Automatic White Blood Cell Classification Using Biased-Output Neural Networks with Morphological Features

Nipon Theera-Umpon
Department of Electrical Engineering,
Faculty of Engineering, Chiang Mai University,
Chiang Mai 50200 Thailand
E-mail: DrNipon@chiangmai.ac.th

Abstract

Numbers of white blood cells in different classes help doctors to diagnose patients. A new set of features based on the mathematical morphology in the white blood cell classification problem is proposed in this paper. The proposed features are the maximum value of a pattern spectrum, the location where the maximum value of a pattern spectrum occurs, the first and second granulometric moments. We also propose a method to unbias neural networks by biasing the desired output using *a priori* information of the number of samples in each class. Regular artificial neural networks and the biased-output neural networks are applied in the experiments using the five-fold cross validation as the testing method. The results show the good performances of classifiers using our biased-output neural networks and our proposed morphology-based features.

Keywords: Automatic white blood cell classification, Feature extraction, Biased-output neural network, Morphological features, Granulometric moments

1. Introduction

There are two methods to count white blood cells in bone marrow namely the total countand the differential counts. The total count is the total number of white blood cells without any classification. In contrast, the differential counts are the counts of different cell classes in bone marrow. One important step to achieve the white blood cell differential counting is to classify each white blood cell in bone marrow.

The differential counting provides invaluable information to doctors in diagnosis of diseases such as leukemia or cancers. The traditional method for an expert to achieve the differential counting is to look through a microscope to select an area of interest in a bone marrow slide, detect a white blood cell, classify it based on his knowledge, and increase the count of the corresponding cell class. Conducting all of these processes manually would require a trained expert. Moreover, it is a very tedious job.

White blood cells in bone marrow are classified according to their maturation stages. When a white blood cell becomes older, its size, nuclei shape and many other features change. White blood cells in the myelocytic series can be classified into six classes, i.e., myeloblast, promyelocyte, myelocyte, metamyelocyte, band, and polymorphonuclear (PMN) ordered from the youngest to the oldest cells [1,2]. Figure 1 shows samples of white blood cells in this series.

Although there are some commercial automatic systems available for counting white blood cells in peripheral blood, there is no such system for bone marrow. There have been several attempts proposed to solve the problem. Most methods follow the traditional manual maneuver, i.e., to detect a cell, extract its features, classify the cell, and then update the count [3–7]. Some are based on neural

A part of the content has appeared and been discussed at the 25th Electrical Engineering Conference (EECON-25), on 21-22 Nov.2002 held at Prince of Songkla University, Hat Yai, Songkla, Thailand.

networks with features such as area of cell, area of nuclei, ratio of area of nuclei to cytoplasm, Fourier descriptors of nuclei, some textural features, etc [7]. Some are developed under the mixing theories of the mathematical morphology [8–10]. Some develop new training algorithm for neural networks in order to count numbers of different cell classes, without classification [11,12].

In this paper, we apply artificial neural networks to the white blood cell classification for single-cell images under the assumption that the cell segmentation is available. We propose new features of a white blood cell based on morphological granulometries. We also apply the bias to the desired outputs of neural networks based on the prior information of number of cells in each class in a training set.

This paper is organized as follows. Section 2 introduces the mathematical morphology, neural networks, and how to extract the proposed features. The data set is described in section 3. A description of the experiments and their results are shown in section 4. Section 5 concludes this paper.

2. Methodology

In this research, artificial neural networks are used as our classifiers in the six-class problem. The input features are mainly extracted from pattern spectra of nucleus. To be more specific, there are six features – two are area-based, the remaining four are morphology-based.

2.1 Mathematical Morphology

Mathematical morphology is a branch of nonlinear image processing and analysis. It was first introduced by Matheron in the context of random sets [13,14]. Morphological methods are used in many ways in image processing, for example, enhancement, segmentation, restoration, edge detection, texture analysis, shape analysis, etc. [15,16]. It is also applied to several research areas, such as, medical imaging, remote sensing, military applications, etc.

(a) (b) (c)

(d) (e) (f)

Fig. 1 Cell samples in the myelocytic series: (a) Myeloblast; (b) Promyelocyte; (c) Myelocyte; (d) Metamyelocyte; (e) Band; and (f) PMN.

