View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

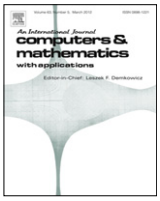
provided by Elsevier - Publisher Connector

Computers and Mathematics with Applications 64 (2012) 1153–1162

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

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



Support vector machine for breast MR image classification

Chien-Shun Lo [a], Chuin-Mu Wang [b,*]

[a] Department of Multimedia Design, National Formosa University, 64, Wunhua Rd, Huwei Township, Yunlin, 63201, Taiwan

[b] Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung, 411, Taiwan

ARTICLE INFO

Keywords:
SVM
C-means
Breast MR image
Classification

ABSTRACT

MR images have been used extensively in clinical trials in recent years because they are harmless to the human body and can obtain detailed information by scanning the same slice with various frequencies and parameters. In this paper, we want to detect the breast tissues within multi-spectral MR images. In the image classification, we apply a support vector machine (SVM) to breast multi-spectral magnetic resonance images to classify the tissues of the breast. In order to verify the feasibility and efficiency of this method, evaluations using classification rate and likelihood ratios are adopted based on manifold assessment and a series of experiments are conducted and compared with the commonly used C-means (CM) for performance evaluation. The results show that the SVM method is a promising and effective spectral technique for MR image classification.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Breast cancer most often occurs in 40- to 50-year-old women. There are a number of contributing factors such as: (1) a familial history of breast cancer, especially with breast cancer in her mother or sister; (2) the woman's first childbirth is after 30 years of age; (3) the woman ingests foods with high fat or calorie content; (4) the woman's chest has been subjected to radioactivity several times; and (5) the woman has had ovarian or endometrial cancer. If a woman conforms to any of the conditions above, she must take precautions and watch out for the development of breast cancer.

According to today's medical technology, if breast cancer is found at an early stage and the patient commences treatment, the patient's survival rate is relatively high. For instance, with breast cancer treated in its initial stage, the survival rate is 95%. The patients have to enhance their vigilance, because breast cancer is occurring earlier today than previously documented.

Although medical technology is progressing rapidly today, preventing the development of breast cancer remains an important goal. It is crucial to choose an instrument with simple and fast scanning which can find breast cancer accurately and surpass the deficiencies of self-examination. The current detecting instruments are as follows: Magnetic Resonance Imaging, Ultrasonography, Computer Tomography (CT), Digital Radiography, Immune Imaging, etc.

The primary merits of Magnetic Resonance Imaging (MRI) are its lack of invasiveness and lack of radioactivity which can harm the human body. It can provide abundant information on tissues, so MRI has become an important imaging instrument in medical clinical diagnosis in recent years. Additionally, MRI provides unparalleled capability of revealing soft tissue characterization as well as 3-D visualization. It produces a sequence of multiple spectral images of tissues with a variety of contrasts using three magnetic resonance parameters: spin–lattice (T1), spin–spin (T2), and dual echo–echo proton density (PD) [1]. Much research has been reported in recent years [2–11] and many approaches to biomedical image processing of the breasts have been proposed. In this paper, we present a "Support Vector Machine" (SVM) application to the multi-spectral breast MR image classification. The SVM is based on the idea of the hyper-plane classifier, and it tries

* Corresponding author.
E-mail address: cmwang@ncut.edu.tw (C.-M. Wang).

0898-1221/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2012.03.033

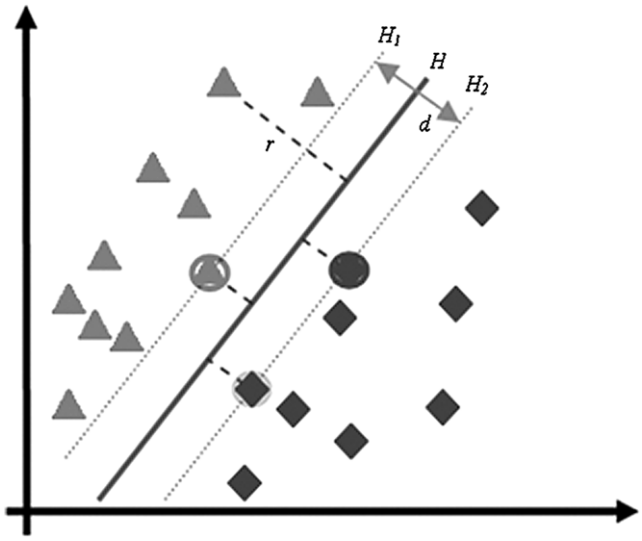


Fig. 1. Linearly separable data.

to look for the hyper-plane that maximizes the margin between two classes. The SVM has been widely utilized in other fields [12–16]. The SVM can classify the target objects from the breast image and the classification result of each tissue is shown as a binary image. In order to evaluate the performance of the SVM, a series of experiments are conducted where the commonly used C-means method is used for comparison.

The remainder of this paper is organized as follows: Section 2 presents the SVM approach; Section 3 briefly describes the CM method to be implemented in this paper; Section 4 illustrates a series of experiments to evaluate the effectiveness of the SVM method in classification performance and also compares the results to those produced by the CM method; and Section 5 is the conclusion.

