

# Computer Aided System for Nuclear Stained Breast Cancer Cell Counting

Pornchai Phukpattaranont, Somchai Limsiroratana, Kanita Kayasut  
and Pleumjit Boonyaphiphat  
*Prince of Songkla University  
Thailand*

## 1. Introduction

### 1.1 Image characteristics

Immunohistochemistry is a technique used for detecting *in situ* a tissue antigen by a specific antibody. An antigen-antibody reaction is visualized by the color development of specific dye and can be seen by light microscope. The tissue antigen is presented at any part of the cell, i.e., cell membrane, cytoplasm or nucleus. Therefore, it is a useful technique to demonstrate the protein markers including cancer cell. Estrogen receptor (ER) and progesterone receptor (PR) are prognostic markers for breast cancer detected by this method. Evaluation of ER and PR positive cells are useful for hormonal therapy.

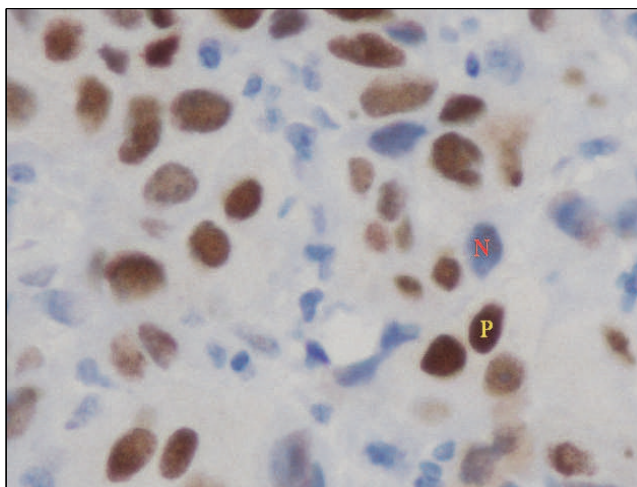


Fig. 1. An example of a stained cancer cell image. The brown and blue nuclei with the added labels are representative samples of positive and negative staining of estrogen receptor of breast cancer cells, respectively.

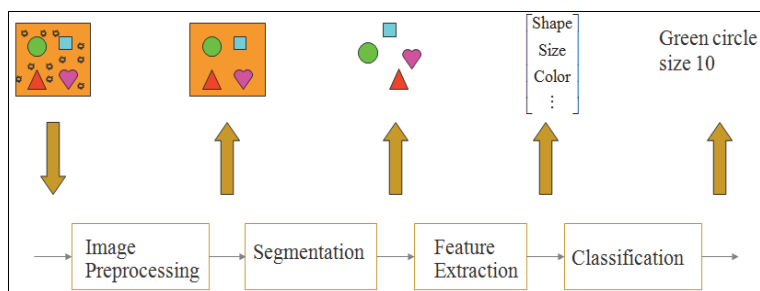


Fig. 2. The principle of image analysis in computer aided system for cancer cell classification.

Figure 1 shows an example of stained cancer cell image from microscope with a magnification of 40x. This staining procedure is utilized to demonstrate the existing of estrogen or progesterone receptors in the breast cancer cells. In other words, stained cancer cells are classified into two categories according to their nuclear color contents, i.e. brown and blue. The brown color indicates a positive (P) staining while the blue one demonstrates a negative (N) result. The brown and blue cells with the added labels shown in Figure 1 are representative samples of positive and negative staining of estrogen receptor of cancer cells, respectively. The ratio of the total number of positive cancer cells to the total number of cancer cells in the whole image is used by a doctor for medical planning and treatment.

Traditionally, the percentage of positive cells of those markers is semi-quantitatively counted. However, it is time consuming, costly, subjective and tedious. To overcome these problems, image analysis that previously requires manual operations is performed on the basis of the developments in computer capabilities and image processing algorithms (Fang et al., 2003; Petushi et al., 2004; Thiran & Macq, 1996). There are a number of benefits that result from image analysis with computer aided system. These include an acceleration of the process, a reduction in cost for image analysis, as well as a decrement in a false inspection due to fatigue. Additionally, the computer aided analysis provides a quantitative description. Based on this quantitative measurement, the analysis result is objective. Furthermore, the correlation of the quantitative categorization with patient symptoms may allow for an automated diagnostic system (O’Gorman et al., 1985). However, it is not expected that computer aided image analysis will replace pathologist’s experience. It is only an aid to the pathologist for the repeated routine work and yields quantitative results that complement and enhance interpretations by pathologists. Visual examination by the pathologist is still required where the objects that the method is not trained to deal with are encountered.

## 1.2 Image analysis

The principle of image analysis in computer aided system for classifying cancer cells is composed of four stages as shown in Figure 2. After the preprocessing stage, the image is segmented in order to keep the interesting parts and remove the undesirable components. Next, the feature extraction process is applied in order to extract the useful information from the segmented objects. Finally, the classification is performed using the characteristics extracted from the previous stage. The ultimate goal of our research project is to develop an automatic algorithm for counting positive and negative cancer cells on immunohistological stained slides from breast cancer tissue.

One of the most important stages for the image analysis is cell image segmentation, which is the main focus of this work. Cell image segmentation methods for pathology are largely relied on two image processing techniques: thresholding and region growing. Thresholding identifies each pixel in the image into either object or background based on its intensity. Image thresholding is a simple yet often effective means for obtaining a segmentation of images in which cells of interest situate in uniform background intensity across the image. Consequently, it is often used as a part in a sequence of image processing operations. For example, thresholding is one of important steps in segmenting live cell image (Wu et al., 1995). For the image with uneven background, local adaptive thresholding method is an effective technique because the threshold value can be adapted to the background intensity variations. In other words, the threshold value is derived in each subregion, which is considered to be composed of both background and objects. Examples of medical image analysis that contain the local adaptive thresholding algorithm are as follows: (1) Segmentation of fluorescence tumor cells from tissue sections (Fang et al., 2003), (2) Segmentation of cell nuclei from the stroma and fat like regions (Petushi et al., 2004), (3) Segmentation of tissue components in liver tissue (O’Gorman et al., 1985), (4) Segmentation of dead and live hepatocytes (liver cells) in cultures from microscopic images (Refai et al., 2003), and (5) Segmentation of nuclei from a breast tissue image (Zhao & Ong, 1998). One of disadvantages of the thresholding technique is that it does not take into account the spatial characteristics of the image. This causes it to be sensitive to noise and intensity heterogeneity. Therefore, further image processing algorithms that consider spatial modeling of the image need to be incorporated (Li et al., 1995).

