INFO- 2083
Primljeno / Received: 2012-09-19
UDK : 159.95.3:681.3:007
Izvorni znanstveni rad / Original Scientific Paper

EXTREME LEARNING MACHINE FOR CLASSIFICATION OF BRAIN TUMOR IN 3D MR IMAGES

ELM ZA KLASIFIKACIJU TUMORA MOZGA KOD 3D MR SNIMAKA

S.N.Deepa, B.Arunadevi
Anna University, Coimbatore, India
Sveučilište Anna, Coimbatore, Indija

Abstract

Extreme Learning machine (ELM) a widely adopted algorithm in machine learning field is proposed for the use of pattern classification model using 3D MRI images for identifying tissue abnormalities in brain histology. The four class classification includes gray matter, white matter, cerebrospinal-fluid and tumor. The 3D MRI assessed by a pathologist indicates the ROI and the images are normalized. Texture features for each of the sub-regions is based on the Run-length Matrix, Co-occurence Matrix, Intensity, Euclidean distance, Gradient vector and neighbourhood statistics. Genetic Algorithm is custom designed to extract and sub-select a decisive optimal bank of features which are then used to model the ELM classifier and best selection of ELM algorithm parameters to handle sparse image data. The algorithm is explored using different activation function and the effect of number of neurons in the hidden layer by using different ratios of the number of features in the training and test data. The ELM classification outperformed in terms of accuracy, sensitivity and specificity as 93.20 %, 91.6 %, and 97.98% for discrimination of brain and pathological tumor tissue classification against state-of-the-art feature extraction methods and classifiers in the literature for publicly available SPL dataset.

Sažetak

ELM, široko prihvaćen algoritam strojnog učenja se predlaže za korištenje u uzorkovanju pomoću klasifikacijskog modela 3D MRI slika za identifikaciju abnormalnosti tkiva u histologiji mozga. Četiri klase obuhvaćaju sive, bijele tvari, cerebrospinalne tekućine-i tumore. 3D MRI koji ocjenjuje patolog, ukazuje na ROI, a slike su normalizirane. Značajke tekstura za svaku od podregija se temelje na Run-length matrici, ponovnom pojavljivanju matrice, intenzitet, euklidska udaljenost, gradijent vektora i statistike susjedstva. Genetski algoritam je obično dizajniran za izdvajanje i sub-optimalan odabir odlučujući o značajkama koje se onda koriste za model ELM klasifikatora i najbolji izbor ELM parametra algoritama za obradu rijetkih slikovnih podataka. Algoritam se istražuje koristeći različite aktivacijske funkcije i utjecaj broja neurona u skrivenom sloju pomoću različitih omjera broja značajki kod trening i test podataka. ELM klasifikacija je nadmašila u smislu točnosti, osjetljivosti i specifičnosti, kao 93,20%, 91,6% i 97,98% za diskriminaciju mozga i patološki kod tumora i sistematizacije metode za prikupljanje podataka i klasifikatore u literaturi za javno dostupne SPL skup podataka.

1. Introduction

Numerous studies had been made to develop automated tumor image classification on 3D models with little knowledge on workforce of pathologists, with the similar knowledge of interpreting data and giving out accurate results. Since decades, many statistical and machine learning perceptions have been identified for medical image classification. Both first and second generation artificial neural networks turned its dominancy. The major deployment of ANN is due to independency in the underlying process from any functional form, even when no prior assumptions can be made and when only data is available. The realization is from the universal approximation property enabling to approximate any continuous function to a desired level of accuracy. Yet, these networks are been considered as 'black box' models and consequently very difficult to interpret in their trained state /**1**/. Among the various learning algorithms, the support vector machine is one of the most important and widely used algorithms in medical image classification. Several studies have reported that use of support vector machines , in case of both binary and multi SVM , shows better performance, than traditional neural compensating high computational complexity and more time to select.

In this state-of-art, Extreme Learning Machine (ELM) is a competitively good solution for such complex tasks. The Extreme learning Machine (ELM), a recent second generation neural network algorithm, is identified as to achieve high quality performance in multifaceted problems and reduced computation time compared with other machine learning algorithms /**2**/. ELM based model presented by Huang et al. in /**3**/ provides a very fast learning phase for relatively large data sets , that does not require iterative tuning which is dominant in other neural networks.

Warfield et al. /**4**/ suggested a model based on elastic atlas warping for brain extraction and statistical pattern recognition for brain interior structures. The intensity feature was augmented by a distance from the boundary feature to account for overlapping probability density functions. This method was found to be successful for simple-shaped tumors with homogeneous texture. Cobzas et al. /**5**/ proposed a variational MRI tumor segmentation method that incorporates both atlas-based priors and learned statistical models for tumor and healthy tissue. A statistical model is designed for MRI brain tumor segmentation task using high multi scale feature set. The results were tested on data from 9 patients having either a grade 2 astrocytoma, an anaplastic astrocytoma or a glioblastoma multiforme .

Georgiadis et al. /**6**/ put forth a model to evaluate the efficiency of three dimensional textural features using a pattern recognition system in the task of discriminating primary from metastatic brain tissues on T1 post-contrast MRI series on 67 brain data set. Devos et al. /**7**/ modelled a SVM-based classification system discriminating gliomas and meningiomas with 95% overall accuracy, employing as features image intensities from four MR sequences (T1, T2, PD and GD). In combining the features derived from MR spectroscopy, classification accuracy reached 99.8%.

