International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-3 Issue-4, April 2014

Hybridized Classification of Brain MRI using PSO & SVM

Amita Kumari, Rajesh Mehra

Abstract- Magnetic resonance imaging (MRI) provides detailed anatomic information of any part of the body. In this method a hybrid approach for classification of brain tissue in MRI based on Particle Swarm Optimization (PSO) and Support Vector Machine (SVM) wavelet based texture feature are extracted from normal and tumor region by using HAAR wavelet. These features are given as input to the SVM classifier which classified them into normal & abnormal brain neoplasm. The algorithm incorporates steps for pre-processing, image segmentation and image classification using SVM classifier.

Keywords- MRI, Classification *PSO, SVM, HAAR wavelet*

1. Introduction

Brain neoplasm/tumor is defined as any abnormal growth of cells in the brain. Basically brain tumors have variety of shapes and sizes. It can occur at any location and in different intensities. It can be Benign and Malignant. Benign tumor is not cancerous; it does not invade nearby healthy tissue or spread to other parts of the body. They may be monitored radiological or surgically removed and they grow back. Malignant tumor is cancerous and had heterogeneous structure. Glioblastoma Multiform (GBM) is the most common and most aggressive malignant primary brain tumor in humans, involving glial cells and accounting for 52% of all functional tissue brain tumor cases and 20% of all intracranial tumors. Despite being the most prevalent form of primary brain cancer, GBM incidence is only2-3 cases per 100000 people in Europe and North America. According to the WHO classification of the tumors of the central nervous, the standard name for this brain tumor is "glioblastoma". It presents two variant giant cells Glioblastoma and Gliosarcoma. Treatment can involve chemotherapy, radiation, radio surgery, corticosteroids, antangiogenic therapy, surgery and experimental an approaches such as gene transfer. MRI is efficient in supplying the location and size of the tumor but it is very difficult to classify the tumor type, so a process of biopsy is known for classifying the images. Biopsy is a very painful process. This inability requires development of new analysis techniques that aim at improving diagnostic ability of MR.

Brain tumor can be defined by grade .The grade of a tumor refers to the way the cells look under a microscope: The tissue is benign and the cells look nearly like normal brain cells and they grow slowly is grade I. The tissue is malignant and the cells look less like normal cells than do the cells in a grade I tumor are known as grade II. The malignant tissue has cells that look very different from normal cells.

Manuscript published on 30 April 2014.
***** Correspondence Author (s)
Amita Kumari, ECE Department, NITTTR, Chandigarh, India
Rajesh Mehra, ECE Department, NITTTR, Chandigarh, India

The abnormal cells are actively growing (anaplastic) defined by the grade III. The grade IV are malignant tissue has cells that look most abnormal and tend to grow quickly. Cells from low grade tumors (grade I and grade II) look more normally and generally grow more slowly than cells form high grade tumors (grade III and grade IV). These are different images of brain MRI showing the different grades of tumor.

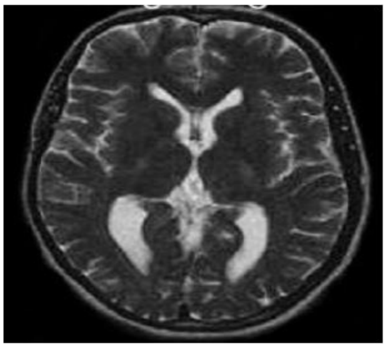

Fig 1. Normal brain

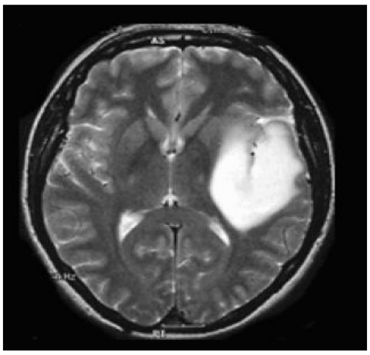

Fig 2. Benign tumor

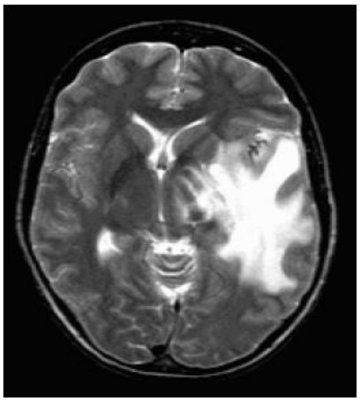

Fig 3. Malignant tumor

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

2. Related Work

Many techniques have been developed for the classification of brain MRI. A hybrid approach for classification of brain tissues in magnetic resonance images had been proposed in [1] using GA &SVM. A wavelet based texture feature set is derived. The optimal texture features extracted from normal and tumor regions by using Spatial Gray level Dependence Method (SGLDM). These features are given as input to the SVM classifier. The choice of feature, which constitutes a big problem in classification technique, is solved by using GA.

