An Efficient Brain Tumor MRI Segmentation and Classification Using GLCM Texture Features and Feed Forward Neural Networks

P. Kumar and B. Vijayakumar

Centre for Information Technology and Engineering, M.S. University, Tirunelveli, Tamilnadu, India
PSN Engineering College, Tirunelveli, Tamilnadu, India

Abstract: Brain tumor detection and segmentation in magnetic Resonance Images (MRI) is vital in medical diagnosis as it provides information related to anatomical structures as well as potential abnormal tissues obligatory to treatment planning and patient record. But regrettably it is a very decisive task. We have proposed a new method. It provides a competent and fast approach for diagnosis of the brain tumor. The Proposed system consists of manifold phases. First phase consists of Preprocessing and segmentation, the second phased consists of first order and second order GLCM (Gray level Co-occurrence Matrix) based features extraction from segmented brain MR images. Third phase classify brain images into tumor and non-tumors using Feed Forwarded Artificial neural network based classifier. After classification tumor region is extracted from those images which are classified as malignant using two stage segmentation process. Experiments have exposed that the method was more robust to initialization, faster and accurate.

Key words: Segmentation · Classification · Texture features · Magnetic Resonance imaging (MRI) · Feed forward Neural Network

INTRODUCTION

Automatic brain tumor segmentation from MR images is a intricate task that involves various discipline face pathology, MRI physics, radiologist’s perception and image analysis based on intensity and shape. Brain tumors might be of any size, may have a variety of shapes, may possibly appear at any location and may appear in different image intensities. Some tumors also warp other structures and appear together with edema that changes intensity properties of the nearby region. For many human experts, manual segmentation is a difficult and time consuming task, which makes an automated brain tumor segmentation method enviable. There are several possible applications of an automated technique, it preserve be used for surgical planning, treatment planning and vascular analysis. It has been shown that blood vessels in the brain reveal assured characteristics within pathological regions. An objective and reproducible segmentation procedure joined with vascular analysis would allow us to study the relation between pathologies and blood vessels and may function as a new diagnostic measure [1, 2]. The Proposed system flow diagram is shown in Fig. 1. Dr. H. B. Kekre et al. [3] have proposed a vector quantization segmentation technique to identify cancerous mass from MRI images. In order to improve the radiologist’s diagnostic performance, computer-aided diagnosis (CAD) scheme has been introduced to enhance the recognition of primary signatures of this disease: masses and micro calcifications. As well, to engage in the class distinguish ability as well as feature space sparseness and solution space intricacy problems in multivariate image segmentation, a Markov random field (MRF) based multivariate segmentation algorithm called “multivariate iterative region growing using semantics” (MIRGS) has been proposed by Jue Wu et al. [4]. Hybrid intelligent system plays a vital role in survival prediction of a breast cancer patient and it is highly significant in decision making for treatments and medications. The primary objective of a hybrid intelligent system is to take the advantages of its constituent models and at the same time lessen their limitations.

This chapter is an attempt to highlight the reliability of infrared breast thermographs and hybrid intelligent system in breast cancer detection and diagnosis [16]. The methods gives the better result are histogram
Fig. 1: Proposed system flow diagram

equalization for enhancement, probabilistic neural network for feature extraction [18]. A kernel based SVM classifier is exploited to amalgamate the hesitant images into mundane and exotic. The espoused tentative are gauged using the metric similarity index (SI), overlap fraction (OF) and extra fraction (EF). For comparison, the enactment of the propositioned technique has colossally exalted the tumor divulging in scrupulousness with other neural network corroborated classifier [19].