Akselrod-Ballin et al. /**9**/ modelled a novel automatic approach for the recognition of anatomical brain structures in 3D MRI. The method incorporates a combination of multi-channel 3D segmentation algorithm using full hierarchy of segments of pyramid along with a support vector machine (SVM) based classifier. The unique approach combined a rich and tunable set of features, emerging from statistical measurements at all scales. Its was experimented on 20 real MRI data set of normal subjects with GM,WM and CSF expert segmentations provided by the Internet Brain Segmentation Repository (IBSR), after been positionally normalized. The results were highly promising.

The paper illustrates the generalization performance of ELM algorithm for bench mark dataset image classification. Appropriate selection of the input weights, hidden bias and number of hidden neurons for minimal image data pose a significant effect on the classification performance due to fewer training samples. In this paper, a rich feature extraction bank and tumour classification for 3D MRI images is characterized by sparse nominal data samples. The tissues of MRI brain reported as White matter (WM), Gray matter (GM) and Cerebrospinal fluid (CSF) , along with the tumour is simultaneously classified . Genetic Algorithm is used for feature sub-selection.

Recent studies on ELM proved enhanced performance in comparison to other classifiers for larger training samples. Custom selection of parameters randomly with suitable feature extraction technique , feature sub selection model obtains better classification accuracy for smaller training samples .To solve the feature sub-selection , genetic algorithm is designed , where the genetic operators select the most relevant set of features for good classification accuracy .The selected features are employed for tumor classification using ELM.

Performance evaluation of the proposed approach with existing literature has been carried out with Harvard SPL benchmark database .The performance is analyzed of the proposed approach on the database and the results are compared with methods reported in the literature. Study has been carried out clearly indicating the classification performance improvement with the other methods . The proposed approach identifies the most promising features to be extracted where intensive computations are reduced for both contrast and non – contrast T1 images. This led to an attempt to investigate on the features extracted , which requires not much expert knowledge. The present study compares the performance of extreme learning machine with a back propagation neural network and support vector machines.

This paper will focus on three segments of classification model for tumour on 3D MR brain images, which are Image Processing method, that is focusing on segmentation and feature extraction model ; and secondly on feature sub-selection for in terms of spectral and intensity characteristics using Genetic Algorithm ; and last on ELM used to build a robust classifier model for healthy and tumour images. The final section discusses the performance of the proposed approach over existing techniques and conclusions based on this study.

2. Segmentation and Feature Extraction Model

2.1 Image Segmentation

Image segmentation in 3D MRI is required to delineate the boundaries of the ROIs ensuring, in the proposed approach, that tumors are outlined and labeled consistently across subjects . It entails finding suitable neighbourhood features that can be distinguished from other objects and from the background. Individual pixel is to be checked to match whether it fits to an object of interest or not. In the proposed approach segmentation process is performed in three stages after pre-processing (i) WM and GM segmentation (ii) CSF segmentation (iii) Tumor segmentation.

The pre-processing stage involves probabilistically classifying the pixels of the input image based on intensity and second order grey level statistics obtained from the co-occurrence matrix. Thresholding is the selection of pixels/voxels which have a gray value between two given borders g_{min} and g_{max} , determines the intensity value which separated the desired class. Let I be the image and I_{thresh} be the result after thresholding .

$$I_{thresh}(x,y,z) = \{ \begin{array}{ll} 1 & \text{if } g_{min} \leq I(x,y,z) \leq g_{max} \\ 0 & \text{otherwise} \end{array} \quad (1)$$

Only the local grey value is taken into account, but no neighbourhoods. So for a 3D possible class of segmentation, edge based segmentation is applied for WM and GM segmentation, which includes a combination of laplace filter and sobel gradient for detection. Edge-based techniques rely on discontinuities in image values between distinct regions, and the goal of the segmentation algorithm is to accurately demarcate the boundary separating these regions. Segmentation of cerebrospinal fluid (CSF) from the brain MRI image is done by Orthogonal Polynomial Transform (OPT), which is applied to the gray scaled image I_g , given as :

$$I_{CSF} = \sin \frac{{I_{g(i)}}^3}{100} + (0.05 * \text{rand} * (|I_g|)) \quad (2)$$

The pathological tissue segmentation of the tumor part is segmented using the edge and point based segmentation from the abnormal images with the features extracted. Point based technique aims at distinction and definition of different regions in the image, which assigns every pair of voxels, x, y, z in the image, a real number between 0 and 1 , determining typical grey values for each structure .