2. Support vector machine

The Support Vector Machine (SVM), as a novel type of learning machine based on statistical learning theory [17,18], was proposed by Vapnik in 1995. It was developed on the basis of the Structural Risk Minimization (SRM) principle [19] of statistics. The SVM is based on the idea of hyper-plane classifiers, and it tries to look for the hyper-plane that maximizes the margin between two classes. The Linear SVM for linear separating case (as Fig. 1) is described as follows:

Suppose we have N training data $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_N, y_N)\}$ where $\mathbf{x}_i \in R^d$ and $y_i \in \{1, -1\}$. We want to establish a linear separating hyper-plane classifier as follows.

$$f(\mathbf{x}) = sign(\mathbf{w} \cdot \mathbf{x} + b) \tag{1}$$

where $sign() \in \{1, -1\}$, also called sgn or signum, -1 is for a negative number (i.e., one with a minus sign "−"), or $+1$ for a positive number (i.e., one with a plus sign "+"), $\mathbf{w}$ is a weight, $\mathbf{x}$ is the training data, and b is a bias of this hyper-plane. We would like to maximize the margin between two classes [18], so $\|\mathbf{w}\|$ should be minimized, subject to $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$. The optimization problem is as follows:

$$\underset{\mathbf{w},b}{\text{Min}} \frac{1}{2}\mathbf{w}^T\mathbf{w}, \quad \text{Subject to } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1,\ i = 1, \ldots, N \tag{2}$$

where min indicates to minimize the value of $\frac{1}{2}\mathbf{w}^T\mathbf{w}$,$\mathbf{x}_i \in R^d$ *and* $y_i \in \{1, -1\}$, $\mathbf{w}$ is a weight, and b is a bias of this hyper-plane. For this problem, we should introduce the Lagrange multiplier $\alpha_1, \alpha_2, \ldots, \alpha_N \geq 0$ to solve this convex, thus quadratic programming follows:

$$L(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{N} \alpha_i y_i(\mathbf{w} \cdot \mathbf{x}_i + b) + \sum_{i=1}^{N} \alpha_i. \tag{3}$$

Finally, the dual problem becomes [18]:

$$\underset{\boldsymbol{\alpha}}{\text{Max}}\, L(\boldsymbol{\alpha}) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i\alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j, \quad \text{Subject to } \sum_{i=1}^{N} \alpha_i \cdot y_i = 0,\ \alpha_i \geq 0. \tag{4}$$

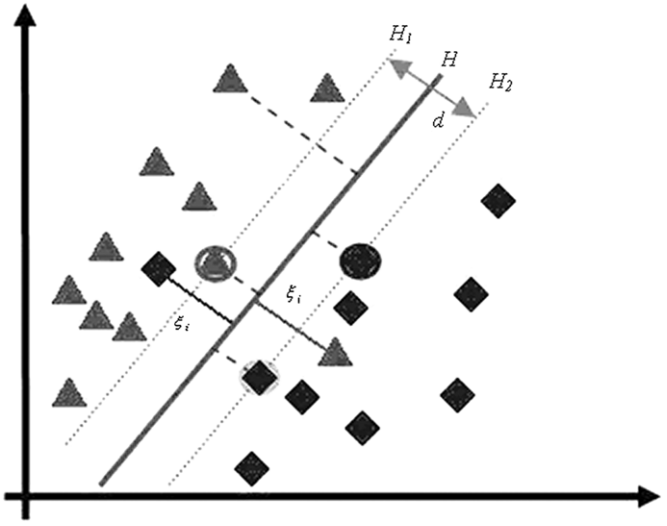


Fig. 2. Linearly non-separable data.

If α_i^* is the optimal solution, the optimal weight and bias can be obtained as follows:

$$\mathbf{w}^* = \sum_{i=1}^{N} y_i \alpha_i^* \mathbf{x}_i \tag{5}$$

$$b^* = -\frac{1}{2}\mathbf{w}^*[\mathbf{x}_p + \mathbf{x}_n] \tag{6}$$

where $\mathbf{x}_p$ and $\mathbf{x}_n$ are support vectors.

In fact, we cannot guarantee that the training data is always linearly separable, so we need to consider the non-separable case shown in Fig. 2.

For this case, we introduce a non-negative slack variable $\xi_i \geq 0$, so the condition $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$ becomes $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) + \xi_i \geq 1$ [18], and then the optimization equation becomes:

$$\underset{\mathbf{w},b,\xi_i}{\text{Min}} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\left(\sum_{i=1}^{N} \xi_i\right) \tag{7}$$

subject to $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)+\xi_i-1 \geq 0$ where $\xi_i \geq 0, i = 1 \ldots N$ and C is a positive parameter if the data $\mathbf{x}_i$ is in the correct region $0 \leq \xi_i < 1$. Otherwise, $\xi_i > 1$ [14]. By a method similar to (2), we introduce the Lagrange multiplier $\alpha_1, \alpha_2, \ldots, \alpha_N \geq 0$ to solve this problem.

$$\underset{\boldsymbol{\alpha}}{\text{Max}}\, L(\boldsymbol{\alpha}) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i\alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \tag{8}$$

subject to $\sum_{i=1}^{N} \alpha_i \cdot y_i = 0, 0 \leq \alpha_i \leq C$, where N is the number of support vectors, and the only difference with the linearly separable case is that α_i has an upper bound of C. Assuming α_i^* is the optimal solution, we can then obtain:

$$\mathbf{w}^* = \sum_{i=1}^{N} y_i \alpha_i^* \mathbf{x}_i \tag{9}$$

and b^* is obtained by satisfying the Karush–Kuhn–Tucker conditions [20,21].