2.1.1 Morphological Operations

Morphological operations are non-linear, translation invariant transformations. This paper describes binary morphological operations only. Binary images can be considered as functions on two-dimensional grids with values of 0 or 1 or, equivalently, as characteristic functions of subsets of the two-dimensional plane. The concept of structuring an element is fundamental in morphology; it is the analogue of a convolution mask in linear image processing. The basic morphological operations involving an image S and a structuring element E are

$$\textit{erosion}: S \ominus E = \cap \{S - e: e \in E\}$$
$$\textit{dilation}: S \oplus E = \cup \{E + s: s \in S\},$$

where $\cap$ and $\cup$ denote the set intersection and union, respectively. $A + x$ denotes the translation of a set A by a point x, i.e.

$$A + x = \{a + x: a \in A\}.$$

The closing and opening operations, derived from the erosion and dilation, are defined by

$$\textit{closing}: S \bullet E = (S \oplus (-E)) \ominus (-E)$$
$$\textit{opening}: S \circ E = (S \ominus E) \oplus E$$

where $-E = \{-e: e \in E\}$ denotes the 180° rotation of E about the origin.

The examples of an image S, a structuring element E, and outputs of the erosion, dilation, closing and opening operators are shown in Figure 2.

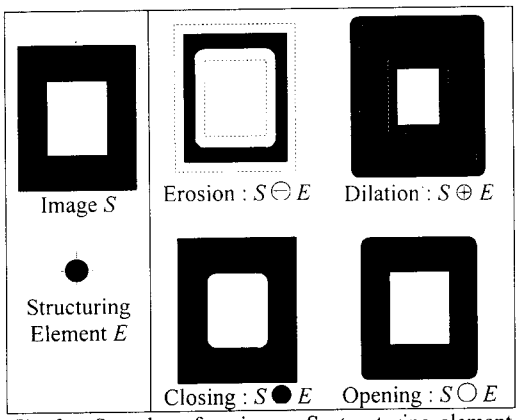

Fig. 2 Samples of an image S, structuring element E, and outputs of the erosion, dilation, closing and opening operators.

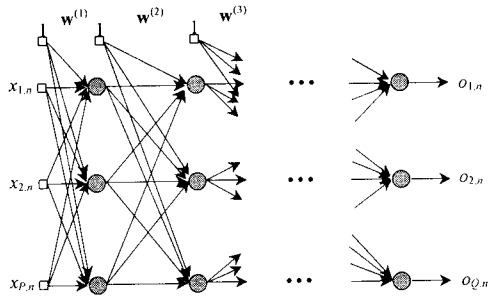

Fig. 3 Feed forward neural network.

2.1.2 Pattern Spectrum

The opening operation is successively applied to an image, increase the size of structuring element in order to diminish the image. Let $\Omega(t)$ be area of $S \bigcirc tE$ where t is a real number and $\Omega(0)$ is area of S. $\Omega(t)$ is called a *size distribution*. The normalized size distribution $\Phi(t) = 1 - \Omega(t)/\Omega(0)$, and $d\Phi(t)/dt$ are called *granulometric size distribution* or *pattern spectrum* of S. This issue will be discussed further in section 2.3.

2.2 Artificial Neural Networks

Artificial neural networks are well-described in literature [17], they are only briefly describe them in this paper. Neural networks have been applied to several areas of research such as military, medicine, business, etc. It can be considered as a universal approximator. The typical structure of a feed forward neural network is depicted in Figure 3.

The goal is to find the best set of weights (**w**) so that the outputs $o_{j,n}$ are as close to the desired outputs $d_{j,n}$ as possible for a given input pattern $x_{i,n}$, $i = 1, \ldots, P$ and $j = 1, \ldots, Q$. P and Q are the number of input features and the number of classes, respectively. There are many approaches used to find the best set of weights. However, in this research, the Levenberg-Marquardt (LM) algorithm was chosen because it provides faster convergence [18].

2.3 Feature Extraction

This study focuses its feature extraction on the morphology-based features. Hence, their derivations are introduced here. For a random set S, $\Omega(t)$ is a random function. The normalized size distribution $\Phi(t) = 1 - \Omega(t)/\Omega(0)$, the so-called *pattern spectrum* of S, is a probability distribution function. Its moments, $\mu^{(1)}(S)$, $\mu^{(2)}(S)$,..., are therefore random variables namely *granulometric moments*. In this research, consider nuclei as an object of interest. A pattern spectrum of each cell's nuclei is calculated. Then, the first and second granulometric moments of the pattern spectrum are calculated to achieve the features.

