Mixed Kernel Function SVM for Pulmonary Nodule Recognition

Yang Li[1,2], Dunwei Wen[3], Ke Wang[1], and A'lin Hou[2]

[1] College of Communication Engineering, Jilin University, Changchun, China
l_y09@mails.jlu.edu.cn, wangke@jlu.edu.cn
[2] College of Computer Science and Engineering,
Changchun University of Technology, Changchun, China
houalin@mail.ccut.edu.cn
[3] School of Computing and Information Systems,
Athabasca University, Alberta, Canada
dunweiw@athabascau.ca

Abstract. Automatic pulmonary nodule detection in computed tomography (CT) images has been a challenging problem in computer aided diagnosis (CAD). Most recent recognition methods based on support vector machines (SVMs) have shown difficulty in achieving balanced sensitivity and accuracy. To improve overall performance of SVM based pulmonary nodule detection, a mixed kernel SVM method is proposed for recognizing pulmonary nodules in CT images by combining both Gaussian and polynomial kernel functions. The proposed mixed kernel SVM, together with a grid search for parameters optimization, can be tuned to seek a balance between sensitivity and accuracy so as to meet the CADs need, and eventually to improve learning and generalization ability of the SVM at the same time. In our experiments, thirteen features were extracted from the candidate regions of interest (ROIs) preprocessed from a set of real CT samples, and the mixed kernel SVM was trained to recognize the nodules in the ROIs. The results show that the proposed method takes into account both the sensitivity and accuracy compared to single kernel SVMs. The sensitivity and accuracy of the proposed method achieve 92.59% and 92% respectively.

Keywords: image recognition, mixed kernel function, support vector machine, pulmonary nodule.

1 Introduction

A pulmonary nodule usually refers to intrapulmonary round dense shadow, the diameter of which is not more than 3 cm, and it is the characterization of lung cancer in early stage. Computed Tomography (CT) is the most important medical imaging technology for obtaining images of pulmonary nodules, which can be further used by either doctors or Computer Aided Diagnosis (CAD) systems for lung cancer early detection and forecast.

A. Petrosino (Ed.): ICIAP 2013, Part II, LNCS 8157, pp. 449–458, 2013.

Despite the difficulty of this problem, many methods have been developed for detection and recognition of pulmonary nodules in CT images for CAD. To reduce the number of omissions and decrease the examination time in radiologist scan, Cascio et al. [1] proposed a method for automatic detection of pulmonary nodules in lung CT scans, by using a 3D MassSpring Model (MSM) for segmentation of sus-pected nodular lesions in CT images, and a neural network classification for distinguishing between true positive (TP) and false positive (FP) candidates after a double-threshold cut to reduce FPs. The detection rate of the system reached 97%. Also, to reduce FPs and increase detection rate, Keserci et al. [2] combined morphological features with the wavelet snake method for automated detection of lung nodules in digital chest radiographs. Surez-Cuenca et al. [3] applied an iris filter and linear discriminate analysis to discriminating between nodules and FP findings, and their test results yielded a sensitivity of 80% at 7.7 FPs per scan.

As a powerful classification model, Support Vector Machines (SVMs) have been introduced in pulmonary nodules detection. Zhang et al. [4] combined rule-based method with SVM for lung nodule identification, and reached an accuracy of 84.39%. Aiming to solve the feature extraction problem that the details of ROI of 3D lung nodule are often ignored with a 2D method, Liu et al. [5] proposed method that combines KL (Karhunen-Loeve) transform with SVM for pulmonary nodules identification, achieved a 94.33% identification accuracy. Considering the low sensitivity caused by the imbalance between positive and negative samples, Campadelli et al. [6] and Liu et al. [7] used cost-sensitive SVMs with different penalty coefficients for nodule and non-nodule samples to improve the detection sensitivity, and they achieved as high as a sensitivity of 90%. Liu et al. [8] proposed a CAD system for qualitative diagnosis of solitary pulmonary nodules in chest CT images, which effectively represented the pathological characteristics of solitary pulmonary nodules with image features, and rapidly and accurately diagnosed solitary pulmonary nodules as benign or malignant. Their experimental results showed a sensitivity of 73.33%, and an accuracy of 71.67%.

