IJCTA, 9(26) 2016, pp. 325-333

An Enhanced Classifier for Brain Tumor Classification

Jokin Arul Raj[1] and Sathees Kumar[2]

ABSTRACT

Brain tumor is life threatening and the major reason for the death in the present world. The growth of abnormal cells in brain leads to brain tumor. Brain tumor is classified into benign tumor and malignant tumor. Benign tumor is non-cancerous whereas malignant is cancerous. Diagnosing at earlier stage can save the person. It is really a great challenge to find the brain tumor and classifying its type. The proposed system has many stages like segmentation, noise removal, textural feature enhancement, feature selection and Classification phase. The proposed system is more accurate in classifying the tumor regions.

Keywords: Brain tumor, Segmentation, Pre-processing, Textural Feature extraction, PSO based Feature Selection, Classification.

1. INTRODUCTION

Brain tumor is an abnormal growth of cancerous or non-cancerous cells in the brain. The cause for the brain tumor is unpredictable. But two major reasons are radiations, and rare genetic condition. Brain tumors are commonly detected by CT (Computed Tomography), MRI (Magnetic Resonance Imaging) and PET (Positron Emission Tomography) [19]. Among these three, MRI is the most commonly used test to diagnosis. Doctors evaluate the tumor from the imaging tests and plan for the treatment. Treatment for the brain tumor is based on the type, size and location of the tumor cells. Benign tumor is non-cancerous and it will not invade the neighbor cells. It can be removed but it may grow again. Malignant brain tumor grows rapidly and affect the nearby cells. There are four grades of tumors from Grade I to IV. Grade I is benign and Grade II is malignant. Grade III is malignant as well as look very different from the normal cells. Grade IV is abnormal cell which grows very fast [21].

Diagnosing the brain tumor is still a difficult problem. Finding the exact location of the tumor part is a great challenge in the medical imaging test. Image mining techniques help to find the exact tumor cells by undergoing many process. It helps to recover the relevant information and reveals the image patterns that are noteworthy from the given dataset. Proposed system uses the dataset obtained from The World Brain Atlas. Pre-processing, Feature extraction, feature selection, segmentation, and classifications are the steps involved in image mining. Image pattern representation, feature selection and pattern visualization are the major tasks as the details in an image has to be presented with accuracy for the diagnosis.

1.1 Paper Organization

The paper is organized in the following manner:

Section 2 contains related work.

[1] Research Scholar, Computer Science & Bishop Heber College, Tamil Nadu, India

[2] Assistant Professor, Computer Science & Bishop Heber College, Tamil Nadu, India

Section 3 contains the process involved in the proposed system.

Section 4 represents the Result. Section 5 has the conclusions and the future enhancement.

2. RELATED WORK

Quratul Ain [19] has proposed a system, where FDCT is performed for removal of noise and then features are extracted from that noise free images. Classification phase uses ensemble base classifier using SVM. Segmentation has multi-step phase. It extracts the tumor region using FCM from the brain portion.

Jan Luts [5], in his paper investigated the use of automated pattern recognition techniques on MRI data with a goal to help the clinicians in diagnosis of tumor. Magnetic Resonance Spectroscopic image has provided metabolic information of the brain. Multiclass classifier has been built from the image intensities and spectroscopic information. Using LS-SVM along with class probabilities and feature selection, tumor regions are recognized.

Saima Anwar Lashari[14] proposed a framework for medical image classification with six phases. Data acquisition is first phase, followed by data pre-processing. Then data partition, soft set classifier, data analysis and finally performance. Dataset was categorized into training and test set and classification process was performed.

To classify the brain CT scan images into normal, benign and malignant, P.Rajendran proposed a technique which is based on the associative classification scheme [7]. Using prune association rule with MARI algorithm, he developed an improved image mining technique for brain tumor classification. Histogram equalization and hybrid median filter were used to enhance the image quality in preprocessing. Feature extraction has done through the Spatial Gray Level Dependent Feature (SGLDF).

3. PROPOSED METHOD

The proposed system segments and categorizes the brain tumor in multiple stages. First stage is segmentation where the brain portion is extracted separately from the skull part and all other portion present in an image. In Pre-processing noise present in the image is removed. Afterhis, features are extracted and among those features, feature selection process involves. Finally classification method classifies the tumorous region.

