Research Journal of Applied Sciences, Engineering and Technology 10(2): 177-187, 2015
DOI:10.19026/rjaset.10.2570
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: October 10, 2014 Accepted: December 18, 2014 Published: May 20, 2015

Research Article

Brain Tumor Detection and Classification Using Deep Learning Classifier on MRI Images

[1]V.P. Gladis Pushpa Rathi and [2]S. Palani
[1]Department of Computer Science and Engineering, Sudharsan Engineering College,
[2]Sudharsan Engineering College, Sathiyamangalam, Pudukkottai, India

Abstract: Magnetic Resonance Imaging (MRI) has become an effective tool for clinical research in recent years and has found itself in applications such as brain tumour detection. In this study, tumor classification using multiple kernel-based probabilistic clustering and deep learning classifier is proposed. The proposed technique consists of three modules, namely segmentation module, feature extraction module and classification module. Initially, the MRI image is pre-processed to make it fit for segmentation and de-noising process is carried out using median filter. Then, pre-processed image is segmented using Multiple Kernel based Probabilistic Clustering (MKPC). Subsequently, features are extracted for every segment based on the shape, texture and intensity. After features extraction, important features will be selected using Linear Discriminant Analysis (LDA) for classification purpose. Finally, deep learning classifier is employed for classification into tumor or non-tumor. The proposed technique is evaluated using sensitivity, specificity and accuracy. The proposed technique results are also compared with existing technique which uses Feed-Forward Back Propagation Network (FFBN). The proposed technique achieved an average sensitivity, specificity and accuracy of 0.88, 0.80 and 0.83, respectively with the highest values as about 1, 0.85 and 0.94. Improved results show the efficiency of the proposed technique.

Keywords: Deep learning classifier, Linear Discriminant Analysis (LDA), MRI image segmentation, Multiple Kernel based Probabilistic Clustering (MKPC), shape and texture based features, tumour detection

INTRODUCTION

These days, MRI (Magnetic resonance Imaging) image classification has surfaced as challenging function mainly in view of the divergence and intricacy of tumours. The brain tumor is essentially an intracranial solid neoplasm or anomalous surge of cells parked in nucleus of the brain or the central spinal canal. As a matter of fact, Detection of the brain tumor in the nascent phases goes a long way in the effective and critical treatment of this ailment (Bankman, 2009). Commonly, it can be said that premature phase brain tumor diagnose essentially encompasses Computed Tomography (CT) scan, Nerve test, Biopsy etc (Sapra *et al*., 2013). Brain cancer, in fact, is triggered by means of a malignant brain tumor. Though, certain kinds of brain tumors are found to be benign (non-cancerous). In this regard, the two important kinds of brain cancer embrace primary brain tumor and secondary brain tumor (Bankman, 2009).

Incidentally, feature extraction and choice are the most important stages in breast tumor detection and classification. An optimum feature set must have efficient and discerning features, simultaneously diminishing in almost cases the excess of features pace to eliminate the "curse of dimensionality" dilemma Acir *et al*. (2006). In this stage, an optimal subset of features which are essential and enough for resolving an issue is shortlisted (Karabatak and Ince, 2009). These features are mined by means of image processing techniques. Diverse kinds of feature extraction doing elegant rounds are from digital mammograms integrating position feature, shape feature and texture feature etc., (Sharma *et al*., 2012). Texture measures are divided to two general kinds such as first order and second order (Georgiardis *et al*., 2008).

Various classification techniques from statistical and machine learning domain have been performed on cancer classification, which is a fundamental function in data analysis and pattern recognition entailing the assembly of a Classifier. Fisher linear Discriminat analysis (Sun *et al*., 2012), k-nearest neighbour (Wang *et al*., 2012) decision tree, multilayer perceptron (Gholami *et al*., 2013) and support vector machine (Sridhar and Krishna, 2013), Artificial Neural Network (ANN) (Kharat *et al*., 2012). There have been many literatures present for tumor detection in medical image processing. Sun *et al*. (2012) have zealously launched the tumor classification by means of Eigen gene-based classifier committee learning technique. In this regard, Eigen gene mined by Independent Component Analysis (ICA) was regarded as one of the most excellent and

Corresponding Author: V.P. Gladis Pushpa Rathi, Department of Computer Science and Engineering, Sudharsan Engineering College, Sathiyamangalam, Pudukkottai, India

effectual feature for tumor classification. They effectively employed Eigen gene together support vector machine dependent Classifier Committee Learning (CCL) method.

Wang *et al*. (2012) have shrewdly spelt out the tumor classification approach in accordance with the correlation filters to recognize the whole design of tumor subtype concealed in characteristically expressed genes. Concretely, two correlation filters such as Minimum Average Correlation Energy (MACE) and Optimal Tradeoff Synthetic Discriminant Function (OTSDF), were launched to assess whether a test sample tally with the templates combined for each subclass. Sridhar and Krishna (2013) have significantly manipulated the tumor Classification by means of the Probabilistic Neural Network in tandem with Discrete Cosine Transformation. Gholami *et al*. (2013) has brilliantly given details of the Statistical Modeling Approach for Tumor-Type Recognition in Surgical Neuropathology by means of Tissue Mass Spectrometry Imaging. Especially, mass spectrometry imaging was employed to achieve the chemical structure of a tissue component and, therefore, furnished a frame to examine the molecular structure of the sample while upholding the morphological features in the tissue.