We can see that there are relatively large differences between the experimental results of the above mentioned methods—some of them emphasized on accuracy, while the others on sensitivity. Also, some of the aforementioned pulmonary nodule detection methods adopted SVMs, but all of them used only single kernel function. The performance of the SVMs, while adopted with different kernel functions and different parameters, are quite different. Due to the fixed format of the single kernel function and the relatively small adapting space, the parameter optimization methods for single-kernel functions have poor generalization ability and robustness.

As different kernel functions have different strengths and weaknesses, one of the keys to improving the performance of an SVM is to design a suitable kernel function or a combination of kernel functions specific to the given problem. In recent years, various forms of mixed kernel functions have been applied to different areas [9–12]. Though mixed kernel functions have been applied in different areas, they have not yet been taken seriously enough in the detection of

pulmonary nodules. In this paper, we propose a mixed kernel SVM method for pulmonary nodule recognition.

2 SVMs with Mixed Kernel Functions

2.1 Selection of Kernel Functions

SVMs can be used to effectively solve pattern classification problems with small samples by a trade-off between complexity and learning ability to obtain better generalization ability, i.e., the ability to correctly classify unseen samples.

This paper deals with the binary classification problem for pulmonary nodule recognition, with nodules and non-nodules as the two classes. Let $T = \{(x_i, y_i)\}\ (i = 1, 2, \cdots, l)$ be training samples, where $x_i \in R^N$ and $y_i \in \{-1, +1\}$ are the input feature and category/class label of sample input i, respectively. Let $y_i = 1$ correspond to nodule category and $y_i = -1$ correspond to non-nodule category. The binary classification algorithm of single and/or mixed kernel function SVMs can be expressed as the primal form of SVM:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} \xi_i \tag{1}$$

$$s.t. \quad y_i\left((w \cdot \Phi(x_i)) + b\right) \geq 1 - \xi_i, \quad i = 1, 2, \ldots, l \tag{2}$$

$$\xi_i \geq 0, \quad i = 1, 2, \ldots, l \tag{3}$$

where w is a weight vector to be determined, C and ξ_i are penalty constant and slack variables respectively.

To achieve nonlinear SVM classification, the input data X_i are mapped to high dimensional feature space Z by nonlinear transformation functions $\Phi(X)$, and the optimal maximum-margin classification hyperplane is constructed in Z, and a kernel function $K(x_i, x_j)$ is introduced to represent the inner product of $\Phi(x_i)$ and $\Phi(x_j)$ after the transformation as follows

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \tag{4}$$

In practice, instead of using $\Phi(x_i)$ and $\Phi(x_j)$ separately, the function $K(x_i, x_j)$ is used as a whole.

The problem in primal form can be solved as a convex quadratic programming optimization problem by introducing Lagrange multiplier $\{\alpha_i\}\ (i = 1, 2, \cdots, l)$, and can finally be transformed as its dual form as follows

$$\min_{\alpha} \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^{l} \alpha_i \tag{5}$$

$$s.t. \quad \sum_{i=1}^{l} y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \ldots, l. \tag{6}$$

As in (1), C is the penalty parameter. The bias b will be solved by the following formula,

$$b = y_j - \sum_{i=1}^{l} y_i \alpha_i K\left(x_i, x_j\right) \tag{7}$$

Construct a decision function

$$f\left(x\right) = \operatorname{sgn}\left(g\left(x\right)\right) \tag{8}$$

where

$$g\left(x\right) = \sum_{i=1}^{l} \alpha_i y_i K\left(x_i, x\right) + b \tag{9}$$

Frequently used kernel functions are polynomial kernel function K_{poly}, RBF (radial basis function) kernel function K_{rbf} as well as linear kernel function K_{linear}, as expressed in formula (10) - (12):

$$K_{poly}\left(x, x'\right) = \left(x \cdot x' + 1\right)^d \tag{10}$$

$$K_{rbf}\left(x, x'\right) = \exp\left(-\left\|x - x'\right\|^2 / 2s^2\right) \tag{11}$$

$$K_{linear} = x \cdot x' \tag{12}$$

where parameter d is the number of the order of the polynomial kernel and s is the width of RBF kernel, both of which are given in advance.