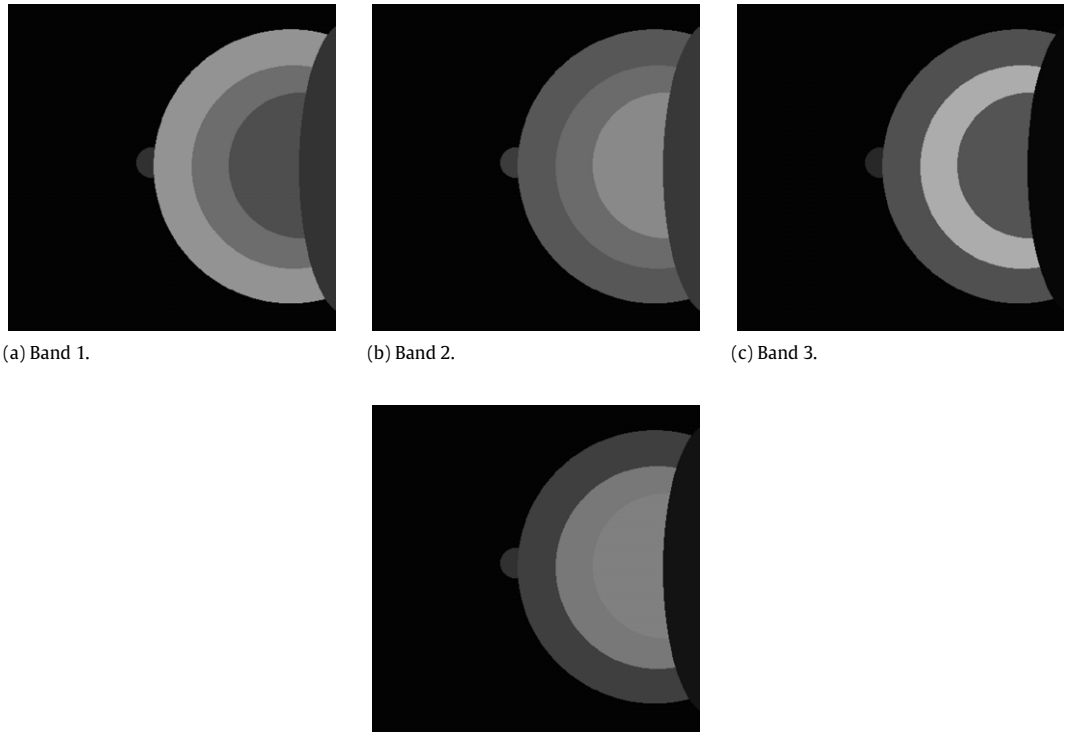
3. C-means (CM)

In order to evaluate the performance of the SVM approach, the widely used CM method [19] (also known as k-means in [22]) is used for comparative analysis. The reason to select the CM method is twofold. One is that it allows us to generate background signatures in an unsupervised manner for classification. Another is that it is basically a spatial-based pattern classification technique. In order to make a fair comparison, the CM method used here includes in its clustering procedure the same knowledge of objects of interest that is required by the SVM approach.

The CM method to be implemented in this paper for the experiment is a modified version of the commonly used CM method, which is also referred to as ISODATA in [19,23,22].

Table 1
The parameters used by the MRI pulse sequence and the gray level values of the tissues of each band used in the experiments.

Band #	MRI	Background	Fatty	Glandular	Tumor	Muscle
Band 1	Flash 2D	3	147	109	78	51
Band 2	T1	3	87	106	138	56
Band 3	T2	3	79	172	85	6
Band 4	PD	3	62	121	128	19



(a) Band 1. (b) Band 2. (c) Band 3.

(d) Band 4.

Fig. 3. Four band test phantoms for simulation study.

4. Experimental results

In this section, we present two sets of experiments — one set is computer-generated phantom images and another set is authentic magnetic resonance images. The phantom image experiments enable us to conduct a quantitative study and error analysis for the SVM approach, while the authentic MRI experiments allow us to assess its utility and effectiveness in medical diagnosis. In all experiments, the training data were selected by a 3×3 mask from corresponding tissues. Then the one vs. all multi-class SVMs were applied to training SVMs.

4.1. Computer simulations for phantom experiments

In this section, a series of computer simulations is performed to conduct a quantitative study and performance analysis of the SVM approach in comparison with the CM method described in Section 3 with the number of classes ($c = 5$) representing four classes of fatty, glandular, tumor and muscle tissue, as well as the image background. The computer-generated phantom images shown in Fig. 3 have four bands, each of which has the same size of 419×419 (total number of image pixels is 175,561). The semicircle structures represent the areas of the four breast tissues of interest (fatty, glandular, tumor, and muscle). The gray level values of these areas in each band were simulated in such a fashion that these values reflect the average values of their respective tissues in the real MR images shown in Fig. 8. Table 1 shows the values of the parameters used by the MRI pulse sequence and the gray level values of the tissues of each band used in the experiments. Zero-mean Gaussian noise was added to the phantom images in Fig. 3 so as to achieve different levels of signal-to-noise ratios (SNRs) ranging from 5 dB to 20 dB. Despite the fact that such MR phantom images may be unrealistic, they only serve the purpose of illustrating the proposed technique and demonstration of its advantages.