In [2] computer based method for defining tumor region in the brain using MRI images using two different segmentation steps canny edge detection and adaptive threshold. The comparison between SVM with and without the implementation of principle component analysis has been proposed by in [3]. A genetic fuzzy system for modeling different tissue in brain MRI and proposed a statistical pixel classification based on maximum likelihood (ML) and Bayesian classifier has been defined in [4]. SVM is an effective tool in sonography for diagnosis of breast cancer. A SVM is a machine learning system developed using statistical learning theories to classify data points into two classes. It has been applied extensively for classification, image recognition and bioinformatics [5]. Classification of brain MRI using Back Propagation neural network (BPN) and Radial Basis Function (RBFN) has been demonstrated in [6].it is classified in the axial and coronal images. Both testing and training phase gives the percentage of accuracy on each parameter in neural network. Another method [7] is using Support Vector Machine with Recursive Feature Elimination (SVM RFE). A knowledge based technique [8] and neural based network[10] has also been developed .

3. Proposed Technique

The different stages of proposed technique are shown in Fig 4.firstle the acquisition of image take place in which images are acquired from different equipments like MRI, PET,CT. Next step is the wavelet based feature extraction. In this technique, Haar is used . At last feature selection and the classification.

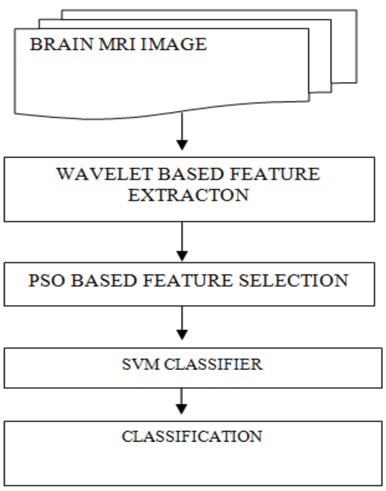

Fig 4.The Image Analysis Process

4. Support Vector Machine

Support vector machine (SVM) was first heard in 1992, introduced by Boser, Guyon and Vapnik in COLT-92. It is a set of related supervised learning methods used for classification and regression.[13]. They belong to a family of generalized linear classifier. It is also defined as a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy while automatically avoiding over fit to the data. It can be defined as the system which uses hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. It gained popularity due to better empirical performance. The formulation uses the structural risk minimization (SRM) principle, used by conventional neural network. SRM minimizes an upper bound on the expected risk, where as ERM minimizes the error on the training data. It is difference which equips SVM with a greater ability to generalize, which is the main goal in the statistical learning. There are many common kernels such : Linear, Polynomial of degree and Radial Basis Function (RBF) .Here the linear kernel is used. SVM classifier is used to distinguish between the normal and the abnormal brain using the confusion matrix. The feature extraction is done using the wavelet based using the HAAR function.

The statistical learning theory provides a framework for studying the problem of gaining knowledge, making predictions, making decisions from a set of data. In simple terms, it enables the choosing of the hyper plane space such a way that it closely represents the underlying function in the target space.

In statistical learning theory the problem of supervised learning is formulated as follows. We are given a set of training data $\{(\mathbf{x}_1,y_1)\ldots(\mathbf{x}_l,y_l)\}$ in $R^n \times R$ sampled according to unknown probability distribution $P(\mathbf{x},y)$ and a loss function $V(y,f(\mathbf{x}))$ that measures the error, for a given $\mathbf{x}$, $f(\mathbf{x})$ is "predicted" instead of the actual value y. The problem consists in finding a function f that minimizes the expectation of the error on new data that is, finding a function f that minimizes the expected error:

$$\int V(y, f(\mathbf{x}))\, P(\mathbf{x}, y)\, d\mathbf{x}\, dy \qquad (1)$$

In statistical modeling, choose a model from the hypothesis space, which is closest (with respect to some error measure) to the underlying function in the target space

4.1 Learning and Generalization

Early machine learning algorithms aimed to learn representations of simple functions. Hence, the goal of learning was to output a hypothesis that performed the correct classification of the training data and learning algorithms were designed such that it accurately fit in to the data .