2.2 Feature Extraction

Feature extraction a unique form of dimensionality reduction .It Transforms the input data into a set of features. In this study, the heterogeneous regions of brain tumors by combining imaging features from several sequences is undergone. A bank of rich feature set is characterized to aggregate in requisites of intensity, texture, shape, and location. All these features were chosen in analysis with expert radiologists. Image is represented by a function f(x,y,z) of three space variables x ,y and z, x=0,1,N-1 , y=0,1,…, M-1. and z=0,1,…L-1.The function f(x,y z) can take any value i=0,1,…,G-1 where G is total number of intensity levels in the image. A bank of features extracted in this proposed method includes :

a) **Image Intensity / gray value** - All pixels with a given intensity range belongs to the region
- Mean - Standard Deviation -Intensity

b) **Geometric features**
- Gray Level Run Length Matrix

c) **Spectral features**
- Gray level gradient of neighbourhoods (sobel) - Laplace

d) **Texture Features**
- Gray level Co-occurrence

The spectral characteristics are obtained by valuation of Sobel gradient and Laplace in the filtered images. The run length statistics encoding allows in capturing the coarse characteristics of the texture in 3D image data . The run length matrix /**9**/ allows in capturing the coarse characteristics of texture features ,which is obtained along directions of 0°, 45°, 90°, and 135°.11 texture features are calculated /**10**/ to determine the vector that characterizes the texture for each sub- region.

Spatial Gray Level Co-occurence Matrices captures the spatial dependence of gray-level values across multiple slices .The matrix is calculated for 0, 45, 90 and 135 degrees for θ and a distance scale of 1 . From these co-occurrence matrices, nine Haralick texture features /**11**/ were calculated in order to quantify the spatial dependence of gray-level values . Two intensity features, Euclidean Distance, laplace and three gray level gradient based features are extracted (sobel) .These features amplify the significant differences between class areas. A rich bank of the 27 texture descriptors are calculated . These features are computed as part of the segmentation process.

3. Feature Sub- selection model

The aim is to locate the best minimum subset of the original features, rather than transforming the data to an entirely new set of dimensions. Feature selection approach typically incorporates a search strategy for exploring the space of feature subsets. Using the wrapper approach, which selects the most relevant features by evaluating random subsets of the original features, based on an optimization criterion, Genetic Algorithm is evaluated /**12**/. The Genetic Algorithm is custom designed with the parameters as in table 1. The Genetic algorithm reduces the dimensionality into 8 most decisive features out of 27 features extracted. Fig. x denotes the features selection with the most decisive features, selecting amongst features for best feature space via GA.

Table 1: Parameters for Genetic Algorithm for feature selection

Parameters	**Description**	**Value**
P_m	Mutation Probability	0.02
P_c	Cross-over probability	0.1
M_g	Maximum no of generations	50
N	Population size	30

4. Extreme Learning Machine Classification

Extreme Learning Machine (ELM) is a single-hidden layer feed forward neural network (SLFNN) which randomly selects input weights and hidden neuron biases without training. The outputs weights are analytically are analytically determined using the norm least-square solution and Moore-Penrose inverse of a general linear system, thus allowing a significant training time reduction. The activation function like sine , gaussian , sigmoidal etc., can be chosen for hidden neuron layer and linear activation functions for the output neurons .The SLFNN evaluated here uses additive neuron design instead of kernel based , hence random parameter selection. SLFNs are considered as a linear system .

4.1 Approximation of ELM with additive neurons

Let the given training set N= {(x_i,t_i) }, where the training sample $\boldsymbol{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$ and the corresponding target value $\boldsymbol{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbf{R}^m$. SLFNs with $\tilde{N}$ hidden neurons and activation function *f(x)* are mathematically modelled as :

$$\sum_{i=1}^{\tilde{N}} \beta_i f(w_i \cdot x_j + b_i) = o_j \quad , j = 1,\dots\dots N \tag{3}$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector connecting the *i*th hidden neuron and the input neurons, b_i denotes the bias of *i*th hidden neuron; $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in}]^T$ is the weight vector connecting the *i*th hidden neuron and the output neurons . $w_i \cdot x_j$ denotes the inner product of w_i and x_j . The fact that standard SLFNs with $\tilde{N}$ hidden neurons each with activation function *f(x)* can approximate these *N* samples with zero error, means $\sum_{i=1}^{\tilde{H}} || o_j - t_j || = 0$.i.e there exist β_i , w_i,, and b_i such that

$$\sum_{i=1}^{\tilde{N}} \beta_i f(w_i \cdot x_j + b_i) = t_j \quad , j = 1, \dots ,N$$

The above *N* equations can be rewritten as:

$$\boldsymbol{H\beta} = \boldsymbol{T} \tag{5}$$

where $\boldsymbol{H}(w_1,\dots,w_{\tilde{N}}, b_1,\dots,b_{\tilde{N}}, x_1,\dots,x_N) =$

$$\begin{bmatrix} f(w_1 \cdot x_1 + b_1) & f(w_{\tilde{N}} \cdot x_1 + b_{\tilde{N}}) \\ & \\ f(w_1 \cdot x_N + b_1) & f(w_{\tilde{N}} \cdot x_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \cdot\cdot \\ \cdot\cdot \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N}\times m}$$

$$T = \begin{bmatrix} t_1^T \\ \cdot\cdot \\ \cdot\cdot \\ t_N^T \end{bmatrix}_{N\times m}$$

H is the hidden layer output matrix of the neural network; the *i* th column of **H** is the *i* th hidden neuron's output vector with respect to inputs x_1, $x_2, \dots, x_N$. Fig. 1 shows the learning procedure and structure in ELM. As shown in Fig.2, the ELM consists of single-hidden layer feed forward networks (SLFNs).