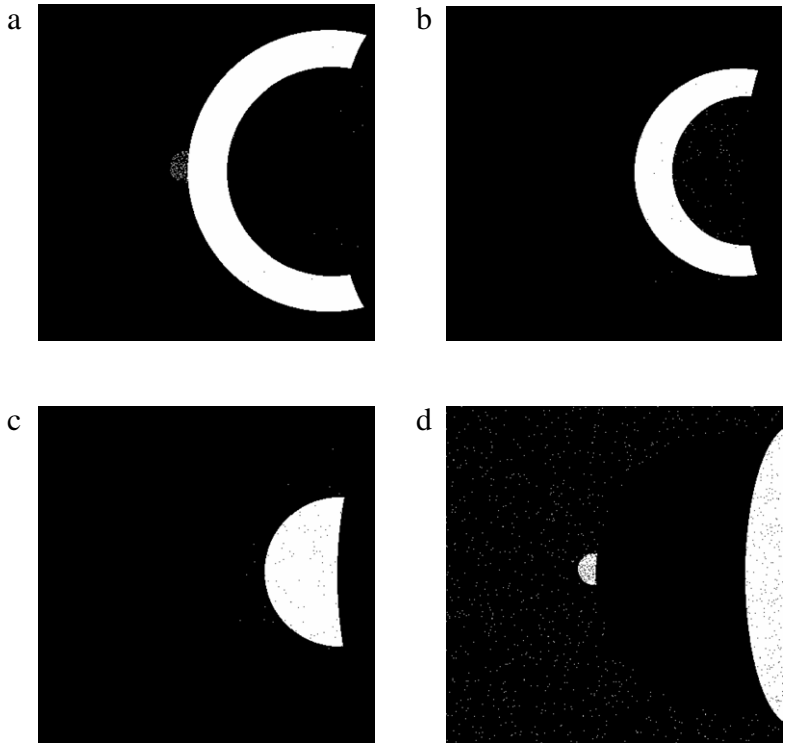


Fig. 4. Classification results produced by the SVM using the four images (SNR = 10 dB) in Fig. 3: (a) fatty; (b) glandular; (c) tumor; and (d) muscle tissues.

The classification results of the proposed method are tissues to be evaluated. Since there are considerable changes in performance between SNR = 5 dB and SNR = 10 dB, Figs. 4 and 5 show the SVM-classification results of background, fatty, glandular, tumor, and muscle tissues for SNR = 5 dB and 10 dB, respectively. Similarly, Figs. 6 and 7 show the CM-classification results of background, fatty, glandular, tumor, and muscle tissues for SNR = 5 dB and 10 dB, respectively.

4.1.1. Assessment method

This experiment's assessment method uses statistical theory analysis to assess the classification results [24,25]. Initially, we must get the classification results of the breast MR images and represent them in two-dimensional data form. In this section, we use statistical theory to compute the performance of SVM for the experiment with the phantom image and give a comparison with the results of CM.

To begin, let $\{p_i\}_{i=1}^n$ be a set of objects of interest which we would like to classify. We use $N(p_i)$ to denote the total number of pixels specified by the ith object signature p_i, $N_D(p_i)$ as the total number of pixels specified by the ith object signature, and actually detected as the p_i by SVM, $N_F(p_i)$ as the total number of false alarm pixels not specified by the ith object signature p_i but detected as the ith object by SVM or C-means, and N as the total number of pixels of a band image. Furthermore, we define the detection rate $R_d(p_i)$, false rate $R_f(p_i)$, and mean detection rate R_F respectively, as follow:

$$R_d(p_i) = \frac{N_D(p_i)}{N(p_i)} \tag{10}$$

$$R_f(p_i) = \frac{N_F(p_i)}{N - N(p_i)} \tag{11}$$

$$R_D = \sum_{i=1}^{n} R_d(p_i) w(p_i) \tag{12}$$

$$R_F = \sum_{i=1}^{n} R_f(p_i) w(p_i) \tag{13}$$

where $w(p_i) = N(p_i) / \sum_{i=1}^{n} N(p_i)$. It is worth noting that the mean detection rate defined by (12) is the mean of detection rates over the detected objects. Similarly, the mean false alarm rate defined by (13) is the mean of false alarm rates over the detected objects.

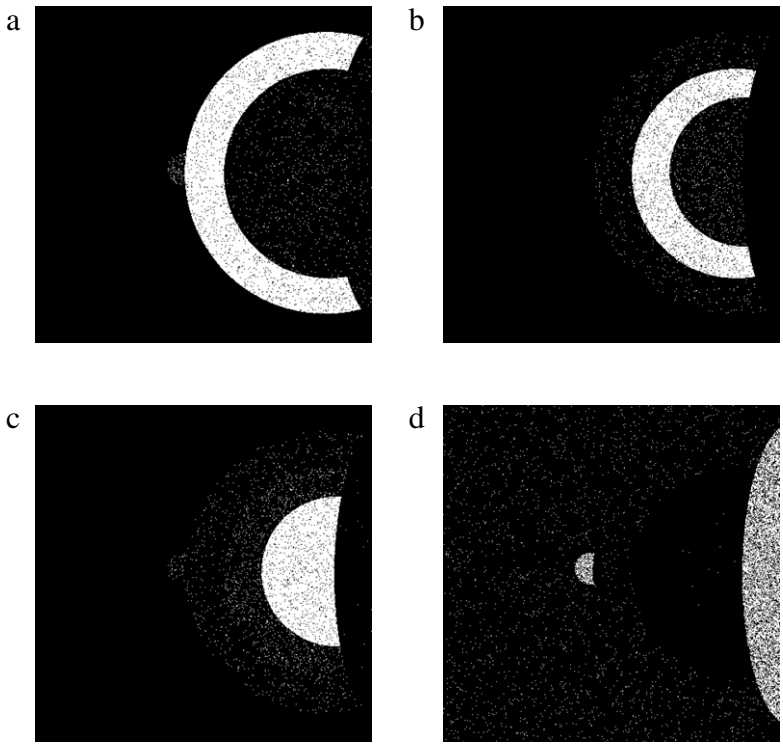


Fig. 5. Classification results produced by the SVM using the four images (SNR = 5 dB) in Fig. 3: (a) fatty; (b) glandular; (c) tumor; and (d) muscle tissues.

