Quadratic Program Optimization using Support Vector Machine for CT Brain Image Classification

J Umamaheswari[1] and Dr.G.Radhamani[2]

**[1] Research Scholar, Department of Computer Science,
Dr. G.R.D. College of Science, Coimbatore, Tamilnadu, India**

**[2] Director, Department of Computer Science,
Dr. G.R.D. College of Science, Coimbatore, Tamilnadu, India**

Abstract

In this paper, an efficient Computer Tomography (CT) image classification using Support Vector Machine (SVM) with optimized quadratic programming methodology is proposed. Due to manual interpretation of brain images based on visual examination by radiologist/physician that cause incorrect diagnosis, when a large number of CT images are analyzed. To avoid the human error, an automated optimized classification system is proposed for abnormal CT image identification. This is an automated system for content based image retrieval with better classifier accuracy and prediction time. SVM classifier can accurately train up the data's as normal and abnormal brains interpreted manually by the user. The system can retrieve more number of images present in the query data base. The proposed classifier is analyzed with existing Sequential Minimal Optimization (SMO) and K Nearest Neighbour classifier KNN). From the experimental analysis, the proposed classifier outperforms all other classifier taken for examination.

Keywords: *CT Brain image classification, Quadratic programming, Linear SVM, SMO, KNN.*

1. Introduction

Medical imaging is in need of automatic processing for efficientdiagnosis with higher accuracy. Medical images are normally attained by X-rays, Magnetic Resonance (MRI) Imaging and Computer Tomography (CT). CT is used to analyze the diseases in medical environment. [1, 2]. For diagnosis and monitoring, the patient's images plays important tools in medical imaging. A systematic analysis for diagnosis monitoring the disease is one of the most important tools in medicine since it provides a method for diagnosis, monitoring disease of patients having the advantage of being a very fast detection. As new image acquisition devices are constantly being developed, to increase efficiency and produce more accurate information, and data storage, a steady growth of the number of medical images produced can be easily inferred [6].

The lack of systematic research on features extracted and their ole to the classification results forces researchers to select features arbitrarily as input to their systems. Content Based Image Retrieval (CBIR), a query with an MRI image may retrieve some PET images too, although the user [14]. In the last couple of years, several useful research prototypes in the medical domain have been successfully implemented with each work catering to specific areas [3,4]. The effectiveness of a CBIR system depends on the choice Improved Content Based Image Retrieval using SMO and SVM [5, 6].

In this paper, Quadratic Program optimization (QPO) technique is used to classify the brain images using SVM. The normal and abnormal image feature is extracted using MATLAB. The system is used for separating the data collection containing both normal and abnormal images. This results in significant cost and time saving by avoiding the classification task for large number of databases. The performance of the classifier is examined using linear SVM, SMO and KNN method [7,11,14]. The motivation behind this paper is to develop a classification process for evaluating the classification performance of different classifiers in terms of statistical measures.

The paper is organized as follows. Section2 discusses the proposed methodology using QPO classification using SVM for MRI image classification. Implementation of the proposed approach and the performance of different classifiers re-discussed in Section 3. The paper is concluded in Section 4.

2. Methodology

The image classification refers to the grouping of images into certain categories. This problem seems to be difficult for the machines to classify. Therefore, image classification using CBIR system is a challenging task for physician to diagnosis purpose [8]. The proposed method consists of Feature extraction and classification. Figure 1 shows the frame work of the proposed system.

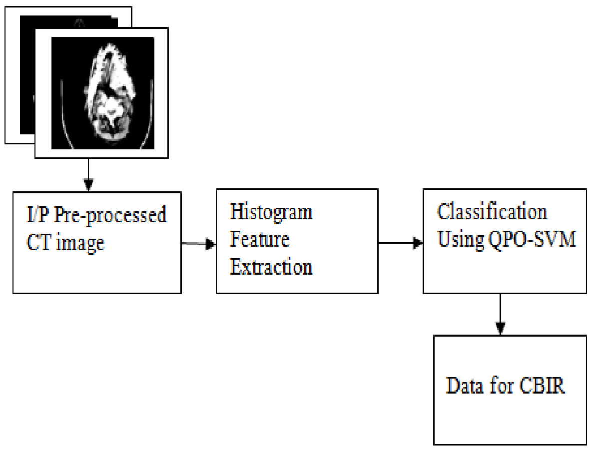

Fig .1. Frame work of the proposed system

2.1 Feature Extraction

To analyze the set of features, the process of transforming the input data into the set of features is known as Feature extraction. The extraction of the feature is the most important step for representing all the relevant information from the input images to accomplish good quality result and to proceed further process. In our paper the process of feature extraction is followed accordingly: 1. To identify the boundary regions by contour method,
2. To extract diagnostic data from the boundary regions. To find the irregularity of the image, the classification is a critical task to find the accurate results. So to improve the accuracy, the related features are extracted for better classification. The intensity histogram features are analyzed using mean, variance, skewness, kurtosis, entropy and energy. The categorization of images into normal and abnormal is done using statistical features of images such as co-occurrence based textural features of images such as energy, entropy, difference moment, inverse difference moment and correlation. The extracted features are then given to SVM for quadratic program problem optimization.

