

# Combining SVM and Rule-Based classifiers for optimal classification in breast cancer diagnosis

I. ANDREADIS, G. SPYROU, A. ANTARAKI, G.ZOGRAFOS, D. KOULOHERI, G. GIANNAKOPOULOU, K. NIKITA & P. LIGOMENIDES

*Abstract*-- Mammography is accepted as the most effective method to detect breast cancer. Breast microcalcifications are considered very important findings, which may be associated to the existence or not of breast cancer. It has been proven that in some cases the evaluation of their characteristics contributes to the early diagnosis of breast cancer. A computer aided diagnosis (CAD) system has been already developed in the lab based on detailed analysis and evaluation of related features of microcalcifications and a subsequent decision tree classification scheme. This system has very good sensitivity while suffering from low specificity. In this paper, we present a binary classifier which combines the aforementioned rule-based classifier with certain classification schemes based on the support vector machines (SVM) methodology. We evaluated the performance of the proposed binary classifier and we compared it with the existing CAD system using a database of 195 (153 benign and 42 malignant) clinical mammograms provided from collaborating diagnostic centers focused on breast examination. All 195 cases have undergone biopsy in regions where there were clustered microcalcifications, thus there is histological verification of their status. In our experiments, the binary classifier reinforces the computer-aided diagnosis system. While a sensitivity as high as 94.12% and a specificity equal to 31.51% was achieved by the decision tree classification scheme, the binary classifier provided a sensitivity equal to 82.35%, but a higher specificity equal to 65.75%, classifying correctly more benign cases. In other words, the binary classifier is able to reinforce the diagnosis of the CAD system by highly improving the specificity of the system with a little cost in terms of sensitivity.

*Index Terms*-- Computer-aided diagnosis, mammography, microcalcifications, support vector machines.

## I. INTRODUCTION

Breast cancer is the second leading cause of cancer deaths in women today (after lung cancer) and is the most common form of cancer among women worldwide, occurring in nearly one out of nine women. The key to surviving breast cancer is early detection and treatment [1], [2].

Mammography is nowadays accepted as the most effective method to detect breast cancer since with this technique many important findings, that may be associated to the existence of breast cancer, can be revealed. Among these important findings, microcalcifications are the smallest structures identified on a mammogram and they are either easily or hardly distinguished on the

Manuscript received February 20, 2007. I. Andreadis (email: [iandr@biosim.ntua.gr](mailto:iandr@biosim.ntua.gr)) and K. Nikita are with School of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou Str., 15780, Zografos, Athens, Greece

G. Spyrou (corresponding author, email: [gspyrou@bioacademy.gr](mailto:gspyrou@bioacademy.gr), tel. +302106597151) is with Biomedical Research Foundation of the Academy of Athens, Soranou Efessiou 4, 115 27 Athens, Greece.

A. Antarakis and P. Ligomenides are with Informatics Laboratory, Academy of Athens, Soranou Efessiou 4, 115 27 Athens, Greece.

mammograms depending on the existing tissue background [3]-[5].

The subtle nature of these radiographic findings accompanied with other factors such as poor image quality and oversight by the radiologist, may lead to missed detections of breast cancer or misclassifications. In general, the interpretation of a mammogram is a difficult diagnostic task, especially for not experienced radiologists [6], [7]. The successful development of computer aided diagnosis (CAD) systems would be of great value, if these systems could provide a reliable second opinion to the radiologist [8]-[10].

A system, called “Hippocrates-mst”, has been already developed in the lab and is based on detailed analysis and evaluation of related features of microcalcifications (individually and in clusters) [11]-[14]. After the detection of the existing microcalcifications in a selected region of the breast, a rule-based decision tree classification scheme is applied for the final risk assessment. This system has very good sensitivity while it suffers from low specificity.

In this paper, we present a binary classifier for the classification and characterization of clustered microcalcifications in digitized mammograms that reinforces the aforementioned CAD system. This classifier combines the rule-based classifier, used in “Hippocrates-mst” system, with certain classification schemes based on the support vector machines (SVM) binary methodology [15]-[18].