Kharat *et al*. (2012) has competently dealt with the two Neural Network techniques for the classification in respect of the magnetic resonance human brain images. In the first stage, they achieved the features connected with MRI images by means of Discrete Wavelet Transformation (DWT). In the subsequent stage, the features of Magnetic Resonance Images (MRI) were cut down with the execution of Principles Component Analysis (PCA) on the further important features. In the classification stage, two classifiers with foundation on supervised machine learning were devised. Ali *et al*. (2013) have astoundingly advocated the brain tumor extraction by means of clustering and morphological operation methods. The domain of the mined tumor regions was estimated. The investigation demonstrated that the four performed methods fruitfully recognized and mined the brain tumor, thereby extending an olive branch to the doctors in recognizing the dimension and domain of the tumor.

Tumor segmentation and classification using multiple kernel-based probabilistic clustering and deep learning classifier is proposed in this study. The proposed technique consists of three modules, namely segmentation module, feature extraction module and classification module. The input MRI (Magnetic resonance Imaging) images are pre-processed and segmented using Multiple Kernel based Probabilistic Clustering (MKPC) in segmentation module. Features are extracted for every segment based on the shape (circularity, irregularity, Area, Perimeter, Shape Index). Additionally, texture (Contrast, Correlation, Entropy, Energy, Homogeneity, cluster shade, sum of square variance) and intensity (Mean, Variance, Standard Variance, Median Intensity, Skewness and Kurtosis) from segmented regions. After features extraction, important features will be selected using Linear Discriminant Analysis (LDA) to classification purpose. In training phase, deep learning classifier is trained with the features of training data and in testing, the features from the segmented image are fed into the trained Deep learning classifier to detect whether the region has brain tumour or not.

MATERIALS AND METHODS

Tumor segmentation and classification using multiple kernel-based probabilistic clustering and deep

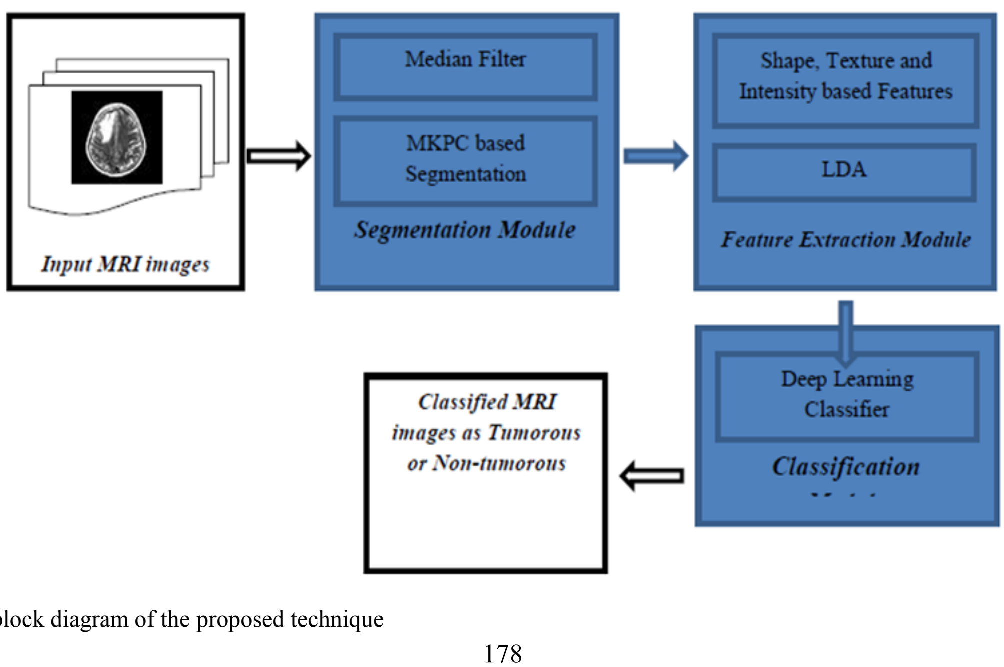

Fig. 1: The block diagram of the proposed technique

learning classifier is described here. The proposed technique consists of three modules, namely segmentation module, feature extraction module and classification module. In segmentation module, initially the image is pre-processed using median filter and then segmented using Multiple Kernel based Probabilistic Clustering (MKPC). Subsequently, shape, texture and intensity based features are extracted in feature extraction module. From these, important features are selected using Linear Discriminant Analysis (LDA). Finally in classification module, deep learning classifier is employed having two important processes of training phase and testing phase. The block diagram of the proposed technique is given in Fig. 1.

Segmentation module: The input MRI (Magnetic resonance Imaging) images are pre-processed and segmented using Multiple Kernel based Probabilistic Clustering (MKPC). Pre-processing is carried out to make the input images fit for segmentation by removing noise. It is carried out by the use of linear smoothing filters such as median filter.

Median filtering: Median filter which is nonlinear digital filtering technique is used for noise reduction of the input MRI image. It is used as the pre-processing step to improve the image and make it fit for later processing. Median filter has the advantage that it preserves edges while removing noise. The median filter works by moving through the image pixel replacing each pixel with the median of neighbouring pixels for each considered pixel. The pattern of neighbours is termed as the window, which slides, pixel by pixel over the entire image. The median is calculated by first sorting all the pixel values from the window into numerical order and then replacing the pixel being considered with the middle (median) pixel value.

Suppose, let the image pixels in the window be represented as $\{p_1, p_2, p_3, p_4, p_5\}$. In order to find out the median, initially sorting is carried out to have the sorted list given by:

$$\begin{aligned} &Sorted\,List = \{p_3, p_1, p_5, p_4, p_2\}\ given \\ &p_3 < p_1 < p_5 < p_4 < p_2;\ Then\ Median = p_5 \end{aligned} \tag{1}$$

We can see that the median is p_5 as it falls in the middle of the sorted pixels. The employment of median filter results in de-noising the input image and makes it fit for segmentation and feature extraction modules.