Region growing separates objects of interest from background in the image based on some predetermined criteria, i.e. intensity and/or edges (Haralik & Shapiro, 1985). While edge-based methods are sensitive to noise and artifacts, the intensity-based algorithms are usually more computationally expensive. Additionally, one of main disadvantage of region growing is the requirement of manual interaction to obtain the seed point (Adams & Bischof, 1994). Similar to thresholding, region growing is partially used in a set of image processing operations (Beveridge & et al., 1989; Wani & Batchelor, 1994). An example of region growing used in medical image analysis includes the extraction of noisy cell contour as appear in (Wu & Barba, 1994). In addition to thresholding and region growing, other techniques are the segmentation of white blood cells based on morphological granulometries (Theera-Umpon & Gader, 2000) and the segmentation of white blood cells based on the principle of least commitment (Park & Keller, 1997).

We propose two image segmentation methods for breast cancer cell counting. Firstly, we present the segmentation algorithm for breast cancer cell image based on local adaptive thresholding (Phukpattaranont & Boonyaphiphat, 2006). The method is appropriate for microscopic images with low histological noise, i.e., low variations on background color and intensity. However, the degree of histological noise in breast cancer images varies. For the image with high histological noise, the local adaptive thresholding approach is sensitive to noise and intensity heterogeneity. In addition, the computational time for the approach is quite lengthy. To address these problems, we propose a strategy for segmenting cancer cells in a microscopic image of immunohistological nuclear staining of breast cancer tissue based on the color of pixel (Phukpattaranont & Boonyaphiphat, 2007; Phukpattaranont et al., 2009). This is motivated from the way that a pathologist determines the positive and negative of tumor cells by using their color contents manually. As a result, we propose the use of pixel color partitioning based on a neural network classifier for segmenting cancer cells microscopically as the second algorithm. The remainder of this chapter is organized as

follows: Section 2 describes the cell counting technique based on local adaptive thresholding; Section 3 presents the cell counting technique based on neural network; and Section 4 gives conclusions and recommendations for future work.

## 2. Cell counting based on local adaptive thresholding

### 2.1 Method

In this section, we present an approach on segmenting cancer cells from a microscopic tissue image of breast cancer. Color space changing and anisotropic diffusion filtering for noise removal are performed in the preprocessing stage. Subsequently, the preprocessing result is segmented using local adaptive thresholding, morphological operations, and cell size considerations. Finally, cell types are classified based on their color content.

#### 2.1.1 Image preprocessing

Two processes are performed in the preprocessing stage. Firstly, we transform the red-green-blue (RGB) color space to CIE Lab space. The CIE Lab space is chosen due to the close correlation between its Euclidean distances and human perception of colors. The CIE Lab space can be defined by (Trussell et al., 2005)

$$L^* = 116 \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \quad (1)$$

$$a^* = 500 \left[ \left( \frac{X}{X_n} \right)^{\frac{1}{3}} - \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} \right] \quad (2)$$

$$b^* = 200 \left[ \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left( \frac{Z}{Z_n} \right)^{\frac{1}{3}} \right] \quad (3)$$

for  $\frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n} > 0.01$ . The values  $X_n, Y_n, Z_n$  are the CIE (Commission Internationale de l'éclairage) tristimulus values of the reference white under the reference illumination, and  $X, Y, Z$  are the tristimulus values, which are mapped to the CIE color space. While the  $L^*$  component represents intensity, the  $a^*$  and  $b^*$  components are proportional to red-green and yellow-blue color contents, respectively.

Secondly, we apply the anisotropic diffusion with the  $L^*$  component from the first step. The objective of this operation is to smooth regions inside cancer cells while still preserve the edge and contrast at sharp intensity gradients. The image resulting from this step significantly facilitates the segmentation algorithm in the next stage. The anisotropic diffusion equation can be expressed as (Perona & Malik, 1990)

$$I_t = \text{div}(c(x, y, t) \nabla I), \quad (4)$$

where  $\text{div}$  is the divergence operator,  $c(x, y, t)$  is the diffusion coefficient, and  $\nabla I$  is the gradient operator. In order to smooth the area within an object of interest and simultaneously preserve high gradient boundaries, the diffusion coefficient is given by

$$c(x, y, t) = g(\|\nabla I(x, y, t)\|) = e^{-(\|\nabla I(x, y, t)\|/\kappa)^2}, \quad (5)$$

where  $\kappa$  is a constant that controls conduction. We implement the numerical solution of Equation (4) using the algorithm provided in (Perona & Malik, 1990), which is given by

$$I_{i,j}^{t+1} = I_{i,j}^t + \lambda [c_N \cdot \nabla_N I + c_S \cdot \nabla_S I + c_E \cdot \nabla_E I + c_W \cdot \nabla_W I], \quad (6)$$

where  $0 \leq \lambda \leq 1/4$  for numerical scheme to be stable. Please see (Perona & Malik, 1990) for more details.

### 2.1.2 Segmentation

We use the combination of local adaptive thresholding, morphological operations, and cell prior knowledge in our segmentation algorithm. For local adaptive thresholding, we apply  $M$ -by- $M$  sliding blocks with the output image from the anisotropic diffusion. The threshold of each local block is determined using Otsu's method (Otsu, 1979). After finishing adaptive thresholding for all local blocks, we process the black and white image using morphological opening in combination with a cell size consideration. These two operations are successively used to eliminate spike noise, fill holes, and separate touching cancer cells.