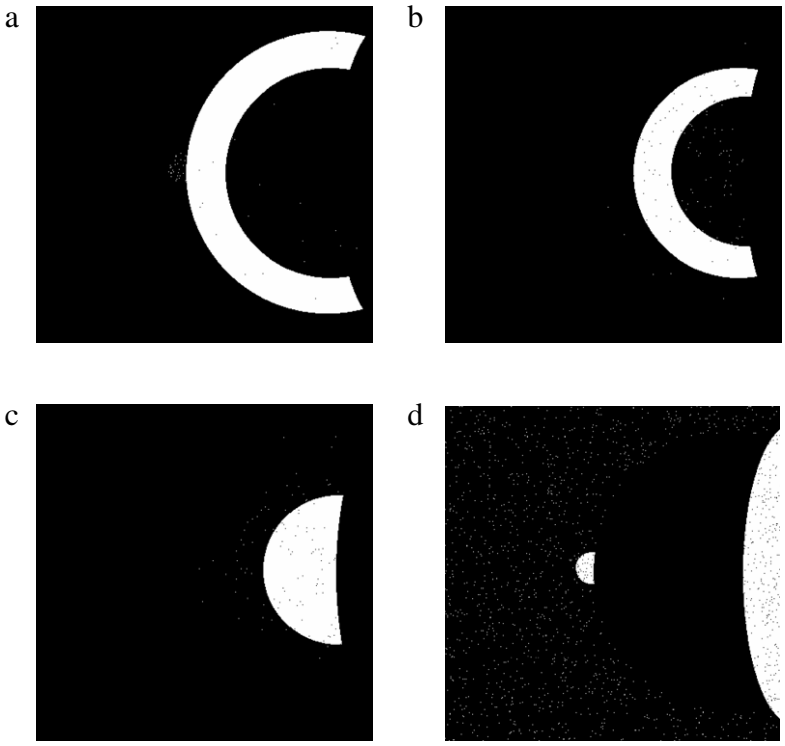


Fig. 6. Classification results produced by the CM using the four images (SNR = 10 dB) in Fig. 3: (a) fatty; (b) glandular; (c) tumor; and (d) muscle tissues.

After introducing the evaluation standard of assessment method, we then continue to compare the correct rate of SVM classification with CM classification. Tables 2 and 3 are the correct rate of fatty, glandular, tumor, and muscle tissues of SVM

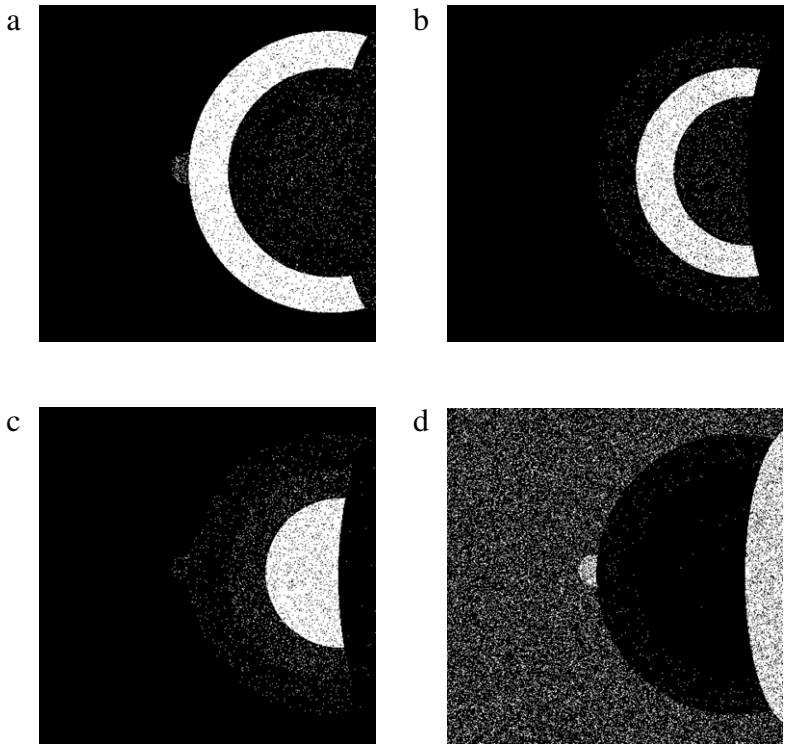
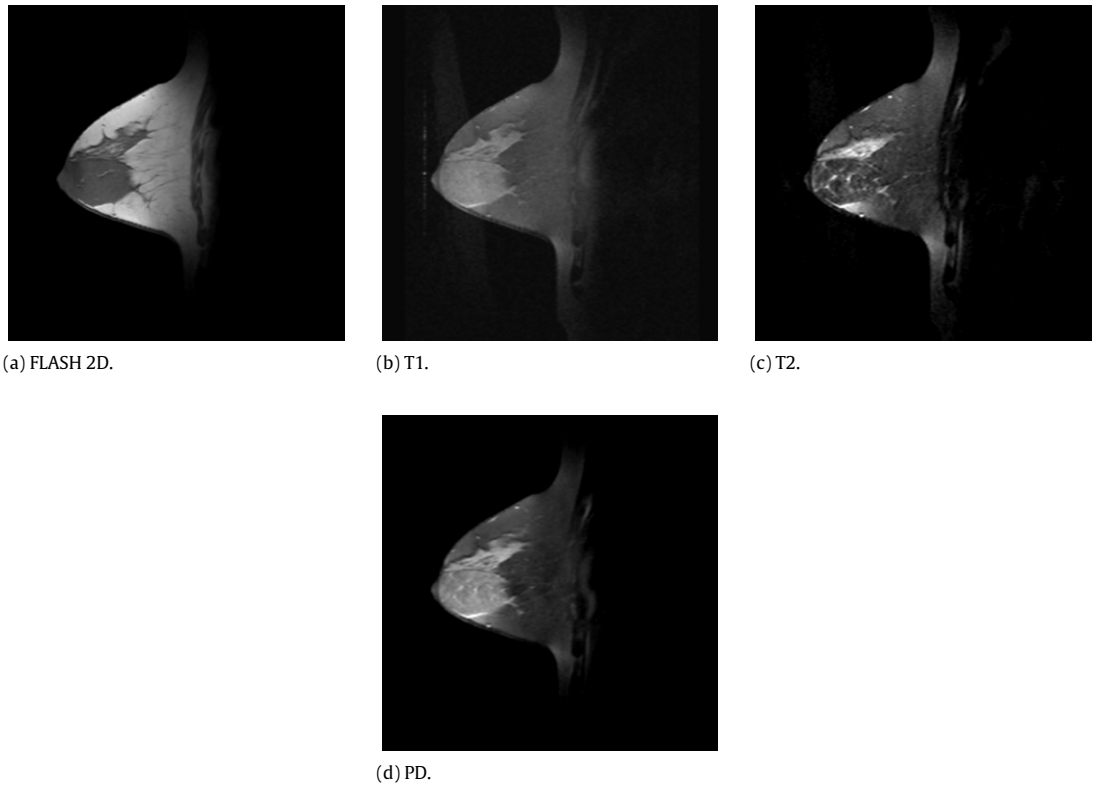