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The ability of a hypothesis to correctly classify data not in the training set is known as its generalization. SVM performs better in term of not over generalization, when the neural networks might end up over generalizing easily. Another thing to observe is to find where to make the best trade-off in trading complexity with the number of epochs. The SVM is defined as the mapping the non linear inseparable data into a linear high dimensional feature space F by using transformation, Φ ; RN → F;

Then the optimal hyper plane H:f(x)=(w.ϕ(x)+ b) can be obtained by solving optimization problem as
:

$$\min(w,\varepsilon)= \frac{1}{2}||w||_2 + c\sum_{i=1}^{l}\varepsilon_i \qquad (2)$$

$$\text{s.t} \quad y_1((w.\phi(x)+ b)\geq 1-\varepsilon_i$$

w – coefficient vector of the hyperplane in feature space
b- hyperplane threshold value
ε- slack factor introduced for errors in classification
c- Penalty factor for error

The dimension of the feature space F is usually very large, so direct calculation may lead to "dimension disaster" so $w=\sum_{i=1}^{l} \alpha\, y\, \phi(x)$. All operation of SVM in feature space are only dot product operation. Then

$$K(x_i, x_j) = \phi(x_i).\phi(x_j) \qquad (3)$$

A kernel function can transform the dot product operation in high dimension space into kernel function operation in input operation in input space as long as it satisfies the mercer condition as in[10].The common kernel function are Linear kernel, Polynomial kernel, Gaussian radial basis function (RBF), Sigmoid neural network.

5. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling[12]. PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithm (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation.

In PSO, the potential solutions, called particles, fly through the problem Space by following the current optimum particles. Each particle keeps track of its coordinates in the problem space which are associated with the best solution (fitness) it has achieved so far. This is called the pbest. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the neighbors or the particle. This location is called lbest, when a particle takes all the population as its topological neighbors, the best value is a global best and is called gbest.

PSO has been used for approaches that can be used across a wide range of application focused on a specific requirement. At each iteration, each particle adjusts its velocity vector, based on its momentum, influences of its best solution of its neighbors, and then computes a new point to be evaluated. The displacement of a particle is influenced by three components:

1. Physical component: The particle tends to keep its current direction of displacement.
2. Cognitive component: The particle tends to move towards the best site that it has explored until now
3. Social component: The particle tends to rely on the experience of its congeners, then moves towards the best site already explored by its neighbors.

The collective and the social behavior of living creatures is known as the swarm intelligence. The main properties of collective behavior are homogeneity, locality, collision avoidance , velocity matching and flock centering.

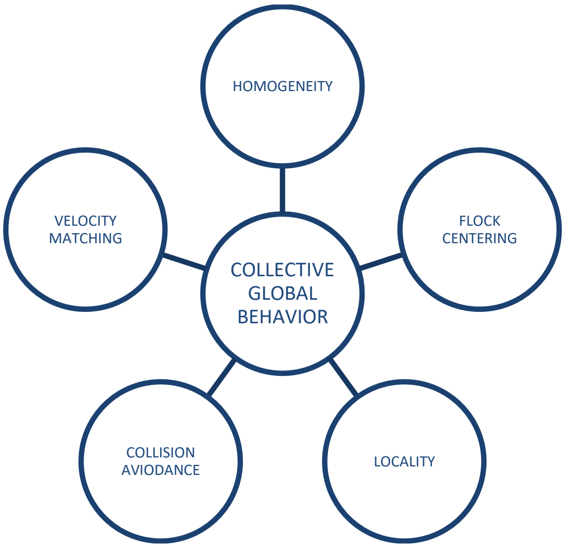

Fig 5. Properties of collective behavior

5. PSO based SVM

The particle swarm optimization concept consists of, at each time step, changing the velocity of (accelerating) each particle toward its p*best* and *lbest* locations (local version of PSO). Acceleration is weighted by a random term, with separate random numbers being generated for acceleration toward *pbest* and *lbest* locations.

The algorithm keeps track of three global variables:
1. Target value or condition
2. Global best (gbest) value indicating which particle's data is currently closest to the target.
3. Stopping value indicating when the algorithm should stop if the target isn't found.

The swarm size is denoted by s, and the search space is n-dimensional. In general, the particles have three attributes: the current position is given $X_i = x_{i,1}, x_{i,j}, \dots \dots x_{i,n}$, the current velocity vector is $V_i = v_{i,1}, v_{i,2} \dots \dots \dots . v_{i,n}$ and the past best position is given $pbest_i = p_{i,1}, p_{i,2} \dots \dots . p_{i,n}$. These attributes are used with the global best position $Gbest = (g_1, g_2 \dots \dots g_n)$, of the swarm, to update iteratively the state of each particle in the swarm. The objective function to be minimized is denoted by f. The new velocity vector Vi of each particle is updated as follows:

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$v_{i,j}(t+1) = v_{i,j}(t) + c_1 r_{1i,j}(t)[pbest_{i,j}(t) - x_{i,j}(t)] + c_2 r_{2i,j}(t)[gbest_j(t) - x_{i,j}(t)] \quad (4)$$

$v_{i,j}$ is the velocity of the ith particle (i ϵ 1, 2, . . . , s) of the j th dimension (j ϵ1, 2, . . . , n) where: c1, c2 are the learning factors that will be fixed throughout the whole process, called acceleration coefficients. r_1, r_2 are two random numbers in the range [0, 1] selected uniformly for each dimension. At each iteration.$v_{i,j}(t)$ is the physical component $c_1 r_{1i,j}(t)[pbest_{i,j}(t) - x_{i,j}(t)]$ is the cognitive component, where C_1 controls the cognitive behaviour of the particle. $c_2 r_{2i,j}(t)[gbest_j(t) - x_{i,j}(t)]$is the social component, where c_2 controls the social behaviour of the particle.The new position Xi of each particle is calculated as follows

$$pbest_i x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad (5)$$

Where

$x_{i,j}$ is the position of the ith particle (i ϵ 1, 2, ... s) of the j th dimension (j ϵ1, 2, . . . , n). In case of minimization of f, the past best position $Pbest_i$ of each particle is updated as:

$$Pbest_i\,(t+1) = \begin{cases} Pbest_i(t), if\ f\ (Xi\ (t\ +1)) \ \geq\ Pbest_i\ (t) \\ Xi\ (t\ +1), otherwise \end{cases} \quad (6)$$

The global best position Gbest, found by the evaluations of the particles during each Generation, is defined as

$$Gbest_i\ (t+1) = argmin_{Pbest_i}\, f\,(Pbest_i\ (t+1)), \quad \text{where } 1 \le i \le s \quad (7)$$

In the global version of PSO, the best particle Gbest is chosen among the whole population.

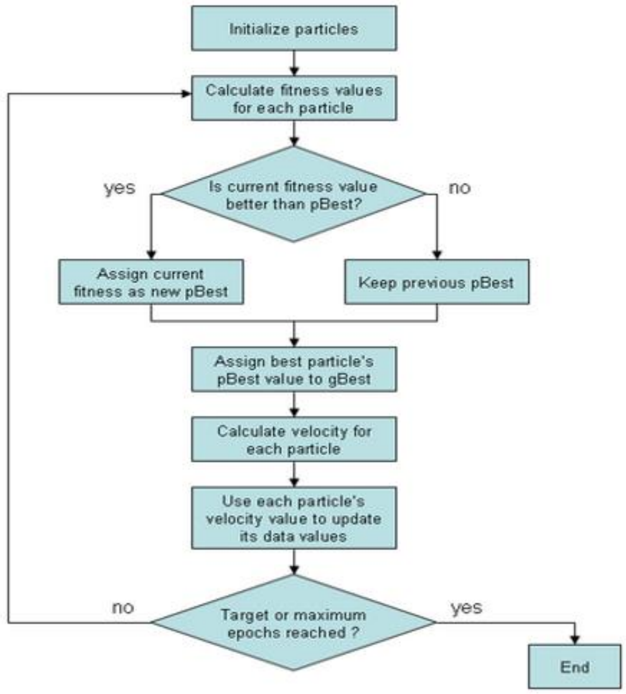

Fig 6. Flow diagram illustrating the PSO

6 .PSO-SVM experimental result

Our proposed hybrid technique is implemented on a real data set consisting of transaxial images of brain MRI. It consists of 246 images: 82 are normal, 82 are benign and 82 images are taken as malignant tumor suffering from a low grade glioma, meningioma. These normal and pathological images are axial, T2-weighted of 256*256 sizes and acquired at several positions of transaxial plane, images are taken from [11] .In this section, performance is evaluated by three factors that is sensitivity, specificity and accuracy.

Sensitivity= TP/ (TP+FN)*100% (6)

Specificity= TN/ (TN+FN)*100% (7)

Accuracy=TP+TN/ (TP+TN+FP+FN)*100% (8)

Where
TP (true positive) = correctly classified positive cases
TN (true negative) = correctly classified negative cases
FP (false positive) = incorrectly classified positive cases
FN (false negative) = incorrectly classified negative cases

Accuracy is defined as the degree of closeness of measurement of quantity actual value. Sensitivity is called true positive cases. It means proportion of actual true cases, which are correctly classified as true cases. Specificity is known as true negative cases. It defined as proportion of negative cases which are correctly classified as true cases. In this case total no images 247 are used for training purpose. Overall 20 images are normal, 18 are benign and 20 are malignant are given for testing. Table 1 shows the classification rates for performing the proposed hybrid approach and comparison with different methods.