### 2.1.3 Feature extraction and cell classification

In this step, we use the features based on color components to classify cancer cells. The average values of  $L^*$ ,  $a^*$  and  $b^*$  from the representative regions of light brown (P1), brown (P2), dark brown (P3), and blue (N) cells are selected to serve as the reference markers. All segmented cancer cells are categorized using the following two steps. First, we classify each N cancer cell from the others, i.e. P1, P2, and P3, in the image. The Euclidean distances between the average values of  $a^*$  and  $b^*$  of the considering cell and those of all reference markers are compared. The smallest distance will tell us that the cell most closely matches that reference marker. For example, if the distance between the considering cell and the N reference marker is the smallest, then it would be labeled as the N cancer cell. Second, we separate the remaining P1, P2, and P3 cancer cells using the average value of  $L^*$  component. Also similar to the first step, the minimum Euclidean distance is used to determine the class of cancer cells.

## 2.2 Results and discussion

### 2.2.1 Image preprocessing

Figure 3(a) shows the original color image used in demonstrating the proposed algorithm. Figure 3(b) shows the intensity image resulting from the  $L^*$  component of CIE Lab color space of stained cancer cell image. We can see that most of cancer cells stay separately and have the round shape. Figure 3 (c) shows output images from anisotropic diffusion. Parameters used for the numerical solution of anisotropic diffusion are as follows: speed of diffusion ( $\lambda$ ) = 0.2, conduction coefficient ( $\kappa$ ) = 20, and number of iterations = 50. It can be seen from the image that the anisotropic diffusion filtering successfully removes undesirable noise while still preserving sharp edges of cancer cells.

### 2.2.2 Segmentation

Segmented images resulting from the local adaptive thresholding with block sizes 7, 11, 15, and 19 are investigated. Results show that the pixel elements of cells from the output image with block size 7 is not uniform and not well connected. Cells from the output image with increasing block sizes, i.e. 11, 15 and 19, show more homogenous and completed shapes. However, cell sizes also appear larger and any two cells that stay closely may merge together. As a result, the local adaptive thresholding that operates with block size 11 is the most appropriate for image segmentation. When the 11-by-11 sliding block is used for local adaptive thresholding in the demonstrating image, the output image is shown in Figure 3(d). In order to eliminate spike noise, the binary image is processed using

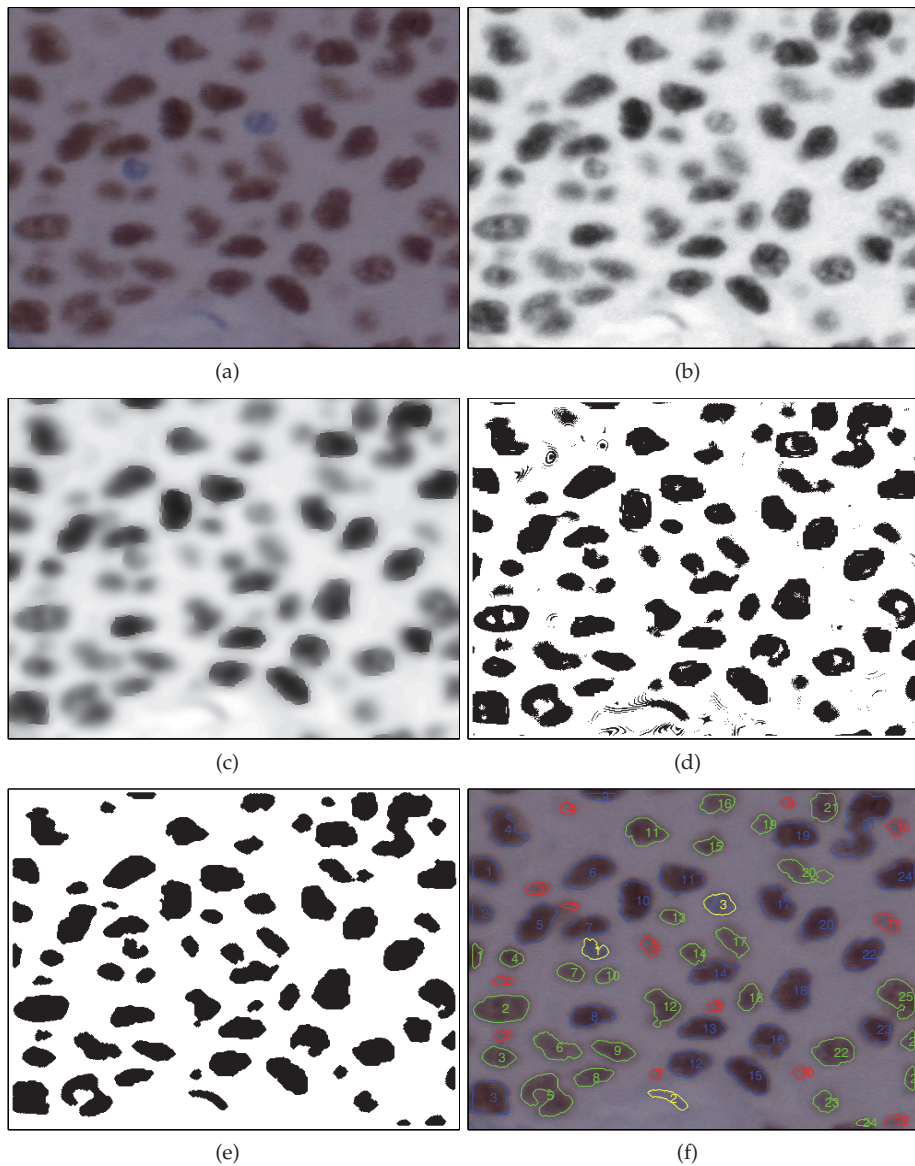


Fig. 3. Images resulting from the algorithm based on local adaptive thresholding. (a) Original color image. (b) Intensity image of stained cancer cell image. (c) Image from the anisotropic diffusion filtering. (d) Image from local adaptive thresholding. (e) Image from morphological operations. (f) Cell classification results.

