Implementation of Vertical Search technique for diagnosis of Abnormal Tumor Region from Brain CT Images

M. Ragavi¹, T. L. Nija Shining Gold²

¹PG Student, Department of CSE, Ponjesly College of Engineering, Nagercoil, Tamilnadu, India
²Assistant Professor, Department of CSE, Ponjesly College of Engineering, Nagercoil, Tamilnadu, India

Abstract— Automatic segmentation of brain computed tomography images is an important task. Automating this process is challenging due to high diversity in appearance of tumor tissue among different patients and in many cases, similarity between tumor and normal tissue. After the classification and segmentation of a brain computed tomography images vertical search is implemented on the images. Diagnosis of tumor region from the brain computed tomography image can be done by using vertical search implementation. A Dominant gray level run length texture feature set is derived from the ROI of the image is to be selected. We construct the SVM based classifier and evaluate the performance by comparing the classification results. To improve the computing efficiency it select the most suitable feature extraction method that can be used for classification and segmentation of brain tumor in computed tomography images efficiently and accurately. Vertical search method entirely searches the segmented region space.

Keywords—Support Vector Machine (SVM); Genetic Algorithm (GA); Region of Interest (ROI); CT (Computed Tomography);

I. INTRODUCTION

In recent years, medical CT Images have been applied in clinical diagnosis widely. That can assist physicians to detect and locate Pathological changes with more accuracy. Computed tomography (CT) uses special X-Ray equipment to obtain many images from different angles and join them together to produce multiple cross sectional images of the head. In recent years, medical CT Images have been applied in clinical diagnosis widely. That can assist physicians to detect and locate Pathological changes with more accuracy. Computed Tomography images can be distinguished for different tissues according to their different gray levels. The images, if processed appropriately can offer a wealth of information which is significant to assist doctors in medical diagnosis. A lot of research efforts have been directed towards the field of medical image analysis with the aim to assist in diagnosis and clinical studies [1]. Pathologies are clearly identified using automated CAD system [2]. It also helps the radiologist in analyzing the digital images to bring out the possible outcomes of the diseases. The medical images are obtained from different imaging systems such as MRI scan, CT scan, Ultra sound B scans.

The computerized tomography has been found to be the most reliable method for early detection of tumors because this modality is the mostly used in radio therapy planning for two main reasons. The first reason is that scanner images contain anatomical information which offers the possibility to plan the direction and the entry points of radio therapy rays which have to target only the tumor region and to avoid other organs. The second reason is that CT scan images are obtained using rays, which is same principle as radio therapy. This is very important because the intensity of radio therapy rays have been computed from the scanned image.

Advantages of using CT include good detection of calcification, hemorrhage and bony detail plus lower cost, short imaging times and widespread availability. The situations include patient who are too large for MRI scanner, claustrophobic patients, patients with metallic or electrical implant and patients unable to remain motionless for the duration of the examination due to age, pain or medical condition. For these reasons, this study aims to explore methods for classifying and segmenting brain CT images. Image segmentation is the process of partitioning a digital image into set of pixels. Accurate, fast and reproducible image segmentation techniques are required in various applications. The results of the segmentation are significant for classification and analysis purposes. The limitations for CT scanning of head images are due to partial volume effects which affect the edges produce low brain tissue contrast and yield different objects within the same range of intensity. All these limitations have made the segmentation more difficult. Therefore, the challenges for automatic segmentation of the CT brain images have many different approaches.

A. Related works

Kai ping et al [5] introduced the effective Particle Swarm optimization algorithm to segment the brain images into Cerebro spinal fluid (CSF) and suspicious abnormal regions but without the annotation of the abnormal regions. Dubravko et al and Matesin et al [6] [7] proposed the rule based approach to label the abnormal regions such as calcification, hemorrhage and stroke lesion.
Ruthmann et al [8] proposed to segment Cerebro spinal fluid from computed tomography images using local thresholding technique based on maximum entropy principle. Loncaric et al proposed [9] to segment CT images into background, skull, brain, ICH, calcifications by using a combination of K means clustering and neural networks. Tong et al proposed [10] to segment CT images into CSF, matter and detection of abnormal regions using unsupervised clustering of two stages. Clark et al [11] proposed to segment the brain tumor automatically using knowledge based techniques. From the above literature survey shows that intensity based statistical features are the straightest forward and have been widely used, but due to the complexity of the pathology in human brain and the high quality required by clinical diagnosis, only intensity features cannot achieve acceptable result. In such applications, segmentation based on textural feature methods gives more reliable results. Therefore texture based analysis has been presented for tumor segmentation such as SGLDM method and wavelet based texture features are used and achieve promising results. Based on the above literature, better classification accuracy can be achieved using wavelet based statistical texture features.

