Genetic Algorithm optimized SVM for Tumor Prediction in Mammogram

P.Ezhilarasu[1], Dr.J.Suganthi[2]

[1.] Research Scholar, Department of CSE, Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India-641 032. ezhilarasu29@gmail.com
[2.] Professor & Head, Department of Computer Science and Engineering, Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India-641 032.

Abstract: Screening mammography is the best available radiological technique for early breast cancer detection. But due to the huge number of mammograms requiring analysis, radiologists can make false detections. Hence, new solutions regarding automatic detection pertaining to analysis problems should be explored. Microcalcification detection/segmentation helps computerized mammogram screening to classify clusters as either malign or benign. In this paper, Gabor filter with Walsh Hadamard Transform (WHT) is applied to extract microcalcification features from mammograms. This was tested through the use of Mammographic Image Analysis Society (MIAS) mammographic databases. The mammograms were classified using a genetic-based SVM (GA-SVM) model that can automatically determine the optimal parameters, C and Gamma, of SVM with the highest predictive accuracy and generalization ability simultaneously.
[P.Ezhilarasu, Dr.J.Suganthi. **Genetic Algorithm optimized SVM for Tumor Prediction in Mammogram**. *Life Sci J* 2013;10(7s): 460-465] (ISSN:1097-8135). http://www.lifesciencesite.com.

Keywords: Breast cancer; Mammogram; Microcalcification; Gabor filter; Walsh Hadamard Transform; MIAS Database; Genetic Algorithm

1. Introduction

Breast cancer – which cannot be easily detected visually - is a major cause of death among women [1]. Thus, the quality of radiologist' judgment as to whether the suspected region is normal, benign or malignant is crucial. Till date, screening mammography is the best radiological technique available for early breast cancer detection [2]. But as huge volumes of mammograms which require analysis, radiologists could make false detections. Breast region identification is important to improve analysis. Hence, micro-calcification and macro-calcifications in mammograms appear with different shape characteristics and distribution. Thus, detecting the region gives an idea about the nature of diagnosis. There has however been tremendous evolution in mammography process [3-7] over the past few years.

Mammography is characterized by low radiation dose and is the only imaging method accepted for breast cancer screening as it allows radiologists to perform both screening and diagnostic examinations. Screening mammography detects breast cancer in an asymptomatic population, whereas diagnostic mammography ensures an examination of a patient with abnormal clinical elements like breast mass or other disease signs/symptoms. Diagnostic mammography is usually a follow up examination to abnormal screening mammograms. Adoption of mammographic examinations, specially screening mammography, has proved to both increase cancer detection rate and to reduce morbidity/mortality rates.

Mammograms are examined for malignant masses, skin thickening, and microcalcifications. Masses occur often in breast-tissue-dense areas with many shapes like circumscribed, spiculated, lobulated, or ill defined. Circumscribed masses have distinct boundaries and high radiopaque density; spiculated masses have rough, star- shaped boundaries while lobulated masses are irregular. Round, low-density masses with smooth, sharply defined margins are usually benign whereas high-density, stellate, spiculated masses having poorly defined margins are malignant [8].

Tiny calcium deposits accumulated in breast tissue are microcalcifications. They usually appear in mammograms as small bright spots in an inhomogeneous background [9]. Size, shape, density, distribution pattern, and number of microcalcifications are analyzed in the benign and malignant classification phase [10]. Malignant micro-calcifications usually are less than 0.5 mm in diameter and are fine, linear-branching, stellate-modeled, varying in size and shape. Generally, their distribution pattern is as clusters of more than 3 microcalcifications [11]. Though micro-calcifications have high inherent attenuation, their small size makes detection difficult, especially when images are of poor quality. This difficulty increases when analysing young women's mammograms, as they have high-density breast tissue, for predominance of fibro glandular tissues. Breasts that are particularly dense, leads to mammography insensitivity as early malignancy detection is lowered due to the effort

required to locate the cancer within an opaque, uniform background [12]. In such cases, difficulty is mainly because of lowered contrast between micro-calcifications and surrounding dense tissues. Radiologists miss around 10%–30% breast cancer cases. To overcome this, investigators developed computer-aided diagnosis (CAD) schemes [13, 14] to identify potential microcalcification cluster regions in mammograms.