Multiple Kernel based Probabilistic Clustering (MKPC): Segmentation of the pre-processed image to Tumorous and Non-Tumorous regions is carried out with the use of Multiple Kernel Based Probabilistic Clustering (MKPC). The proposed MKPC is a modification to the KFCM with the addition of probability operator.

Fuzzy C-Means and KFCM: Fuzzy C-means (FCM) is one of the many clustering algorithms but gives more accurate clustering results with the inclusion of fuzzy concept when compared to other techniques. Let us assume that the input data (which in our case is image pixel values) be denoted by x and number of input data is denoted by n. Let the weighting co-efficient be represented by w which is a real number greater than one and the center of the cluster is represented by c. Let the degree of membership of x_i in the cluster j be represented by μ_{ij} and the number of clusters be denoted by $numc$. The minimization objective function (Fun_{min}) of Fuzzy C- Means (FCM) clustering is defined as:

$$Fun_{\min} = \sum_{i=1}^{n} \sum_{j=1}^{numc} \mu_{ij}^{w} \parallel x_i - c_j \parallel^2 \tag{2}$$

At first, random data points are considered as centroids and afterwards, membership values of the data points are calculated using the formula:

$$\mu_{ij} = 1 \Bigg/ \sum_{m=1}^{numc} \left(\frac{\parallel x_i - c_i \parallel}{\parallel x_i - c_m \parallel} \right)^{\frac{2}{w-1}} \tag{3}$$

Subsequently, the updated centroid values are calculated with the help of calculated membership values and are given by:

$$c_j = \sum_{i=1}^{n} \mu_{ij}^{w} x_i \Bigg/ \sum_{i=1}^{n} \mu_{ij}^{w} \tag{4}$$

Using the updated centroid values, membership values are again calculated and this loop procedure is continued until it satisfies the equation:

$$\max imum_{ij} \{ | \mu_{ij}^{w=1} - \mu_{ij}^{w} | < \lambda \} \tag{5}$$

where, λ has the value between 0 and 1 and eventually, FCM would converge to a local minimum or a saddle point of $Fun_{\min}$. The clustering accuracy is improved with the employment of kernel functions to FCM to have KFCM. In KFCM, input data (x) is mapped to a higher dimensional space (S) denoted by non-linear feature map function φ: $x \rightarrow \varphi(x) \in X$. The objective function of KFCM can be defined as:

$$Fun_{\min} = \sum_{i=1}^{n} \sum_{j=1}^{numc} \mu_{ij}^{w} \parallel \varphi(x_i) - \varphi(c_j) \parallel^2 \tag{6}$$

where,

$$\| \varphi(x_i) - \varphi(c_j) \|^2 = K(x_i, x_i) + K(c_j, c_j) - 2K(x_i, c_j) \quad (7)$$

where, $K(a, b) = \varphi(a)^T \varphi(b)$ is the inner product kernel function and for Gaussian kernel function:

$$K(a,b) = e^{-\frac{\|a-b\|^2}{\sigma^2}}, \text{ hence } K(a,a) = 1; G(x_i, x_i) = G(c_j, c_j) = 1 \quad (8)$$

The objective function can be redrafted as:

$$Fun_{\min} = 2\sum_{i=1}^{n} \sum_{j=1}^{numc} \mu_{ij}^{w} [1 - K(x_i, c_j)] \quad (9)$$

The updation equations for finding membership value μ_{ij} and centroids x_j is given as:

$$\mu_{ij} = \frac{\left(\frac{1}{(1-K(x_i,c_j))} \right)^{\frac{1}{(w-1)}}}{\sum_{m=1}^{n} \left(\frac{1}{(1-K(x_i,c_m))} \right)^{\frac{1}{(w-1)}}}, \quad c_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{w}.K(x_i,c_j)x_i}{\sum_{i=1}^{n} \mu_{ij}^{w}.K(x_i,c_j)} \quad (10)$$

Multiple Kernel based Probabilistic Clustering (MKPC): Employment of Multiple Kernel based Probabilistic Clustering results in having more accurate results. Multiple Kernel based Probabilistic Clustering (MKPC) is an extension to KFCM by incorporating probability operator. The operator is employed to find the respective cluster for the pixel concerned. Initially, two clusters are formed by choosing arbitrary centroids and assigning the pixels to respective cluster. In the next iteration, one more centroid is calculated (to form a new cluster) and then assign the pixels to the centroids so as to form three clusters in total. This process is continued to have N number of clusters at the N^{th} iteration.

In our proposed MKPC, in each iteration, the probability factor is calculated which decides to which cluster the considered pixel belongs to rather than membership function values. The inclusion of probability improves the cluster assignment. The probability factor is based on the membership values of the pixel to all clusters. Suppose the clusters be represented as $CL = \{cl_1, cl_2, \ldots, cl_N\}$. Then the membership value μ_{ij} of a pixel to the j^{th} cluster μ_{ij} is given in Eq. (10).