In this paper, we would like to propose a wavelet based statistical texture analysis method and vertical search method to segment the soft tissues and automatically diagnosis abnormal tumor region from brain CT images. The proposed method is illustrated in Figure a. This system uses the classifier SVM [12][13] to classify and segment the abnormal tumor region from brain CT images and gives relatively good segmentation results as compared to the literature discussed above. At first by applying 2 level Discrete Wavelet Transform (DWT), the image is represented by one approximation and three detail sub bands and the co-occurrence matrix [14, 15] is derived for detail sub bands. Then from these co-occurrence matrices, the statistical texture features are extracted using the SGLDM method. The extracted texture features are optimized by Genetic Algorithm (GA) [16] for improving the classification accuracy and reducing the overall complexity.

**II. SYSTEM DESIGN**

Wavelet based statistical texture features is adopted to classify and segment the abnormal and tumor region. The proposed system is divided into (a) Discrete Wavelet Decomposition (b) Feature extraction (c) Feature selection (d) Classification and Evaluation (e) vertical search.

In the proposed system for feature extraction, we have discovered two methods which are wavelet based statistical texture feature extraction method, SGLDM method without wavelet transform.

Firstly the two level wavelet decomposition is performed to decompose the image into one approximation and three detail images and the co-occurrence matrix is derived for 2nd level detail images. Then from these co-occurrence matrices, the statistical texture features are extracted using the SGLDM method. Once all the features are extracted, then for feature selection, we use Genetic Algorithm (GA) to select the optimal texture features. The selected optimal texture features are given as input to the SVM classifier to classify and segment the abnormal tumor region from brain CT images.

**A. Image Preprocessing**

Brain CT images are noisy, inconsistent and incomplete, thus preprocessing phase is needed to improve the image quality and make the segmentation results more accurate. The cropping operation can be performed to remove the background.

Brain CT images are noisy, inconsistent and incomplete, thus preprocessing phase is needed to improve the image quality and make the segmentation results more accurate. The cropping operation can be performed to remove the background. The Contrast Limited Adaptive Histogram Equalization (CLAHE) can be used to enhance the contrast within the soft tissues of the brain images and hybrid median filtering technique can also be used to improve the image.

**B. Discrete Wavelet Decomposition**

A two level wavelet decomposition of region of interest (ROI) is performed which results in four sub bands. Daubechies wavelet filter of order two is used. In 2D wavelet decomposition the image is represented by one approximation and three detail images representing low and high frequency contents image respectively.

The approximation can be further to produce one approximation and three detail images at the next level of decomposition, wavelet decomposition process is shown in Figure a.

**Figure a. Two level discrete wavelet decomposition**
In this Figure a, A1 and A2 represent the wavelet approximations at 1st and 2nd level respectively, and are low frequency part of the images. H1, V1, D1, H2, V2, D2 represent the details of horizontal, vertical and diagonal directions at 1st and 2nd level respectively, and are high frequency part of the images.

Among the high frequency sub bands, the one whose histogram presents the maximum variance is the sub band that represents the clearest appearance of the changes between the different textures. The textures features are extracted from these high frequency sub bands are useful to classify and segment the abnormal tumor region from brain CT images.

In the Figure b, it clearly explains the architecture diagram of the process. First CT scan brain images are taken and perform image processing to remove the noise in the image. Then the features are extracted and selected by using an SVM classifier and the vertical search method is implemented on them for entire search on the image.

C. Feature Extraction

Texture analysis is a quantitative method that can be used to quantify and detect structural abnormalities in different tissues. As the tissues present in brain are difficult to classify using shape or intensity level of information, the texture feature extraction is found to be very important for further classification.

The purpose of feature extraction is to reduce original data set by measuring certain features that distinguish one region of interest from another. The analysis and characterization of textures present in the medical images can be done by using wavelet based statistical feature extraction method. Each sub image is taken from top left corner of the original image is decomposed using two level DWT and co-occurrence matrices are derived for detail or high frequency sub bands (i.e., H2, V2, D2 sub bands).

a) Wavelet based extraction method: A two level discrete wavelet decomposition of region of interest is performed which results in four sub bands. From these co-occurrence matrices, the Wavelet Co-occurrence Texture features (WCT) are computed. Algorithm for feature extraction is as follows

- Obtain the sub-image blocks, starting from the top left corner.
- Decompose sub-image blocks using 2-D DWT.
- Derive SGLDM or Co-Occurrence matrices for detail sub-bands of DWT with 1 for distance and 0, 45, 90 and 135 degrees for θ and averaged.
- From these co-occurrence matrices, the nine Haralick texture features called wavelet Co-occurrence Texture features (WCT) are extracted.