This paper used Gabor filter with Walsh Hadamard Transform to extract microcalcification features from mammograms. This was evaluated through the use of Mammographic Image Analysis Society (MIAS) mammographic databases. The mammograms were classified using a genetic-based SVM (GA-SVM) model that can automatically determine the optimal parameters, C and Gamma, of SVM with the highest predictive accuracy and generalization ability simultaneously.

2. Related Works

Wei et al [15] suggested a microcalcification classification scheme, assisted by content-based mammogram retrieval, for breast cancer diagnosis. A machine learning approach for mammogram retrieval was recently developed where similarity measure between two lesion mammograms was modelled after expert observers. The proposed method was based on adaptively incorporating local proximity information into a classifier which helped improve classification accuracy, resulting in an improved "second opinion" to radiologists. Experimental results on a mammogram database proved that the proposed retrieval-driven approach with an adaptive support vector machine (SVM) improved classification performance from 0.78 to 0.82 with regard to the area under the ROC curve.

Jona et al [16] proposed optimization of the feature set using hybrid of Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) techniques called Genetical Swarm Optimization (GSO) in Digital Mammograms. Though PSO is a good optimization technique, it could be trapped in local minima leading to premature convergence. So, genetic operators are used in PSO to offset such difficulties. Feature selection is of great importance in mammogram diagnosis. Level Co-occurance Matrix (GLCM) texture features are extracted from mammogram. All extracted features do not help detect abnormality in mammograms, and hence feature sets have to be reduced to improve classification accuracy. In this work, experiments are conducted on MiniMIAS database with Support Vector Machine (SVM) classifying mammograms as normal and abnormal. GSO performance is compared with GA and PSO through the use of a Receiver Operating Characteristic (ROC) curve. Results show that GSO convergence is better than both PSO and GA; GSO based SVM (GSO-SVM) classifier showed superior performance with accuracy of 94% that was approximately 1% higher than GA based SVM (GA-SVM) and PSO based SVM (PSO-SVM) classification.

Uma Maheswari et al [17] proposed a hybrid approach for DICOM image classification consisting of feature extraction and classification. The classification includes Multi Linear Discriminant Analysis (MLDA) and Support Vector Machine (SVM). Classification is based on parameter extracted by Gray Level Co-occurrence Matrix (GLCM) and histogram texture feature extraction. Feature is selected through the use of fuzzy rough set and Genetic Algorithm (GA). The proposed approach showed the capability for high approximation and much faster convergence.

Papadopoulos et al [18] suggested an automated method for microcalcification clusters characterization in digitized mammograms, implemented in three stages: (a) cluster detection stage for identification of microcalcifications clusters, (b) feature extraction stage to compare each clusters important features and (c) classification stage. In the classification stage, a rule-based system, a neural network and a SVM were implemented and evaluated the process using receiver operating characteristic analysis. The original feature set was enhanced by adding four rule-based features. In Nijmegen dataset, SVM performance was Az = 0.79 and 0.77 for the original and enhanced feature sets respectively, while in MIAS datasets corresponding characterization scores were Az = 0.81and 0.80. Using neural network classification, Nijmegen dataset's corresponding performance was Az = 0.70 and 0.76 and that of the MIAS dataset, Az = 0.73 and 0.78.

Al Mutaz et al [19] based breast cancer detection on second order statistics. Extraction of textural features of segmented region of interest (ROI) is through gray level co-occurrence matrices (GLCM) extracted from four spatial orientations; horizontal, left diagonal, vertical and right diagonal corresponding to (0o, 45o, 90o and 135o) and two pixel distances for three different block size windows (8x8, 16x16 and 32x32). Results reveal that GLCM at 0o, 45o, 90o and 135o with a window size of 8X8 produces informative features to classify between masses and non-masses. This method achieved accuracy of 91.67% sensitivity and 84.17% specificity comparable to what was reported using state-of-the-art Computer-Aided Detection system.

Hajare et al [20] focused on identifying of relevant, representative and important, discriminate

image features for breast cancer analysis. Gabor wavelets extracted mammogram image features representing normal tissues or benign/malign tumours. These large dimensions (1024x1024) features are applied to Principal Component Analysis (PCA) to lower data dimensionality and then converted into 140x140 pixel size images. Finally, extracted features are classified by proximal support vector machines and features with orientations of 0, π/4, 3π/4, and π/2 and with Gabor filters orientations combine with low and high frequency filters to be compared for recognition rate. Compared to others, Gabor filter obtained rate with low frequency and total orientation is the highest at 84.37%.