Similarly, the membership value of the pixel to all the N clusters is found out to have $\mu_{ij} = \{\mu_{i1}, \mu_{i2}, \mu_{i3}, \ldots, \mu_{iN}\}$. Subsequently, membership based probability is calculated from these value. The probability of pixel i belonging to the j^{th} cluster is calculated as:

$$P_{ij} = \frac{\mu_{ij}}{\sum_{k=1}^{N} \mu_{ik}} \quad (11)$$

The probability factor P_{ij} (for $0<j\leq N$) is calculated for all the clusters with respect to the pixel in consideration and the pixel is assigned to the cluster having maximum probability factor. The iteration process is continued to have the clusters formed based on the probability factor and these formed clusters form the segmented areas.

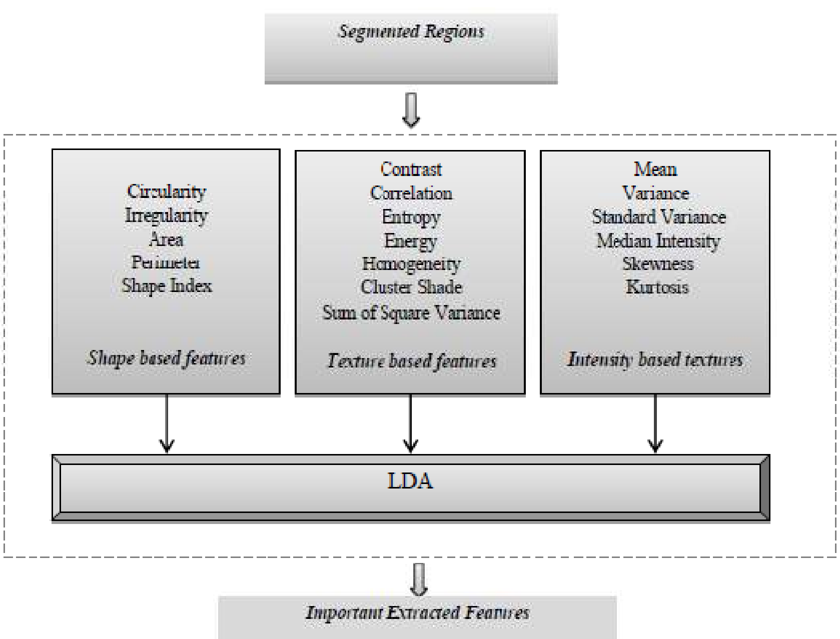

Fig. 2: Illustration of feature extraction module

Feature extraction module: Various features are extracted from the segmented regions and subsequently, important features are selected using Linear Discriminant Analysis (LDA). The features extracted are shape based features, texture based features and intensity based textures. All together of eighteen features are extracted and from these only the important ones are selected with the use of LDA. Illustration of the feature extraction module is given in Fig. 2.

Shape based features: Shape based features such as circularity, irregularity, area, perimeter, shape index are extracted from the segmented regions.

Circularity is the measure of how closely the shape of an object approaches that of a circle. Circularity is dominated by the shape's large-scale features rather than the sharpness of its edges and corners, or the surface roughness of a manufactured object. A smooth ellipse can have low circularity, if its eccentricity is large. Regular polygons increase their circularity with increasing numbers of sides, even though they are still sharp-edged. Circularity applies in two dimensions. Here, the circularity of the segmented region is extracted as the feature. Irregularity is the quality of not being regular in shape or form.

Area of the region in an image can be described as the number of pixels confined inside the segmented region including the boundary. Hence, the area extracted from a segment is the number of pixels inside the segmented region. Perimeter is a path that surrounds a two-dimensional shape in consideration. Perimeter of the segmented region can be described as the number of pixels in the boundary line of the region.

Shape Index (SI) is a statistic used to quantify the shape of any unit of area. It can be expressed in mathematical terms as $SI = 1.27*A*L$, where A is the area of shape and L is the length of the longest axis in the region. A value of 1.0 expresses maximum compaction, where the shape is circular. As the shape is elongated, the less compact is the slope and the lower the value of the index.

Texture based features: Contrast, Correlation, Entropy, Energy, Homogeneity, cluster shade and sum of square variance are the texture based features extracted from the segmented region.

Contrast can be defined as the difference in luminance and/or colour of the image. Contrast is determined by the difference in the colour and brightness of the object and other objects. Various definitions of contrast are used in different situations and it represents a ratio of the type of luminous difference to average luminance. Correlation refers to any of a broad class of statistical relationships involving dependence. Correlation of the image is defined the ratio between covariance and the standard deviation given by: $\rho_{X,Y} = q_{jk}/\sigma_X\sigma_Y$, where $\rho_{X,Y}$ is the correlation and σ is the standard deviation. Here, covariance is a measure of how much two variables change together. The covariance between two real-valued random variables Y and Z is given by:

$$Cov(Y,Z) = E[(Y-E(Y))\ (Z-E(Z))]$$

where,

E(Y) = The expected value of Y

E(Z) = The expected value of Z

Entropy (en) can be described as a measure of unpredictability or information content and is given by:

$$en = -\sum_i \Pr_i \log_2 \Pr_i \tag{12}$$

Here, $\Pr_i$ is the probability that the difference between 2 adjacent pixels is equal to i and Log_2 is the base 2 logarithms.

Energy (eg) is used to describe a measure of information present in the segmented region. It can be defined as the mean of squared intensity values of the pixels. Let the intensity of the pixels be represented by In_i, where $0 < i \leq N$ and N is the number of pixels. The energy can be given by:

$$eg = \frac{\sum_i In^2{}_i}{N} \tag{13}$$

Homogeneity is the state of being homogeneous. Pertaining to the sciences, it is a substance where all the constituents are of the same nature consisting of similar parts or of elements of the like nature. It can be given by the generalised formula:

$$Homogenity = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng} P(i,j)^2 \tag{14}$$

where, Ng is the number of grey levels and $P(i, j)$ is the pixel intensity. Cluster shade is given by the formula:

$$Cluster\,Shade = \sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}\{i + j - \mu_x - \mu_y\}^3 P(i,j) \tag{15}$$

The calculation of a sample variance or standard deviation is typically stated as a fraction. The numerator of this fraction involves a sum of squared deviations from the mean. The formula for this total sum of squares is given by $\Sigma(X-\mu)^2$.