To form an input feature vector to a neural network, six features are extracted from each cell, i.e.,

- the area of cell,
- the nuclei-to-cytoplasm ratio,
- the maximum value of a pattern spectrum,
- the location where the maximum value of a pattern spectrum occurs,
- the first granulometric moments and
- the second granulometric moments.

A small digital disc is selected as the structuring element in there experiments. The structuring element is shown in Figure 4. By applying this structuring element, sample pattern spectra of nucleus for all six cell classes are shown in Figure 5.

0	1	1	0
1	1	1	1
1	1	1	1
0	1	1	0

Fig. 4 Structuring element used in the experiments.

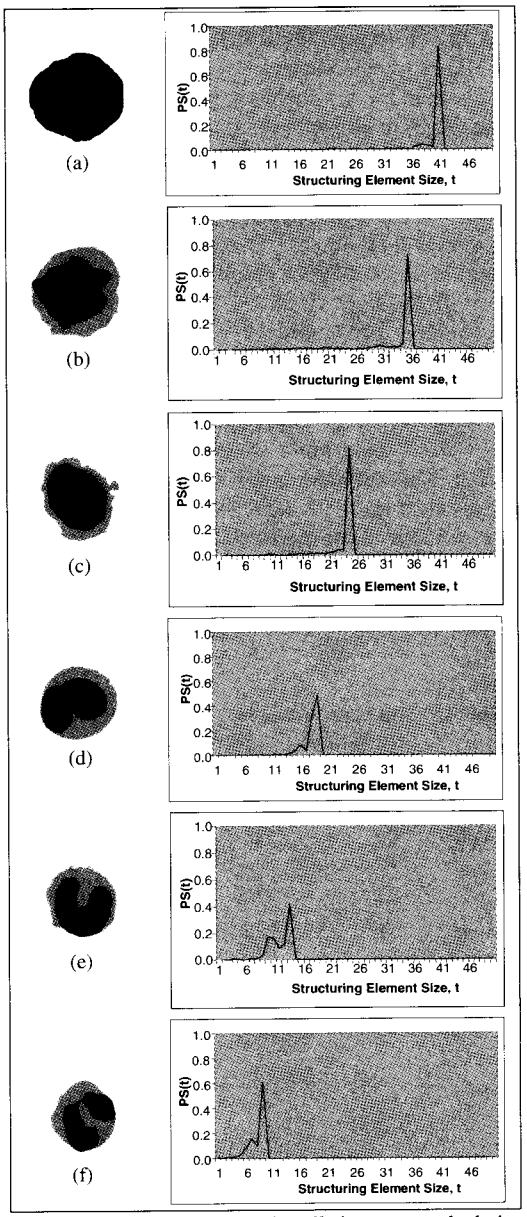

Fig. 5 Sample segmented cell images and their pattern spectra of all six cell classes corresponding to the structuring element in Figure 4.

3. Data Description

In the experiments, bone marrow images collected at the University of Missouri Ellis-Fischel Cancer Center is used. Each white blood cell image is cropped manually to form a single-cell image. Then, a single-cell image is segmented manually into nucleus, cytoplasm, and background regions. The images were manually classified by Dr. C. William Caldwell, Professor of Pathology and Director of the Pathology Laboratory at the Ellis-Fischel Cancer Center. The data set consists of six classes of white blood cells — myeloblast, promyelocyte, myelocyte, metamyelocyte, band, and PMN. There are 20, 9, 139, 33, 45, and 185 hand-segmented images for all six cell classes, respectively. Each hand-segmented image is composed of three regions — nucleus, cytoplasm, and background — with gray level = 0, 176, and 255, respectively. Samples of cells and their corresponding segmented images of all six cell classes are shown in Figure 6.

4. Experimental Results and Discussion

As mentioned in section 3 that the data set is not divided into training and test sets. However, training and test sets are required to train and test a neural network to evaluate its performance and generalization. Cross validation is a standard method to solve this problem.

The following will briefly describe the cross validation. The experiments are performed using the five-fold cross validation. That means data points in each class are divided into five groups. Each group contains the same or comparative amount of data points. Then, the data in the first group of each class to form a test set are selected.