3.1 Segmentation

Segmentation is the process to divide the image into many different parts. Sharp discontinuities in an image is classified and processed. It helps to recognize the portions or appropriate information from the digital images. So segmentation helps to identify the normal and the cancerous tissues in the brain image which is obtained through MRI. It has various methods. Otsu's method, *K*-means clustering, watershed segmentation, texture filter are few segmentation methods. Otsu's method is a Thresholding method. In that method, image is imported and background of the image is analyzed and estimated. The contrast of the image is increased and undergo into threshold process. The object or the part is identified and examined. Statistical data is generated from the identified parts. Color-based segmentation using *K*-means clustering is another method. It converts RGB color image into CIELAB which is defined by international commission where '*L*' represents luminosity layer, '*A*' represents chromaticity layer where the color comes in the red-green axis and '*B*' indicates chromaticity-layer where the color lies along the blue-yellow axis. Then the colors are identified and segmented. Marker-Controlled Watershed Segmentation has few basic procedures. It computes the foreground markers and background markers. It uses gradient magnitude as segmentation function.

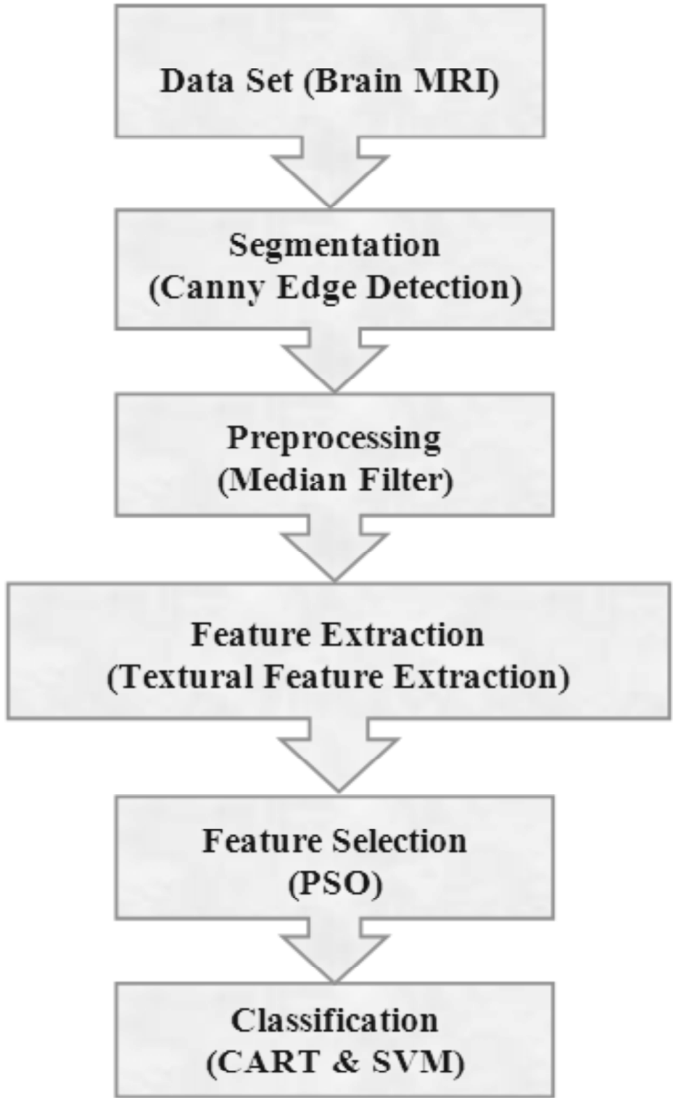


Figure 1: Proposed System

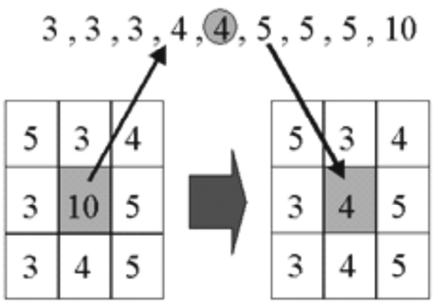
Although we got many segmentation process, accuracy is still a great challenge. In proposed system Canny Edge Detection is used which was developed by John F. Canny (JFC) in the year 1986. Detecting real edge points in the image are maximized whereas falsely detected non-edges are minimized. The detected edges have to be as close as possible to the real edges.