2.2 Mixed Kernel Function

Based on the properties of kernel function, we can construct a new kernel function by using some known kernel functions as the basis. If the basic kernel functions are selected properly, then the new kernel function K can have greater generalization ability. Moreover, multiple kernel functions can form mixed kernel functions by their weighted sum. The expression of this combination is as follows:

$$\begin{aligned} K\left(x, x'\right) &= \sum_{p=1}^{U} m_p K_p\left(x, x'\right) \\ \sum_{p=1}^{U} m_p &= 1, \quad p = 1, \ldots, U \end{aligned} \tag{13}$$

where K_p is the p-*th* base kernel function, mp is the weight for p-*th* base kernel function in the mixed kernel function, used to control the proportion of each kernel function in the mixed kernel. The sum of the weights of all the U base kernel functions should be 1. It is easy to verify that the kernel function expressed

by formula (13) satisfies Mercer conditions, and thus can be used in SVM training and classification.

In this paper, we use mixed kernel functions to perform the non-linear transformation, train the corresponding mixed kernel SVM classifiers, and to produce the experimental results for pulmonary nodules recognition. We can see that the parameters of mixed kernel functions are much more than that of single kernel functions. Even for the least case that only two kernel functions are considered (i.e., $U = 2$), we need an extra set of parameters than the same kind of single kernel, plus an additional weight coefficient m_1 (now the other weight is $1 - m_1$, according to the sum to 1 condition). Note that, it is precisely by this parameter that we can freely choose a suitable proportion of the kernels in the mixed kernel function and balance the ability of learning and generalization.

3 Detection of Pulmonary Nodules

In general, a CAD system for pulmonary nodules includes several tasks such as preprocessing of CT images, segmentation of pulmonary parenchyma, segmentation of ROI, extraction of a variety of features, and distinguishing between benign and malignant nodules [13]. Because of the differences in the shape and structure of the pulmonary nodules and the gray characteristics of their CT images, it is easy to cause false and missing detection. Detection of pulmonary nodules is the key module of a CAD system and the suitability of the recognition algorithms will directly affect the test results.

3.1 Image Preprocessing and Feature Extraction

In this paper, image preprocessing and feature extraction, including segmentation of pulmonary parenchyma, segmentation of ROIs and feature extraction from the ROIs, are the same as that in [13, 14], in which a total of thirteen features were eventually extracted from an ROI, including seven morphological features, two gray-scale features as well as four texture features, as did in [4].

3.2 Mixed Kernel SVM for Pulmonary Nodule Detection

Although SVMs has been applied to the detection of pulmonary nodules as mentioned above, most of the works focused on ROI segmentation and feature extraction for direct nodule detection by using SVM [4, 8], while a few of them tried to improve SVM classification by parameters adjustment and optimization. For instance, [6, 7] increased positive penalty coefficient C for improving sensitivity. Another work combined SVM with other algorithms in order to improve SVMs performance [8]. As far as we know, however, only single kernel functions have been involved in SVM-based lung nodule recognition despite their limitations, which makes it hard to simultaneously improve both sensitivity and accuracy for nodule detection.

RBF kernel is the most widely used kernel function. It is normalized kernel and has strong learning ability, and many base kernel functions (e.g., linear kernel, Sigmoid kernel) have similar nature with it [16, 17]. On the other hand, low-order polynomial kernel has strong generalization ability relative to other kernels. When the dimension of the feature space is very high, however, the amount of computation cost will quickly increase and may exceed the systems capacity. A better idea is to use the weighted sum of both polynomial and RBF kernel function to form mixed kernels as mentioned above, so as to keep their advantages and obtain good generalization ability and learning ability at the same time, yet at an acceptable computation cost. According to (13), a mixed kernel function including both polynomial kernel and RBF kernel can be expressed as follows:

$$K(x, x') = mK_{poly}(x, x') + (1 - m) K_{rbf}(x, x') \tag{14}$$

which will be used in our SVM classifier in the hope that both the accuracy and sensitivity will be improved.