We developed three different SVM classifiers, conducting three experiments by using each time different training set and technique (simple test method, Cross-Validation, Leave-One Out). We evaluated their performances on a dataset of 90 mammograms. The results from the classification of this dataset indicated that all the three SVM classifiers provided values of sensitivity and specificity in the range 65%-85%. On the other hand, the “Hippocrates-mst”

classification scheme achieved the highest sensitivity and the lowest specificity. Subsequently, we developed a new classifier which will exploit the different behaviour of each of the four individual classifiers (3 SVM classifiers and “Hippocrates-mst”) and combine them in an appropriate way in order to improve the diagnostic accuracy of the existing CAD system and especially the low levels of specificity.

Our aim was to investigate the effectiveness of this new classifier and its potentiality to optimize the final diagnosis phase of the “Hippocrates-mst” CAD system, which is mainly suffering from low specificity. The proposed classification scheme can be beneficial for the CAD system, by reducing the number of false positive diagnoses, achieving in that way greater levels of specificity, and maintaining at the same time the sensitivity at high levels.

## II. METHODS AND MATERIALS

### A. Description of the “Hippocrates-mst” CAD system

The computer aided diagnosis system we have already developed and tested, named “Hippocrates-mst”, is based on detailed analysis and evaluation of related features of individual microcalcifications and of formed clusters [11]-[14].

Through this system, every selected image can be displayed inside a form with various digital tools that can be applied either to the whole image or to a Region Of Interest (ROI). Furthermore, the user may use the detection algorithm and reveal the microcalcifications existing in the selected region. After the detection of the microcalcifications, image analysis is needed for the microcalcification feature extraction and quantification. The considered microcalcification features are: (i) size, (ii) circularity, (iii) existence of dark center, (iv) level of brightness, (v) irregularity regarding the shape, (vi) level of branching and (vii) circumvolution.

The above features are calculated for every microcalcification inside a sub region of the

initial ROI selected by the user, comprising a cluster. Then, the system uses a two stage procedure for the risk assessment of the selected ROI. In the first stage, a hand-crafted rule based decision tree, exploiting the seven aforementioned characteristics, is used to estimate the risk for each microcalcification. The design of the decision tree is based on mathematical modeling of rules coming from the related literature as well as from selected medical experts. In the second stage, the final assessment is obtained from a risk estimation, which is based on the evaluation of the findings and specifically on the following parameters: (i) the risk distribution of microcalcifications, (ii) the number of microcalcifications at high risk, (iii) the cluster polymorphism.

Preliminary laboratory tests have already been performed in order to evaluate the diagnostic accuracy of “Hippocrates-mst” system. Moreover, during the last year, the system is undergoing an evaluation procedure in Hippocraton Hospital of Athens in Greece. The results of these evaluation tests indicate that “Hippocrates-mst” CAD system estimates the risk of breast cancer towards the right direction. However, although it presents high sensitivity, it appears to suffer from low specificity. This indicates that there is an overestimation of the risk, driven from the attempt to minimize the false negative results. Thus, there is space for optimization and refinement of the procedure, in order to improve the classification accuracy. Our primary goal is to improve the specificity of the system, while maintaining the sensitivity at high levels.

### *B. SVM classifiers*

SVM are learning machines used in pattern recognition [15]-[18]. The method is based on intuitive geometric principles, as it aims to the definition of an optimal hyperplane (maximal margin hyperplane, MMH), which linearly separates the training data so that minimum expected risk is achieved. The SVM methodology was used for the development of various classification

schemes that are focused on the diagnosis of clusters of microcalcifications. As in the classification scheme of “Hippocrates-mst”, the SVM classification schemes that we are going to develop follow a two-layer procedure. In the first layer, the input vectors are the seven morphological features of each microcalcification in the selected cluster, which indicate whether one microcalcification is suspicious or not. In the second layer, the input vectors are the number of suspicious microcalcifications as calculated from the first layer and other parameters such as the density and the polymorphism that are used for the final estimation for the existence or not of breast cancer.

For the development and evaluation of the classification schemes a dataset of 195 clinical mammograms was used. All cases have undergone biopsy in regions where there were clustered microcalcifications, thus there is histological verification of their status. The collection of mammograms was carried out mainly at the Hippocraton hospital of Athens and the mammograms were provided to us into two datasets. The first one was constituted of 105 digitised mammograms accompanied with the corresponding biopsy test results. We exploited this first dataset for the training and evaluation of new classifiers based on the SVM methodology. A few months later we were supplied a second dataset of the rest 90 mammograms which was used to test the performance of the previously developed SVM classifiers and compare them with “Hippocrates-mst” classification technique (rule-based decision tree).