(6)

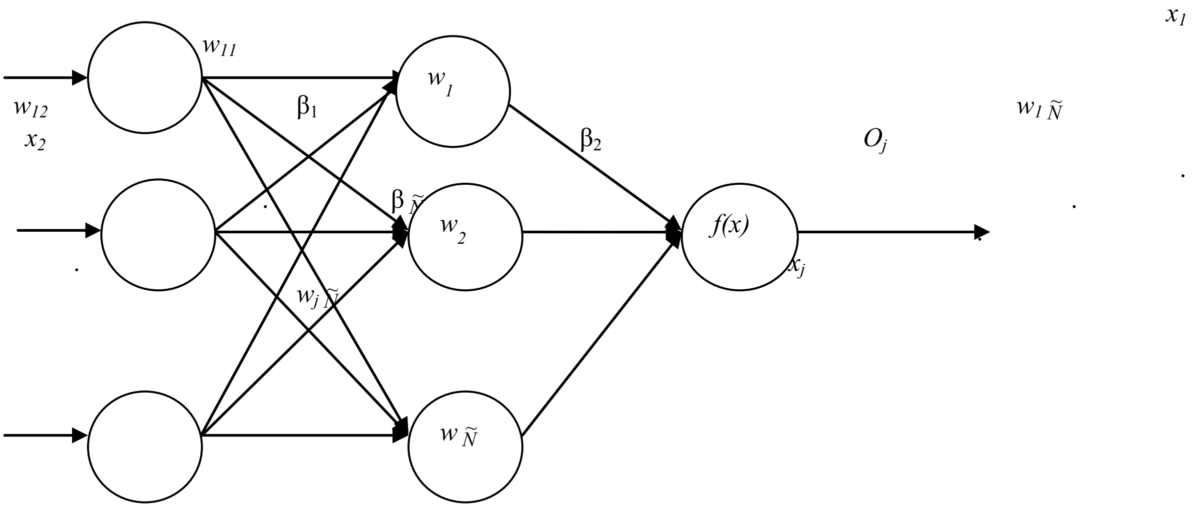


Figure 1: ELM Architecture

4.2 ELM Learning Algorithm

The SLFN can be solved by using a gradient based solution by finding suitable values of w', b' and β' satisfying the model as :

$$\| H(w'_1,\dots,w'_H, b'_1,\dots,b'_H)\,\beta' - T \| = \min_{w_i,b_i,\beta} \| H(w_1,\dots,w_H, b_1,\dots,b_H)\,\beta - T \| \qquad (7)$$

A gradient based learning algorithm can be used to minimise the $\boldsymbol{H\beta = T}$ by adjusting the parameters w_i, b_i and β_i , when the H hidden layer matrix is unknown iteratively . Studies carried out by Huang et al., /**13**/ proved that single layer feed forward neural network with randomly assigned input weights and hidden layer biases and with almost any nonzero activation function can universally approximate any continuous functions on any input data sets. Huang et al., /**14**/ suggested an alternate way to train a SHLFN by finding a least square solution β' of the linear system represented by (5) .The unique minimum norm least-square (LS) solution is modelled as

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T} \qquad (8)$$

where $\mathbf{H}^{\dagger}$ is the MP generalized inverse of matrix **H**. As analyzed by Huang, ELM using such MP inverse method tends to obtain good generalization performance with dramatically increased learning speed. The summarization of the ELM algorithm can be as:

Given a training set N = { ($\mathbf{x}_i$, $\mathbf{t}_i$)| $\mathbf{x}_i \in \mathbf{R}^n$, $\mathbf{t}_i \in \mathbf{R}^m$, $i = 1, \dots, N$ }, kernel function f(x), and hidden neuron $\tilde{N}$.

Step 1: Select suitable activation function and number of hidden neurons $\tilde{N}$ for the given problem .

Step 2 : Assign arbitrary input weight w_i and bias b_i , $i= 1, ...H$

Step 3 : Calculate the output matrix H at the hidden layer

$H = f. (\mathbf{w} \oplus \mathbf{x}+ \mathbf{b})$

Step 4: Calculate the output weight β .

$\hat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T}$

5. Performance Comparison with existing techniques in literature

Studies were carried on Harvard benchmark image datasets. The tumor area was manually segmented in each image by an expert radiologist. Two classification methods, Back propagation networks and Support vector machines were compared with the proposed Extreme Learning Machine. MRI was performed at 1.5 tesla and images consisted of contrast T1-weighted with matrix size of 256x128 for T1-weighted images both contrast and non-contrast images.This method is experimented on 3D axial slices of Harvard benchmark dataset /**15**/ which has 10 different patient data sets with expert segmentation as validation. The data was divided in two sets, training on 7 patients and testing on the other 3 patients. The output of the image is compared with the ground truth (target). Ground truth was obtained from the boundary drawings of the radiologist.

Internal segmentation /**16**/ evaluation were based on three validation measures based on True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) were computed:

$$\text{Sensitivity} = TP / (TP + FN) \qquad (9)$$

$$\text{Specificity} = TN /(TN + FP) \qquad (10)$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \qquad (11)$$

where, TP, TN, FP, and FN are the number of True Positive cases (abnormal cases correctly classified), the number of True Negatives (normal cases correctly classified), the number of False Positives (normal cases classified as abnormal), and the number of False Negatives (abnormal cases classified as normal) respectively. Accuracy is the proportion of correctly diagnosed cases from the total number of cases. Sensitivity measures the ability of the proposed method to identify abnormal cases. Specificity measures the ability of the method to identify normal cases.

