Journal of Theoretical and Applied Information Technology
20th September 2013. Vol. 55 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

OPTIMIZING OF SVM FOR CT CLASSIFICATION

[1.] N.T.RENUKADEVI [2] Dr.P.THANGARAJ

[1]Asst. Prof. Dept. of CT-UG, Kongu Engineering College, Perundurai , Tamil nadu,
[2]Prof. and Head, Dept. of CSE, Bannariamman Institute of Technology Trust
[1]E-mail: renuka.kec@gmail.com

ABSTRACT

An automated classification of Computerized tomography (CT) images method uses Coiflet wavelets to extract features as input for the classifiers. Support Vector Machine (SVM) module is used to classify the images into different classes. Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) are used to optimize the parameters C and gamma of the RBF kernel in SVM. Parameter selection is thought of as an optimization problem where search techniques are used to maximize SVM performance. The observed results are considerably better than the results achieved by employing Support Vector Machine.

Keywords: *Image Classification, Computed Tomography (CT), Coiflet wavelets, Support Vector Machine (SVM), Radial Basis Function (RBF), Particle Swarm Optimization (PSO), Genetic Algorithm (GA)*

1. INTRODUCTION

Computer-aided diagnosis (CAD) is a major research subject in medical imaging and diagnostic radiology. A Computerized tomography (CT) scan produces images of body structures like internal organs, blood vessels, bones and tumours [2]. Various CT scan types which investigate specific body areas include: Head scans, Abdominal scans, Vascular scans, Bone scans and others. CT scans are a rapid imaging modality with excellent image resolution, ensuring quicker and highly accurate diagnosis of patients over many clinical indications. Data from one scan can later be manipulated to provide multi planar/3D reconstructions.

This paper presents a method for automatic CT image classification of various types which has 3 major steps: 1. Feature extraction using Coiflet wavelets is from CT images. 2. Extracted features are classified with Support Vector Machine; 3. SVM parameters are optimized. The paper investigates optimization effect using Particle Swarm Optimization (PSO) and genetic algorithm (GA).

SVM is a data classification [3, 4] technique. Though Neural Networks are considered easier to uses, it leads to occasional unsatisfactory results. Classification usually based on training/testing data consisting of data instances. Each training set instance has a target value and many attributes. SVM aims to produce a model to predict data instances target value in testing sets given as attributes [5] alone.

SVM classification is a Supervised Learning example. Known labels indicate whether the system performs correctly. This information has a chosen response which validates system accuracy or helps it to perform better. SVM classification involves identification intimately connected to known classes, called feature selection/feature extraction. Feature selection and SVM classification combined are useful even when it is unnecessary to predict samples. They identify key sets involved in class distinguishing processes.

SVM's strength is that it is easy to train. Unlike neural networks it has no local optimal. It scales to high dimensional data relatively well, and the trade-off between classifier complexity and error is controllable. Its weakness includes requiring a good kernel function. Though SVM accomplishments are governed by kernel parameter choice and regularization of parameters. Here, parameter selection is an optimization problem where a search technique is used to maximize the SVM performance [6-8] to optimal parameters. Radial Basis Function (RBF) networks were studied due to their good generalization and universal approximation through RBF nodes use in the hidden layer. Such techniques have many advantages. The approach is systematic and theoretically motivated properly. Learning machine is constructed using most informative data patterns. Because of clear

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

data dependence it is easier to explain/interpret the model. Also, data cleaning [9] can improve performance. Learning process involves a cost function optimization that is provably convex. This is in contrast to the neural network approaches where presence of false local minima in the error function complicates the learning process.

SVM parameters are selected through calculation of various parameter combinations and using that which achieves the best performance for a specific dataset. To automatize search, many search and optimization techniques are used [10, 11]. The present work implements PSO and GA for SVM's RBF kernel parameter selection. "Swarm intelligence" is generally used in optimization to maximize/minimize cost function by searching for variables set which is called optimization. Swarm optimization is dependent on the collective behaviour of tees/ants, or social behavior during bird flocking and fish schooling. Particle Swarm Optimization algorithm is a population-based stochastic search algorithm efficient in solving complex non-linear optimization problems [12]. PSO is popular as it can be implemented easily and is inexpensive computationally.