The hybrid technique	Sensitivity (%)	Specificity (%)	Accuracy (%)
SGLDM+GA+SVM	91.87	100	94.44
WT+SGLDM+GA+SVM	94.6	100	96.29
SVM+PSO(PROPOSED METHOD)	97.5	100	97.5

Table 1 .classification rates (in %) for the proposed technique.

7. Conclusion

This technique provides an efficient method for the image classification in comparison to other technique. This method consisting of wavelet transform using, PSO (Particle Swarm Optimization) and SVM (Support Vector Machine). This system helps doctor to take the decision for classifying the tumorous brain into normal and abnormal classes.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

This technique is accurate, robust easy to operate, non invasive and inexpensive. But it necessitates fresh training each time whenever there is change in database. This technique provides the high sensitivity, specificity and accuracy for classifying the brain MRI images.

REFERENCES

[1] A .Kharrat, K Gasmi , M Ben Messaoud , N Benamrane and M Abid . " A hybrid approach for automatic classif ication of brain MRI using genetic algorithm and support vector machine", Leonardo journal of sciences, Issue 17, pp 71-82, July-Dec ,2010.

[2] E F Badran, E G Mahmoud, N Hamdy "An algorithm for brain tumor in MRI images", International Conference on Communication Computer Engineering and Systems (ICCES), Networking & Broadcasting, pp 368-373, June, 2010.

[3] N Abdullah, Lee Wee Chuen; U K Ngah, KA Ahmad, "Improvement of MRI brain classification using Principles Component Analysis", IEEE International Conference on Control System, Computing and Engineering (ICCSCE), ,pp 567-571, March, 2011.

[4] M Hasanzadeh ,S Kasaei."Multispectral brain MRI segmentation using genetic fuzzy systems" international conference on communication ,networking and broadcasting(ISSPA),pp 1-4,June,2001

[5] V.vapnik "the nature of statistical learning theory, springer-verlag, newyork, 1995

[6] S N Deepa, B A Devi ."Artificial neural networks design for classification of brain tumor" IEEE International Conference on Computer Communication and Informatics(ICCII), pp 1-6, 2012.

[7] E. I.Zacharaki ,Sumei Wang ,S Chawla, Dong Soo Yoo, R Wolf ,E.R Melhem, C.Davatzikos."MRI based classification of brain tumor type and grade using SVMRFE".IEEE Symposium on bioengineering,pp-1035-1038,April ,2009.

[8] M .C Clark, L. O Hall, D.B.Goldgof, R.Velthuizen, F.R. Murtagh, and M.S Silbiger. "Automatic tumor segmentation using knowledge based technique, IEEE transaction on medical imaging, vol 17,no 2,pp-187-192,April,1998.

[9] W. E. Reddick , J.O Glass, E. N. Cook , T. D Elkin ,R. J. Deaton. "Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks", IEEE transaction on medical imaging, vol 16, no 6,pp 911-918 ,1997.

[10] Huazhu Song , Zichun Ding, Cuicui Guo , Zhe Li, Hongxia Xia. " Research on combination kernel function of Support Vector Machine ." International Conference on Computer science and software engineering, pp-345-356, 2008

[11] Mcconnell Brain Imaging centre (june 2006) stimulated data base [online] available http://www.bic .mni.mcgill.ca /brainweb).

[12] Y.J .Kennedy ,R.Eberhart. " Particle Swarm Optimization." IEEE International Conference on Neural Network,pp-1942- 1948, 1995.

[13] Vladimir N. Vapnik " The Nature of Statistical Learning Theory" New York: Springer –Verlag,2000

Mr. Rajesh Mehra is currently Associate Professor at National Institute of Technical Teachers Training & Research Chandigarh, India. He is pursuing his PhD from Punjab University Chandigarh, India .He has completed his M.E from NITTTR Chandigarh, India and B.Tech from NIT, Jalandhar, India. Mr. Mehra has more than 17 years of academic experience. He has authored more than 100 research paper including more than 50 in journal. Mr. Mehra's interest areas are VLSI Design, Embedded System Design, and Advanced Digital Signal Processing. Mr. Mehra is member of IEEE & ISTE.

Mrs. Amita is currently pursuing M.E from National Institute of Technical Teachers Training & Research Chandigarh, India. She has completed B.Tech from Dr B S.B.R.A College of Agricultural Engineering and Technology, Etawah (u .p). She is having four years of teaching experience in Moradabad Institute of Technology, Moradabad. Mrs. Amita's interest areas are Image processing, VLSI, wireless and mobile communication and digital electronics.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)