Table 2 denotes the classification results for cerebrospinal fluid (CSF), White matter(WM) , Gray matter (GM) and Tumor (excluding oedema) for the GA –ELM approach. Table 5 denotes the summary of tumor segmentation results as compared to the SPL segmentation results and ITK segmentation. (CT is the Connected Threshold, NC is the Neighborhood Connected, and CC is the Confidence Connected algorithm /**17**/). The comparative analysis of the proposed method and the existing algorithms are shown in table 3 and 4.The optimization problem is solved using the MATLAB optimization toolbox R 2011b.

In ELM, initially the connection weights of the network are assigned randomly. After using sigmoidal activation function, the weight set is updated by applying a pseudo inverse matrix process. The ELM does not automatically select how many of hidden neurons may work well for the learning process. So employing incremental learning on ELM , the process can be done by adding one or more hidden neuron in each iteration and each time the performance of the learning should be tested. The best learning parameters are selected after maximum number of iterations.

Table 2 : Validation classification result on Tumor detection for GA –ELM approach

Parameter	**CSF**	**WM**	**GM**	**Tumor**
Sensitivity	89.06 %	78.31 %	72.04 %	94.17 %
Specificity	96.67 %	84.09 %	82.32 %	96.4 %

Table 3: Comparative Analysis of the Classifiers

Classifiers	**Sensitivity**	**Specificity**	**Accuracy w/o Feature sub - selection**	**Accuracy with Feature sub-selection**
BPN	82.3%	88.23%	76.19 %	82.25%
GA-SVM	94.17%	98.45%	92%	95.14 %
GA-ELM	94.77	97.98	92.8%	97.55%

Table 4 : ELM Experiment Results

Algorithm	MSE Error	Training Efficiency%		Testing Efficiency%		hidden neurons	Accuracy %	Time [a] (sec)
		Mean	STD	Mean	STD			
BPN	0.866	83.9	8.02	88.2	7.12	34	86.25	65
SVM- GA	0.909	93.9	8.02	88.2	7.12	8[b]	91.2%	75
ELM-GA	0.966	91.92	6.05	91.27	6.89	60	93.2%	45

[a] Computational time includes selection of number of hidden neurons

[b] Number of support vectors

Table 5: Summary of tumor segmentation results as compared to the SPL segmentation and ITK segmentation results

Case (SPL)	**CT**	**NC**	**CC**	**SPL**	**SVM -GA**	**ELM-GA**
1	0.94	0.97	0.97	0.98	0.97	0.97
2	0.94	0.91	0.86	0.91	0.89	0.90
3	0.97	0.95	0.97	0.96	0.96	0.94
4	0.87	0.91	0.88	0.91	0.88	0.89
5	0.68	0.71	0.70	0.85	0.65	0.70
6	0.96	0.94	0.87	0.98	0.91	0.90
7	0.57	0.69	0.93	0.84	0.90	0.90
8	0.98	0.96	0.96	0.88	0.93	0.92
9	0.93	0.94	0.94	0.97	0.93	0.94
10	0.92	0.94	0.94	0.96	0.93	0.94
Mean	0.87	0.86	0.90	0.91	0.895	0.9
STD	0.14	0.14	0.08	0.053	0.09	0.074