Table I

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Second Order WCT Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Entropy (Measure the disorder of an image)</td>
</tr>
<tr>
<td>2</td>
<td>Energy (Measure the Textural uniformity)</td>
</tr>
<tr>
<td>3</td>
<td>Contrast (Measure the local contrast in an image)</td>
</tr>
<tr>
<td>4</td>
<td>Sum Average (Average of the Gray levels)</td>
</tr>
<tr>
<td>5</td>
<td>Variance (Heterogeneity of an image)</td>
</tr>
<tr>
<td>6</td>
<td>Correlation (Correlation of pixel pairs)</td>
</tr>
<tr>
<td>7</td>
<td>Max Probability (The most prominent pixel pair)</td>
</tr>
<tr>
<td>8</td>
<td>Inverse Difference Moment (Homogeneity)</td>
</tr>
<tr>
<td>9</td>
<td>Cluster Tendency (Measure the grouping of Pixels with similar values)</td>
</tr>
</tbody>
</table>

b) Spatial Gray Level Dependence Matrix method: For each ROI, the SGLDM method can be used to extract the second order statistical texture features for the better diagnosis. The estimated value for these probability density functions can be represented as

\[
\phi (d, \theta) = \left[ P (i, j/d, \theta) \right] 0 < (i, j) \leq N
\]
Where N is the maximum gray level. In this method, four gray co-occurrence matrices for four different directions are obtained for the given distance and the following haralick statistical texture features are calculated for each gray level co-occurrence matrix.

D. Feature Selection

Feature selection is the process of choosing subset of features relevant to particular application and improves classification by searching for the best feature subset, from the fixed set of original features according to a given feature evaluation criterion(i.e., classification accuracy). Optimized feature selection reduces data dimensionalities and computational time and increase the classification accuracy.

The feature selection problem involves the selection of a subset of ‘d’ features from a total of ‘D’ features, based on a given optimization criterion. The D features are denoted uniquely by distinct numbers from 1 to D, so that the total set of D features can be written as S = {1, 2 …D}. X denotes the subset of selected features and Y denotes the set of remaining features.

So S = X U Y at any time. J(X) denotes a function evaluating the performance of X. J depends on the particular application. Here J(X) denotes the classification performance of classifying and segmenting abnormal tumor region from brain CT images using the set of features in X. In this work, Genetic Algorithm (GA) technique is used.

Genetic Algorithm

We consider the standard GA to begin by randomly creating its initial population. Solutions are combined via a crossover operator to produce offspring, thus expanding the current population of solutions. The individuals in the population are then evaluated via a fitness function, and the less fit individuals are eliminated to return the population to its original size. The process of crossover, evaluation, and selection is repeated for a predetermined number of generations or until a satisfactory solution has been found. A mutation operator is generally applied to each generation in order to increase variation. In the feature selection formulation of the genetic algorithm, individuals are composed of bit strings: a 1 in bit position indicates that feature should be selected; 0 indicates this feature should not be selected. As an example bit string 00101000 means the 3rd and 5th features are selected. That is the chromosome represents X={3,5} and Y={1,2,4,6,7,8}. Fitness function for the string X is defined as

\[
\text{Fitness}(X) = J(X) - \text{penalty}(X)
\]

Where X is the corresponding feature subset and penalty(X) = w * (|X| -d) with a penalty coefficient w.

The size value d is taken as a constraint and a penalty is imposed on chromosomes breaking this constraint. The chromosome selection for the next generation is done on the basis of fitness. The fitness value decides whether the chromosome is good or bad in a population. The selection mechanism should ensure that fitter chromosomes have a higher probability survival. So, the design adopts the rank-based roulette-wheel selection scheme. If the mutated chromosome is superior to both parents, it replaces the similar parent. If it is in between the two parents, it replaces the inferior parent; otherwise, the most inferior chromosome in the population is replaced. The selected optimal feature set based on the test data set is used to train the SVM classifier to classify and segment the abnormal tumor region from brain CT images.

D. SVM Classifier

Classification is the process where a given test sample is assigned a class on the basis of knowledge gained by the classifier during training. Support Vector Machine (SVM) performs the robust non-linear classification with kernel trick. SVM is independent of the dimensionality of the feature space and that the results obtained are very accurate. It outperforms other classifiers even with small numbers of available training samples.