Histogram equalization is used in enhancement stage, gray level co-occurrence matrix and wavelets are used as features and these extracted features are trained and classified using Support Vector Machine (SVM) classifier. The tumor region is detected using morphological operations. The performance of the proposed algorithm is analyzed in terms of sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV). The proposed system achieved 0.95% of sensitivity rate, 0.96% of specificity rate, 0.94% of accuracy rate, 0.78% of PPV and 0.87% of NPV, respectively [20]. Experiments were carried out on a dataset consisting of T1-weighted post-contrast and T2-weighted MR images of 550 patients. The developed CAD system was tested using the leave-one-out method. The experimental results showed that the proposed segmentation technique achieves good agreement with the gold standard and the ensemble classifier is highly effective in the diagnosis of brain tumor with an accuracy of 99.09% (sensitivity 100% and specificity 98.21%).

Thus, the proposed system can assist radiologists in an accurate diagnosis of brain tumors [24]. The input image is preprocessed using wiener filter and Contrast Limited Adaptive Histogram Equalization (CLAHE). The image is then quantized and aggregated to get a reduced image data. The reduced image is then segmented into four regions such as gray matter, white matter, cerebrospinal fluid and high intensity tumor cluster using Fuzzy C Means (FCM) algorithm. The tumor region is then extracted using the intensity metric. A contour is evolved over the identified tumor region using Active Contour model (ACM) to extract exact tumor segment [25].

In MIRGS, the impact of intra-class variation and computational cost has been minimized by means of the MRF spatial context model integrated with adaptive edge penalty and applied to regions. To restrain the initialization sensitivity, a region-level means (RKM) based initialization technique has been utilized, which always provides exact initial conditions at low computational cost. Experiments have demonstrated the pre-eminence of RKM relative to two frequently used initialization techniques. The Block diagram of the proposed technique is shown schematically in Fig 1.

For image segmentation, Chunming Li et al. [5] have proposed a region-based technique, which has the potential to deal with intensity in homogeneities in the segmentation. Initially, based on the model of images with intensity in homogeneities, the local clustering criterion function for the image intensities in a neighborhood of
each point has been defined. Subsequently, this local clustering criterion function has been included with respect to the neighborhood centre to provide a global criterion of image segmentation.

In a level set formulation, this criterion defines an energy based on the level set functions that depict a partition of the image domain and a bias field that accounts for the intensity in homogeneity of the image. Thus, by reducing this energy, the technique was able to segment the image and estimate the bias field simultaneously and the estimated bias field has been employed for intensity in homogeneity correction (or bias correction). The method has been validated on artificial and real images of different modalities and obtained enviable performance in the presence of intensity in homogeneities. The main contributions of our proposed technique are: Preprocessing the input MR images using anisotropic filter. Tumor regions are identified using region props algorithm and segmented the region using morphological operation. Texture Features are extracted from the segmented MR Images using first order statistical features and second order GLCM. Selected features are input to Feed Forwarded Artificial Neural Network for training, finally classify the tumor and non-tumor images.

Used evaluation matrices parameters of sensitivity, specificity and accuracy to evaluate the performance of the proposed technique of tumor detection. The rest of the paper is organized as follows: The proposed segmentation technique is presented in Section 2. The First order and second order Gray level Co-occurrence matrix based Feature extractions are presented in Section 3. The detailed experimental results and discussions are given in Section 4. The conclusions are summed up in Section 5.

**Proposed Segmentation Method:** Segmentation is achieved in three stages: In first stage removal of the background using intensity histograms, generation of an initial mask that determines the intracranial boundary with a nonlinear anisotropic diffusion filter and final segmentation with an active contour model [6]. The fact that the brain has a relatively higher intensity than other tissue in MR images constitutes the second type of prior information. Using the anisotropic diffusion filter on T2 (or PD) images, the majority of the tissue other than the brain can be darkened, allowing a simple threshold to be used subsequently for segmentation.