The concepts of evolution, selection and mutation helped develop evolutionary programming. Holland introduced Genetic Algorithm (GA) concept as a principle of the Darwinian theory of evolution through natural biology [13]. GA learning methods are computational model based via natural adaptation and evolution. Procedures modeling population genetics and survival of the fittest help improve performance. GA has applications in solving problems that requiring effective/efficient search in business, scientific and engineering like neural networks architecture synthesis, travelling salesman problem, scheduling, numerical optimization and pattern recognition and image processing.

2. RELATED WORKS

Padma et al [14] compared dominant grey level run length feature extraction method with wavelet based texture feature extraction and SGLDM methods. A high gray level run length texture feature set is derived from a to-be-selected image's region of interest (ROI). Optimal texture features are chosen through a GA. Selected optimal run length texture features are fed to SVM classifier to both classify and segment a tumor from brain CT images. The method is applied 120 images CT data with normal and abnormal tumor images. Results are compared with radiologist's truth. A quantitative analysis between this and segmented tumour is presented regarding classification accuracy. From analysis/performance measures like classification accuracy, brain tumor classification/segmentation is best with a SVM which has a dominant run length feature extraction method. It is better than using a SVM with wavelet based texture feature extraction method and SVM with SGLDM methods. In a bid to improve computing efficiency an attempt has been made to use SVM as it chooses a suitable feature extraction method for accurate brain tumour classification/segmentation in CT images. Average accuracy of above 97% was achieved through this classification/segmentation algorithm.

Jiang et al [15] proposed A liver cancer identification method based on PSO-SVM was proposed by Jiang et al [15]. The region of interest (ROI) is determined first by Lazy-Snapping, and texture features are extracted from ROI. Later, F-score algorithm selects relevant features, with the liver cancer classifier being designed by combining parallel Support Vector Machine (SVM) with Particle Swarm Optimization (PSO) algorithm. PSO automatically chooses SVM parameters and its advantage is that it makes parameter choice more objective avoiding traditional SVM's randomicity and subjectivity and where parameters are based on trial and error methods. Experiment on real-world datasets proved that parallel PSO-SVM training algorithm improved liver cancer prediction accuracy..

A hybrid classification approach for brain tissues in magnetic resonance images (MRI) based on GA and SVM was proposed by Kharratet al [16]. A feature set based on wavelet is first derived. Spatial gray level dependence method (SGLDM) extracts optimal texture features which become input to a SVM classifier. Features choice which is a big issue in classification is solved through user of GA. Optimal features classify brain tissues into normal, benign or malignant categories. The algorithm's performance is evaluated on a many brain tumor images.

An adaptive chaotic particle swarm optimization (ACPSO) for optimizing parameters was presented by Zhang et al [17]. The methodology classifies MR brain image as normal/abnormal. Wavelet transforms extract features and principal component analysis (PCA) reduces features dimensions. Feed forward neural network then classifies features. K-fold stratified cross validation is applied to enhance generalization. The proposed method was evaluated

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A total of 160 images (20 normal, 140 abnormal) were used for evaluation resulting in 98.75% classification accuracy.

Wu et al [18] simultaneously combined feature selection and parameter setting in this study where ultrasound breast tumors underwent automatic segmentation by a level set method. Texture features auto-covariance and morphologic features were extracted through the use of a GA to detect significant features and determine near-optimal SVM parameters to identify the tumor as benign/malignant.The proposed CAD system differentiates benign from malignant breast tumors with great accuracy and reduced feature extraction time. Based on results the proposed CAD system's accuracy for classifying breast tumors was 95.24%, and computing time to calculate breast tumor image features was only 8% of that of features selection-less system. Also, the time to locate a (near) optimal classification model is greatly lesser than that of grid search. It is useful to reduce biopsies of benign lesions and offers a second reading to ensure inexperienced physicians avoid wrong diagnosis.