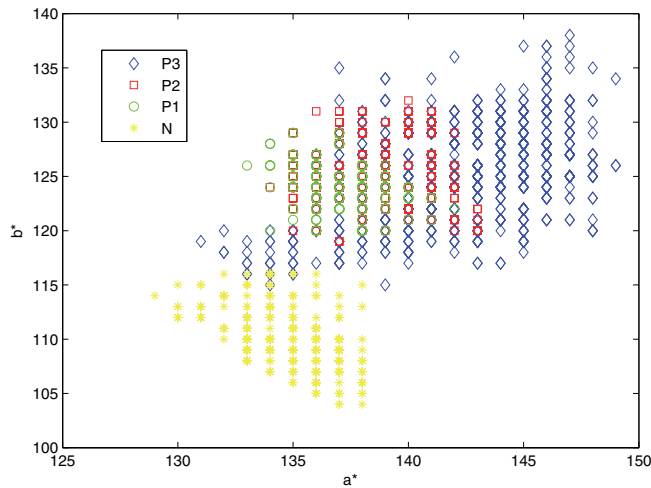


Fig. 4. Scatter plot between the  $a^*$  and  $b^*$  components of pixels from four different classes of reference cancer cells. Diamond: Light brown (P1). Square: Brown (P2). Circle: Dark brown (P3). Point: Light blue (N).

morphological opening with the disk-shaped structuring element. The algorithm based on morphological reconstruction is subsequently used to fill holes in the image. Additionally, the size of cancer cells under consideration must be greater than 140. A segmentation result after morphological operations is shown in Figure 3(e). We can clearly see that the boundaries of segmented cancer cells are in agreement with those of their original images shown in Figure 3(a).

**2.2.3 Feature extraction and cell classification**

The scatter plot between the  $a^*$  and  $b^*$  components of pixels from four different classes of reference cancer cells in demonstrating image is displayed in Figure 4. Color components of pixels from P1, P2, P3, and N cancer cells are shown using diamond, square, circle, and point markers, respectively. The scatter plot shows a good separation of the  $a^*$  and  $b^*$  components of N cancer cells from those of P1, P2, and P3.

Cell type	$L^*$	$a^*$	$b^*$
P1	102 ± 4	137 ± 2	124 ± 2
P2	87 ± 5	138 ± 2	125 ± 2
P3	42 ± 8	139 ± 3	121 ± 3
N	95 ± 5	134 ± 2	110 ± 2

Table 1. Average values ± standard deviation (SD) of color components from reference cancer cells.

Table I reports the average values and corresponding standard deviation (SD) of the  $L^*$ ,  $a^*$  and  $b^*$  components of P1, P2, P3, and N reference cancer cells. The average values shown in the table are used as the reference markers for classifying all cancer cells in the image. We can clearly see from Table I that the separation degree of average values of  $a^*$  and  $b^*$  agrees with the scatter plots shown in Figure 4. Moreover, values of  $L^*$  component from CIE Lab color

space in Table I suggest that we can use intensity of pixels to classify P1, P2, and P3 cancer cells due to their degree of separation.

Classification result of the cancer cell image using minimum Euclidean distances of color components is shown in Figure 3(f). While N cancer cells are shown using yellow contour plots, P1, P2, and P3 cancer cells are shown in red, green, and blue contour plots, respectively. A total number of cancer cells in the image are 67. The number of P1, P2, P3, and N cancer cells are 13, 27, 24, and 3, respectively. This classification result is consistent with the manual result from a specialist. However, we need to develop an additional segmentation algorithm for overlapping and irregular-shaped cancer cells such as the cells shown in the bottom left and top right corners of Figure 3(a).

#### 2.2.4 Image with high histological noise

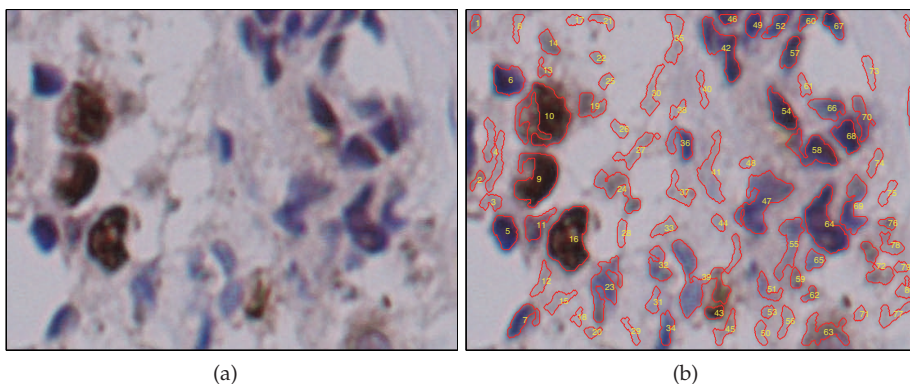


Fig. 5. (a) Breast cancer cell image with high histological noises. (b) Results of image analysis based on local adaptive thresholding.

The cell counting based on local adaptive thresholding is appropriate for microscopic images with low histological noise, i.e., low variations on background color and background intensity. However, the degree of histological noise in breast cancer images is varied image by image. Figure 5(a) shows an example of breast cancer cell with high histological noise, i.e., high variations on background color and background intensity. When the image is applied with the local adaptive thresholding algorithm, the result is shown in Figure 5(b). Not only the cancer cells but also other artifacts are detected. This causes the so called over-segmentation problem. In addition, the computational time for the approach is quite lengthy. To address these problems, we propose a strategy for segmenting cancer cells in a microscopic image of immunohistological nuclear staining of breast cancer tissue based on the color of pixel in the next section.