**Background Removal:** Considering the fact that MR scanners typically generate normally distributed white noise, the best threshold for separating background noise is determined with the technique of Brummer et al. [7]. In reconstructed MR data, background noise has a Rayleigh distribution given by:

\[ p_{\text{noise}}(f) = \frac{f}{\sigma^2} \exp \left( -\frac{f^2}{2\sigma^2} \right) \]

where \( f \) is the intensity and \( \sigma \) is the standard deviation of the white noise. This distribution is observed in the lower intensities of the uncorrected histogram of MR volumes as illustrated in a bimodal distribution \( g(f) \) is obtained if the best fit Rayleigh curve, \( r(f) \), is subtracted from the volume histogram, \( h(f) \):

\[ g(f) = h(f) - r(f) = \exp \left( -\frac{\left( \nabla I(x,t) \right)^2}{2\kappa} \right) \]

We can obtain a minimum error threshold, \( \tau \), by minimizing an error term, \( \varepsilon \), given by:

\[ \varepsilon = \sum_{f=0}^{\tau} g(f) + \sum_{f=\tau}^{\infty} r(f) \]

The process that generates the initial brain mask has three steps. First it smooth’s the brain image using 2D nonlinear anisotropic diffusion and attenuates narrow non-brain regions. Then, it sets an automated threshold to the diffused MR volume and produces a binary mask. Third, it removes misclassified non-brain regions such as the eyes from the binary mask based on morphology and spatial information obtained from the head mask.

**Non-Linear Anisotropic Diffusion:** Nonlinear anisotropic diffusion filters introduced by Perona and Malik [8] are tunable iterative filters that can be used to enhance MR images. Nonlinear anisotropic diffusion filters can be used also to enhance and detect object edges.

The anisotropic diffusion filter is a diffusion process that facilitates intraregional smoothing and inhibits interregional smoothing:

\[ \frac{\partial}{\partial t} I(x,t) = \nabla \cdot \left( c(x,t) \nabla I(x,t) \right) \]
Consider \( I(x, t) \) to be the MR image where \( x \) represents the image coordinates (i.e. \( x,y \)), \( t \) is the iteration step and \( c(x, t) \), the diffusion function is a monotonically decreasing function of the image gradient magnitude. Edges can be selectively smoothed or enhanced according to the diffusion function. An effective diffusion function is [8]:

\[
E(v) = \alpha E_{cont}(v) + \beta E_{bal}(v) + \gamma E_{int}(v) + \kappa E_{grad}(v)
\]

where \( K \) is the diffusion or flow constant that dictates the behavior of the filter. Good choices of parameters that produce an appropriately blurred image for thresholding are \( K = 128 \) with 25 iterations and a time step value of just under 0.2. Filtering can be fairly sensitive to these three parameters [22], however, for all the PD, T2 and T1 - weighted data sets displayed axially or coronally, the above parameter settings provide good initial brain segmentation.

**Final Brain Mask:** The final boundary between the brain and the intracranial cavity is obtained with an active contour model algorithm that uses the initial brain mask as its initial condition. The active contour model, extended from the "Snakes" algorithm introduced by Kass et al. [9], gradually deforms the contour of the initial brain mask to lock onto the edge of the brain. The active contour, described in Chapter Segmentation 4, is defined as an ordered collection of \( n \) points in the image plane such that:

\[
V = \{v_1, \ldots, v_n\} \\
v_i = (x_i, y_i), \quad i = 1, \ldots, n
\]

An energy minimization criterion iteratively brings the points of the contour closer to the intracranial boundary. For each point, \( v_i \), an energy matrix, \( E(v_i) \), is computed:

\[
E(v_i) = \alpha E_{cont}(v_i) + \beta E_{bal}(v_i) + \gamma E_{int}(v_i) + \kappa E_{grad}(v_i)
\]

where \( E_{cont}(v_i) \) is a "continuity" energy function that forces the contour to take a smooth shape, \( E_{bal}(v_i) \) is an adaptive "balloon" force used to push the contour outward until a strong gradient is encountered [9], \( E_{int}(v_i) \) is an "intensity" energy function, computed from the PD-weighted MRI volume, that tends to move the contour towards low intensity regions and \( E_{grad}(v_i) \) is a "gradient" energy function, computed from the diffused MRI volume, that draws the contour towards regions where the gradient is high. Relative weights of the energy terms are provided by the scalar constants \( \alpha, \beta, \gamma \) and \( \kappa \). This procedure moves each \( v_i \) to the point of minimum energy in its neighborhood. The active contour model algorithm finds the intracranial boundary in all image slices using the same relative energy weightings for the combination of energy functions described above.