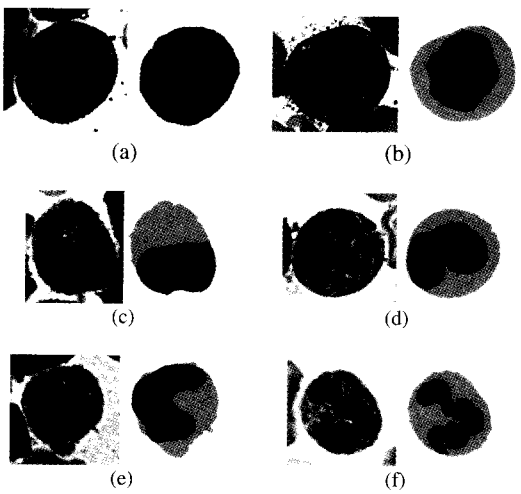

Fig. 6 Sample gray scale and corresponding hand-segmented images of white blood cell:
(a) Myeloblast, (b) Promyelocyte, (c) Myelocyte, (d) Metamyelocyte, (e) Band, and (f) PMN.

The remaining is considered a training set. Therefore, the training set contains {2,3,4,5} and the test set contains {1}, where the numbers in the brackets denote the group numbers. A neural network is trained and tested using these two sets and the network's performance is evaluated. That is the end of the first fold. In the second fold, the training set contains {1,3,4,5} and the test set contains {2}. Then keep doing this until completion of all five folds. Therefore, in each fold, use about 80% of data points as the training set and about 20% of those as the test set. After finishing validation of all five folds, each data point (or each cell image) is once used as test data.

The results are presented as confusion matrices of total classified outputs of all five folds, i.e., the sum of all five confusion matrices in training and the sum of all five confusion matrices in testing. To save the space, the confusion matrix of each fold is not shown here. It should be noted that the number of cells in each class in the total confusion matrix on test sets is exactly the same as the number of cells in that class in the data set because each cell image is used in testing only once. In contrast, the number of cells in each class in the total confusion matrix on training sets is quadruple of the number of cells in that class in the data set because each cell image is used four times in training.

4.1 Results of regular neural networks

The number of hidden neurons in a neural network is one parameter to choose. In these experiments, a feed forward neural network with one hidden layer consisting of ten hidden neurons is chosen. The number of neurons of ten is large enough to be used to approximate a function of six inputs and six outputs. This provides $(6\times10) + (10\times6) = 120$ weight parameters, excluding $6+10 = 16$ more weight parameters of biases to neurons in the input and hidden layers. Ten is also not too large so that we will not lose the generalization. The desired output is set to 0.9 for the output neuron corresponding to a given class, and 0.1 for the other output neurons. The training using the Levenberg-Marquardt (LM) algorithm would stop when the maximum epochs reaches 100 or the mean square error is less than 10^{-6}

The total confusion matrices on the training and test sets are shown in Tables 1 and 2, respectively. The classification rates of 90.66% and 72.16% are achieved in training and testing, respectively. However, when considering the numbers in the confusion matrices, they show a problem of neural networks, i.e., there are biases to the classes those have larger number of samples (myelocyte and PMN, in this case.) In other words, they consider these classes more important than others. Given a cell image, the classifiers would more likely decide that the cell come from these classes. This is obvious because it is preferable to minimize the mean square error in training of neural networks. To increase the chance of correct classification, the networks would give more probability to classes those have more input samples (or numbers of cells.) This problem leads to the loss of generalization of a neural network. It can be seen in Table 2 that the testing cells are more likely to be either myelocyte or PMN according to the classifiers' decisions.

Table 1 Total confusion matrix on training sets (neural networks, no *a priori* information).

Alg → Actual↓	Blast	Pro	Myelo	Meta	Band	PMN
Blast	80	0	0	0	0	0
Pro	2	26	7	1	0	0
Myelo	0	2	535	8	1	10
Meta	0	5	45	68	5	9
Band	0	0	1	0	140	39
PMN	0	0	12	6	8	714
Classification rate (Train) = 90.66 %						

Table 2 Total confusion matrix on test sets (neural networks, no *a priori* information).

Alg → Actual↓	Blast	Pro	Myelo	Meta	Band	PMN
Blast	16	0	1	0	3	0
Pro	1	1	7	0	0	0
Myelo	0	6	116	8	0	9
Meta	0	0	17	9	3	4
Band	0	0	0	4	17	24
PMN	0	0	7	6	20	152
Classification rate (Test) = 72.16%						

Table 3 Total confusion matrix on training sets (biased-output neural networks).