2.2 Classification using Support Vector Machine

SVM often provides better classification results that are widely used for pattern recognition methods, such as the maximum likelihood and neural network classifiers. SVM

are set of related supervised learning methods used for classification and regression. The property of SVM is, minimization of experimental classification error and maximization of geometric margin. So it is also called Maximum Margin Classifiers. The hyperplane which separates the data erect two parallel hyperplane on each side. Each hyperplane separated maximize the distance occurred. Quadratic Programming (QP) is solved by Radial Basis Function and Multi-Layer Perceptron classifiers.

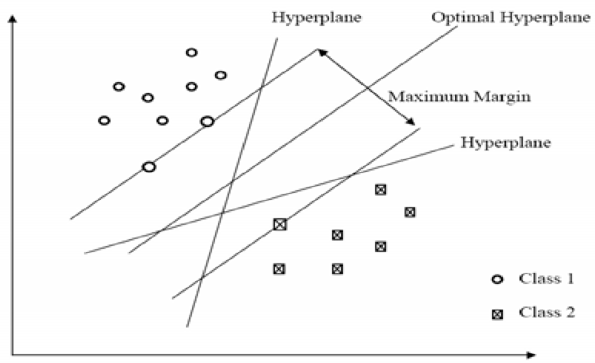

Fig.2. Structure of the optimized SVM

The learning problem setting for SVM is as follows: there is some unknown and nonlinear dependency (mapping, function) $y = f(x)$ between some high-dimensional input vector x and scalar output y (or the vector output y as in the case of multiclass SVM). The classification problem can be considered as a two-class problem without loss of generality. The goal is to separate the normal and abnormal images by a function which is induced from available brain database. The linear classifier is shown in figure 2. This linear classifier is termed as the optimal separating hyperplane.

2.2.1 Motivation

Classifying data is a common task in machine learning. A group of data present in the database is classified into two classes accordingly to find the abnormal brain for diagnosis. In the case of support vector machines, a data point is viewed as a pdimensional vector (a list of p numbers), and to separate such points with a (p − 1)-dimensional hyperplane. This is called a linear classifier. To classify the data, many hyperplanes are used. If hyberplane represents the largest separation between two

classes then it is called as best hyperplane. Choosing of data from
nearest data to an each side should be maximized. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier known as maximum margin classifier; or equivalently, the perceptron of optimal stability. The following figure 3 shows the three hyperplane which separates the data in to two classes. The H3 does not separate the two classes, H1 does, with a small margin and H2 with the maximum margin.

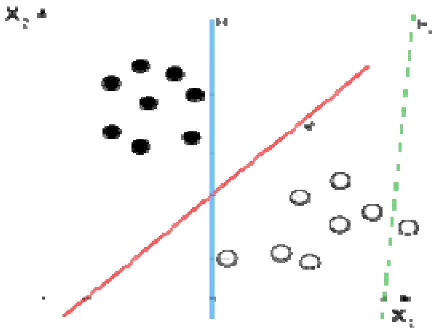

Fig. 3 H3 (doesn't separate the two classes). H1 (separates with a small margin) and H2 (separates with the maximum margin).

2.2.2 Linear SVM

Some training data D, a set of *n* points of the form is given to a linear SVM classifier,

$$D = \left\{ \left(x_i, y_i \right) \middle| x_i \in R^p, y_i \in \{-1,1\} \right\}_{i=1}^{n}$$

where the *yi* is either 1 or −1, indicating the class to which the point Xi belongs. Each Xi is a p-dimensional real vector. To find the maximum-margin hyperplane that divides the points having yi= 1, from those having, yi = -1. X is a hyperplane which is shown in figure 4.