There are four parameters to optimize in the mixed kernel function shown in (14), i.e., the order of the polynomial d, the width of the RBF kernel s, the weight coefficient m and the penalty parameter C. We restrict each parameter to a certain range given in advance for simplifying the optimization process. More specifically, C is selected between 2^{-9} and 2^{9}, and s between 2^{-7} and 2^{7}. When the order d of polynomial kernel function is small, the generalization ability is relatively strong. Hence we select d as a positive integer between 2 and 3. The weight m of mixed kernel function is a very important parameter as well, directly affecting the share of each component kernel in the mixed kernel SVM. We search it between $[0, 1]$, and a search process with a step size 0.01 is undertaken to find the most appropriate value of it.

We incoporated a grid search algorithm [18] for parameters optimization in our experiments. Basically, we first fixed the weight m, then chose an arbitrary integer value for d between 2 and 3 and sought the optimal values for s and C by the grid search in a progressive way with their value multiplied by 2 each time within the ranges described above. Then we sought the best m value between $[0, 1]$ by the grid search with the step size 0.01 for a fixed d. By iterating the process and comparing the test results for different combination of the four parameters, we picked out the best parameter set as the search result for final recognition test.

4 Experimental Results and Analysis

Our experimental data are from a hospital, including 700 CT images from 20 groups of cases, and each group is equipped with doctors diagnosis in standard text format. The size of each piece of the CT images is 512×512 pixels and the slice thickness of the CT scanning is 5.0 mm. With the image preprocessing and feature extraction mentioned in Section 3, a total of 270 ROIs including 80 nodules and 190 false positives were segmented and their features extracted, and

all the ROI samples were randomly divided into two groups, i.e., 170 training samples and 100 test samples.

The experiments were performed on MATLAB platform, and SVM related implementation, training and test were based on libsvm toolbox [18]. All the experimental samples were first normalized to $(0, 1)$. Both accuracy (ACC) and sensitivity (SEN) were used in the experiments as the evaluation criteria, which represent overall correct rate and true positive rate of pulmonary nodule recognition respectively, defined as follows,

$$ACC = \frac{(TP + TN)}{TP + TN + FP + FN} \tag{15}$$

$$SEN = \frac{TP}{(TP + FN)} \tag{16}$$

where, TP represents true positives, the number of nodules that are correctly recognized; FP represents false positive, the number of non-nodules that are falsely recognized as nodules; FN is false negative, the number of nodules that are incorrectly recognized as non-nodules; and TN is true negative, the number of non-nodules that are correctly recognized as are. In medical practice, more emphasis is placed on higher sensitivity, to prevent from missing true nodules so that patients can get necessary treatment. Therefore, during our model selection stage for parameter optimization, where a 5-fold cross-validation was used, we collected possible set(s) of parameters corresponding to highest average ACC, and whenever more than one such set appeared, we selected the parameter set that yielded the highest SEN as our optimal parameters.

To evaluate the effect of the proposed mixed kernel SVM method on pulmonary nodules detection, four kinds of kernels, i.e., linear kernel, RBF kernel, polynomial kernel and a mixture of RBF and polynomial kernel, were tested in our pulmonary nodule detection experiments. Table 1 shows the trained parameters through the grid search method validated by the 5-fold cross-validation, and test results on test samples using the trained (optimal) parameters.

From the final test results under the optimal parameters, we can see that the mixed kernel SVM yields a 92.59% SEN, the highest value among all kernel modes under test, also 3.7% higher than the SEN of RBF kernel. Moreover, it yields a 92% ACC at the same time—although just lower than the best one produced by RBF kernel by 1%, in return it obtains high SEN. This demonstrates that the mixed kernel SVM has better overall performance and stronger learning and generalization ability compared to other single kernel modes.