Figure : 2(a) Training Efficiency

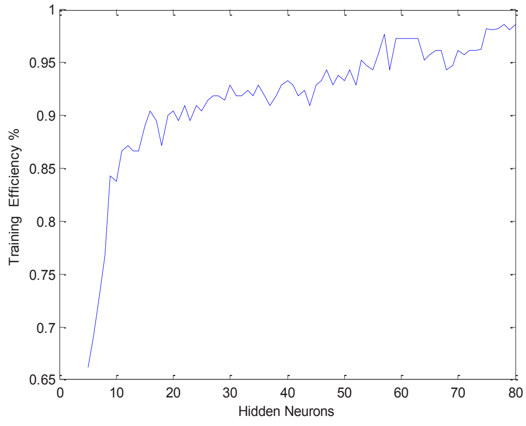


Figure: 2(b) Testing Efficiency

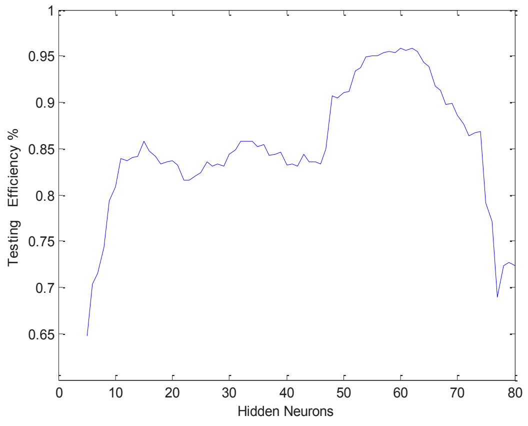


Figure: 3(a) Feature extraction

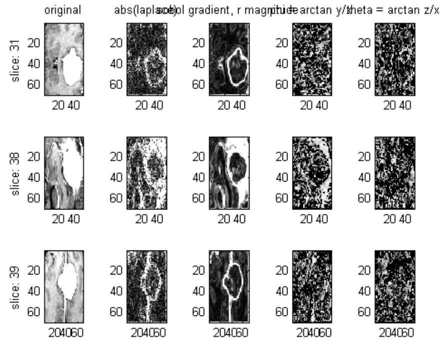


Figure: 3(b) Segmented tumor

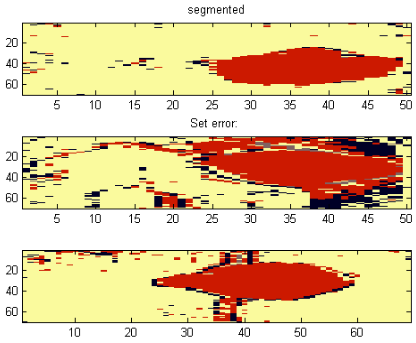


Figure: 4(a) Histogram plot of tumor

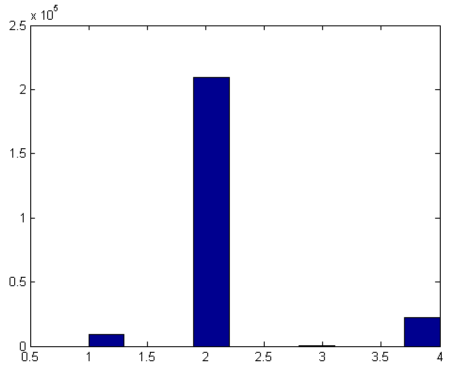
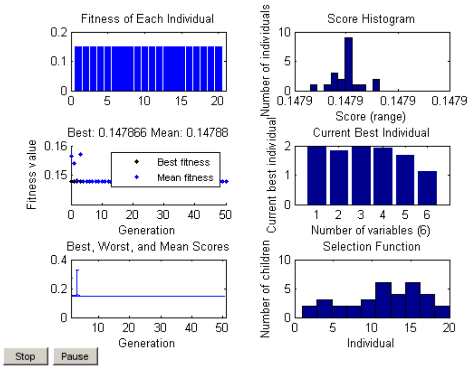


Figure: 4(b) Genetic Algorithm Generations



6. **Discussion**

The strong correlation between the expert defined ROI and segmentation manually and the algorithmically segmented tumors suggests that the algorithm is capable of detecting and following change relative to the standard of reference. The ability to track change in tumor contrast enhancement is important. Manually segmented images served as a standard of reference against which the performances of the algorithms were judged. T1-weighted contrast-enhanced images were analyzed. The approach includes pathology tissue tumour types classification of astrocytoma (AA), low grade glioma (GA) and meningioma (MA) along with brain tissue classification. Abnormal structures of brain and anatomical brain structures are identified with the 3D MRI image. Focus is on segmenting normal brain into four classes white matter, gray matter , cerobinospinal fluid and tumor /**18**/. The difference is a binary value for each pixel. It is 0 if the expert and the algorithm classification is same for the given pixel and it is 1 if they differ. The training set is about a few hundred pixels and test region is in the hundred thousand orders of magnitude. The analysis of how the raw intensity input is for the images of thresholded MRI is the most crucial consideration. The intensity values shows an exponential distribution .So a lower threshold below which everything is black and upper threshold above which everything is white is accounted. By considering mean of the object and background gray values as threshold, bias is avoided. Combining the conceptually simplest edge and pixel based segmentation. Since the original image is used for the segmentation process, the features represent a single pixel not the neighbourhood. Hence the complexity in computational time is avoided. The feature selection by genetic algorithm GA provides excellent selection capabilities for any data set. It employs a population of competing solutions evolved over time by cross over mutation and selection to converge to an optimal solution is employed .Table 1 reports the study parameters of Genetic Algorithm used . The Genetic algorithm produces a set of solutions instead of a single solution so as to avoid getting trapped in a local optimum. It is a promising method for feature selection over a high dimension space. In Figure. 4b , the last generation run of GA , the score distribution indicates the score being the result from the classifiers objective function. When generations proceeds and attains the last generation , population error distributution has diminished and score diversity disappears and the global minima reaches for all children. The individual best shows the value 2 and thus included in classification. The design of extreme learning machine requires setting of one user-defined parameter i.e. number of hidden nodes in hidden layer. A number of experiments were carried out by using the training and test data set with varying the hidden nodes from 5 to 80. Results suggests that extreme learning machine achieves highest classification accuracy of 97% with a total of 60 hidden nodes. The input data is normalized between values 0 and 1. The initial weights and bias between ±1.The unipolar sigmoidal activation function f (x) = 1 / (1 + e^{-x}) produces better results for the proposed approach. The mean square error (mse) is 0.9666. The mean standard deviation of training efficiency is 91.790 % and 0.587 .Similarly mean and standard deviation of testing efficiency is 0.8528 and 0.690. The testing and the training efficiency for various initial parameters are calculated for different number of hidden neurons. Fig 2(a) illustrates that the training efficiency increases with increase with number of hidden neurons .The efficiency is reached maximum with 75 hidden neurons .Similarly testing efficiency reaches a maximum when the hidden nodes are between 55 to 65as in Fig 2(b).The performance of the ELM network with 60 hidden neurons , increases considerably with respect to initial parameters , and training efficiency also reaches maximum at some random runs during this interval . The rich aggregate bank of features of the tissue image were calculated by using the gray-level run-length matrix (GL-RLM) method , GLCM , gradients and intensity . In the proposed system, twenty- seven features were extracted and used for obtaining optimized feature set as in Fig 3a. The optimal feature set with 8 features was obtained using GA. The most decisive features exploited to the ELM to obtain highest classification accuracy , include Low Gray-Level Run Emphasis(LGRE), High Gray-Level Run Emphasis (HGRE), run percentage (RPC), sobel gradient , neighbourhood statistics , mean and cluster tendency and entropy. The features were estimated and fed to the genetic algorithm for feature selection optimization. The strong practical need for image analysis methods that derive quantitative measures about tissue characteristics to incorporated into statistics comparing clinical features to magnetic resonance imaging (MRI)is required /**19**/. The genetic algorithm selects the best and the most promising 8 features after maximum iterations of 50 . The need to increase the sensitivity and specificity of co-occurrence features is required in order to enable recognition of textural distinction associated to various normal and pathological structures /**20**/ . Fig .3b shows the tumour extracted part and Fig.4a addresses the same with identification of tumour. The ELM-GA although are relatively simplistic using pixel and edge , increase the sensitivity of features extracted by reducing the number of large feature bank. The classification result with most selected primitive features is increased when compared to existing literature methods .