Steps involved in Canny Edge Detection:

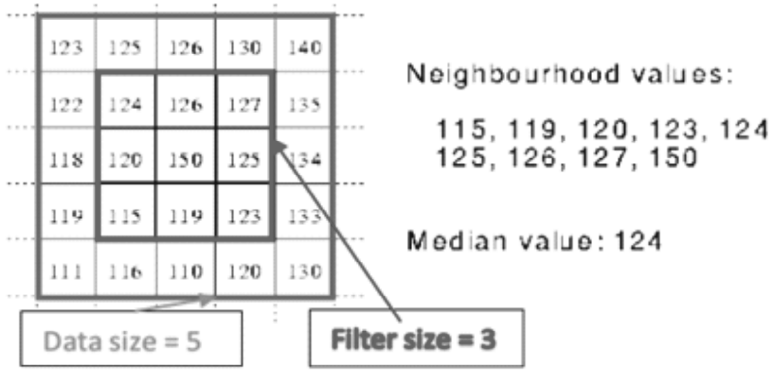
1. Smoothening makes the MRI to blur and this helps in removing the noise.
2. Finding gradient in the brain image finds the normal and tumorous cells by detecting the edges.
3. Blurred edges are then converted into sharp edges by non-maximum suppression.
4. Double threshold to determine the potential edges.
5. All edges that are not connected to a very certain or strong edges are suppressed to determine the final edges.

It is implemented by BLOB (Binary large object) analysis. BLOB containing the strong edge pixel is preserved and all other BLOB are suppressed. This is edge tracking hysteresis.

After the brain image going through all the above steps, the tumor part is segmented. It identifies the tumor region and separates the normal cell area in the MRI images.



(a) 3 × 3



(b) 5 × 5

Figure 2: Illustration to calculate the median value

3.2 Preprocessing

The proposed system uses median filter to remove noise from the MRI. Median filter has robustness capacity as well as preserving capacity. Two procedures are applied. They are noise lessening and fundamental image enhancement. It reduces “salt and pepper” noise as it is non-linear operation. It also smoothens an image, hence the blurring of edges are reduced. The idea applied is to replace the current point in the image by the median value of the brightness of the nearby pixels *i.e.*, replacing each entry with the median value of the neighbor value. Median value do not affected by the individual noise spikes. Noise produced by film artifact is removed. Hence it eliminates the impulse noise. This noise free image is further carried to the next phase.

3.3 Feature Extraction

Feature extraction is the process of transforming an image into its set of features. They are obtained from color, texture and shape.Good features should have Informativeness/Distinctiveness, Locality, Accuracy, and Repeatability, Quantity, Robustness, and Efficiency.These features are extracted for classification process. But it is quite challenging to extract the main and useful feature. Many feature extraction techniques are available. Few are gabor features, texture feature, decision boundary feature extraction, discriminant analysis. Texture feature is used in the proposed system.

3.3.1 Texture Features

Proposed system used two methods for texture feature extraction. They are first order histogram and second order texture feature.

A. First order histogram

Histogram reveals the summary of the statistical information present in an image. So through histogram, first order statistical information can be obtained from an image. Dividing the value of intensity level histogram by total number of pixels present in an image gives the probability density occurrence of the intensity level.

$$P(i) = h(i)/NM, \ i = 0,1, \ldots, G - 1 \tag{1}$$

Where N = no. of resolution cells in the horizontal spatial domain

M = no. of resolution cells in the vertical spatial domain

G = total gray level

Following features can be obtained from the first order histogram:

$$\text{Mean: } \mu = \sum_{I=0}^{G-1} ip(i) \tag{2}$$

[Average value of the intensity of an image]

Variance:

$$\sigma^2 = \sum_{I=0}^{G-1} (i-\mu)^2 p(i) \tag{3}$$

[Intensity variation of the mean value]

Skewness:

$$\mu_3 = \sigma^{-3} \sum_{I=0}^{G-1} (i-\mu)^3 p(i) \tag{4}$$

[Symmetriness of the histogram around the mean]

Kurtosis:

$$\mu_4 = \sigma^{-4} \sum_{I=0}^{G-1} (i-\mu)^4 p(i) - 3 \tag{5}$$

[Flatness of the histogram]

Energy:

$$E = \sum_{I=0}^{G-1} [p(i)]^2 \tag{6}$$

[Recognizes the disorders in textures]

Entropy:

$$H = -\sum_{I=0}^{G-1} p(i) \log_2 [p(i)] \tag{7}$$

[Uniformaty of the histogram]

B. Second order texture feature - Co-occurrence matrix based features

Histogram based features does not consider spatial information. It is local in nature. So second order histogram based features are defined which is otherwise known as grey level co-occurrence matrix $h_d(i, j)$ based features. Joint probability distribution of pairs of pixels are calculated with Distance d and angle θ within given neighborhood pixels. For calculations, $d = 1, 2$ and $\theta = 0^o, 45^o, 90^o, 135^o$ are used.