SVM is a supervised learning method and is used for one class and n class classification problems. It combines linear algorithms with linear or non-linear kernel functions that make it a powerful tool in the machine learning community with applications such as data mining and medical imaging applications. To apply SVM into non linear data distributions, the data can be implicitly transformed to a high dimensional feature space where a separation might become possible. For a binary classification given a set of separable data set with N samples X = {Xi}, i = 1, 2 … N, labeled as Yi = ± 1.

a) Classifier: Classification is the process where the given test sample is assigned a class on the basis of knowledge gained by the classifier during training.

b) Support Vector Machine (SVM): SVM is a supervised learning method and is used for one-class and n-class classification problems. Its basic idea is to transform a non-linear dividing problem into a linear problem by some kind of kernel function. To apply SVM into non-linear data distributions, the data can be implicitly transformed to a high-dimensional feature space where a separation might become possible.

SVM has the following advantages: First, It is designed to use limited samples to obtain optimum solution for the practical problems. Second, it can ensure to find the global rather than the local optimum solution.
Third, the generalization ability of support vector machine is very good relatively less computational complexity. One major problem is that the computation is complex, the idea of the kernel function solve this difficulty. The problem is a non linear classification one; therefore we use a non-linear SVM that uses a Gaussian Kernel Function

\[ K(X, X) = \exp \left[ -\sigma \| x - x_i \|^2 \right] \]  

\[(3)\]

There are many statistical methods for evaluating and estimating the performance of the classifier. The goal of the work is to obtain classification accuracy of the system. Detection is important in selecting the sub band of the image to be decomposed. The process is done by applying the 2 level 2D DWT, the image is decomposed into four sub bands. After decomposition, SGLDM or Co-occurrence matrices is derived on detail sub bands. From these co-occurrence matrices, WCT features are extracted as given in the feature extraction algorithm and the optimal texture feature set is selected by GA based on the classification performance of SVM classifier.

From the experiments conducted for feature selection, it is found that the optimal feature set which gives good classification performance are the second order WCT features like energy, entropy, variance and inverse difference moment. The four texture features from detail sub bands form the feature vectors or feature set. Efficiency or accuracy of the classifiers for each texture analysis methods are evaluated based on the error rate (i.e.) All tests could have an error rate. This error rate can be described by the terms true and false positive and true and false negative as follows

- Sensitivity = \( \frac{TP}{(TP+FN) \times 100} \)  
- Specificity = \( \frac{TN}{(TN+FP) \times 100} \)  
- Accuracy = \( \frac{(TP+TN)}{(TP+TN+FP+FN)} \)  

Where TN is the number of benign cases truly classified as negative, TP is the number of malignant cases truly classified as positive, FN, malignant cases falsely classified as negative and FP, benign cases falsely classified as positive.

Sensitivity is the ability of the method to identify malignant cases. Specificity is the ability of the method to identify benign cases.

Accuracy is the proportion of correctly diagnosed cases from the total number of cases. To make the classification results comparable the classification performance is evaluated by the round robin method, in each step one dataset is left out and the classifier is trained using the rest and the classifier is applied to the left out dataset. This procedure is repeated such that each dataset is left out once.

F. Vertical search technique

In biomedical application full image clarification is must in order to find the disease. This technique is implemented for to find the directions of pixels. If we found the directions of pixels (i.e.) horizontal and vertical search of the pixels even the small part of the images can be implemented. The result of this technique is explained in the next section.

III. RESULTS AND DISCUSSION

This section describes the comparative study of classification performance of SVM classifier for different texture analysis methods used for classification and segmentation of tumor from brain CT images. The texture feature extracted from each ROI of the image to be selected by pixel based intensity method. The texture feature extracted with same set of images and is obtained from 16 bit gray level images. The SVM is used as a classifier. The results from the SVM classifier for all the texture analysis method are evaluated by using the statistical analysis. Our proposed method is implemented on real human brain CT dataset based on proposed flow diagram. The input data set consists of 100 images: 40 images are normal, 60 images are abnormal. For each texture analysis method, input data set is partitioned into training and test sets which are classified using SVM classifiers. This section describes the wavelet based texture analysis method of classifying and segmenting abnormal tumor region of CT images.