Intensity based features: Mean, Variance, Standard Deviation, Median Intensity, Skewness and Kurtosis are intensity based features that are extracted from segmented regions. After features extraction, important features will be selected using Linear Discriminant Analysis (LDA) to classification purpose.

Mean (μ) is simply the average of the objects in consideration. Mean of the region is found out by adding all the pixel values of the region divided by the number of pixels in the region. Suppose there are Np number of pixels in the ith region each having a pixel value P_i, then mean of the ith region is given by:

$$\mu_i = \sum P_i \big/ Np$$

Variance (σ^2) measures how far a set of pixels of the image are spread out. A variance of zero indicates that all the values are identical:

$$\sigma^2 = E(X^2) - (E(X))^2 \tag{16}$$

where, $E(.)$ is the expectation operation and $E(X) = \mu$ (mean). Standard deviation (σ) is the square root of the variance. It also measures the amount of variation from the average. A low standard deviation indicates that the data points tend to be very close to the mean.

Median is the numerical value separating the higher half of pixel values from the lower half. The median can be found by arranging all the observations from lowest value to highest value in order and then picking the middle one. Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive or negative, or even undefined. Kurtosis is any measure of the peak of the probability distribution of a real-valued random variable. One common measure of kurtosis is based on a scaled version of the fourth moment of the given region pixels.

Once all the eighteen features are extracted, only the important one needs to be taken for further processing. This is accomplished by the use of Linear Discriminant Analysis (LDA). Linear Discriminant Analysis is commonly used technique for dimensionality reduction.

Classification module: After feature extraction, the image is classified into tumorous or non-tumorous with the use of deep learning based neural network. For classification, two important processes are, training phase and testing phase. In training phase, deep learning classifier is trained with the features of training data and in testing, the features from the segmented image are fed into the trained Deep learning classifier to detect whether the region has brain tumour or not. Artificial Neural Networks provide a powerful tool to analyze, model and predict. The benefit is that neural networks are data driven self-adaptive methods. In general, the neural network consists of three layers named as input layer, hidden layer and the output layer. The neural network works making use of two phases, one is the training phase and the other is the testing phase.

Commonly used neural network uses back propagation algorithm. But it is not adequate for training neural networks with many hidden layers on large amounts of data. Over the last few years, advances in both machine learning algorithms and computer hardware have led to more efficient methods like Deep Neural Networks that contain many layers of nonlinear hidden units and a very large output layer. Deep neural networks have deep architectures which have the capacity to learn more complex models than shallow ones. The extra layers give it added levels of abstraction, thus increasing its modelling capability.

The DNN is a feed-forward, artificial neural network that has more than one layer of hidden units between its inputs and its outputs. Each hidden unit i, typically uses the logistic function to map its total input from the layer below (h_i) to the scalar state (z_i), that it sends to the layer above:

$$z_i = \text{log}istic(h_i) = \frac{1}{1+e^{-h_i}},\ h_i = b_i + \sum_k z_k \varpi_{ki} \tag{17}$$

where, b_i is the bias of unit, i, k is an index over units in the layer below and ϖ_{ki} is the weight on a connection to unit i from unit k in the layer below. For multiclass classification, output i unit converts its total input, h_i, into a class probability, P_i by:

$$P_i = \frac{e^{h_i}}{\sum_k e^k} \tag{18}$$

DNN can be discriminatively trained with the standard back propagation algorithm. The weight updates can be done via stochastic gradient descent using the following equation:

$$\Delta\varpi_{ki}(t+1) = \Delta\varpi_{ki}(t) + \Lambda\frac{\partial\gamma}{\partial\varpi_{ki}} \tag{19}$$

where, Λ is the learning rate and γ is the cost function. Here, Initially the DNN is trained using various images and subsequently test image is fed to the trained image to classify the image as timorous or not.

RESULTS AND DISCUSSION

Real Time MRI images are collected from Metro Scans and Laboratory, a unit of Trivancore Healthcare

Pvt Ltd, Trivandrum, India. The proposed brain tumour detection technique is analysed with the help of experimental results.

Experimental set up and evaluation metrics: The proposed technique is implemented using MATLAB on a system having the configuration of 6 GB RAM and 2.8 GHz Intel i-7 processor. The evaluation metrics used to evaluate the proposed technique consists of sensitivity, specificity and accuracy. In order to find these metrics, we first compute some of the terms of True Positive (TP), True Negative (TN), False Negative (FN) and False Positive (FP) based on the definitions given in Table 1.

The evaluation metrics of sensitivity, specificity and accuracy can be expressed in terms of TP, FP, FN and TN. Sensitivity is the proportion of true positives that are correctly identified by a diagnostic test. It shows how good the test is at detecting a disease:

$$\text{Sensitivity} = TP/(TP + FN) \quad (20)$$

Specificity is the proportion of the true negatives correctly identified by a diagnostic test. It suggests how good the test is at identifying normal (negative) condition:

$$Specificity = TN/(TN + FP) \quad (21)$$

Accuracy is the proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition:

Table 1: Table defining the terms TP, FP, FN, TN

Experimental outcome	Condition as determined by the standard of truth	Definition
Positive	Positive	True Positive (TP)
Positive	Negative	False Positive (FP)
Negative	Positive	False Negative (FN)
Negative	Negative	True Negative (TN)

$$Accuracy = (TN + TP)/(TN + TP + FN + FP) \quad (22)$$

Experimental results: The experimental results achieved for the proposed technique are given. The MRI image data set used to evaluate the proposed image technique is taken from the publicly available sources. Figure 3 shows some of the input MRI images and segmented output obtained in each case.