Fig. 7. Classification results produced by the CM using the four images (SNR = 5 dB) in Fig. 3: (a) fatty; (b) glandular; (c) tumor; and (d) muscle tissues.



(a) FLASH 2D. (b) T1. (c) T2.

(d) PD.

Fig. 8. Four spectral bands of authentic MR breast images.

Table 2
Classification results of fatty, glandular, tumor, and muscle tissues for SNR = 10 dB.

		Fatty	Glandular	Tumor	Muscle
SNR = 10 dB	CM	1	1	1	0.96
	SVM	1	1	1	0.987

Table 3
Classification results of fatty, glandular, tumor, and muscle tissues for SNR = 5 dB.

		Fatty	Glandular	Tumor	Muscle
SNR = 5 dB	CM	0.906	0.907	0.924	0.77
	SVM	0.944	0.928	0.933	0.823

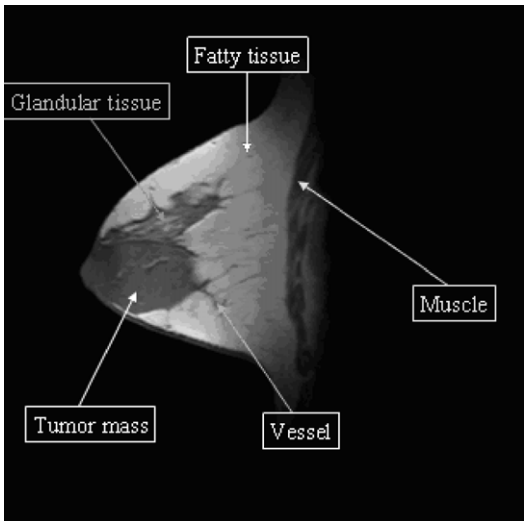


Fig. 9. Five tissues of the breast were selected and confirmed by the doctor.

and CM for SNR = 10 and 5 dB, respectively, each of the values being the means implemented from 20 runs. To compare classification results of SVM and CM, we can see that the correct rates of fatty, glandular, and tumor tissues lack obvious differences, but in the classification result of muscle tissue, the correct rate of SVM classification is better than CM. In the case of SNR = 5 dB, the correct rate of each tissue has decreased distinctly, especially for muscle tissue. After comparing the SVM-classification results with CM-classification results for SNR = 5 dB, we can see that the correct rate of SVM is higher than CM. Finally, we reveal that SVM has better performance for multi-spectral MR image segmentation and better robustness against noise.

4.2. Real MR image experiment

In this section, we use a set of a breast MR images to evaluate the performance of SVM. Band 1 is the FLASH 2D spectral image. Band 2 is the T1-weighted spectral image acquired by the pulse sequence TR/TE = 780 ms/20 ms. Band 3 is the T2-weighted spectral image acquired by the pulse sequence TR/TE = 4000 ms/105 ms. Band 4 is the PD-weighted spectral image acquired by the pulse sequence TR/TE = 4000 ms/15 ms. The time parameters of an MRI are repetition time (TR) and echo time (TE). The slice thickness of all the MR images is 3 mm and axial sections were taken from a Siemens 1.5 T scanner. Before acquisition of the MR images, the scanner was adjusted to prevent artifacts caused by magnetic fields from static, radio-frequencies and gradients.

