extracted to represent

☆ This paper has been recommended for acceptance by Nappi Michele.

* Corresponding author.
E-mail address: phoebe.chen@latrobe.edu.au (Y.-P.P. Chen).

http://dx.doi.org/10.1016/j.patrec.2014.06.010
0167-8655/

the visual information about the red blood cells, our results showed that the optimum feature space is highly dependent by the samples used. Hence, we proposed a clustered feature analysis approach for classification. The entire collections of features are divided into clusters according to the property of the features.

In this paper, we performed the red blood cell classification based on the single cell images extracted from the microscopic blood smear images. We divided the entire collections of features into two clusters according to the properties of the features – shape and texture. Separate input layers were used to process individual clusters of features. We proposed a hybrid neural network architecture – incorporate both parallel and cascade topologies for red blood cell classification.

2. Problem description and methodology

This study considered the data samples associated with normal and various abnormal red blood cells. Studies have shown that red blood cells can mutate due to genetic inheritance, such as sickle cell anemia (SCA) [12]. A condition usually presents in childhood, and is more commonly in people or their descendants from the tropical and subtropical regions where malaria is common. SCA is characterized by red blood cells that appear abnormal, rigid sickle shape, due to a mutation in the hemoglobin gene. This sickling condition decreases the cells' flexibility and may results in various complications such as stroke, and chronic renal failure [1]. Spur cell anemia is a form of anemia where red blood cells have unusual spiked appearance due to liver disease. The condition shortens the life span of the red blood cells, and if new red cells are not made fast enough to replenish the red blood cells lost, anemia will occur. Malaria is caused by an infection of blood parasites transmitted by mosquito bites. The infected blood cells circulate to the liver and allow the parasites to invade the liver cells causing liver and kidney failures [1,20]. In the clinical approach, a blood smear test can be carried out to analyze any blood abnormalities under a microscope [1,14,19,20,25]. Our study focused on the visual feature analysis of the red blood cell images to distinguish different red blood cell diseases.

3. Blood cells segmentation and separation

One of the challenging aspects in blood cell classification is to delineate the individual blood cells accurately from the blood smear image [17,18]. Individual samples of the red blood cell are delineated from the background by the Otsu's method, by searching for the threshold that minimizes the intra-class variance, defined as the weighted sum of variances of the two classes of the image (foreground and background):

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \tag{1}$$

where the weights w_i are the probabilities of the two classes separated by a threshold t and σ_i^2 is the variances of the two classes.

The Otsu threshold can be obtained by maximizing the variance between the classes, $\sigma_b^2(t)$, defined by:

$$\sigma_b^2(t) = w_1(t)w_2(t)(\mu_1(t) - \mu_2(t))^2 \tag{2}$$

which is an expression of the class probabilities, w_i and class means μ_i, and the process can be implemented by an iterative process which updates $\sigma_b^2(t)$ iteratively. However the Otsu segmentation is highly effected by the illumination changes in the image, hence to address this issue, the images have been normalized before the segmentation process.

The image of the red blood cell is then super-imposed on to the resulting Otsu segmentation output, to retain the complete segmentation of the red blood cell. Visual features describe the shapes and textures of the red blood cells extracted for this study. Most of the red cell abnormalities resulting in the deformation of the cells, such as sickle cell anemia and spur cell anemia, and some associated with changes in the cell texture due to infections, as in the case of malaria. In our study, we extract various shape and texture features to classify normal and abnormal red blood cells.

In the case of overlapping red cells, the segmentation result based on the Otsu algorithm failed to obtain an accurate shape of the red cell. Due to the similarity in the color and intensity, it will be difficult to separate red blood cells by a threshold level. To separate two overlapping cells, we first enhance the edge of image by canny edge algorithm. The second task is to separate the two blood cells. The upper cell will be extracted and used for the classification in this paper, and the lower cell will be discarded due to incomplete boundary information.

To determine the overlapping order of the cells, we determine if the overlapping edge is forming a concave or convex arch to each of the cells, by connecting the end-points of the overlapping edge curve. Fig. 1(a) depicts the situation when the two cells overlapped, it is clearly that the line falls on the cell 1, hence cell 1 is on top of cell 2.

The individual cell is then extracted by applying image mask at the enclosed regions defined by the Canny edge Fig. 1(c). After the cells been separated, it shows the original shape of the upper cell (Fig. 1(d)) was retained after the separation, and while the lower cell (Fig. 1(e)) was distorted at the overlapping region.

In this paper, our aim is to classify different red blood cell abnormalities by the visual information, (as described in Section 4), hence the distorted cell was discarded.

4. Feature extraction

4.1. Shape feature cluster

A healthy red blood cell resembles a thin circular disk. To quantitatively analyze the shape of the red blood cells, we extract the following shape features: cell circularity, medial axis ratio, cell deform ratio, eccentricity and Hausdorff distance.

Cell circularity:

The circularity measurement of the cell is defined as the ratio between the cell area and the square of its perimeter

$$Cir = \frac{4\pi A}{p^2} \tag{3}$$

where A is the total area of the cell, and p represents the cell perimeter.

The cell roundness measures how close the shape of a given cell is close to a circle, which resembles a normal healthy red blood cell. The closer the roundness value to unity, the better it is describing a normal cell.

Medial axes ratio (MAR):

$$MAR = \frac{L_{minor}}{L_{major}} \tag{4}$$

where L_{minor} represents the length of the minor principal axis, and L_{major} represents the length of the major principal axis.

The medial axes ratio describes the property that the shape of the red blood cell has been stretched. The closer this feature is to unity, the less of the shape distortion of the cell.

Cell deformation ratio (DR):

$$DR = \frac{A_{mx} - A_{cell}}{A_{mx}} \tag{5}$$

where A_{mx} denotes the area of the circle bounded by the cell, with the diameter equals the length of the major p axis of the cell. A_{cell} denotes the actual area of the cell.