#### *1) Development of three SVM classifiers*

As we have already mentioned, we used an initial dataset of 105 real cases for the training of three new classifiers. Among the 105 clusters of microcalcifications, there have been 80 benign and 25 malignant cases. One of the major problems that we had to face was the unbalanced set due to the small number of the available malignant cases. For this reason, we conducted three

different experiments, using each time different training set and technique, in order to investigate the diagnostic accuracy of the developed system and its generalization ability. Each experiment leads to a different classifier and thus to different classification results.

In the first experiment, we divided arbitrarily the 105 cases into two sets: the training set and the evaluation set. The training dataset consists of 30 cases (15 malignant, 15 benign). For the evaluation of the system, we examined the rest 65 benign clusters and 10 malignant. We applied single training method, using all the different kernels presented in the previous section, for values of parameters  $C \in \{2^{-2}, 2^{-1} \dots 2^{13}\}$ ,  $\gamma \in \{2^{-12}, 2^{-11}, \dots 2^3\}$  and  $p \in \{2, 3, \dots 9\}$ . Our aim was to find the combination which leads to minimization of the number of training errors and optimization of the classification accuracy of the training dataset. Finally, the Gaussian RBF kernel was selected with  $C=8192 (=2^{13})$  and  $\gamma=8 (=2^3)$ . After the training phase, the 75 cases of the testing set were used to evaluate the performance of this first classifier, achieving that way a sensitivity as high as 90.00% and a specificity equal to 70.77%.

The final selection of the first classifier aimed to the optimization of the classification accuracy of the training dataset. However, a learning algorithm may perform well on the training data, but achieve poor performance when applied to data outside the training dataset. Therefore, we used an alternative training method, which could provide useful information about the generalization ability of the classifier. The training method we used in the second experiment is the cross-validation. The datasets are the same with those of the first experiment. We trained the classifier with the dataset of 30 cases with the cross-validation procedure for the range of parameters we mentioned above. The combination that provided the best classification accuracy was the Gaussian RBF kernel with  $C=4 (=2^2)$  and  $\gamma=2$ . At the end, we tested the performance of the classifier on the remaining 75 cases. The sensitivity achieved was 80.00% and the specificity

52.31%. The comparison between the two first classifiers is straightforward, as they have been trained and tested on the same datasets. Although we expected that the second classifier would provide better classification results due to the use of the cross-validation procedure, the first classifier seems to outperform the second one for the specific dataset of 75 cases, as both the indices of sensitivity and specificity have greater values in the first case.

In the third experiment, we exploited all the available 105 cases in order to investigate the generalization ability of the resulting classifier. The basic problem in this experiment is that the dataset is unbalanced, as the number of benign cases is almost three times greater than the number of malignant cases, leading to a behavior of the classifier biased towards the benign cases. Actually, during the training phase, we conducted a comparative study between the four given kernels, applying the Leave One Out method. The selection of the Gaussian RBF kernel with  $C=8192$  ( $=2^{13}$ ) and  $\gamma=0.5$  led to the best classification accuracy for the whole dataset. The classifier of this experiment provided specificity as high as 87.50%, while the sensitivity was only 48%.

The classification results on the training dataset obtained from each classifier are listed in Table I.

TABLE I  
COMPARISON BETWEEN THE THREE SVM CLASSIFIERS

	Accuracy(%)	Specificity(%)	Sensitivity (%)
<i>Classifier #1</i>	73.33	70.77	90.00
<i>Classifier #2</i>	56.00	52.31	80.00
<i>Classifier #3</i>	78.10	87.50	48.00

## 2) Comparative evaluation

After the initial development of the three different classifiers, we had to test their performance

on a new dataset. As we have already mentioned, this dataset consists of 90 new real cases, which have not been used at all during the training phase. We also test the performance of “Hippocrates-mst” CAD system, whose classification methodology is based on a decision tree. We have already described the four virtual zones that the CAD system uses for the classification of a microcalcifications cluster. When the final risk percentage is greater than 35% the patient is referred for a surgical biopsy. Otherwise, it is considered that there is no evidence for malignancy and biopsy test is discouraged. On the contrary, the SVM approach is a binary procedure. An SVM classifier does not give risk estimation, but decides whether there is malignancy or not. Only when the prediction for malignancy is positive, a biopsy test is encouraged. The performance evaluation and comparison between the aforementioned classifiers was based on the calculation of three objective indices: accuracy, sensitivity and specificity.