In breast MR images, four breast tissues (fatty, glandular, tumor and muscle) are the major points of interest and their location can be generally obtained directly from the images as in Fig. 9. In our experiments, the spectral signatures of fatty, glandular, tumor, and muscle tissues, as well as the background used for the SVM were extracted directly from the MR images and verified by experienced radiologists. Fig. 10(a)–(d) shows the detection results of the SVM for fatty (a), glandular (b), tumor (c), and muscle (d) tissues produced. For comparison, we also applied the CM method to Fig. 8(a)–(d) to produce Fig. 11(a)–(d) with $c = 5$ for the four tissue classes: fatty (a), glandular (b), tumor (c), and muscle (d) and the background. Compared to Fig. 11(a)–(d), the SVM performed significantly better than the CM method. It should be noted that all experimental results presented here were verified by experienced radiologists.

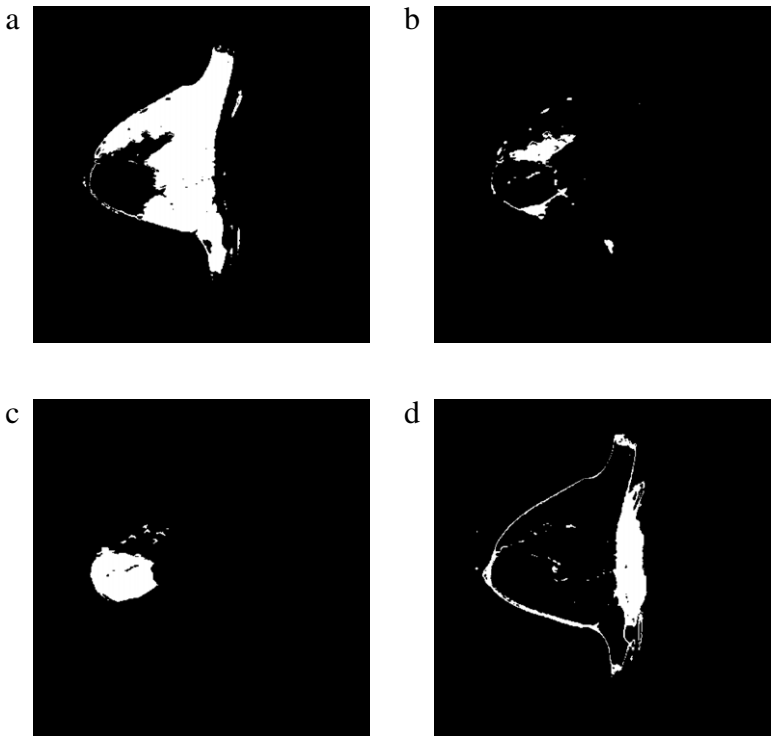


Fig. 10. Classification results produced by the SVM using the four images in Fig. 8: (a) fatty; (b) glandular; (c) tumor; and (d) muscle tissues.

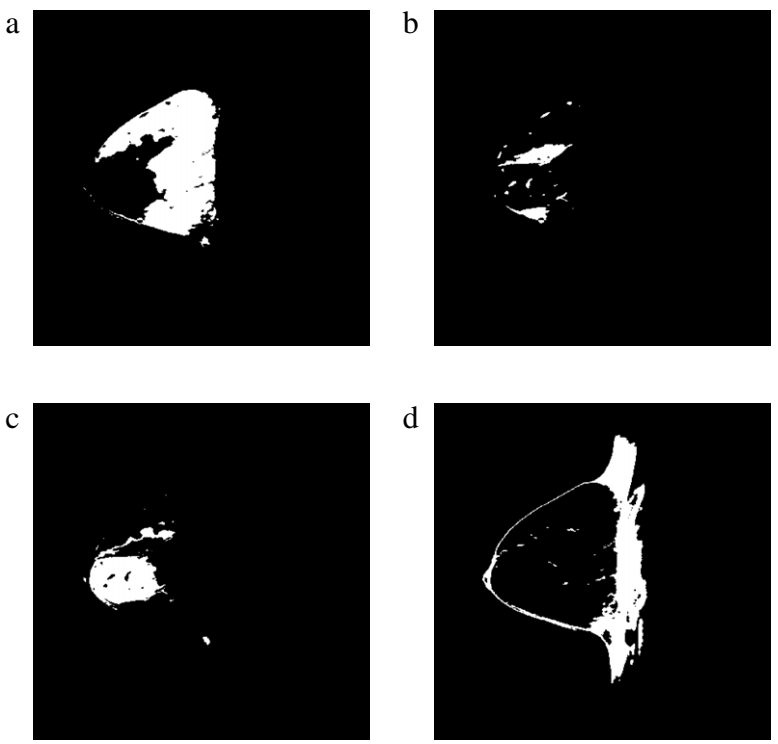


Fig. 11. Classification results produced by the CM using the four images in Fig. 8: (a) fatty; (b) glandular; (c) tumor; and (d) muscle tissues.

5. Conclusions

Support vector machine (SVM)-based approaches have been widely applied to many fields. In this paper, we present a new application of SVM in breast MR image classification. The SVM is based on the idea of a hyper-plane classifier, and it looks for the hyper-plane that maximizes the margin between two classes. Therefore, SVM-based classifiers can be utilized for MR image classification. In the experiment section, we present two sets of experiments: one set consists of computer-generated phantom images and the other set uses real MR images. From the experimental results, the correct rate of SVM classification is significantly better than CM in the case of SNR $= 5$ dB. Accordingly, we can know that the SVM has the capability for multi-spectral MR image segmentation and robustness against noise. For the authentic MR images, classification results can be provided to doctors as a basis of more accurate diagnosis and judgment of the patient's condition.