**Feature Extraction:** The transformation of an image into its set of features is known as feature extraction. Useful features of the image are extracted from the image for classification purpose. It is a challenging task to extract good feature set for classification. There are many techniques for feature extraction e.g. texture Features [10,11], Gabor features [12], feature based on wavelet transform [13], principal component analysis, minimum noise fraction transform, discriminate analysis, decision boundary feature extraction, non-parametric weighted feature extraction and spectral mixture analysis [14]. We are using texture feature for our proposed system.

**First-Order Histogram Based Features:** Histogram of the image gives summary of the statistical information about the image. So first order statistical information of the image can be obtained using histogram of the image. Probability density of occurrence of the intensity levels can be obtained by dividing the value of intensity level histogram with total number of pixels in the image.

\[
P(i) = \frac{h(i)}{NM}, \quad i = 0, 1, G-1
\]

**Inverse Difference Moment (IDM):**

\[
IDM = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \frac{p(i,j)}{1 + (i - j)^2}, \quad i \neq j
\]

where \( N \) is number of the resolution cells in the horizontal spatial domain and \( M \) is the number of resolution cells in the vertical spatial domain. \( G \) is the total gray level of an image. For quantitatively describing the first order statistical features of the image, useful features of the image can obtained from the histogram. Mean is the average value of the intensity of the image. Variance tells the intensity variation around the mean. Skewness is the measure which tells the symmetries of the histogram around the mean. Kurtosis is the flatness of the histogram. Uniformity of the histogram is represented by the entropy.

**Co-occurrence Matrix Based Features:** Histogram based features are local in nature. These features do not consider spatial information into consideration. So for this purpose gray-level spatial concurrence matrix \( h(i,j) \) based features are defined which are known as second order histogram based features. These features are based on
Angular Second Moment (ASM):

$$ASM = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} p(i,j)^2$$

where ASM measures the image homogeneity and $p(i,j)$ denotes probability of occurrence of pixel pair $(i,j)$

Entropy:

$$Entropy = -\sum_{i=0}^{k-1} \sum_{j=0}^{k-1} p(i,j) \log(p(i,j))$$

where entropy is a measure of non-uniformity in the image based on the probability of co-occurrence values.

Inverse Difference Moment: Where IDM is a measure of local homogeneity

Difference Moment: It is a measure of contrast or gray-level variations between the reference pixel and its neighbor.

Contrast:

$$Contrast = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} (i - j)^2 p(i,j)$$

Maximum Probability: It is simply the largest entry in the matrix and corresponds to the strongest response.

$$Maxprobability = \max p(i,j)$$

Final Classification: After feature extraction process, in order to detect the presence of the tumor in the input MRI image, we perform the final classification step. Here we use the Feed Forward neural network classifier to classify the image into tumors or not. A three layer Neural network was created with 500 nodes in the first (input) layer, 1 to 50 nodes in the hidden layer and 1 node as the output layer. We varied the number of nodes in the hidden layer in a simulation in order to determine the optimal number of hidden nodes. This was to avoid over fitting or under fitting the data. Due to hardware limitations, ten nodes in the hidden layer were selected to run the final simulation.

Figure 2 shows the design of the Feed Forward Neural networks used in this research. The 500 data points extracted from each subject were then used as inputs of the neural networks. The output node resulted in either a 0 or 1, for control or patient data respectively. Since the nodes in the input layer could take in values from a large range, a transfer function was used to transform data first, before sending it to the hidden layer and then was transformed with another transfer function before sending it to the output layer. In this case, a tan sigmoid transfer function was used between the input and hidden layer and a log sigmoid function was used between the hidden layer and the output layer.
The weights in the hidden node needed to be set using “training” data. Therefore, subjects were divided into training and testing datasets. Out of the 69 subjects, 2 random patients and 2 random controls were selected as “test data”, while the rest of the dataset was used for training. Training data was used to feed into the neural networks as inputs and then knowing the output, the weights of the hidden nodes were calculated using back propagation algorithm. 120 trials were performed on the same Neural Network, selecting 65 subjects randomly every time for retraining and 4 remaining subjects for testing to find accuracy of Neural network prediction.