Table 1 also shows the optimal parameters obtained by the grid search method. As the mixed kernel function has more parameters than a single kernel functions, to reduce the computation cost caused by more parameters, the constant term of polynomial in the mixed kernel is set to 1. On the one hand, parameters of mixed kernel function are harder to be optimized; on the other hand, they make the mixed kernel SVM more flexible for producing ideal results—there is a trade-off between computation cost and performance. A suitable balance can be sought, depending on the need of an application domain. In our pulmonary

Table 1. Comparison between different kernels

Kernal Mode	Optimal Parameters	SEN	ACC
Linear	$C = 2^8$	0.8889	0.91
RBF	$C = 2^9$, $s = 2^{-3}$	0.8889	0.91
Polynomial	$C = 2^9$, $d = 2$	0.8519	0.89
Mixed Kernel	$C = 2^4$, $s = 2^{-8}$, $d = 3$, $m = 0.57$	0.9259	0.92

nodule detection, the need for a relatively high sensitivity and the search for balanced learning and generation ability has led us to choose and construct suitable mixed kernel SVMs for better recognition performance.

To further investigate the performance, the receiver operating curve (ROC) were depicted in Fig. 1 based on our test results by the trained linear, RBF, polynomial and mixed kernel SVMs. The area under curve (AUC) for each case were calculated and presented in Table 2.

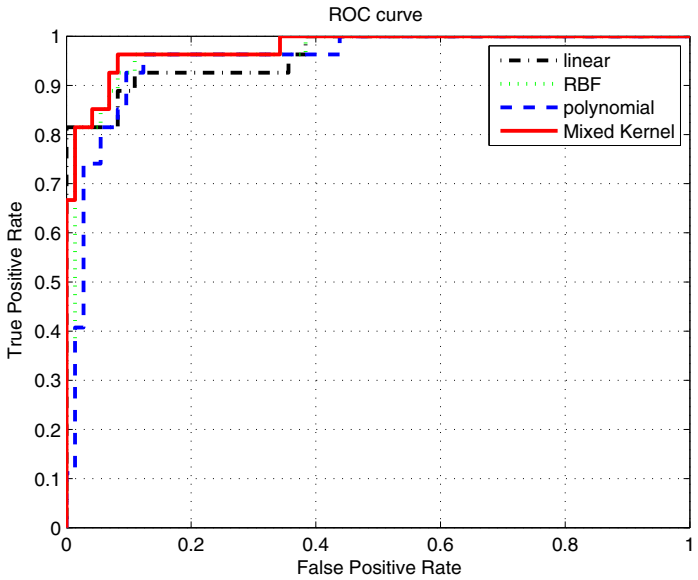

Fig. 1. ROC Curve of the SVMs with different kernels

The model whose ROC curve has bigger AUC value is considered to be more accurate. As shown in Table 2, the proposed mixed kernel method has the biggest value (0.9756). This can also be seen from Fig. 1, where the curve of mixed kernel SVM is closer to $(0, 1)$ point than the other curves. Both of them demonstrate that the proposed mixed kernel SVM is superior to the other three methods.

Table 2. Area under the ROC curves

Mode of kernel	Linear	RBF	Polynomial	Mixed kernel
AUC	0.9680	0.9625	0.9518	0.9756

To sum up, by using the multiple kernel functions and the grid-search based parameter optimization (including both the kernel function parameters and weight coefficients), the most suitable parameter set can be found. This way, the mixed kernel SVM gains the advantages from both RBF kernel and polynomial kernel functions, and the weight coefficients can be used to adjust the proportion of RBF kernel to polynomial kernel in the mixed kernel, so as to balance the sensitivity and accuracy of the trained SVM for obtaining acceptable detection performance and meeting the need of certain applications. The test results can verify that the obtained mixed kernel SVM has balanced sensitivity and accuracy, and good generalization and learning ability at the same time.

5 Conclusions

In this paper, we have proposed a mixed kernel SVM based recognition method for pulmonary nodule detection. The experimental results of the mixed kernel SVM, whose optimal parameters are trained by 5-fold cross-validation, show that the method yields 92.59% sensitivity and 92% accuracy. Compared with other single kernel SVMs, it can better balance between these two evaluation criteria. While our experiments present its strength to recognize pulmonary nodules, there are issues to be resolved:

- Parameter optimization: Grid search algorithm is workable for only small size problems. Some simplification and automation of parameter optimization process are expected to speed up training and ease optimal parameter seeking.
- Sensitivity: There is still much room to further improve sensitivity; cost-sensitive SVMs may be helpful on this regard.
- Mixed kernel: It can certainly be extended to other forms of mixed kernels by incorporating different components and combinations of the existing basic kernel functions.