Performance analysis: The performance of the proposed technique is analysed with the use of evaluation metrics of sensitivity, specificity and accuracy. The proposed technique results are also compared with the existing technique which uses Feed-Forward Back propagation Network (FFBN).

Table 2 shows the performance of the proposed technique using various evaluation metrics. The results are taken by varying the number of hidden neurons. It is observed that the improved results are obtained by increasing the hidden layer neurons. The Highest values of sensitivity, specificity and accuracy are obtained as 1, 0.85 and 0.916, respectively. High evaluation metric

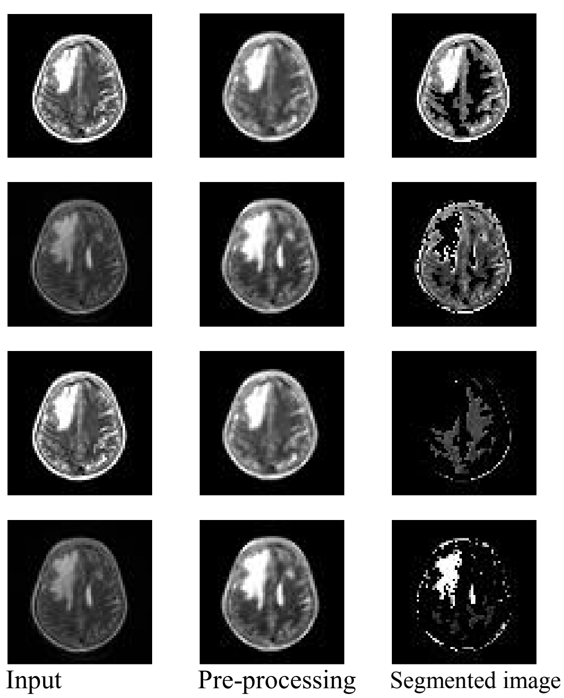

Fig. 3: Input MRI images and segmented output

Table 2: Performance of the proposed technique

Hidden neurons	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
25	3	6	1	2	0.6	0.857142857	0.75
50	4	6	1	1	0.8	0.857142857	0.833333333
75	5	5	2	0	1	0.714285714	0.833333333
100	5	5	2	0	1	0.714285714	0.833333333
150	5	6	1	0	1	0.857142857	0.916666667

Table 3: Evaluation metric values for varying number of clusters

Cluster Size	TP	FP	TN	FN	Sensitivity	Specificity	Accuracy
2	4	1	1	4	0.50	0.50	0.50
3	4	1	3	2	0.667	0.75	0.70
4	5	0	5	0	1	1	1
5	5	0	5	0	1	1	1

Table 4: Evaluation metric values under noise conditions

	Sensitivity	Specificity	Accuracy
Gaussian noise	0.67	0.72	0.695
Salt and pepper noise	0.62	0.78	0.692
Speckle noise	0.72	0.72	0.722

values indicate the effectiveness of the proposed technique.

Robustness analysis: Robustness analysis is carried out by varying the parameters like number of clusters under various noise conditions. From Table 3 as cluster size is increased the sensitivity, specificity and accuracy are increased. The evaluation metrics value become 1 when the ciuster size is 4 and 5. Table 4 gives the performance under various noise conditions such as gaussion noise, salt and pepper noise and speckle noise. Here it is observed that the sensitivity is higher for speckle noise and lower for salt and pepper noise. The specificity is higher for salt and pepper noise while it remains the same and lower for Gaussian noise and speckle noise. Further it is observed the accuracy is higher for speckle noise followed by Gaussian noise and then salt and pepper noise

The proposed technique is also tested under various noise conditions of salt and pepper, Gaussian and speckle noise. We can infer that the proposed technique has performed well under noise conditions by attaining good evaluate on metric values.

Comparative analysis:

Comparing with existing neural network algorithm: The evaluation metric results obtained by the proposed technique using Deep Learning (DL) and also by the existing technique (FFBN) are given here. In the past Feed Forward Back Probagation Network (FFBN) has been extensively used to evaluate the metrics. However a new algorithm using Deep Learning (DL) is used to evaluate the metrics by varying the hidden neurons from 25 to 150 in step of 25. Figure 4 to 6 give the comparative graphs of sensitivity, specificity and accuracy respectively. The graphs are taken by varying hidden neuron in the neural networks.

Figure 4 shows the sensitivity metric obtained by DL and FFBN methods. Further it is noticed that the results obtained using DL method is much higher than those obtained using FFBN method. At all levels of hidden neurons the sensitivity increases as a hidden neuron is increased except at the point the hidden neuron is 125. It is also observed that highest sensitivity values of 1 is obtained by the proposed method whereas 0.8 is obtained using the existing method.

Figure 5 gives the specificity when the hidden neuron is varied using the existing FFBN algorithm and the proposed DL algorithm. Here also the hidden neuron is varied from 25 to 150 in step of 25. The specificity value is obtained as 0.85 by both the existing method and the proposed method.