Acknowledgments

The financial support of this research by the National Science Council of the R.O.C., under Grant No. NSC 98-2221-E-167-014 is greatly appreciated. The authors would like to thank Dr. C.-I. Chang and Chi-Chang Chen of the Department of Radiology, Taichung Veterans General Hospital for their suggestions.

References

[1] C.M. Wang, S.C. Yang, P.C. Chung, Y.N. Chung, C.C. Chen, C.W. Yang, C.H. Wen, Orthogonal subspace projection-base approach to classification of MR image sequences, Computerized Medical Imaging and Graphics 25 (6) (2001) 465–476.
[2] D. Jia, F. Han, J. Yang, Y. Zhang, D. Zhao, G. Yu, A synchronization algorithm of MRI denoising and contrast enhancement based on PM-CLAHE model, JDCTA 4 (6) (2010) 144–149.
[3] S. Chandra, R. Bhat, H. Singh, D.S. Chauhan, Detection of brain tumors from MRI using gaussian RBF kernel based support vector machine, IJACT 1 (1) (2009) 46–51.
[4] S.D. Salman, A.A. Bahrani, Segmentation of tumor tissue in gray medical images using watershed transformation method, IJACT 2 (4) (2010) 123–127.
[5] G. Cardenosa, Breast Imaging Companion, second ed., Williams and Wilkins, 2001.
[6] Y. Kita, E. Tohno, R.P. Highnam, M. Brady, A CAD system for the 3D location of lesions in mammograms, Medical Image Analysis 6 (3) (2002) 267–273.
[7] M.J. Stoutjesdijk, C. Boetes, Magnetic resonance imaging and mammography in women with a hereditary risk of breast cancer, Journal of the National Cancer Institute 93 (14) (2001) 1095–1102.
[8] F.A. Cardillo, A. Starita, D. Caramella, A. Cilotti, A Neural tool for Breast cancer detection and classification in MRI, in: Proceedings of the 23rd Annual EMBS International Conference, vol. 3, 2001, pp. 2733–2736.
[9] Yibao Li, Junseok Kim, Multiphase image segmentation using a phase-field model Original Research Article, Computers & Mathematics with Applications 62 (3) (2011) 737–745.
[10] Pablo Irarrazaval, Carlos Lizama, Vicente Parot, Carlos Sing-Long, Cristian Tejos, The fractional fourier transform and quadratic field magnetic resonance imaging, Computers & Mathematics with Applications 62 (3) (2011) 1576–1590.
[11] Weishi Chen, Tiejun Liu, Baofa Wang, Ultrasonic image classification based on support vector machine with two independent component features, Computers & Mathematics with Applications 62 (7) (2011) 2696–2703.
[12] Navid Razmjooy, B. Somayeh Mousavi, F. Soleymani, A real-time mathematical computer method for potato inspection using machine vision, Computers & Mathematics with Applications 63 (1) (2012) 268–279.
[13] J. Li, N. Allinson, D. Tao, X. Li, Multitraining support vector machine for image retrieval, IEEE Transaction on Image Processing 15 (2006) 3597–3601.
[14] M. Pontil, A. Verri, Support vector machines for 3D object recognition, IEEE Transaction on PAMI 20 (6) (1998) 637–646.
[15] L. Nannia, A. Luminia, S. Brahnamb, Local binary patterns variants as texture descriptors for medical image analysis, Artificial Intelligence in Medicine 49 (3) (2010) 117–125.
[16] D. Bariamisa, D. Maroulisa, D.K. Iakovidis, Unsupervised SVM-based gridding for DNA microarray images, Computerized Medical Imaging and Graphics 34 (6) (2010) 418–425.
[17] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, NY, 1995.
[18] V.N. Vapnik, Estimation of Dependences Based on Empirical Data, Springer-Verlag, NY, 1982.
[19] R. Schalkoff, Pattern Recognition: Statistical Structural and Neural Approaches, Wiley, NY, 1992.
[20] Y. Sekine, A. Yokoyama, K. Yasuda, Y. Hayashi, R. Tanabe, H. Okamoto, Y. Tada, Optimal power flow in power systems, Japan Electric Association (2002).
[21] R. Tanabe, Y. Tada, A practical algorithm for multi-area state estimation using phasor measurement units in: Proceeding of the the International Conference on Electrical Engineering. 2008, pp. 1–6.
[22] C.M. Wang, C.C. Chen, Y.N. Chung, S.C. Yang, P.C. Chung, C.W. Yang, C.I Chang, Detection of spectral signatures in multispectral MR images for classification, IEEE Transactions on Medical Imaging 22 (1) (2003) 50–61.
[23] H.D. Chenga, J. Shana, W. Jua, Y. Guoa, L. Zhangb, Automated breast cancer detection and classification using ultrasound images: a survey, Pattern Recognition 43 (1) (2010) 299–317.
[24] Yingzi Du, C.-I. Chang, 3D combinational curves for accuracy and performance analysis of positive biometrics identification, Optics and Lasers in Engineering 46 (6) (2008) 477–490.
[25] G.C. Lin, C.C. Kang, W.J. Wang, C.M. Wang, Multispectral MR images segmentation based on fuzzy knowledge and modified seeded region growing, Magnetic Resonance Imaging 30 (3) (2012) 230–246.