The above issues will also serve as our future research focuses.

Acknowledgments. This work was jointly supported by Science and Technology Development Plan of Jilin Province (201201129), Scientific Research Development Fund of Changchun University of Technology (2011LG04), and Short-term Foreign Expert Program of Jilin University, 2012.

References

1. Cascio, D., Magro, R., Fauci, F., Iacomi, M., Raso, G.: Automatic detection of lung nodules in CT datasets based on stable 3D massspring models. Computers in Biology and Medicine, 1098–1109 (2012)
2. Keserci, B., Yoshida, H.: Computerized Detection of Pulmonary Nodules in Chest Radiographs Based on Morphological Features and Wavelet Snake Model. Medical Image Analysis 6, 431–447 (2002)
3. Suárez-Cuenca, J.J., Tahoces, P.G., Souto, M., Lado, M.J., Remy-Jardin, M., Remy, J., Vidal, J.J.: Application of the Iris Filter for Automatic Detection of Pulmonary Nodules on Computed Tomography Images. Computers in Biology and Medicine 39, 921–933 (2009)
4. Zhang, J., Li, B., Tian, L.-F., et al.: Lung Nodule Recognition Combining Rule-Based Method and SVM. Journal of South China University of Technology (Natural Science Edition) 39, 125–129 (2011)
5. Liu, Y., Zhao, D., Liu, J.: Recognition of 3-D Lung Nodules Based on K-L Transform and Support Vector Machine. Journal of Northeastern University (Natural Science) 30, 1249–1252 (2009)
6. Campadelli, P., Casiraghi, E., Valentini, G.: Support Vector Machines for Candidate Nodules Classification. Neurocomputing 68, 281–288 (2005)
7. Liu, L., Liu, W.: A Method of Pulmonary Nodules Detection with Support Vector Machines. In: Proceedings of 8th International Conference on Intelligent Systems Design and Applications, ISDA, pp. 32–35 (2008)
8. Liu, L., Liu, W., Chu, C., et al.: Fast Classification of Benign and Malignant Solitary Pulmonary Nodules in CT image. Optics and Precision Engineering 17, 2060–2067 (2009)
9. Xie, J.: Optimal Control of Chaotic System Based on LS-SVM with Mixed Kernel. In: Proceedings of 3rd International Symposium on Intelligent Information Technology Application, pp. 622–625 (2009)
10. Lu, Y.-L., Li, L., Zhou, M., Tian, G.: A New Fuzzy Support Vector Machine Based on Mixed Kernel Function. In: Proceedings of 2009 International Conference on Machine Learning and Cybernetics, pp. 526–531 (2009)
11. Wang, H., Sun, F., et al.: On Multiple Kernel Learning Methods. Acta Automatica Sinica 36, 1037–1047 (2010) (in Chinese)
12. Gönen, M., Alpaydin, E.: Multiple Kernel Learning Algorithms. Journal of Machine Learning Research, 2211–2268 (2011)
13. Wang, Q.: Detection of Lung Nodules in CT Images Based on 3D SVMs. PhD Dissertation, Changchun: Jilin University (2011)
14. Li, Q., Sone, S., Doi, K.: Selective Enhancement Filters for Nodules, Vessels, and Airway Walls in Two and Three-dimensional CT Scans. Medical Physics 30, 2040–2051 (2003)
15. Jia, T., Zhao, D., Wei, Y., et al.: Computer-Aided Lung Nodule Detection Based on CT images. In: Proceedings of IEEE/ICME International Conference on Complex Medical Engineering, pp. 816–819. IEEE, Beijing (2007)
16. Keerthi, S.S., Lin, C.-J.: Asymptotic Behavior of Support Vector Machines with Gaussian Kernel. Neural Computation 15, 1667–1689 (2003)
17. Chen, P.-H., Fan, R.-E., Lin, C.-J.: A Study on SMO-type Decomposition Methods for Support Vector Machines. IEEE Trans. on Neural Networks 17, 893–908 (2006)
18. Chang, C.-C., Lin, C.-J.: LIBSVM: A Library for Support Vector Machines. ACM Trans. on Intelligent Systems and Technology 2, 27:1–27:27 (2011), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`