3. METHODOLOGIES

Coiflets are discrete wavelets which have scaling functions and vanishing moments [19]. Wavelet is near symmetric with N/3 vanishing moments and N/3-1 scaling functions. The image uses 2nd order Coiflet wavelet transformation with four decomposition levels to transform. Every decomposition level ensures that wavelet transformation divides signals into approximation signals and detail coefficients.

$$h(z)=\sum_k h_k z^{-k} \text{ and } g(z)=zh\left(-z^{-1}\right)$$

where h and g be the wavelet decomposition (analysis) filters.

Given a features set is represented in space, features are mapped by SVM non-linearly into n dimensional feature space with a features set represented in space. When a high computation kernel is introduced, the algorithm uses inputs as scalar products. Classification is solved by translating the issue into a convex quadratic optimization problem with convexity ensuring a clear solution [20]. An attribute is a predictor variable in a SVM and feature a transformed attribute. A feature set describing an example is a vector and it also defines a hyperplane. SVM locates an optimal hyperplane separating vector clusters with an attributes class on the plane's one side with the other on the other side. The distance between hyperplane and support vectors is the margin. SVM analysis orients margin so that space between it and support vectors is maximized.

Given a training set of $(x_i, y_i), i1,2,....l$ where $x_i \in R^n$ and $y \in \{1,-1\}^l$, SVM has to solve the optimization problem [21] of:

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

Subject to $y_i\left(w^T \phi(x_i)+b\right) \geq 1-\xi_i$ and $\xi_i \geq 0$.

The function ϕ maps the vectors x_i in higher dimensional space. C>0 is penalty parameter of the error term.

A kernel function is defined as $K\left(x_i, x_j\right)=\phi(x_i)^T \phi\left(x_j\right)$. The Radial Basis function is given as follows:

$$K\left(x_i, x_j\right)=\exp\left(-\gamma\left\|x_i-x_j\right\|^2\right) ,\gamma>0$$

SVM classification accuracy is improved by proper parameter setting. SVM model with RBF kernel determines 2 parameters: C and gamma (γ). C and γ values influence SVM learning performance. To optimize C and γ parameters, PSO [12] are executes a search for optimal combination (C, γ). The objective of evaluating parameter quality combination is based on Root Mean Squared Error (RMSE) achieved by SVM in a 10-fold cross validation experiment. Hence, PSO finds parameters combination with lowest RMSE.

Each particle *i* represents a parameter combination indicating the particle's position in search space. Particle velocity indicates the direction of particle search. The PSO algorithm regularly updates particle's position and velocity in each iteration, leading to best regions in search space. Particle velocity and position are updated as follows:

$$v_i^d = wv_i^d + c_1 r_1\left(p_i^d - x_i^d\right)+c_2 r_2\left(p_g^d - x_i^d\right)$$

$$x_i^d = x_i^d + v_i^d$$

where w is Inertia weight; d represents dimensions number; i the size of the population; the two "best" values - pbest and gbest - of a particle where 'pbest' $\left(p_i^d\right)$ is the best solution achieved by particle till then and 'gbest' $\left(p_g^d\right)$ is best value obtained till then by any population particle in the population; c_1, c_2 are positive constants and r1 and r2 random values with a value between [0, 1].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

GA are a family of evolution inspired computational models which encode a possible solution to specific problems on a simple chromosome-like data structure applying recombination operators to structures for critical information preservation. GAis usually thought of as function optimizers, though problems to which GA can be applied.

GA implementation begins with a population of chromosomes one of which evaluates structures and allocates reproductive opportunities to ensure that chromosomes representing an improved solution to a problem get more chances to `reproduce' than those offering poorer solutions. A solution's 'goodness' is defined with regard to its current population.