Alg → Actual↓	Blast	Pro	Myelo	Meta	Band	PMN
Blast	80	0	0	0	0	0
Pro	0	27	9	0	0	0
Myelo	0	15	508	16	2	15
Meta	0	0	50	72	7	3
Band	0	0	0	4	139	37
PMN	0	0	22	19	63	636
Classification rate (Train) = 84.80 %						

Table 4 Total confusion matrix on test sets (biased-output neural networks).

Alg → Actual↓	Blast	Pro	Myelo	Meta	Band	PMN
Blast	19	0	0	0	0	1
Pro	1	3	4	1	0	0
Myelo	2	4	114	14	1	4
Meta	0	0	17	10	4	2
Band	0	0	0	3	28	14
PMN	0	0	4	9	20	152
Classification rate (Test) = 75.64 %						

4.2 Results of biased-output neural networks

In this experiment there is a wide variety of numbers of cells in classes. To remedy the classifier's bias, one possible solution is to have the same (or comparative) number of data samples in all six classes to give the same *a priori* probability to all classes. However, this approach is not feasible in this case because some classes have too small number of samples (9 promyelocytes, 20 myeloblasts.) Nine cells could be selected from each of six cell classes, but that will decrease the number of cells in the experiments dramatically.

Hence, the attempt was made to achieve unbiased classifier by biasing the desired output using *a priori* information of the number of samples in each class. We set

$$d_j = 1 - \frac{n_j}{\sum_{j=1}^{6} n_j}, \quad j = 1, \ldots, 6$$

as the desired output of the output neuron j, the desired output of the other output neurons is set to 0.1. Therefore, the output neuron corresponding to the correct class will produce small value if that class has large number of samples. A neural network trained by using this approach is called "biased-output neural network (BONN)".

The confusion matrices of the classifiers trained by using this approach are shown in Tables 3 and 4. The classification rates of 84.80% and 75.64% are achieved in training and testing, respectively. The decreasing of classification rate in training is not surprising. It can be seen from Table 3 that the correct classification of cells from two "big" classes – myelocyte and PMN – decrease compared to those in Table 1 because we decrease their output levels. Even though the classification rate on the training sets decreases, the rate on the test sets increases. To evaluate the generalization of the networks, consider the testing results because the test data are blind to the classifiers. Let us consider the numbers of test cells with correct classification using the regular neural network in Table 2. The numbers are 16, 1, 116, 9, 17, and 152 for the six cell classes, respectively. The numbers of test cells with correct classification using the biased-output neural network in Table 4 are 19, 3, 114, 10, 28, and 152 for the six cell classes, respectively. It can be seen that the number of myelocyte cells with correct classification decreases only by two while the number of PMN cells with correct classification remains the same. The results suggest that the bias does not hurt the "big" classes. Other four classes have better numbers of correct classification. These numbers lead to the better classification rate of the biased-output neural network. That means there are more generalized classifiers by introducing the bias to the desired output of a neural network.

Furthermore, the five-fold cross validation on eight more experiments are applied: four on the regular neural networks and the other four on the biased-output neural networks to verify whether the random initial values of the networks' parameters would affect the results. Because five networks are trained and tested in each experiment, 40 more networks are trained and tested. The total confusion matrices from these experiments are not shown here to save the space. The classification rates of all eight experiments and the rates of the other two experiments described earlier are shown in Table 5.

Table 5 Classification rates on training and test sets using regular and biased-output neural networks in ten experiments of five-fold cross validation.

Net-works	Data Sets	Experiment Number				
		1	2	3	4	5
Regular NN	Train	90.66	89.73	89.97	89.27	91.94
	Test	72.16	71.00	72.85	71.69	72.39
BONN	Train	84.80	83.76	83.76	85.21	84.11
	Test	75.64	75.64	74.25	74.71	75.17

By the extensive experiments, Table 5 emphasizes the validity of the results discussed earlier. The classification rates on the test sets of all five experiments using the biased-output neural networks are better than all five of those using the regular neural networks.

5. Conclusion

The mathematical morphology-based features – the maximum value of a pattern spectrum, the location where the maximum value of a pattern spectrum occurs, the first and second granulometric moments – have been proposed in this research. It has also been demonstrated that these features are useful in the automatic white blood cell classification. The classification rate of about 75 % on test sets is similar to those achieved in [7] with different features and different data set of hand-segmented cell images. However, the results in [7] were carried from ten-fold cross validation those used more training data than five-fold cross validation applied in this research. Moreover, the data set applied in [7] contains six cell classes whose numbers of cells vary much less than the data set that was applied here.