Angular second moment (energy):

$$\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [p(i, j)]^2 \tag{8}$$

Correlation:

$$\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{ij\, p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{9}$$

1	1	2	2
1	1	2	2
4	2	2	3
4	3	3	3

(a) Image Example

i/*j*	1	2	3	4
1	#(1,1)	#(1,2)	#(1,3)	#(1,4)
2	#(2,1)	#(2,2)	#(2,3)	#(2,4)
3	#(3,1)	#(3,2)	#(3,3)	#(3,4)
4	#(4,1)	#(4,2)	#(4,3)	#(4,4)

(b) Construction of Co-occurrence Matrix

4	2	0	0
2	6	1	1
0	1	4	1
0	1	1	0

(c) $h_{1,0^\circ}$

Figure 3: Illustration for spatial co-occurrence matrix calculation

Inertia:

$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2\,p(i,j) \tag{10}$$

Absolute value:

$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}|i-j|\,p(i,j) \tag{11}$$

Inverse difference:

$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{p(i,j)}{1+(i,j)^2} \tag{12}$$

Entropy:

$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)\log_2[p(i)] \tag{13}$$

3.4 Feature Selection

Particle Swarm Optimization is one of the best feature selecting techniques. PSO was developed by Eberhat and Kennedy. It is a successful and suitable algorithm for feature selection as it is easy to encode the feature andglobal search facility and easy to device. It is a population based stochastic optimization method. Every particle represent potential solution for the problem. The search space through which a subset of major components or principal features were discovered and carefully chosen via PSO is principal space.Distributing 0s and 1s randomly generates the swarm of particles. If 1 is the principal component, it is selected and ignored if it is 0 [20].

Personal best *(pbest):* the position that has provided the greatest realization. Local best *(lbest):* the position of the best particle member of the neighbor of a given particle. Global best *(gbest):*position of whole swarm. The velocity is the vector that decides the direction in which a particle needs to move to improve its present position. The particle is moved to better regions of the search space by the leader of the particle.

3.5 Classification

Classification is the process of classifying the items according to its type and pattern. Selecting the suitable classifier results in accuracy and improved performance for various datasets. Here, the tumors are classified as benign and malignant tumor. So the proposed system combines CART classification and ensemble SVM based classification to make the process more efficient than the existing process. It is a discrete hybrid method. Classification tree and regression tree is combined and termed as CART. Classification tree predicts the class to which the data belongs to and the regression tree will predict the exact values in real numbers. SVM helps in clustering the data into groups. So this has increased the accuracy in detecting the exact tumor affected area.

CART Classification

CART uses the learning set for building decision tree. It works on non-linear input. CART algorithm classifies the tumor region by its own. CART uses the binary tree method for classification. So the image classified easily by predicting the sudden change in the intensity value in the brain image.