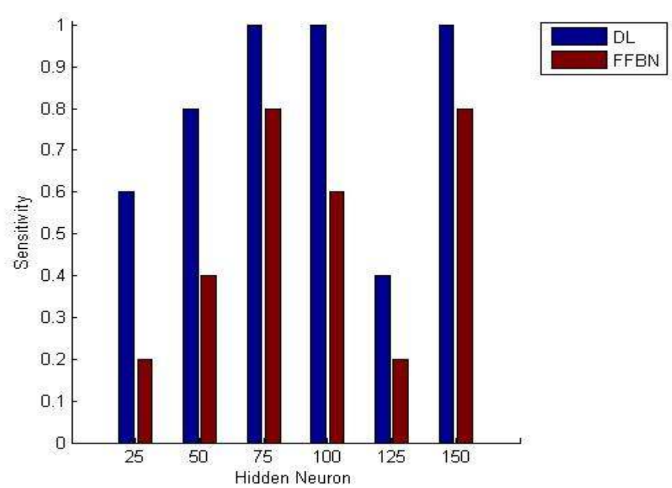

Fig. 4: Comparative chart showing sensitivity for varying hidden neurons

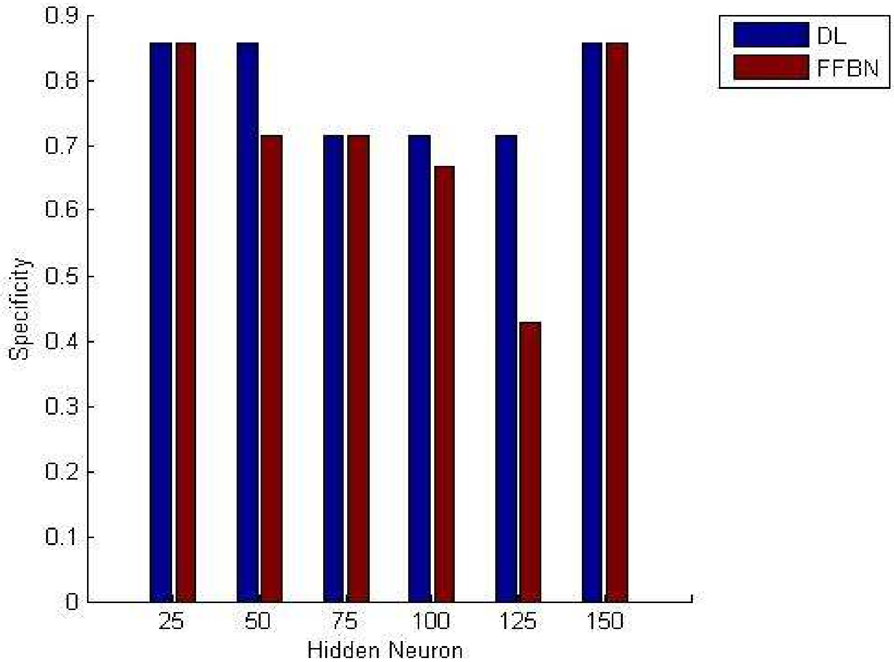

Fig. 5: Comparative chart showing specificity for varying hidden neurons

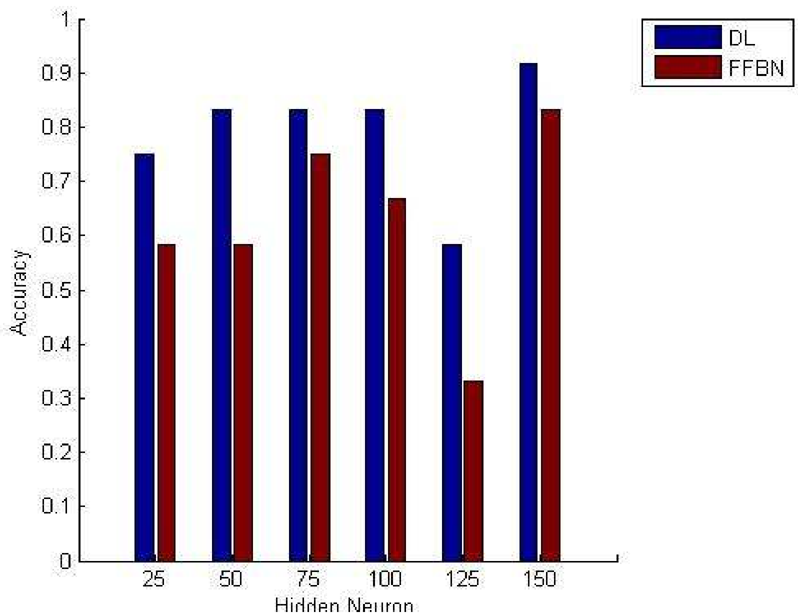

Fig. 6: Comparative chart showing accuracy for varying hidden neurons

Table 5: Evaluation metrics using different segmentation algorithms

	Region growing with NN	Modified region growing with NN	Proposed technique
Sensitivity	0.54	0.72	0.88
Specificity	0.81	0.82	0.80
Accuracy	0.66	0.74	0.83

Figure 6 gives the accuracy when the hidden neuron is varied using the existing FFBN algorithm and the proposed DL algorithm. Here also the hidden neuron is varied from 25 to 150 in step of 25. The highest accuracy value is obtained as 0.94 by the proposed method whereas it is 0.84 using existing method. From the above it is very much evident that by using the proposed method better evaluation metric values are obtained compared to the existing method. Thus the efficiency of the proposed technique is established beyond doubt.