Table II lists the classification results on the test dataset.

TABLE II  
CLASSIFICATION OF BREAST MICROCALCIFICATIONS WITH THE SVM-BASED CLASSIFIERS AND THE EXISTING  
“HIPPOCRATES-MST” SYSTEM

	ACCURACY (%)	SPECIFICITY (%)	SENSITIVITY (%)
<i>SVM classifier #1</i>	77.78	78.08	76.47
<i>SVM classifier #2</i>	70.0	68.49	76.47
<i>SVM classifier #3</i>	76.67	84.93	41.18
<i>Hippocrates-mst</i>	43.33	31.51	94.12

### C. Binary Logical Classifier (BLC)

Analyzing the classification results listed in Table II, we can obtain important information concerning the performance of each classifier.

Concerning the two first classifiers, we observe that both of them have a stable behavior for

both the malignant and benign cases, achieving medium levels of sensitivity and specificity. Although their performances are enough satisfying, they are outperformed by “Hippocrates-mst” as far as sensitivity is concerned.

The third SVM classifier presents a behavior towards the benign cases. Although it achieves the higher specificity, the poor performance in the case of malignant cases deters its use for the classification of unknown cases, as it is expected that many cases of breast cancer may be misclassified.

As far as “Hippocrates-mst” is concerned, we remark that although it achieves the higher sensitivity, it lacks in specificity. That fact was expected, as previous clinical evaluating procedures have indicated that “Hippocrates-mst” suffers from low specificity.

The aim of our work is to propose a new classifier which will combine the decisions of all the previous classifiers in order to optimize the diagnostic performance. Our purpose is to improve the specificity of “Hippocrates-mst” system, maintaining at the same time the sensitivity at high levels. The combination of the three developed SVM classifiers, which are mainly biased towards benignity, may be the substantial factor that will drive to the desirable result.

Our aim was to reinforce the “Hippocrates-mst” classification scheme since it presents high and robust sensitivity while its specificity is low. Thus, we put an effort on keeping the prediction of the rule-based decision tree as the mainstream diagnosis mode, while giving to the aforementioned SVM classifiers the potentiality to act as a reliable second opinion to the “Hippocrates-mst” classifier, which may change the initial diagnosis only under special circumstances where there is strong evidence for benignity.

As a result, we ended in a Binary Logical Classifier, whose topology is shown in Figure 1.

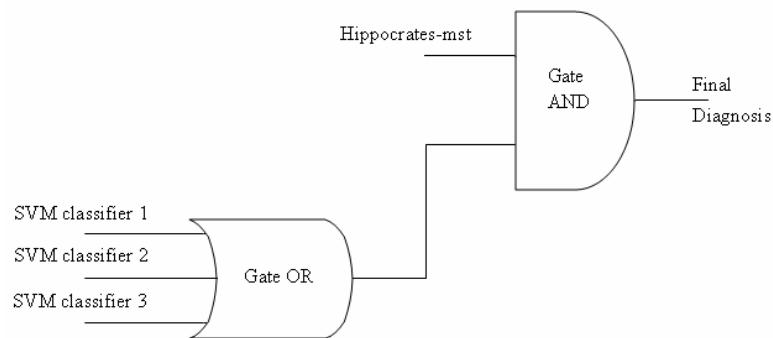


Fig. 1 Binary Logical Classifier (BLC)

We have already mentioned that the SVM approach is a binary problem and, as a result, the predictions of all the SVM classifiers are directly used as inputs to the BLC system. When the prediction for malignancy of an SVM classifier is positive it is considered to be equal to 1. Otherwise, when there is no evidence for malignancy, the SVM prediction is equal to 0. On the contrary, “Hippocrates-mst” calculates a continuous risk percentage for the existence of breast cancer. That risk cannot be used as input to the classifier, as a binary signal is required. For this reason, we convert the risk percentage of “Hippocrates-mst” into a binary variable as follows: when the risk percentage is lower than 35% the patient is not referred for a surgical biopsy and the “Hippocrates-mst” prediction is equal to 0. Otherwise, it is considered that there is important evidence for malignancy and the “Hippocrates-mst” prediction is equal to 1.