![CT Normal and Tumor Images](image)

**Figure c (a, b, c, d) Example of CT normal and tumor images**

The error rate can be described by terms true and false positive and true and false negative as follows:

**True Positive (TP): Abnormal cases correctly classified**
True Negative (TN): Normal cases correctly classified
False Positive (FP): Normal cases classified abnormal
False Negative (FN): Abnormal cases classified normal

Vertical search technique is done after SVM classification. This vertical search technique entirely searches the brain CT images and find out whether there is any abnormal tumor region in the brain. It searches both horizontally and vertically through the pixel directions.

a) Vertical Y region  

b) vertical X region

Figure d (a,b) Results of vertical search on brain CT images

Feature selection is carried out using GA. There are 9 features are extracted from detail sub bands. So totally 9*3= 27 features are extracted. The next step is to determine the relevance of each selected feature to the process of classifying and segmenting abnormal tumor region. During the evaluation process by using GA, some features may be selected many times as the number of generation increases. If the feature was selected more times that feature was given as more important in the feature selection. The number of times the features selected was energy, entropy, variance and inverse difference moment. The parameter set for the GA algorithm is as follows: Population size is 30; Cross Over probability is 1.0; Mutation rate is 0.1; Penalty coefficient is 0.5and stopping condition is 100 generations .Results show that, if the number of sample images increased, we get good classification accuracy for the 10 fold cross validation method.

The table clearly explains the segmentation accuracy of different segmentation methods, from this SVM classifier has better segmentation accuracy.

Segmentation accuracy= (number of pixels matched/total number of tumor pixels in ground truth)*100

Segmentation error= (number of misclassified pixels/number of pixels in segmented tumor image)*100

The segmentation accuracy is calculated as the direct ratio of the number of tumor pixels common for ground truth and the proposed method output to the total ground truth tumor pixels. The segmentation error is calculated as the number of misclassified pixels to the total number of pixels in the segmented tumor region.

A. Comparison of the computation time with other reported results

A computer software system is designed for classification of benign and malignant tumor slices. The computation time to extract the texture features from each single image is evaluated with the existing texture analysis methods and the results of the graph are represented using the graph as shown in Figure 3. The existing methods such as co-occurrence texture features proposed by Sharma et al. [3] and Padma Sukanesh [12], wavelet co-occurrence texture features proposed by Kharrat et al. [10]. From the results, it is observed that SVM with wavelet based combined dominant grey level run length matrix method and grey level co-occurrence matrix method is less computation time when compared with conventional texture analysis method.

During the evaluation process by using GA, some features may be selected many times as the number of generation increases. If the feature was selected more times in the feature set that feature was given as more important in the feature selection.
B. Conclusion and Future Enhancement

A method for combined dominant run length and co-occurrence texture features from wavelet approximation tumor image of each slice and evaluated the performance of SVM classifier in segmenting and classifying the CT image. The algorithm has been designed based on the concept of different types of brain soft tissues (CSF, GM, abnormal tumor region) have different textural features. The selection method of ROI is simple and accurate. The result shows that the segmentation and classification of tumor for the new run length feature extraction method yields better results when compared to the other texture analysis methods based on SVM classifier. It is found that this method gives favorable result with accuracy percentage of above 98% for the images that are being considered. This would be highly useful as a diagnostic tool for radiologists in the automated segmentation of tumor in CT images.

The goal of this work is to compare the classification performance of the SVM classifier based on different texture analysis methods. Use of large data bases is expected to improve the system robustness and ensure the repeatability of the resulted performance. The automation procedure proposed in this work using a SVM enables proper abnormal tumor region detection and segmentation thereby saving time and reducing the complexity involved.

The limitation of this method is that it needs new training for the classifiers whenever there is change in the slice dataset and this method can be applied to brain CT images only. The proposed method has better performance when compared to the other existing methods. Plans for future work include the specific annotation of the abnormal regions such as hemorrhage, calcification and lesion, and other types of imaging such as liver CT imaging, MRI imaging and ultrasound imaging as a future work. The developed segmentation system is expected to provide valuable diagnosis for the physicians.

REFERENCES


AUTHOR’S PROFILE

M. Ragavi B.E.,(M.E.), she received his Bachelor degree in Computer Science and Engineering from Noorul Islam College of Engineering, Thuckalay, and Affiliated to Anna University Chennai in 2012. Now she is doing her Post Graduate degree II nd yr in Computer Science and Engineering in Ponjesly College Of Engineering, Nagercoil and Anna University, Chennai. Her area of interest includes Network security, Image processing.

T. L. Nija Shining Gold M. E., Assistant Professor in Ponjesly College Of engineering, Nagercoil. She received his Bachelor degree from C.S.I Institute of Technology, Thovalai, Tamilnadu, India in 2008. She received his post graduate degree from Saveetha Engineering College, Chennai, Tamilnadu, India in 2010. Her area of processing.

T. L. Nija Shining Gold M. E., Assistant Professor in Ponjesly College Of engineering, Nagercoil. She received his Bachelor degree from C.S.I Institute of Technology, Thovalai, Tamilnadu, India in 2008. She received his post graduate degree from Saveetha Engineering College, Chennai, Tamilnadu, India in 2010. Her area of interest includes Data mining, Image processing.