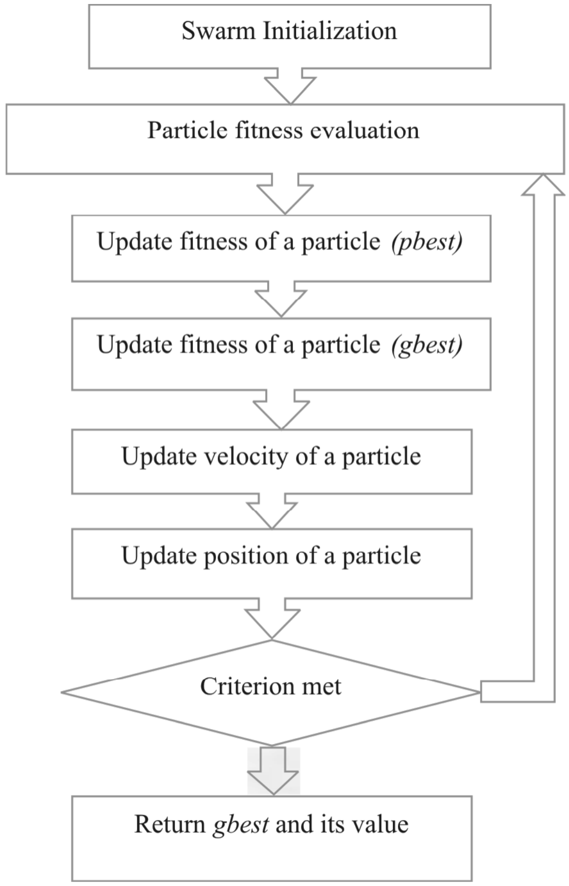
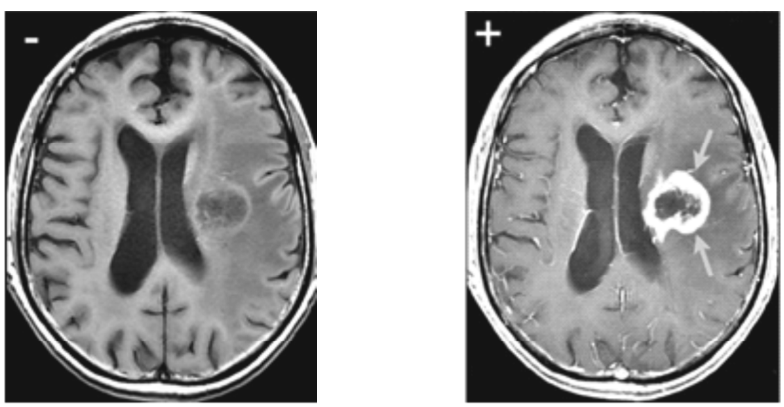


Figure 4: Feature selection process (PSO)



(a) **(b)**

Figure 5. (a) Input Image (b) Processed Image

Ensemble SVM Classification

Ensemble SVM (Support vector machine) is used for classification process. It works on linear data. It is also used to classify many real time problems like text categorization, face recognition, cancer diagnosis and many. It uses divide and conquer strategy. Based models are combined with high predictive performance in ensemble SVM. It is effective in high dimensional regions and classifies the image.

4. RESULTS

The accuracy and the performance of the proposed system can be tested using confusion matrix. Outcomes of the prediction can be True Positive (TP), False Positive (FP), False Negative (FN), and True Negative

(TN). Normal images and tumor images are categorizedcorrectly asTrue Negative(TN) and True Positive(TP).In some cases, test says no tumor but it has tumor, then it is False Negative (FN) and test says tumor is present but it has no tumor, then it is False Positive (FP). 39 images were tested using this confusion matrix. The performance of the binary classification test has two major statistical measures in medical field and they are sensitivity and specificity. Positive values correctly identified is given in percentage by sensitivity. Negative values correctly identified is given in percentage by specificity.

Table 1
Confusion Matrix Measures

	True Positive	*True Negative*
Predicted Positive	29	2
Predicted Negative	1	7

Table 2
Effectiveness of proposed system

Performance Measure	*Result*	*Formula*
Sensitivity	96.67	TP/(TP + FN)
Specificity	77.78	TN/(FP + TN)
Accuracy	92.31	(TP + TN)/(TP + FP + FN + TN)

CONCLUSION

The proposed system has been developed to diagnosis of brain tumor from Magnetic Resonance Imaging (MRI) of the brain. There are many phases to detect the tumor part from the brain image. Segmentation extracts the tumor portion. Median filter is used for noise removal. Textural feature extraction is used in proposed system. Classification uses CART and SVM classifiers which is hybrid method. As CART and SVM is combined,proposed system achieved 92.31% of accuracy. Future enhancement is to perform with the Perfusion based MRI images which is quite complex.

REFERENCES

[1] Robert M. Haralick, K. Shanmugam, Its'hakDinstein,"Textural Features for Image Classification",*IEEE Transactions on Systems, Man, and Cybernetics,* **3(6)**, November 1973.

[2] John Canny, "A computational approach to edge detection. Pattern Analysis and MachineIntelligence", *IEEE Transactions* on, PAMI-8(6): 679–698, Nov. 1986.

[3] A. Materka, M. Strzelecki, "Texture Analysis Methods – A Review", *Technical University of Lodz, Institute of Electronics,* COST B11 report, Brussels 1998

[4] Chung-JuiTu, Li-Yeh Chuang, Jun-Yang Chang, and Cheng-Hong Yang, Member, IAENG, "Feature Selection using PSO-SVM", *International Journal of Computer Science*, 33:1, IJCS_33_1_18, 13 February 2007.