The topology of the BLC system meets the specifications we have discussed above. All the SVM classifiers are connected to a gate OR and then the output of this gate is connected with the “Hippocrates-mst” binary diagnosis to a gate AND. We remind that our goal is to improve the low specificity of “Hippocrates-mst”. As a result, when the diagnosis of “Hippocrates-mst” is negative (0) the output of gate AND is also 0. On the other hand, when the diagnosis of

“Hippocrates-mst” is positive (1) the final diagnosis depends on the SVM classifiers: if one at least SVM prediction is 1 the final diagnosis remains also 1. Otherwise, if all three SVM predictions are 0, the final diagnosis is changed to negative, despite the positive prediction of “Hippocrates-mst”.

It is expected that the proposed topology of the BLC system will lead to improvement of the specificity of the CAD system. However, there is the risk of misclassifying some malignant cases, causing a decrease to the sensitivity’s value. We should emphasize that the proposed binary classifier is not able to improve the sensitivity of the CAD system. The maximum expected sensitivity’s value is that achieved by the initial rule-based classifier.

We exploited the second dataset constituted of 90 mammograms to test the performance of the proposed BLC system. We have already seen in the previous section the performances of all the other classifiers. For each case, each individual prediction is applied as input to the BLC system. We are going to compare the results of the initial rule-based classifier used in “Hippocrates-mst”, as listed in Table II, with those of the proposed classifier. Table III lists the classification results for the two classifiers.

TABLE III

CLASSIFICATION OF BREAST MICROCALCIFICATIONS WITH THE BLC SYSTEM AND THE EXISTING “HIPPOCRATES-MST” SYSTEM

	True Negative(TN)	False Positive(FP)	True Positive(TP)	False Negative(FN)
<i>BLC system</i>	48	25	14	3
<i>“Hippocrates-mst”</i>	23	50	16	1

To estimate performance, we used five objective indices: accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) [5]. The values of these

indices obtained from our experiments for both classifiers are shown in Table IV.

TABLE IV

SUMMARY OF PERFORMANCE FOR BLC SYSTEM AND “HIPPOCRATES-MST” SYSTEM

	Accuracy(%)	Specificity (%)	Sensitivity (%)	PPV(%)	NPV(%)
<i>BLC system</i>	68.89	65.75	82.35	35.89	94.12
<i>“Hippocrates-mst”</i>	43.33	31.51	94.12	24.24	95.83

The results listed in the Tables III, IV reveal the potentiality of the BLC system to improve the specificity of the CAD system. The proposed classifier would discourage biopsy to 48 out of 73 cases, achieving that way a reduction of 65.75% on the unnecessary biopsies, while the corresponding percentage of the “Hippocrates-mst” current classification scheme would be only 31.51%.

Additionally, the BLC system is near to the high levels of sensitivity provided by “Hippocrates-mst”, as two more malignant cases are misclassified as benign, leading to a reduction of the sensitivity’s value. Although we wish to achieve the maximum value of sensitivity, the percentage achieved by the binary classifier is a promising result, as it remains one of the greater if compared with the sensitivities values of the others classification schemes, as listed in Table II.

### III. CONCLUSIONS

The aim of this paper was to investigate the effectiveness of new classification schemes and their potentiality to optimize the final diagnosis phase of an already developed computer aided diagnosis system, called “Hippocrates-mst”. The technique currently used in the “Hippocrates-mst” system is a rule-based decision tree, measuring seven characteristics for each

microcalcification, calculating few cluster parameters and finally producing that way the risk distribution of MCs and other relative parameters which drive to the classification of the cluster.

We propose a binary logical classifier which combines multiples SVM with the rule-based classifier of “Hippocrates-mst” for optimal diagnosis of breast cancer. Each individual classifier makes a prediction about the existence of cancer in a selected region of the breast, evaluating the morphological features of the existing microcalcifications and other cluster parameters (i.e. density, polymorphism). At the next level, all the classifier’s decisions are applied as inputs to the BLC system in order to reach to the final diagnosis.

A dataset of 90 mammograms, containing 73 benign cases and 17 malignant, was used for the evaluation of the performance of the classifiers. Data analysis shows that the BLC’s performance is enough satisfying, achieving high levels of both sensitivity and specificity. Actually, BLC outperforms “Hippocrates-mst” in terms of specificity, discouraging biopsy to 48 out of 73 benign cases, providing specificity as high as 65.75%. The corresponding percentage for “Hippocrates-mst” was only 31.51%. As far as the malignant cases are concerned, BLC’s sensitivity was equal to 82.35%, misclassifying 3 out of 17 malignant cases as benign. “Hippocrates-mst” outperformed BLC in terms of sensitivity, achieving a percentage of 94.12%, misclassifying only one case.