A neural network namely the biased-output neural network has been proposed in this research based on a method to unbias the classifiers. The bias to the desired output using *a priori* information of the number of samples in each class was applied and experimented. The testing results to evaluate the generalization of the neural networks has been analyzed. The biased-output neural network yields better classification performance than the regular neural network has been analyzed. The bias introduced to the desired output does not hurt the classes with large numbers of data points. In the meanwhile it yields better classification in the classes with small numbers of data points. Therefore, the biased-output approach increases the classification rate on the test sets. That means it has provided better generalization of the classifier.

The future work is to incorporate the automatic cell segmentation to this system because the features that have been used in the experiments heavily rely on the hand-segmented images.

6. Acknowledgments

The author is grateful to Dr. C. William Caldwell of Ellis-Fishel Cancer Center, University of Missouri, for providing the data and the ground truth. We acknowledge the contribution of Dr. James Keller and Dr. Paul Gader through many technical discussions. Finally, the author would like to thank Chiang Mai University and Faculty of Engineering, Chiang Mai University for supporting this research.

References

[1] Diggs, L. W., Sturm D., and Bell A., *The Morphology of Human Blood Cells*, Abbott Laboratories, Abbott Park, 1985.

[2] V. Minnich, *Immature Cells in the Granulocytic, Monocytic, and Lymphocytic Series*, American Society of Clinical Pathologists Press, Chicago, 1982.

[3] M. Beksaç, M. S. Beksaç, V. B. Tipi, H. A. Duru, M. U. Karakas, and A. Nur Çakar, An Artificial Intelligent Diagnostic System on Differential Recognition of Hematopoietic Cells From Microscopic Images, *Cytometry*, Vol. 30, pp. 145-150, 1997.

[4] H. Harms, H. Aus, M. Haucke, and U. Gunzer, Segmentation of Stained Blood Cell Images Measured at High Scanning Density With High Magnification and High Numerical Aperture Optics, *Cytometry*, Vol. 7, pp. 522-531, 1986.

[5] J. Park, and J. Keller, Fuzzy Patch Label Relaxation in Bone Marrow Cell Segmentation, *IEEE Int. Conf. Syst., Man,Cybern*, pp. 1133-1138, Orlando, FL, 1997.

[6] S. S. S. Poon, R. K. Ward, and B. Palcic, Automated Image Detection and Segmentation in Blood Smears, *Cytometry*, Vol. 13, pp. 766-774, 1992.

[7] S. Sohn, *Bone Marrow White Blood Cell Classification*, Master's Project, University of Missouri-Columbia, 1999.

[8] N. Theera-Umpon, *Morphological Granulometric Estimation with Random Primitives and Applications to Blood Cell Count-*

ing, Doctoral Dissertation, University of Missouri-Columbia, 2000.

[9] N. Theera-Umpon and P. D. Gader, Counting White Blood Cells Using Morphological Granulometries, *Journal of Electronic Imaging*, Vol. 9, No. 2, pp. 170-177, 2000.

[10] N. Theera-Umpon, E. R. Dougherty, and P. D. Gader, Non-Homothetic Granulometric Mixing Theory with Application to Blood Cell Counting, *Pattern Recognition*, Vol. 34, No. 12, pp. 2547-2560, 2001.

[11] N. Theera-Umpon and P. D. Gader, Training Neural Networks to Count White Blood Cells via a Minimum Counting Error Objective Function, *15th Intl. Conf. on Pattern Recognition*, pp. 299-302, Barcelona, Spain, September 2000.

[12] N. Theera-Umpon, and P. D. Gader, System Level Training of Neural Networks for Counting White Blood Cells, *IEEE Trans. Systems, Man, and Cybernetics Part C: Applications and Reviews*, Vol. 32, No. 1, pp. 48-53, 2002.

[13] G. Matheron, *Random Sets and Integral Geometry*, Wiley, New York, 1975.

[14] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1983.

[15] E. R. Dougherty, *An Introduction to Morphological Image Processing*, SPIE Press, Bellingham, Washington, 1992.

[16] E. R. Dougherty, *Random Processes for Image and Signal Processing*, SPIE Press, Bellingham, Washington, and IEEE Press, New York, 1999.

[17] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd Ed., Prentice-Hall, Upper Saddle River, New Jersey, 1999.

[18] H. Demuth and M. Beale, *Neural Network Toolbox: For Use with MATLAB®*, Mathworks, Natick, 1998.