Major steps involved are generation of population solutions, locating objective and fitness functions and genetic operators' application all of which are described below.

When a problem is encoded in a chromosomal manner and a fitness measure to discriminate between good and bad solutions has been selected, solutions to the search problem start evolving through as follows:

1. *Initialization.* A randomlygenerated initial population of candidate solutions across the search space is created.

2. *Evaluations.* Once a population is initialized/offspring population, candidate solutions fitness values are evaluated.

3. *Selection.* Selection allocates more solution copies with higher fitness values, imposing a survival-of-the-fittest mechanism on candidate solutions.

4. *Recombination.* Recombination combines parts of two or more parental solutions for the creation of new, improved solutions (offspring). Many are the ways to achieve this. Competent performance is based on a proper recombination mechanism.

5. *Mutation.* While two or more parental chromosomes combination operators, mutation offers a local but randomly modified solution. There are also many mutation variations and it usually, involves one or more changes in an individual's trait/traits. In other words, mutation performs a random walk near a candidate solution.

6. *Replacement.* Offspring population created through selection, recombination, and mutation substitute's original parental population.

7. Repeat steps 2–6 till a terminating condition is met.

4. RESULTS AND DISCUSSION

Experiments were conducted with 150 CT scans of brain, chest and colon images. Coiflet wavelet extracted features. Experiments evaluated classification accuracy for SVM-RBF, with PSO and GA. Experiments were undertaken for 10-fold cross validation. Classification accuracy and root mean square error (RMSE) achieved are tabulated in Table 1. Figure 1 reveals classification accuracy and Figure 4 the RMSE.

Table1: Classification Accuracy and RMSE

Classifier	Classification Accuracy %	RMSE
Naïve Bayes	90	0.2582
SVM-RBF	88.67	0.265
SVM, PSO	90.67	0.214
SVM, GA	89.33	0.246

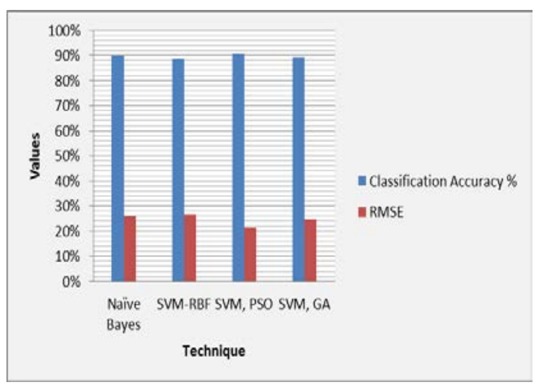

Figure 1: Classification Accuracy and RMSE

It is observed from the Table and Figures that the PSO improves classification accuracy and reduces the RMSE significantly. Table 2 tabulates the precision and recall achieved.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Precision and Recall

Classifier	Precision	Recall
Naïve Bayes	0.900	0.900
SVM-RBF	0.887	0.887
SVM, PSO	0.908	0.907
SVM, GA	0.894	0.893

The precision and recall are high for the PSO optimization when compared to GA.

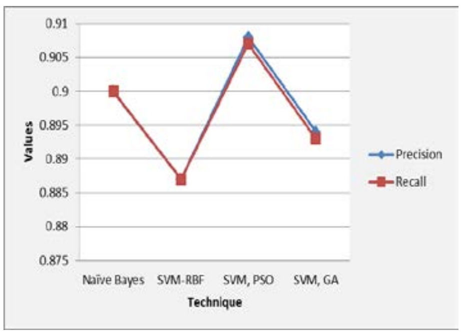

Figure 2: Precision and Recall

5. CONCLUSION

This paper presents results of a comparative study on SVM optimization with Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) to classify CT images. Naïve Bayes and SVM-RBF are used for comparison. To improve SVM-RBF performance to classify CT images, SVM parameters C and Gamma (γ) are optimized. PSO and GA are implemented to select values for 2 SVM parameters for classification problems. Experiments were undertaken for 10-fold cross validation. Classification accuracy and root mean square error (RMSE) for PSO are higher when compared to GA.