Boujelben et al [21] proposed a CADe based on a three-step work flow; detection, analysis and classification. This paper deals with the problem of automatic detection of Region of Interest (ROI) based on Level Set approach depending on edge and region criteria. This approach provides good visual information for radiologists. Then feature extraction through the use of textures characteristics and vector classification using Multilayer Perception (MLP) and k-Nearest Neighbours (KNN) are adopted to distinguish different ACR (American College of Radiology) classifications. Also Digital Database for Screening Mammography (DDSM) is used for experiments and accuracy results varied between 60 % and 70% which were acceptable to radiologists.

3. Methodology

Mini MIAS database

This procedure was evaluated using Mini MIAS database [22] subset, consisting of 161 pairs of medio-lateral-oblique view mammograms. The database images originated from a film-screen mammographic imaging process in the United Kingdom National Breast Screening Program. The films were digitized and corresponding images annotated by radiologists according to breast density, using three classes: Fatty (F) (106 images), Fatty-Glandular (G) (104 images) and Dense-Glandular (D) (112images). Abnormalities were detected and calcifications which were well-defined, speculated or ill-defined masses, architectural distortion or asymmetry were also revealed. Each database image pair is annotated as Symmetric (146 pairs) or Asymmetric (15 pairs) with provisions for each abnormality (benign or malignant) severity.

Walsh Hadamard Transform

Walsh Hadamard Transform (WHT) is an image processing tool. Unser [23] used Hadamard matrices with local transforms like DCT and KLT for texture measurements. Different small size filters and a filter sliding scheme were applied to the spatial domain to evaluate filter effectiveness in texture analysis.

$$W(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}I(x,y)\left[(-1)^{\psi(u,v,x,y)}\right]$$

where I is the image, N is the image size, and ψ determines the transform's parametric kernel function [24]. The kernel function is selected from a diverse set of possibilities.

WHT has many computational advantages. It is a real (not complex) transform, needing addition and subtraction operations alone. When input signal is a set of integer-valued data, it requires only integer operations. There is also a fast algorithm for Walsh transforms through substituting exponential kernel of Fast Fourier transform with the $-1^{\psi(.)}$ kernel of Walsh. The transform matrix, called Hadamard matrix, is saved in binary format leading to lowered memory requirements. WHT is simpler to implement in hardware than other transforms.

Gabor filter

2-D Gabor functions are non-orthogonal wavelets and are a Gaussian modulated by a sinusoid [25]. Gabor filters give optimal resolution in space as well as frequency, thereby simultaneously analysing both domains. A 2-D Gabor filter is defined as follows:

$$J_k(z)=\int\int I(z')\psi_k(z-z')dz'$$

for Image I(z), and $\psi_k(z)$ defined by

$$\psi_k(z)=k^Tk/\sigma^2\exp\left(\left(-k^Tk/2\sigma^2\right)*z^Tz\right)$$
$$\left(\exp\left(ik^Tz\right)-\exp\left(-\sigma^2/2\right)\right)$$

where z = x,y and k is characteristic wave vector.

Support Vector Machine (SVM)

SVM classifier is based on structural risk minimization principle searching for a hyperplane maximizing distance from it to the nearest, each class examples [26]. SVM's can map linearly inseparable data into higher dimensional space when linear separation is possible. SVM tries to locate a decision hyperplane written as

$$w.\phi(x_i)+b=0$$

where w and b are classification model parameters and Φ is mapping a certain higher dimensional space in which x_i undergoes linear separation. Training task for the model as an optimization task is formalized as

$$\min_w \left(\|w\|^2 / 2 \right) \text{ subject to } y_i \left(w.\phi(x_i) + b \right) \geq 1$$

As the task is a convex optimization problem it is rewritten as an optimization formula to a Lagrangian function.