### 3. Cell counting based on neural network

#### 3.1 Method

The originally acquired image is in the red-green-blue (RGB) color space. That is, the color image is formed by the combination of red, green, and blue monochrome images. In the first step, we classify color pixels in the image into one of three categories, i.e. background, positive (P), or negative (N), based on their RGB components. There are many classifiers



that can be used for partitioning color of a pixel. However, a neural network is chosen due to its well known as a successful classifier for many applications (Gelenbe et al., 1996; Heermann & Khazenie, 1992; Reddick et al., 1997). Subsequently, morphology operations are used for addressing the spatial characteristics of cells. Finally, in order to obtain accurate cell counting results, the marker-controlled watershed, a classical method for separating overlapping objects, is applied for separating attached multiple cells into distinct single cells. An algorithm for segmenting cancer cells based on their colors and sizes is as follows.

#### **An algorithm for cancer cell segmentation**

```
Read image file
Classify pixels using neural network
Do thresholding
Do morphology opening
Fill holes
Label objects in the image
Classify each object using its size
IF size < 400 THEN
    Ignore object
ELSE IF  $400 \leq \text{size} \leq 1000$  THEN
    Object is a distinct single cell
    Apply morphology closing (1)
ELSE size > 1000 THEN
    Object is attached multiple cells
    Apply morphology opening
    Apply marker-controlled watershed (2)
ENDIF
Combine the results from (1) and (2)
```

Details of the algorithm are given below.

#### **3.1.1 Neural network**

We use backpropagation neural network to classify pixels in the microscopic image according to their color contents. Backpropagation is created by generalizing the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train the network until it can classify the defined pattern. The training algorithms use the gradient of the performance function to determine how to adjust the weights to maximize performance. The gradient is determined by a technique called backpropagation, which involves performing computations backwards through the network. The backpropagation computation is derived using the chain rule of calculus (Hagan et al., 1996).

Based on our experiences, the number of neural network layer between two and three is appropriate for classifying color of pixels in cancer cell images. Therefore, a backpropagation

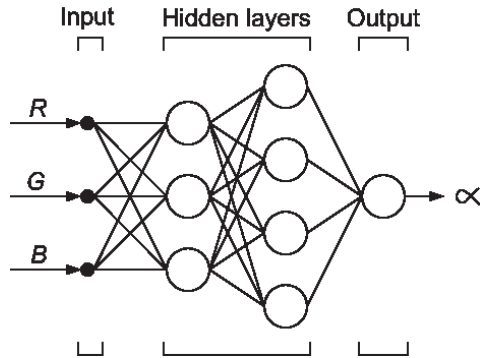


Fig. 6. Diagram of backpropagation neural network.

neural network of three layers shown in Figure 6 is chosen to classify image pixels whether they are from background, P, or N regions. The input vector is composed of 3 elements corresponding to the RGB color vector of pixel. Two hidden layers are determined empirically to be 3 and 4 neurons and the output layer consists of one neuron. In addition, the transfer functions of hidden and output layers are tan-sigmoid and linear, respectively. For the training of neural network, the target is assigned to be -1, 0, and 1, which are corresponding to RGB components from background, P, and N regions, respectively. The networks are trained using the Levenberg-Marquardt (LM) algorithm. The training stops when the maximum number of epochs reaches 100 or the mean square error is less than  $1 \times 10^{-12}$ . The number of pixels used for training neural network from each reference region is 1600.

### 3.1.2 Morphology operations

Mathematical morphology is a nonlinear operator based on set theory operating on object shape. It is a powerful tool to numerous image processing problems, for example, image preprocessing, segmentation using object shape, and quantitative description objects (Gonzalez & Woods, 2002). We utilize mathematical morphology as a tool for noise filtering and shape simplification in our work. Note that the disk-shaped structuring element (SE) with a radius of  $R$  is used for all morphological operations. As a result, the size of SE matrix is  $2R + 1$  by  $2R + 1$ . In addition, all mathematical morphology operations are applied once for each stage.

After finishing color partition for all pixels, the output image from neural network is transformed to a black and white image by thresholding. That is, while the pixels in background region are transformed to zero, the pixels in P and N regions, i.e., objects of interest, are transformed to one. In order to eliminate spike noise, the binary image is processed using morphological opening. The disk-shaped SE with a radius of 1 is used in this stage. The algorithm based on morphological reconstruction is subsequently used to fill holes in the image before performing size consideration.

In the next step, we classify each object in the image into one of three categories according to its size: small, medium, and large. The value of sizes used for cell classification is predetermined from guidance by a specialist. The small object is regarded as noise and is ignored. The

medium object is considered to be a distinct single cell. For the large object, it is determined as attached multiple cells. All distinct single cells are processed further with mathematical closing to complete their shape. It is used as spatial compensation for an uneven distribution of color in the cell. The SE used in this stage has a radius of 4. On the other hand, all attached multiple cells are applied with morphological opening. The radius of the SE in this stage is 4. There are two explanations for doing this performing. First, the multiple cells with small degree of attachment can be kept apart. Second, it can be used as a preparation step before marker-controlled watershed processing.

### 3.1.3 Marker-controlled watershed

In order to separate attached cancer cells into individual objects, we further process the result from last step with marker-controlled watershed. The watershed algorithm is shown to be a powerful tool for dividing attached objects (Vincent, 1993). The marker computation is used as an additional processing because the direct use of watershed transform usually yields the over-segmented result (Fang et al., 2003). The computational procedures for marker-controlled watershed are as follows.

- Step 1. Use the Sobel edge marks to compute the gradient magnitude of all attached cells,
- Step 2. Determine the marker, which is connected blobs of pixels inside each cell, based on the distant transform.
- Step 3. Combine the results of Step 1 and Step 2,
- Step 4. Compute the watershed transform of the result from Step 3.