Concluding, the use of the proposed binary classifier can be beneficial for the CAD system as it provides an effective solution to the problem of the low specificity. This first laboratory evaluation has shown encouraging performance. Our future goals include design and implementation refinement as well as evaluation of the system with larger number of cases.

#### REFERENCES

- [1] World Health Organization, WHO Statistical Information System.
- [2] American Cancer Society, Cancer Facts and Figures 2004.

- [3] M. Le Gal, G. Chavanne, and D. Pellier, "Diagnostic value of clustered microcalcifications discovered by mammography (apropos of 227 cases with histopathological verification and without a palpable breast tumor)", *Bull cancer*, vol. 71(1), pp. 57-64, 1984.
- [4] M. Le Gal, J.C. Durand, M. Laurent, and D. Pellier, "Management following mammography revealing grouped microcalcifications without palpable tumor", *Nouv Presse Med*, vol. 5(26), pp.1623-7, 1976.
- [5] American College of Radiology (ACR). *Breast Imaging Reporting and Data System (BI-RADS)*. 3rd ed. Reston, Va: American College of Radiology, 1998.
- [6] D.J. Vyborny, "Can computers help radiologists read mammograms", *Radiology*, vol. 191, p.p 315-317, 1994.
- [7] M.L. Geiger, "Computer-aided diagnosis, AAPM/RSNA Categorical Course in Diagnostic Radiology Physics: Physical Aspects of Breast Imaging – Current and Future Considerations", Haus A.and Yaffe M.,eds., pp 249-272, 1999.
- [8] S. Lee, C. Lo, C. Wang, P. Chung, C. Chang, C. Yang, P. Hsu, "A computer aided design mammography screening system for detection and classification of microcalcifications", *Int J Med Inf*, vol. 60(1), pp. 29-57, 2000.
- [9] W. Zhang, K. Doi, M.L. Giger, Y. Wu, R.M. Nishkawa, and R.A. Schmidt, "An improved shift - invariant artificial neural networks for computerized detection of clustered microcalcifications in digital mammograms", *Med. Phys.*, vol. 23, pp. 595-601, 1996.
- [10] M.A. Gavrielides, J.Y. Lo, C.E. Jr Floyd, "Parameter optimization of a computer-aided diagnosis scheme for the segmentation of microcalcification clusters in mammograms", *Med Phys*, vol 29(4), pp.475-83, 2002.
- [11] G. Spyrou, M. Nikolaou, M. Koussaris, A. Tsibanis, S. Vassilaros, and P. Ligomenides, "A system for Computer Aided Early Diagnosis of breast cancer based on microcalcifications analysis", *Res-Systemica*, vol. N°2, Special Issue, December 2002.
- [12] G. Spyrou, K. Koufopoulos, S. Vassilaros and P. Ligomenides, "Computer aided image analysis and classification schemes for the early diagnosis of breast cancer", *Hermis International Journal of Computer Mathematics and its Applications*, vol. 4, pp.175-181, 2003.
- [13] A. Frigas, S. Kapsimalakou, G. Spyrou, K. Koufopoulos, S. Vassilaros, A. Chatzimichael, J. Mantas, P. Ligomenides, "Evaluation of a Breast Cancer Computer Aided Diagnosis System", *Stud Health Technol Inform*, vol. 124, pp.631-6,2006
- [14] G. Spyrou, S. Kapsimalakou A. Frigas K. Koufopoulos S. Vassilaros P. Ligomenides "'Hippocrates-mst": A prototype for Computer-Aided microcalcification analysis and risk assessment for breast cancer", *Medical & Biological Engineering & Computing*, submitted for publication.
- [15] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", *Knowledge Discovery Data Mining*, 2:1-43, 1998.
- [16] N. Cristianini, J. Shawa-Taylor, "An introduction to Support Vector Machines and other kernel based learning methods", Cambridge: UK: Cambridge University Press, 2000.
- [17] J. Platt, "Sequential minimal optimization: a fast algorithm for training support vector machines", Microsoft Research, Technical Report MSR-TR-98-14, 1998.
- [18] Chih-Chung Chang and Chih-Jen Lin, LIBSVM: a library for support vector machines, 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.