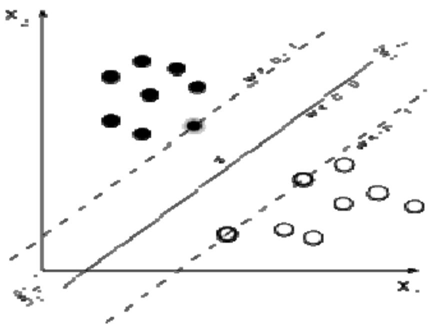

Fig.4. Maximum Margin classifier

Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors.

$$w \cdot x - b = 0$$

where **.** denotes the dot product and W the normal vector to the hyperplane. The parameter determines the offset of the hyperplane from the origin along the normal vector w. one want to choose the w and b to maximize the margin, or distance between the parallel hyperplanes that are as far apart as possible while still separating the data. These hyperplanes can be described by the equations

$$w \cdot x - b = 1$$ and $$w \cdot x - b = -1$$

Note that if the training data are linearly separable, we can select the two hyperplanes of the margin in a way that there are no points between them and then try to maximize their istance. By using geometry, we find the distance between these two hyperplanes is $\frac{2}{\|w\|}$, so we want to minimize $\|w\|$. As we also have to prevent data points from falling into the margin, we add the following constraint: for each i either

$$w \cdot x_i - b \geq 1$$ for the first and second for $$w \cdot x_i - b \leq -1$$

This can be rewritten as: C
We can put this together to get the optimization problem:

Minimize (in w,b) $\|w\|$

subject to (for any $i = 1, \ldots\ldots n$)

$$y_i \left(w \cdot x - b \right) \geq 1$$

2.2.3 Quadratic Equation Problem

The process of minimizing ||**q**|| makes the optimization process difficult. In Quadratic Optimization the ||**q**|| is change to1 2

Minimize (in w,b)

$$\frac{1}{2} \|x\|^2$$

subject to (for any $i = 1, \ldots\ldots n$)

$$y_i \left(w \cdot x - b \right) \geq 1$$

IJCSI
www.IJCSI.org

$$\min_{w,b,\alpha} = \left\{ \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i [y_i(w \cdot x - b) - 1] \right\}$$

The hyperplanes which divide the points; then all $y_i(w \cdot x - b) - 1 \geq 0$. Hence we could find the minimum, and this minimum would be reached for all the members of the family, not only for the best one which can be chosen solving the original problem.

Nevertheless the previous constrained problem can be expressed as

$$\min_{w,b,\alpha} \max_{\alpha \geq 0} = \left\{ \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i [y_i(w \cdot x - b) - 1] \right\}$$

a saddle point. In doing so all the points which can be separated as $y_i(w \cdot x - b) - 1 \geq 0$ do not matter since we must set the corresponding α_i to zero. This problem can now besolved by standard quadratic programming techniques

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

Only a few α_i will be greater than zero. The corresponding x_i are exactly the *support vectors*, which lie on the margin and satisfy $y_i(w \cdot x - b) = 1$. From this one can derive that the support vectors also satisfy

$$w \cdot x_i - b = 1 / y_i = y_i \Leftrightarrow b = w \cdot x_i - y_i$$

which allows one to define the offset b. In practice, it is more robust to average over all $_{NSV}$ vectors:

$$b = \frac{1}{NSV} \sum_{i=1}^{NSV} (w \cdot x_i - y_i)$$

2.2.4 Quadratic Equation Problem

The success of SVM depends on the selection of kernel and its parameters(C). Commonly Gaussian kernel, which has a single parameter γ *with* combination of C is often selected by a grid search with exponentially growing sequences of C and γ. Normally, each combination of parameter is cross validated for best accuracy. The validated data is finally given to the SVM classifier for testing and training process.

2.2 Steps involved in Proposed Methodology

1. Hyperplane acting as the decision surface is defined as

$$\sum_{i=1}^{N} \alpha_i d_i k(x, x_i)$$

Where $K(x,x_i) = \varphi^T(x)\varphi(x_i)$ represents the inner product of two vectors induced in the feature space by the input vector x and input pattern xi pertaining to the ith example. This term is referred to as inner-product kernel [13].
Where

$$\sum_{i=1}^{N} \alpha_i d_i \varphi(x, x_i)$$
$$\varphi(x) = [\varphi_0(x), \varphi_1(x), \ldots\ldots, \varphi_{m_1}(x)]$$

φ_0 (x) = 1 for all x
w_o denotes the bias b

2. The requirement of the kernel $K(x,x_i)$ is to satisfy Mercer's theorem. The kernel function is selected as a polynomial learning machine.