Finally, to obtain whole cancer cells, we combine the image of attached multiple cells with the image of distinct single cells using a logical operator OR.

## 3.2 Results and discussion

### 3.2.1 Neural network

Figure 7(a) shows an original RGB image of cancer cells. Most of cancer cells are located separately, but some of them attach with their neighborhood. We also see an uneven distribution of color and intensity in the background region surrounding cancer cells. It is considered as histological noise. The R-, G- and B-components of pixel values from background, P, and N reference regions (Boxes **B1**, **B2**, and **B3**) shown in Figure 7(a) are used as the input vectors for training neural network. After the backpropagation neural network is trained, network responses of pixel values throughout the image are calculated.

Figure 7(b) shows the output image from neural network. Pixels from background, P, and N regions are shown in white, gray, and black colors, respectively. Results demonstrate that the neural network can classify color pixels very well. That is, pixel values are appropriately categorized into a connected region corresponding to the cancer cells shown in Figure 7(a). Additionally, the results from neural network show that color contents of pixel values for each cancer cell are unevenly distributed. For example, we can notice the mixture appearance of positive and negative colors of the cell situating at the left hand side of box **B1**. This is in agreement with the original RGB image.

### 3.2.2 Morphology operations

Figure 8(a) shows the binary image after thresholding and morphological opening for spike noise removal. Subsequently, hole filling is performed using the algorithm based on morphological reconstruction. Each object in the image is classified into three categories

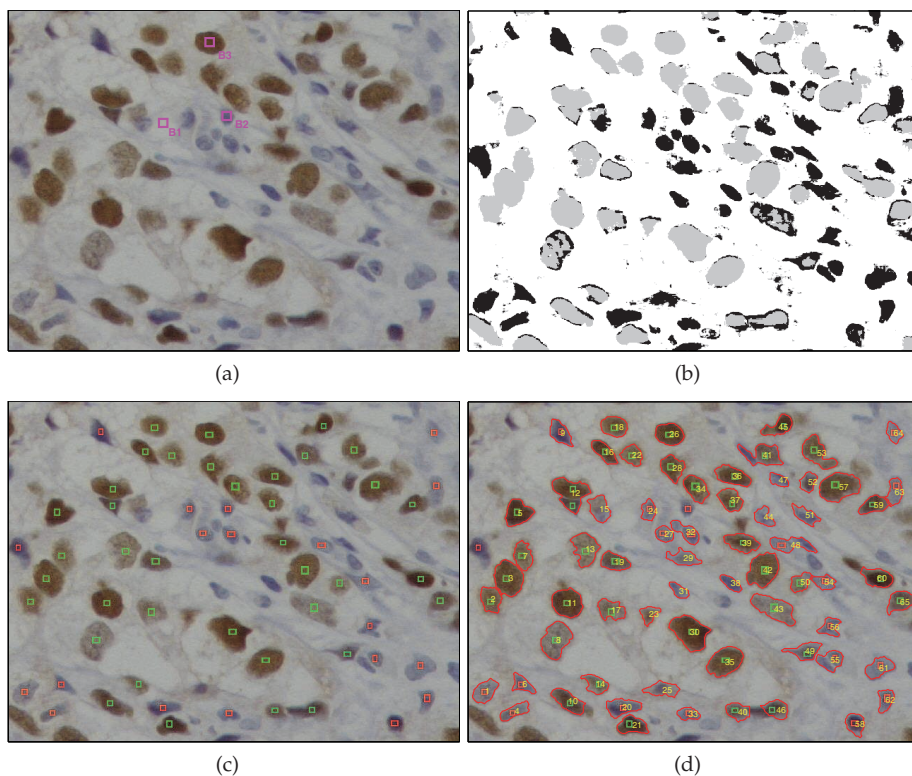


Fig. 7. (a) Original RGB image of stained cancer cells. Boxes **B1**, **B2**, and **B3** indicate reference regions from background, P, and N, respectively. Pixel values in the box regions are used for training neural network. (b) Output image from neural network. Pixels from background, P, and N regions are shown in white, gray, and black colors, respectively. (c) Positive and negative cancer cells marked by an expert. (d) Segmented image superimposed on the original RGB image compared with the result from the expert.

according to its size. Figure 8(b) shows the binary image of distinct single cells after size consideration. One can notice that some cells do not have the complete shape due to an uneven distribution of color and intensity. To compensate for this shortcoming, we perform mathematical closing and show result in Figure 8(c). It is shown that distinct single cells with perfect round shape are obtained after mathematical closing.

Figure 8(d) shows the binary image of attached multiple cells after size consideration. The size of cancer cells in this image is large in terms of area compared with the size of distinct single cell. Figure 8(e) shows the binary image of attached multiple cells after morphological opening. One can see that a slightly attached multiple cell is separated into two distinct single cells as we expect. In addition, the shape of each cell is smoothed and simplified, which make it appropriate for marker-controlled watershed processing.