In Back propagation algorithms modifying the weights and biases of the network in order to minimize a cost function. The cost function always includes an error term a measure of how close the network’s predictions are to the class labels for the examples in the training set. Additionally, it may include a complexity term that reacts a prior distribution over the values that the parameters can take. The activation function considered for each node in the network is the binary sigmoid function defined (with \( s = 1 \)) as output = \( \frac{1}{(1+e^{-x})} \), where \( x \) is the sum of the weighted inputs to that particular node. This is a common function used in many BPN. This function limits the output of all nodes in the network to be between 0 and 1. Note all neural networks are basically trained until the error for each training iteration stopped decreasing. The inputs \( m \) and outputs of the \( j \) hidden layer neurons can be calculated as follows

\[
Step -1 \quad m^h_j = \sum_{i=1}^{N+1} W_{jisi}
\]

\[
Step -2 \quad y_j = f(m^h_j)
\]

Calculate the \( m \) inputs and outputs of the \( k \) output layer neurons are

\[
Step -3 \quad Z_k = f(m^o_k)
\]

\[
Step -4 \quad m^o_k = \sum_{j=1}^{J+1} V_{kjyj}
\]

Updates the weights in the output layer (? k,j pairs)

\[
Step -5 \quad v_{kj} \leftarrow v_{kj} + c \lambda (d_k - Z_k)Z_k (1 - Z_k)y_j
\]

Updates the weights in the hidden layer (? i, j pairs)

\[
Step -6 \quad w_{ji} \leftarrow w_{ji} + c \lambda y_i (1 - y_i) x_i (d_k - Z_k)Z_k (1 - Z_k) v_{kj}
\]

Update the error term

\[
Step -7 \quad E \leftarrow E + \sum_{k=1}^{k} (d_k - Z_k)^2
\]

And repeat from Step 1 until all input patterns have been presented. If \( E \) is below some predefined tolerance level, then stop. Otherwise, reset \( E = 0 \) and repeat from Step 1 for another epoch.

Fig. 3: Tumour extracted images
Experimental Results: The proposed method has been implemented using the Mat lab environment. The proposed system has been tested on the data set available at web [15]. It has also been tested on dataset of real brain MR images consisting of normal and timorous brain images. Fig- 3 shows some of the tumor extracted images. Table 1 shows the results of all experimentations. We first perform classification of the dataset. Once the classification is done, the malignant brain images are further segmented for extraction of tumor region from these brain MR images. Performance of each classifier is measure in terms of sensitivity, Specificity and Accuracy [16]. Sensitivity is a measure which determines the probability of the results that are true positive such that person has the tumor. Specificity is a measure which determines the probability of the results that are true negative such that person does not have the tumor. Accuracy is a measure which determines the probability that how much results are accurately classified.

Sensitivity = TP/(TP + FN)
Specificity = TN/(TN + FP)
Accuracy = (TN + TP)/(TN + TP + FN + FP)

where, TP stands for True Positive, TN stands for True Negative, FN stands for False Negative and FP stands for False Positive. Figure 4 shows the competitive charts of different methods.

CONCLUSION

We have developed a hybrid segmentation method that uses both First order and second texture features of the images to segment the tumor. These extracted features are used for classification. In classification phase proposed system used neural network base classifier (FFNN) for classifying brain images as benign and malignant. Once the images are determined as malignant these are further processed for tumor extraction from them. All experiments show that the proposed system gives exceptionally good results as compared to the recently proposed techniques. We also achieved very accurate results of segmentation which effectively extract the tumor region from brain MR images.

REFERENCES