$$L(w, b, \lambda)$$

and derive its dual form $\tilde{L}(\lambda)$ as

$$L(w, b, \lambda) = \left(\|w\|^2 / 2 \right) - \sum_{i=1}^{N} \lambda_i \left(y_i \left(w.\phi(x_i) + b \right) - 1 \right)$$

$$\tilde{L}(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \lambda_i \lambda_j y_i y_j \phi(x_i).\phi(x_j)$$

subject to the Karush–Kuhn–Tucker conditions

$$\lambda_i \geq 0, \ \lambda_i \left\{ y_i \left(w.\phi(x_i) + b \right) - 1 \right\} = 0$$

where λ_i Lagrangian multipliers are calculated by exploiting quadratic programming techniques or faster heuristic algorithms. After calculation, the model parameters w and b are determined using the fact that

$$\phi(x_i).\phi(x_j) = K(x_i, x_j)$$

K(.,.) is a kernel function. After multipliers and model parameters determination, a new input test example x_{new} is classified by investigating the hyper plane's side it resides in. In summary, this non-linear SVM classifier's overall decision function h can be written as for a predefined kernel function K where sign (.) is the sign function. This paper considers linear and Gaussian radial basis function (RBF) kernels.

To design an effective SVM model, SVM parameters values have to be selected carefully in advance [27, 28]. These parameters include: (1) regularization parameter C, which determines the tradeoff cost between minimizing training error and minimizing model complexity; (2) Kernel function's Gamma parameter that defines non-linear mapping from input space to a high-dimensional feature space.

Proposed SVM-GA

Genetic algorithms (GAs) are successfully applied to varied optimization problems [29, 30]. GA's are appropriate for concurrent manipulation of models with varying resolutions, and structures as they search non-linear solution spaces without needing gradient information or a priori knowledge on model characteristics. The problem of binary coding is in the fact that a long string occupies computer memory though only a few bits are actually involved in crossover/mutation operations. This is particularly so when many parameters are to be adjusted in the same problem and the final result need a higher precision. To overcome inefficient occupation of computer memory, underlying real-valued crossover and mutation algorithms are employed. In contrast to the binary genetic algorithm (BGA), the real-valued genetic algorithm (RGA) uses real value as a chromosome parameter in populations without coding or encoding before the fitness values of individuals is calculated.

In the proposed GA-SVM model, SVM parameters are dynamically optimized by implementing a RGA evolutionary process. The SVM model then performs prediction using optimal values. The RGA attempts to search optimal values to ensure that SVM fits the dataset. SVM's parameter's optimal values are sought by GAs with a randomly generated initial population of chromosomes. The values of C and Gamma parameters are directly coded in chromosomes with real-valued data. The proposed model implements a tournament selection method to select chromosomes. Crossover method and boundary mutation modified the chromosome. Each generation's single best chromosome survives to the next generation.

Selection, crossover, and mutation operators are used to generate the existing population's offspring. Tournament selection is adopted to decide whether a chromosome survives to the next generation. The chromosomes that do are placed in a matting pool for crossover and mutation. Random points are selected in the chromosomes to be crossed over to form an offspring. The mutation operation follows crossover to determine whether a chromosome should be mutated in the next generation. A uniform mutation method is applied in this study. Newly crossed chromosomes are then combined with the remaining chromosomes to create a new population. The mutation operation follows the crossover operation and determines whether a chromosome should be mutated in the next generation. In this study, uniform mutation method is applied.

4. Results

The mammograms from the Mini MIAS database were classified as microcalcified and non-microcalcified. Features are extracted from the mammograms using Gabor filter with WHT. A subset of Mini MIAS containing 84 mammograms was used for evaluation. The following Table 1 shows the summary of the experimental results for the generalized neural network classifier. Figure 1 shows the classification accuracy obtained by various techniques.

Table 1: Summary of the Experimental Results for Various Techniques

Technique Used	Classification Accuracy	RMSE
SVM-RBF C=0.001, Gamma=0.5	0.928571	0.2673
SVM-RBF C=0.001, Gamma=0.4	0.916667	0.287
SVM-RBF C=0.001, Gamma=0.3	0.904762	0.3086
GA-SVM-Polynomial kernel	0.916667	0.2942
GA-SVM-RBF	0.940476	0.2176