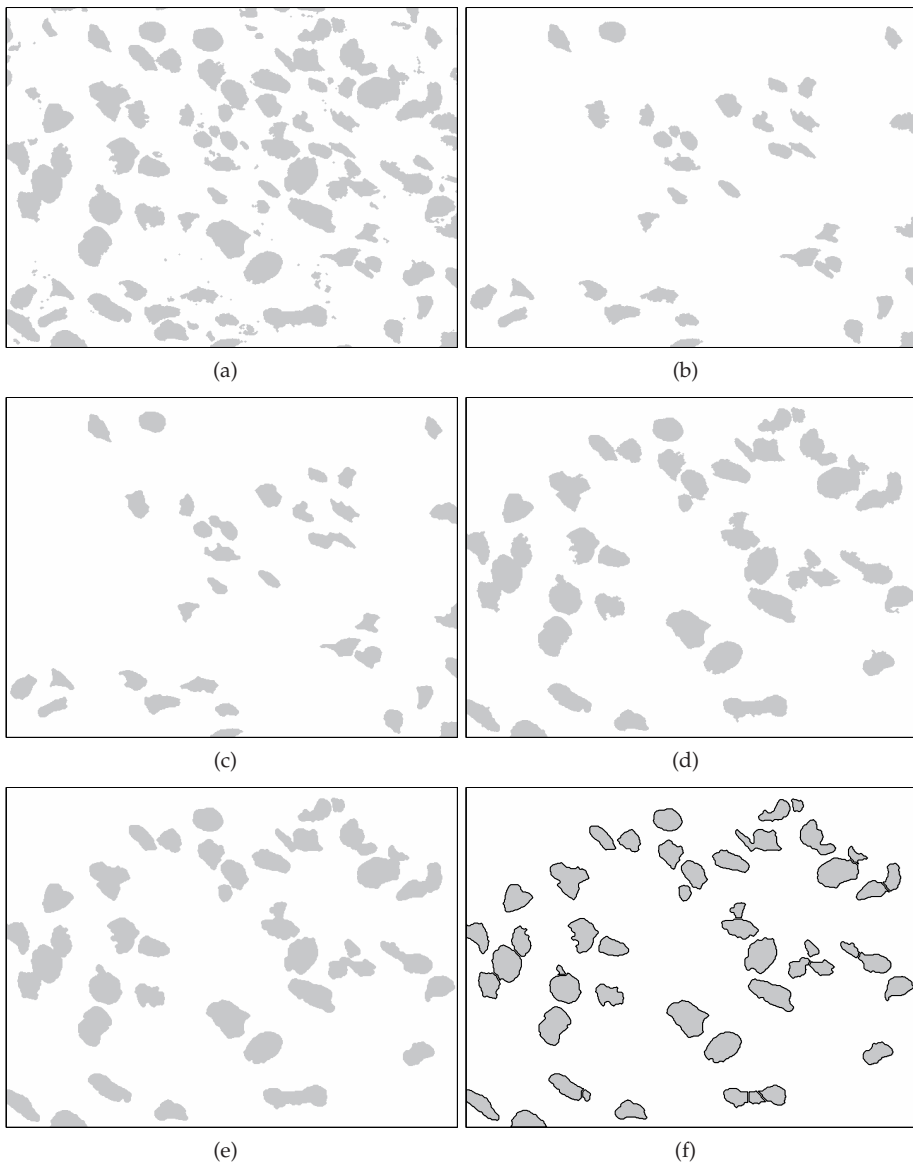


Fig. 8. (a) Binary image after thresholding and morphological operations of the image from neural network. (b) Binary image of distinct single cells before morphological closing. (c) Binary image of distinct single cells after morphological closing. (d) Binary image of attached multiple cells before morphological opening. (e) Binary image of attached multiple cells after morphological opening. (f) Binary image of attached multiple cells after marker-controlled watershed.

### 3.2.3 Marker-controlled watershed

Figure 8(f) shows the segmented image after the application of marker-controlled watershed. A more accurate segmented result is achieved. That is, the attached multiple cells are appropriately separated into the distinct single cells. Figure 7(c) shows positive and negative cancer cells marked by an expert. While the positive nuclei marked by the expert are shown with green rectangular windows, the negative nuclei are shown with red rectangular windows. To compare the results from the proposed algorithm with those from the expert, we superimpose the segmented image on the original RGB image and demonstrate the result in Figure 7(d). The number of segmented cancer cells is 65. It can be clearly seen that the perception of segmented cancer cells is in agreement with their original visualization.

## 4. Conclusions

We present two segmentation methods for nuclear stained breast cancer cell counting. While one is based on a local adaptive thresholding, another is based on the separation of pixel color using neural network. The segmenting results of cancer cells from the background are used as a preliminary step before extracting cell features and classifying cell types. The excellent segmentation results from the proposed algorithm are demonstrated with microscopic images under various histological noise conditions. Quantitative evaluations of the neural network approach compared with the counting results from the expert provide similar agreement to image visualization (Phukpattaranont et al., 2009). In other words, sensitivity and positive predictive value of cell segmentation are 88% and 82%, respectively. Moreover, sensitivity, positive predictive value, specificity, and negative predictive value of color classification are 94%, 99%, 91%, and 78%, respectively. However, to make the method automatic and gain higher accuracy, the following issues have to be addressed:

- The algorithm that can compensate for the selection of training data by a specialist need to be incorporated.
- The more sophisticated algorithm for the separation of overlapping cells is needed.
- More features are necessary such as the texture of cell and its neighborhood and the shape of each cell.

They are ongoing research. Results will be reported in the near future.

## 5. Acknowledgments

This research was supported by Department of Electrical Engineering, Department of Computer Engineering, Faculty of Engineering through contract no. ENG-51-2-7-02-0018-S and Department of Pathology, Faculty of Medicine, Prince of Songkla University.

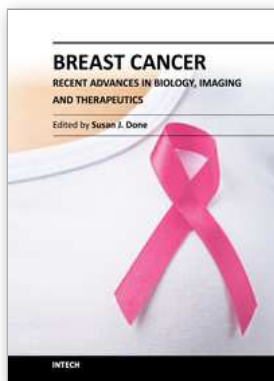
## 6. References

- Adams, R. & Bischof, L. (1994). Seeded region growing, *IEEE Trans. Pattern Anal. Mach. Intell.* 16: 641–647.
- Beveridge, J. R. & et al. (1989). Segmenting images using localized histograms and region merging, *International Journal of Computer Vision* 2: 311–347.
- Fang, B., Hsu, W. & Lee, M. L. (2003). On the accurate counting of tumor cells, *IEEE Trans. Nanobiosci.* 2(2): 94–103.
- Gelenbe, E., Feng, Y. & Krishnan, K. R. R. (1996). Neural network methods for volumetric magnetic resonance imaging of the human brain, *Proc. IEEE* 84(10): 1488–1496.