7. Conclusion

The tumor segmentation results when compared with literature shows promising results. Extreme learning machine has improved classification in contrast to Support Vector Machines, the most commonly used ANN for medical image classification. Tested on the SPL database , Table 2 , Table 3 and Table 5 address the same.Hence Extreme learning Machine provides an improved computational framework for the classification of detection of the healthy and cancerous tissues in the abnormal MRI image is developed(WM, GM, CSF and tumor). The segmentation algorithm automatically identifies intensity values that can be used to define a probability distribution that describes the brain area. The oedema and necrosis part can be further segmented to improve the efficacy of the approach. As a future work different types of tumor grades can be classified and size of tumor growth and more robust fully automatic segmentation approaches can be considered. Real data sets can be extended for future work.

Acknowledgement

This work was supported by the Council of Scientific and Industrial Research (CSIR), India. The authors would like to thank Drs. Simon Warfield, Michael Kaus, Arya Nabavi, Ron Kikinis, Peter Black and Ferenc Jolesz for sharing the tumor database.

Notes

/1/ Kloppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I. : Automatic classification of MR scans in Alzheimer's disease. Brain . 131 ,681-689(2008)

/2/ Huang, G.B., Zhu,Q.Y ., and Siew, C.K . : Extreme learning machine: Theory and applications. Neurocomputing. 70(1), 489-501(2006)

/3/ Galloway, M.M. : Texture Analysis Using Grey Level Run Lengths. Comp. Graph. And Image Proc, .4, 172-179(1975)

/4/ Haralick, R.M., Shanmugam, K., Dinstein, I. : Textural Features for Image Classification. IEEE Trans. Syst. Man Cybern, SMC.3, 610-621(1973)

/5/ Dana Cobzas, Neil Birkbeck , Mark Schmidt , Martin Jagersan : 3D Variational Brain Tumor Segmentation using a High Dimensional Feature Set. Proc .Int. Conf .Computer Vision. 1-8 (2007)

/6/ Georgiadis, P., Cavouras, D., Kalatzis, I., Glotsos, D., Sifaki,K., Malamas, M., Nikiforidis,G., and Solomou, E. : Computer aided discrimination between primary and secondary brain tumors on MRI: From 2D to 3D texture analysis. e-Journal of Science & Technology (e-JST). 9-18(2008)

/7/ Devos, A., Simonetti, A.W., van der Graaf , M., Lukas, L., Suykens, J.A., Vanhamme, L., Buydens, L.M., Heerschap, A., Van Huffel, S. : The use of multivariate MR imaging intensities versus metabolic data from MR spectroscopic imaging for brain tumour classification. J Magn Reson. 73, 218-228 (2005)

/8/ Ayelet Akselrod-Ballin, Meirav Galun, Moshe John Gomor , Ronen Basri, and Achi Brandt : Atlas Guided Identification of Brain Structures by Combining 3D Segmentation and SVM Classification. Proc . Medical Image Computing and Computer Assisted Intervention. Copenhagen, Denmark .209-216,(2006)

/9/ Galloway, M.M. : Texture Analysis Using Grey Level Run Lengths. Comp. Graph. And Image Proc, .4, 172-179(1975)

/10/ Tang, X. : Texture information in run-length matrices. IEEE Transactions on Image Processing. 7(11), 1602-1609(1998)

/11/ Haralick, R.M., Shanmugam, K., Dinstein, I. : Textural Features for Image Classification. IEEE Trans. Syst. Man Cybern, SMC.3, 610-621(1973)

/12/ Siedlecki, W., Sklanky, J. : A note on genetic algorithms for large-scale feature selection. Pattern Recognition Letters.10(5), 335 -347(1989)

/13/ Huang, G.B., Zhu,Q.Y ., and Siew, C.K . : Extreme learning machine: a new learning scheme of feedforward neural networks .Proc .Int .Conf .Neural networks, Budapest, Hungary . 985-990(2004)

/14/ Huang, G.B., Zhu,Q.Y ., and Siew, C.K . : Extreme learning machine: Theory and applications. Neurocomputing. 70(1), 489-501(2006)