$$K(x,x_i) = (1 + x^T x_i)^2$$

3. The Quadratic Program Optimization $\{\alpha_i\}$ for i = 1 to N that minimize the objective function $Q(\alpha)$, denoted by $\alpha_{0,i}$ is determined.

$$Q(\alpha) \sum_{i=1}^{N} = \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(x_i, x_j)$$

Subject to the following constraints

$$\sum_{i=1}^{N} d_i d_j = 0$$
$$0 \leq \alpha_i \leq C,$$

For i= 1,2,…N

4. The linear weight vector w_0 corresponding to the optimum values of the Lagrange multipliers are determined using the following formula:

$$W_0 \sum_{i=1}^{N} \alpha_i d_i k(x, x_i)$$

$\varphi(x_i)$ is the image induced in the feature space due to x_i.

$$W_0 \sum_{i=1}^{N} \alpha_i d_i \kappa(x_i)$$

W_0 represents the optimum bias b_0.

3. Implementation and Results

The parameters of the maximum-margin hyperplane are derived by solving the optimization. There exist several specialized algorithms for quickly solving the QP problem that arises from SVMs, mostly relying on heuristics for breaking the problem down into smaller, more-manageable chunks.

A common method is SMO algorithm, which breaks the problem down into 2-dimensional sub-problems that may be solved analytically, eliminating the need for a numerical optimization algorithm. Instead of solving a sequence of broken down problems, this approach directly solves the problem as a whole. To avoid solving a linear system involving the large kernel matrix, a low rank approximation to the matrix is often used in the kernel trick.

The input to the feature extraction algorithm is the bacterial images. The pattern vectors (features) extracted from the images is given as input to the SVM classifier. Large database are required for the classifier to perform the classification correctly. In this system a sample of 150 CT brain images are collected. 90 images are taken as normal brain and remaining 60 images are taken as abnormal brain. 100 images are used for training phase and remaining 50 images are used for testing. The Classification accuracy and error rate is obtained by using the following formula:

All classification result could have an error rate and on occasion will either fail to identify an abnormality, or identify an abnormality which is not present. It is common to describe this error rate by the terms true and false positive and true and false negative as follows:

True Positive (TP): the classification result is positive in the presence of the clinical abnormality.
True Negative (TN): the classification result is negative in the absence of the clinical abnormality.
False Positive (FP): the classification result is positive in the absence of the clinical abnormality.
False Negative (FN): the classification result is negative in the presence of the clinical abnormality.

Table 1 is the confusion matrix table which defines various terms used to describe the clinical efficiency of a classification based on the terms above and
Sensitivity = TP/ (TP+FN) *100%
Specificity = TN/ (TN+FP) *100%
Accuracy = (TP+TN)/ (TP+TN+FP+FN)*100 % are used to measure the performance of the classifier

Table 1: The confusion matrix table

Actual Data	Predicated Data	
	Normal	Abnormal
Normal	TN	FP
Abnormal	FN	TP

Table 2 Classifier Performance

Method	Sensitivity	Specificity	Accuracy	Time
KNN	99	100	90	158 S
SMO	92	100	95.7	136 S
SVM-QPO	95	100	96.6	133 S

The above table 2 shows the different classifier performance. The proposed SVM-QPO gives the better results when compare to other classifier. The following figure 5 shows the classifier performance

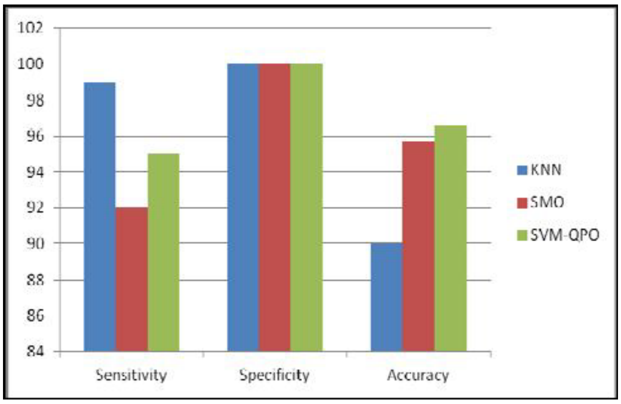

Fig .5 Classifier Performance

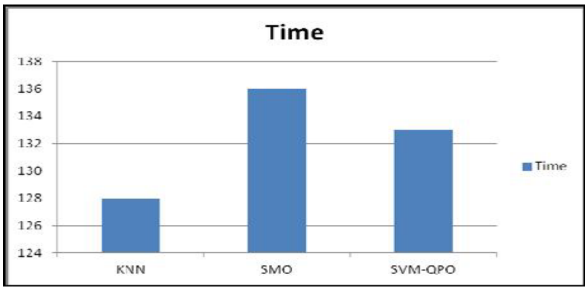

Fig .6 Prediction time

The above figure 6 shows the prediction time results for different classifier. The following table3 shows the number of input images for training phase and testing phase and the correctly classified images by Support Vector Machine. The proposed method gives the efficient results when compare to other classifier.