- Gonzalez, R. C. & Woods, R. E. (2002). *Digital Image Processing*, Prentice-Hall, Inc.
- Hagan, M. T., Demuth, H. B. & Beale, M. H. (1996). *Neural Network Design*, PWS Publishing Company, Boston, MA.
- Haralik, R. M. & Shapiro, L. G. (1985). Image segmentation techniques, *Comput. Vis. Graph. Image Proc.* 29: 100–132.
- Heermann, P. D. & Khazenie, N. (1992). Classification of multispectral remote sensing data using a back-propagation neural network, *IEEE Trans. Geosci. Remote Sensing* 30(1): 81–88.
- Li, H. D., Kallergi, M., Clarke, L. P., Jain, V. K. & Clark, R. A. (1995). Markov random field for tumor detection in digital mammography, *IEEE Trans. Med. Imag.* 14(3): 565–576.
- O’Gorman, L., Sanderson, A. C. & Jr., K. P. (1985). A system for automated liver tissue image analysis: Methods and results, *IEEE Trans. Biomed. Eng.* 32(9): 696–706.
- Otsu, N. (1979). A threshold selection method from gray-level histograms, *IEEE Trans. Syst., Man, Cybern.* 9(1): 62–66.
- Park, J. & Keller, J. (1997). Fuzzy patch label relaxation in bone marrow cell segmentation, *Proc. IEEE Int. Conf. Syst. Man. Cybern.*, Vol. 2, pp. 1133–1138.
- Perona, P. & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Mach. Intell.* 12(7): 629–639.
- Petushi, S., Katsinis, C., Coward, C., Garcia, F. & Tozeren, A. (2004). Automated identification of microstructures on histology slides, *IEEE International Symposium on Biomedical Imaging: Macro to Nano*, Vol. 1, pp. 424–427.
- Phukpattarant, P. & Boonyaphiphat, P. (2006). Automatic classification of cancer cells in microscopic images: Preliminary results, *The 2006 ITC-CSCC International Conference*, Vol. 1, pp. 113–116.
- Phukpattarant, P. & Boonyaphiphat, P. (2007). Color based segmentation of nuclear stained breast cancer cell images, *ECTI Transactions on Electrical Eng. Electronics and Communications* 5(2): 158–164.
- Phukpattarant, P., Limsiroratana, S. & Boonyaphiphat, P. (2009). Computer-aided system for microscopic images: Application to breast cancer nuclei counting, *International Journal of Applied Biomedical Engineering* 2(1): 69–74.
- Reddick, W. E., Glass, G. O., Cook, E. N., Elkin, T. D. & Deaton, R. J. (1997). Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural network, *IEEE Trans. Med. Imag.* 16: 911–918.
- Refai, H., Li, L. & Teague, T. K. (2003). Automatic count of hepatocytes in microscopic images, *International Conference on Image Processing*, Vol. 2, pp. 1101–1104.
- Theera-Umpon, N. & Gader, P. D. (2000). Counting white blood cells using morphological granulometries, *Journal of Electronic Imaging* 9(2): 170–177.
- Thiran, J. P. & Macq, B. (1996). Morphological feature extraction for the classification of digital images of cancerous tissues, *IEEE Trans. Biomed. Eng.* 43(10): 1011–1020.
- Trussell, H. J., Saber, E. & Vrhel, M. (2005). Color image processing (basics and special issue overview), *IEEE Signal Process. Mag.* 22(1): 14–22.
- Vincent, L. (1993). Morphological grayscale reconstruction in image analysis: applications and efficient algorithms, *IEEE Trans. Image Process.* 2(2): 176–201.
- Wani, M. A. & Batchelor, B. G. (1994). Edged-region-based segmentation of range images, *IEEE Trans. Pattern Anal. Mach. Intell.* 16: 314–319–647.
- Wu, H. S. & Barba, J. (1994). An algorithm for noisy cell contour extraction via area merging, *J. Imag. Sci. Technol.* 38: 604–607.

- Wu, K., Gauthier, D. & Levine, M. D. (1995). Live cell image segmentation, *IEEE Trans. Biomed. Eng.* 42(1): 1–12.
- Zhao, X. & Ong, S. H. (1998). Adaptive local thresholding with fuzzy-validity-guided spatial partitioning, *International Conference on Pattern Recognition*, Vol. 2, pp. 988–990.





## **Breast Cancer - Recent Advances in Biology, Imaging and Therapeutics**

Edited by Dr. Susan Done

ISBN 978-953-307-730-7

Hard cover, 428 pages

**Publisher** InTech

**Published online** 14, December, 2011

**Published in print edition** December, 2011

In recent years it has become clear that breast cancer is not a single disease but rather that the term encompasses a number of molecularly distinct tumors arising from the epithelial cells of the breast. There is an urgent need to better understand these distinct subtypes and develop tailored diagnostic approaches and treatments appropriate to each. This book considers breast cancer from many novel and exciting perspectives. New insights into the basic biology of breast cancer are discussed together with high throughput approaches to molecular profiling. Innovative strategies for diagnosis and imaging are presented as well as emerging perspectives on breast cancer treatment. Each of the topics in this volume is addressed by respected experts in their fields and it is hoped that readers will be stimulated and challenged by the contents.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Pornchai Phukpattaranont, Somchai Limsiroratana, Kanita Kayasut and Pleumjit Boonyaphiphat (2011). Computer Aided System for Nuclear Stained Breast Cancer Cell Counting, *Breast Cancer - Recent Advances in Biology, Imaging and Therapeutics*, Dr. Susan Done (Ed.), ISBN: 978-953-307-730-7, InTech, Available from: <http://www.intechopen.com/books/breast-cancer-recent-advances-in-biology-imaging-and-therapeutics/computer-aided-system-for-nuclear-stained-breast-cancer-cell-counting>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.