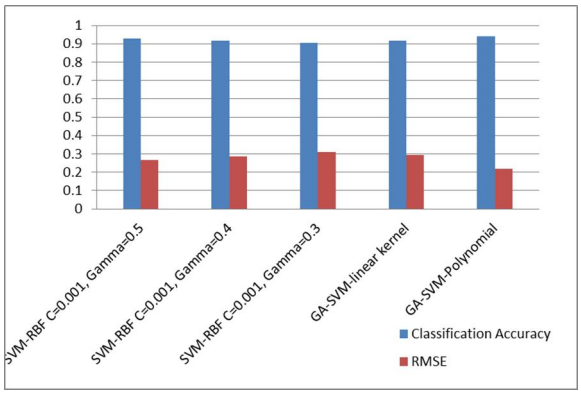

Figure 1: Classification Accuracy and RMSE obtained by various techniques

It is observed from Figure 2 the proposed genetic optimized SVM for RBF achieves the best classification accuracy of 94.05%. The optimizing of the kernel parameters C and Gamma improves the classification.

Table 2 tabulates the precision, recall and f-measure of various techniques. Figure 3 shows the graph of precision and recall.

Table 2: Precision, Recall of Various Kernels

Technique Used	Precision	Recall
SVM-RBF C=0.001, Gamma=0.5	0.944444	0.916667
SVM-RBF C=0.001, Gamma=0.4	0.936364	0.902778
SVM-RBF C=0.001, Gamma=0.3	0.928571	0.888889
GA-SVM-Polynomial kernel	0.927267	0.90625
GA-SVM-RBF	0.95283	0.930556

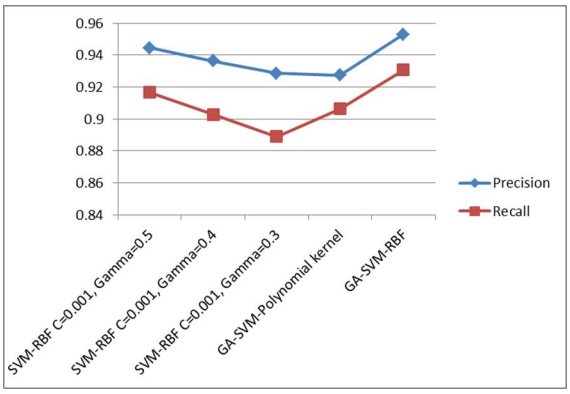

Figure 2: Precision and Recall of Various Techniques

The best precision and recall was achieved for the proposed GA-SVM RBF kernel.

5. Conclusion

A link between the presence of clustered microcalcifications and occurrence of breast cancer was discovered implying that early micro-calcifications detection in mammograms can increase survival chances for those with breast cancer. An automatic tumour detection system is developed to help radiologists provide an accurate diagnosis. Microcalcification presence is crucial to diagnose breast cancer in clinical practice. Detection and diagnosis of breast cancer in its early stage rapidly increase chances for successful treatment and complete patient recovery. This paper investigates the efficiency of genetic optimized SVM for classifying mammograms. In this paper, Gabor filter with Walsh Hadamard Transform is used to extract microcalcification features from mammograms. A subset of Mini MIAS mammograms was used for evaluation. The best precision and recall was achieved for the proposed GA-SVM RBF kernel.

References

1. R. Nishikawa, (2007) "Current status and future directions of computer-aided diagnosis in mammography". In Computerized Medical Imaging and Graphics, vol 31, pp 224-235, 2007.
2. M. Siddiqui, M. Anand& N. Muthur, (2005) "Biomonitoring of organochlorines in women with benign a malignant breast disease". Environmental Research, Tech Rep, pp 250-257, 2005.
3. Mohanty, A. K., Beberta, S., &Lenka, S. K. (2011). Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammogram. International Journal of Engineering Research and Applications, 1, 687-693.
4. McLeod, P., Verma, B., & Park, M. (2009, December). Soft clustering and support vector machine based technique for the classification of