Table 3: SVM-QPO Classifier Performance

Type	No. Of. Input Image		Correctly Classified image		Efficiency	
	Testing	Training	Testing	Training	Testing	Training
1	20	10	20	10	100	100
2	20	10	18	9	96	93
3	20	10	18	10	96	93
4	20	10	19	9	94	97
5	20	10	20	8	97	93
Total	100	50	113	56	96.6	95

4. Conclusions

The proposed approach using SVM as a classifier for classification of brain images provides a good classification efficiency of 96.6% during training phase and 95% efficiency during testing phase. The sensitivity, specificity and accuracy is also improved. The proposed approach is computationally effective and yield good result. The future work is to improve the classification accuracy by extracting more features and increasing the training data set.

References

[1]Epaphrodite Uwimana, Automatic classification of medicalimages for Content Based Image Retrieval Systems CBIR),AMIA Annual Symposium proceedings AMIA Symposium AMIA Symposium, Vol.73, No. 1, pp 1159, 2008.

[2]Chien-Cheng Lee, Sz-Han Chen, and Yu-Chun Chiang,Classification of Liver Disease from CT Images Using a Support Vector Machine, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.11,No.4 , pp. 396-402, 2007

[3] H. Selvaraj, S. Thamarai Selvi, D. Selvathi3, L. Gewali, Brain MRI Slices Classification Using Least Squares Support Vector Machine, Vol. 1, No. 1, pp. 21 -33,2007.

[4]Lahouaoui Lalaoui, Tayeb Mohammad, Mohamed Chemachema and Abdessalem Hocini, Support Vector Machine (SVM) and the Neural Networks for Segmentation the Magnetic Resonance Imaging, proceeding of IEEE 5th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, March 22-26,2009.

[5]Igor F. Amaral, Filipe Coelho, Joaquim F. Pinto da Costa and Jaime S. Cardoso, Hierarchical Medical Image Annotation Using SVM-based Approaches, proceeding of IEEE International conference on Information Technology and Application in biomedicine, pp. 1-5, 2010.

[6]Ramesh Babu Durai C., Balaji V., Duraisamy V., Improved Content Based Image Retrieval using SMO and SVM Classification Technique, European Journal of Scientific Research , Vol.69, No.4, pp. 560-564, 2012.

[7]R.Muralidharan, Dr.C.Chandrasekar, Object Recognition using SVM-KNN based on Geometric Moment Invariant, International Journal of Computer Trends and Technology- July to Aug Issue, pp.215-220,2011.

[8]A.Ramaswamy Reddy, Dr. E.V.Prasad, Dr.L.S.S.Reddy ,Abnormality Detection of Brain MRI Images using a New Spatial FCM Algorithm, International Journal of Engineering Science & Advanced Technology, Vol.2, No. 1, pp. 1-7,2012.

[9]Lei Zhang, Fuzong Lin, Bo Zhang, Support Vector Machine Learning for Image Retrieval, Proceeding on IEEE International conference on Image processing, pp. 721-724,2001.

[10] Issam El-Naqa et.al, A Support Vector Machine Approach for Detection of Microcalcifications, IEEE Transactions On Medical Imaging, Vol. 21, No. 12,pp.1552-1563, Dec. 2002.

[11] Sanghamitra Mohanty, Himadri Nandini Das Bebartta, Performance Comparison of SVM and K-NN for Oriya Character Recognition, International Journal of Advanced Computer Science and Applications, Special Issue on Image Processing and Analysis, pp. 112-116, 2011.

[12] A.Padma, Automatic Classification and Segmentation of Brain Tumor in CT Images using Optimal Dominant Gray level Run length Texture Features, International Journal of advanced Computer Science and Applications, Vol. 2, No. 10, 2011

[13] Fei Peng Kehong Yuan Shu Feng Wufan Chen , Proceeding of International conference on Bioinformatics, pp. 2495 – 2498,2008.

[14] Ali Reza Fallah1, Mohammad Pooyan, Hassan Khotanlou, A new approach for classification of human brain CT images based on morphological operations, J. Biomedica l Science and Engineering,Vol.3, pp. 78-82, 2010.,