REFERENCES

[1]. Doi, K. (2009, October). Computer-aided diagnosis in radiological imaging: current status and future challenges. In Sixth International Symposium on Multispectral Image Processing and Pattern Recognition (pp. 74971A-74971A). International Society for Optics and Photonics.

[2]. Kalender, W. A. (2011). Computed tomography. Wiley-VCH.

[3]. Balaji, V., &Duraisamy, V. (2012). Improved content based image retrieval using SMO and SVM classification technique. European Journal of Scientific Research, 69(4), 560-564.

[4]. Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). A practical guide to support vector classification.

[5]. Ertekin, S., Huang, J., Bottou, L., & Giles, L. (2007, November). Learning on the border: active learning in imbalanced data classification. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management (pp. 127-136). ACM.

[6]. Lorena and A. de Carvalho, "Evolutionary tuning of svm parameter values in multiclass problems," Neurocomputing, vol. 71, pp. 16–18, 2008.

[7]. X. Guo, J. Yang, C. Wu, C. Wang, and Y. Liang, "A novel ls-svms hyper-parameter selection based on particle swarm optimization," Neurocomputing, vol. 71, pp. 3211–3215, 2008.

[8]. S. Lessmann, R. Stahlbock, and S. Crone, "Genetic algorithms for support vector machine model selection," in International Joint Conference on Neural Networks, 2006, pp. 3063–3069.

[9]. Guyon, I., Matic, N., and Vapnik, V. (1996), "Discovering informativepatterns and data cleaning," in Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (Eds.), Advances inKnowledge Discovery and Data Mining, MIT Press, pp. 181-203.

[10]. F. Friedrichs and C. Igel, "Evolutionary tuning of multiple svm parameters," Neurocomputing, vol. 64, pp. 107–117, 2005.

[11]. B. de Souza, A. de Carvalho, and R. Ishii, "Multiclass svm model selection using particle swarm optimization," in Sixth International Conference on Hybrid Intelligent Systems, 2006, pp. 441–446.

[12]. Venter, G., &Sobieszczanski-Sobieski, J. (2003). Particle swarm optimization. AIAA journal, 41(8), 1583-1589.

[13]. Holland, J. H. (1992). Genetic algorithms. Scientific american, 267(1), 66-72.

[14]. Padma, A., &Sukanesh, R. (2011). Automatic classification and segmentation of brain tumor in CT images using optimal dominant gray level run length texture features. Int J Adv Comp SciAppl, 2(10), 53-59.

[15]. Jiang, H., Tang, F., & Zhang, X. (2010, December). Liver cancer identification based on PSO-SVM model. In Control Automation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Robotics & Vision (ICARCV), 2010 11th International Conference on (pp. 2519-2523). IEEE.

[16]. Kharrat, A., Gasmi, K., Messaoud, M. B., Benamrane, N., &Abid, M. (2010). A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine. Leonardo J. Sci, 17, 71-82.

[17]. Zhang, Y., Wang, S., & Wu, L. (2010). A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO. Progress in Electromagnetics Research, 109, 325-343.

[18]. Wu, W. J., Lin, S. W., & Moon, W. K. (2012). Combining support vector machine with genetic algorithm to classify ultrasound breast tumor images. Computerized Medical Imaging and Graphics.

[19]. 16. Ingrid Daubechies, Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, 1992

[20]. B. E. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, pages 144-152. ACM Press, 1992.

[21]. Jrad, N., Congedo, M., Phlypo, R., Rousseau, S., Flamary, R., Yger, F., &Rakotomamonjy, A. (2011). sw-SVM: sensor weighting support vector machines for EEG-based brain–computer interfaces. Journal of neural engineering, 8(5), 056004.