abnormalities in digital mammograms. In Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2009 5th International Conference on (pp. 185-189). IEEE.
5. Shao, Y. Z., Liu, L. Z., Bie, M. J., Li, C. C., Wu, Y. P., Xie, X. M., & Li, L. (2011). Characterizing the Clustered Microcalcifications on Mammograms to Predict the Pathological Classification and Grading: A Mathematical Modeling Approach. Journal of Digital Imaging, 24(5), 764-771.
6. Hussain, M., Khan, S., Muhammad, G., &Bebis, G. (2012, November). A Comparison of Different Gabor features for Mass Classification in Mammography. In Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on (pp. 142-148). IEEE.
7. Jian, W., Sun, X., &Luo, S. (2012). Computer-aided diagnosis of breast microcalcifications based on dual-tree complex wavelet transform. Biomedical engineering online, 11(1), 1-12.
8. Cheng, H. D., Shi, X. J., Min, R., Hu, L. M., Cai, X. P., & Du, H. N. (2006). Approaches for automated detection and classification of masses in mammograms. Pattern recognition, 39(4), 646-668.
9. G. Bocciglione, A. Chainese and A. Picariello, "Computer aided detection of microclassification in digital mammograms," Comp. Biol. Med., 30: 267-286, 2000.
10. Mousa, R., Munib, Q., &Moussa, A. (2005). Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural. Expert systems with Applications, 28(4), 713-723.
11. Sickles, E. A. (1986). Breast calcifications: mammographic evaluation. Radiology, 160(2), 289-293.
12. Yu, S. N., Li, K. Y., & Huang, Y. K. (2006). Detection of microcalcifications in digital mammograms using wavelet filter and Markov random field model. Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society, 30(3), 163.
13. K. Doi, H. MacMahon, S. Katsuragawa, R. M. Nishikawa, and Y. Jiang, "Computer-aided diagnosis in radiology: Potential and pitfall," Eur. J. Radiol., vol. 31, pp. 97–109, 1999.
14. H. P. Chan, K. Doi, C. J. Vyborny, K. L. Lam, and R. A. Schmidt, "Computer-aided detection of microcalcifications in mammograms: Methodology and preliminary clinical study," Invest. Radiol., vol. 23, pp. 664–671, 1988.
15. Wei, L., Yang, Y., & Nishikawa, R. M. (2009). Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. Pattern recognition, 42(6), 1126-1132.
16. Jona, J. B., & Nagaveni, N. (2012). Genetical Swarm Optimization for Feature Set Reduction in Digital Mammograms. European Journal of Scientific Research, 87(2), 227-235.
17. Umamaheswari, J., &Radhamani, G. (2011). A Hybrid Approach for Classification of DICOM Image. World of Computer Science and Information Technology Journal (WCSIT), 1(8), 364-369.
18. Papadopoulos, A., Fotiadis, D. I., &Likas, A. (2005). Characterization of clustered microcalcifications in digitized mammograms using neural networks and support vector machines. Artificial Intelligence in Medicine, 34(2), 141-150.
19. Al Mutaz, M. A., Dress, S., &Zaki, N. (2011). Detection of Masses in Digital Mammogram Using Second Order Statistics and Artificial Neural Network. International Journal of Computer Science & Information Technology (IJCSIT), 3(3).
20. Hajare, P. S., & Dixit, V. V. (2012). Breast Tissue Classification Using Gabor Filter, PCA and Support Vector Machine.
21. Boujelben, A., Tmar, H., Mnif, J., &Abid, M. (2011). Automatic application level set approach in detection calcifications in mammographic image. arXiv preprint arXiv:1109.0138.
22. J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, et al., The mammographic image analysis society digital mammogram database, in: ExerptaMedica. International Congress Series, 1994, pp. 375–378.
23. M. Unser. Local linear transform for texture measurements. Signal Processing, 11:61–79, 1986.
24. K. Beauchamp. Applications of Walsh and related functions. Academic Press, 1984.
25. Weldon, T. P., Higgins, W. E., & Dunn, D. F. (1996). Efficient Gabor filter design for texture segmentation. Pattern Recognition, 29(12), 2005-2015.
26. Keerthi, S. S., Shevade, S. K., Bhattacharyya, C., & Murthy, K. R. K. (2001). Improvements to Platt's SMO algorithm for SVM classifier design. Neural Computation, 13(3), 637-649.
27. Duan, K., Keerthi, S. S., & Poo, A. N. (2003). Evaluation of simple performance measures for tuning SVM hyperparameters. Neurocomputing, 51(1–4), 41–59
28. Min, J. H., & Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Systems with Applications, 28(4), 603–614
29. Goldberg, D. E. (1989). Genetic algorithms in search, optimization &machine learning. Reading, MA: Addision-Wesley.
30. Grefenstette, J. J. (1986). Optimization of control parameters for genetic algorithms. IEEE Transactions on System, Man and Cybernetics, 16(1),122–128

1/8/2013