[5] Jan Luts, ArendHeerschap, Johan A.K. Suykens, Sabine Van Huffel, "A combined MRI and MRSI based multiclass system for brain tumour recognition using LS-SVMs with class probabilities and feature selection",*Artificial Intelligence in Medicine* **40**, 87–102 ELSEVIER. 2007.

[6] Yangqiu Song, Changshui Zhang, Jianguo Lee, Fei Wang, Shiming Xiang, Dan Zhang,"Semi-supervised discriminative classification with application to tumorous tissues segmentation of MR brain images", *Springer-Verlag* London Limited 2008.

[7] P. Rajendran, M. Madheswaran, "An Improved Image Mining Techniques for Brain Tumour Classification Using Efficient Classifier"*(IJCSIS), International Journal of Computer Science and Information Security,* **6(3)**, 2009.

[8] Mr. Salem Saleh Al-amri, Dr. N.V. Kalyankar, Dr. Khamitkar S.D, "Image Segmentation By Using Edge Detection",*(IJCSE) International Journal on Computer Science and Engineering* 2(3), 2010.

[9] Muthukrishnan.R, M.Radha, "Edge Detection Techniques for Image Segmentation", *International Journal of Computer Science & Information Technology (IJCSIT),* **3(6)**, Dec 2011.

[10] Dr.E.Chandra, K.Kanagalakshmi, "Noise Elimination in Fingerprint Image Using Median Filter", *Int. J. Advanced Networking and Applications*, **2(6)**, Pages: 950-955 (2011).

[11] Sudipta Roy, Samir K. Bandyopadhyay, "Detection and Quantification of Brain Tumor from MRI of Brain and it's Symmetric Analysis", *International Journal of Information and Communication Technology Research,* **2(6)**, June 2012.

[12] E. Ben George, M.Karnan, "MRI Brain Image Enhancement Using Filtering Techniques",*International Journal of Computer Science & Engineering Technology (IJCSET),* **3(9)**, Sep 2012.

[13] Indrajit Mandal, Sairam.N, "SVM-PSO based Feature Selection for Improving Medical Diagnosis Reliability using Machine Learning Ensembles",*SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 06,* 267–276, 2012, CS & IT-CSCP 2012 DOI : 10.5121/csit.2012.2326

[14] Saima Anwar Lashari, Rosziati Ibrahim, "A Framework for Medical Image Classification Using Soft Set",*Elsevier, The 4th International Conference on Electrical Engineering and Informatics* (ICEEI 2013).

[15] MeghanaNagori,Madhuri Joshi, "Methods and Algorithms for Extracting Values from MRSGraph for Brain Tumour Detection", *International Conference on Electronic Engineering and Computer Science, ELSEVIER* 2013

[16] S. Sivakumar ,Dr.C.Chandrasekar, "Modified PSO Based Feature Selection for Classification of Lung CT Images", *(IJCSIT) International Journal of Computer Science and Information Technologies,* **5(2)** , 2014.

[17] Kailash D. Kharat, Anil V. Turukmane "Classification of Brain Tumor by Feature Vector Using Cart and C5.0 Algorithms",*JAAST:Material Science (Special Issue),* **1(2)**, 173-178, 2014.

[18] Gaurav Mandloi, "A Survey on Feature Extraction Techniques for Color Images*", (IJCSIT) International Journal of Computer Science and Information Technologies*, **5(3)**, 2014.

[19] QuratulAina, M.ArfanJaffar, Tae-SunChoic,"Fuzzy anisotropic diffusion based segmentation and texture based ensemble classification of brain tumor", *ELSEVIER Applied Soft Computing,* **21**, 330–340, 2014.

[20] Iftikhar Ahmad, "Feature Selection Using Particle Swarm Optimization in Intrusion Detection",*Hindawi Publishing Corporation, International Journal of Distributed Sensor Networks,* 2015.

[21] Strong MJ, Garces J, Vera JC, Mathkour M, Emerson N, et al. (2015) Brain Tumors: Epidemiology and Current Trends in Treatment.*Journal of Brain Tumors &Neurooncol,* 1: 102. doi:10.4172/jbtn.1000102

[22] Mohammad HadiYousofi, Mahdi Esmaeili, Majide Sadat Sharifian, "A Study on Image Mining; Its Importance and Challenges", *American Journal of Software Engineering and Applications,* **5(3-1)**, 5-9, 2016,