/15/ Michael, R., Simon, K., Nabavi, A., Peter, M., Ferenc, A., Kikinis, R.: Automated segmentation of MR images of brain tumors . Radiology. 218, 586-91(2001)

/16/ David W. Shattuck, Stephanie R. Sandor-Leahy, Kirt A. Schaper, David A. Rottenberg and Richard M. Leahy : Magnetic Resonance Image Tissue Classification Using a Partial Volume Model. NeuroImage. 13, 856-876(2001)

/17/ Dirk Padfield and James Ross , Validation Tools for Image segmentation. Proc. of SPIE ,Medical Imaging. Florida , USA , 72594W (2009)

/18/ Zijdenbos, A., Forghani, R., and Evans, A . : Automatic quantification of MS lesions in 3d MRI brain data sets: Validation of INSECT. Proc. Int .Conf. Medical Image Computing and Computer Assisted Intervention. Cambridge MA , USA . 439-448(1998)

/19/ Good, C.D ., Scahill, R.I ., Fox ,N.T. : Automatic differentiation of anatomical patterns in the human brain: Validation with studies of degenerative dementias. Neuroimage. 17(1),29-46(2002)

/20/ Wang, C.M., Wu, M.J., Chen, J.H., Yu, C.Y.: Extension Neural Network Approach to Classification of Brain MRI. Proc. Int.Conf . Intelligent Information Hiding and Multimedia Signal Processing. Kyoto , Japan, 515-517(2009)

Literature

1. Alan Wee-Chung Liew and Hong Yan : Current Methods in the Automatic Tissue Segmentation of 3D Magnetic Resonance Brain Images. Current Medical Imaging Reviews. 2 (1) , 91-103 (2006)
2. Byun, H. , and Lee, S . : Applications of Support Vector Machines for Pattern Recognition: A Survey. Proc. Int .Work . Pattern Recognition with Support Vector Machines (SVM 2002).Niagara Falls, Canada , 213-236 (2002)

3. Chu A., Sehgal C. M., and Greenleaf J. F. : Use of Gray Value Distribution of Run Lengths for Texture Analysis. Pattern Recognition Letters. 11(6), 415-420 (1990)
4. Chris A. Cocosco , Alex P. Zijdenbos, Alan C. Evans : A Fully Automatic and Robust Brain MRI Tissue Classification Method . Medical Image Analysis. 7(4) , 513 -527(2003)
5. Clarke, L.P., Velthuizen,R.P., Camacho,M.A., Heine,J.J., Vaidyanathan,M., Hall,L.O., Thatcher, R.W., Silbiger,M.L. : MRI Segmentation: Methods and Applications. Magnetic Resonance Imaging. 13(3), 343-368 (1995)
6. Deepa ,S.N., and Aruna devi, B., : Second Order Sequential Minimal Optimization for Brain Tumour Classification. European Journal of Scientific Research. 64(3), 377-386 (2011)
7. Deepa ,S.N., and Aruna devi, B ., : A survey on Artificial Intelligence approaches for medical image Classification. Indian Journal of Science and Technology. 4(11), 1583-1595 (2011)
8. Evangelia I. Zacharaki, Sumei Wang, Sanjeev Chawla, Dong Soo Yoo, Ronald Wolf, Elias R. Melhem, and Christos Davatzikos : Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme . Magn Reson Med. 62(6) , 1609-1618 (2009)
9. Mahmoud-Ghoneim, D., Toussaint, G., Constans, J.M., de Certaines , J.D.: Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. Magn Reson Imaging . 21,983-987(2003)
10. Nan-Ying Liang, Paramasivan Saratchandran, Guang-Bin Huang, Narasimhan Sundararajan : Classification of mental tasks from EEG signals using Extreme Learning Machine .International Journal of Neural Systems. 16(1), 2938(2006)
11. Karin Junker, Dusan Sovilj, Ingrid Kroncke, Joachim W. Dippner : Climate induced changes in benthic macrofauna–a non-linear model approach. Journal of Marine Systems (2012).doi: 10.1016/j.jmarsys.2012.02.005 .
12. Li, G., Yang, J., Ye, C., Geng, D .: Degree prediction of malignancy in brain glioma using support vector machines,. Computers in Biology and Medicine. 36 (3), 313-325(2006)
13. Lao, Z., Shen, D., Liu, D .: Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. Acad Radiol. 15 , 300-313(2008)
14. Patil , M.M ., Yardi , A.R. : Classification of 3D Magnetic Resonance Images of Brain using Discrete Wavelet Transform , International Journal of Computer Applications, 31(7), 23-27(2011)
15. Suresh,S., Saraswathi, S., and Sundararajan, N. : Performance enhancement of extreme learning machine for multi-category sparse cancer classification problems. Engineering Applications of Artificial Intelligence . 23, 1149-1157(2012)
16. Suresh, S., Babu, R.V., and Kim, H.J. : No-reference image quality assessment using modified extreme learning machine classifier. Applied Soft Computing . 9, 541-552(2009)
17. Sahoo, P.K., Soltani, S., Wong, A.K.C.: A survey of thresholding techniques . Comp Vis Graph Image Processing . 41, 233-260(1988)
18. Zhu, Q.Y., Qin, A.K., Suganthan, P.N., Huang, G.B. : Evolutionary extreme learning machine. Pattern Recognition. 38(10) ,1759-1763(2005)