Comparing with other segmentation algorithms: Comparison is made with the existing segmentation

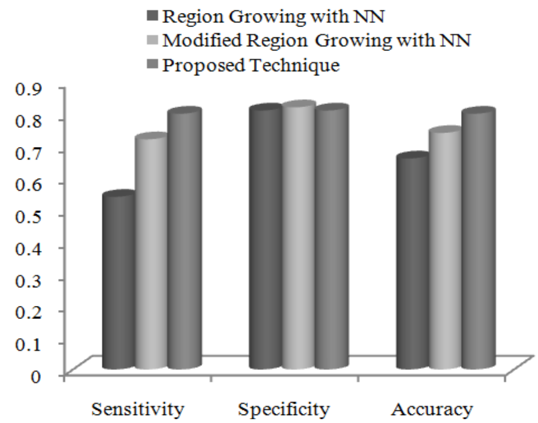

Fig. 7: Comparative analysis chart using segmentation algorithm

techniques of Region growing with Neural Network (NN) and Modified Region growing with Neural Network. Table 5 gives the comparative evaluation metrics and the corresponding figure is given in Fig. 7.

From Fig. 7 and Table 5 it is evident that the highest sensitivity value of 0.88 is obtained using the proposed algorithm whereas using modified region growing with NN and region growing with NN the sensitivity is 0.72 and 0.54, respectively. Similarly highest accuracy of 0.83 is obtained using proposed technique whereas it is 0.74 for modified region growing with NN and 0.66 for region growing with NN. However the specificity is slightly lower for the proposed technique which is 0.8 when compared to 0.82 obtained by modified region growing with NN and 0.81 using region growing with NN.

It is therefore inferred that the proposed segmentation algorithm in this study has got better sensitivity and accuracy with other existing algorithms. However the specificity obtained by the proposed algorithm is more or less same as obtained by using other existing algorithms.

CONCLUSION

In this study, tumor classification using multiple kernel-based probabilistic clustering and deep learning classifier is proposed. The proposed technique consists of three modules, namely segmentation module, feature extraction module and classification module. The image is segmented using Multiple Kernel based Probabilistic Clustering (MKPC) and classified using deep learning classifier. The proposed technique is evaluated using sensitivity, specificity and accuracy. The proposed technique results are also compared with existing technique which uses Feed-Forward Back Propagation Network (FFBN). The proposed technique achieved an average sensitivity, specificity and accuracy of 0.88, 0.80 and 0.83, respectively and highest values of 1, 0.85 and 0.94, which is much higher than the existing method. Similarly, Sensitivity, Specificity and accuracy of 0.88, 0.80 and 0.83 are achieved using a proposed segmentation algorithm which is for better than those obtained by existing segmentation algorithms. Thus the proposed method's supremacy is established.

ACKNOWLEDGMENT

The study done by V.P.Gladis Pushpa Rathi and S.Palani, is supported by All India Council for Technical Education, New Delhi, under Research Promotion Scheme at Sudharsan Engineering College, Sathiyamangalam, Pudukkottai, Tamil Nadu, India.

REFERENCES

Acır, N., O. Ozdamar and C. Guzelis, 2006. Automatic classification of auditory brainstem responses using SVM-based feature selection algorithm for threshold detection. Eng. Appl. Artif. Intel., 19: 209-218.

Ali, S.M., L.K. Abood and R.S. Abdoon, 2013. Brain tumor extraction in MRI images using clustering and morphological operations techniques. Int. J. Geogr. Inform. Syst. Appl. Remote Sens., 4(1).

Bankman, I.H., 2009. Handbook of Medical Image processing and Analysis. Academic Press in Biomedical Engineering, Burlington, MA 01803, USA.

Georgiardis, P., D. Cavouras, I. Kalatzis, A. Daskalakis, G.C. Kagadis, M. Malamas, G. Nikifordis and E. Solomou, 2008. Improving brain tumor characterization on MRI by probabilistic neural networks on non-linear transformation of textural features. Comput. Meth. Prog. Bio., 89: 24-32.

Gholami, B., I. Norton, L.S. Eberlin and N.Y.R. Agar, 2013. A statistical modeling approach for tumor-type identification in surgical neuropathology using tissue mass spectrometry imaging. IEEE J. Biomed. Health Inform., 17(3): 734-744.

Karabatak, M. and M.C. Ince, 2009. An expert system for detection of breast cancer based on association rules and neural network. Expert Syst. Appl., 36: 3465-3469.

Kharat, K.D., P.P. Kulkarni and M.B. Nagori, 2012. Brain tumor classification using neural network based methods. Int. J. Comput. Sci. Inform., 1(4).

Sapra, P., R. Singh and S. Khurana, 2013. Brain TUMOR detection using neural network. Int. J. Sci. Modern Eng. (IJISME), 1(9).

Sharma, M., R.B. Dubey, Sujata and S.K. Gupta, 2012. Feature extraction of mammogram. Int. J. Adv. Comput. Res., 2(5).

Sridhar, D. and M. Krishna, 2013. Brain tumor classification using discrete cosine transform and probabilistic neural network. Proceeding of International Conference on Signal Processing, Image Processing and Pattern Recognition (ICSIPR, 2013), pp: 92-96.

Sun, Z., C. Zheng, Q. Gao, J. Zhang and D. Zhan, 2012. Tumor classification using eigengene-based classifier committee learning algorithm. IEEE Signal Process. Lett., 19(8).

Wang, S.L., Y.H. Zhu, W. Jia and D.S. Huang, 2012. Robust classification method of tumor subtype by using correlation filters. IEEE-ACM T. Comput. Bi., 9(2).