IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Brain Tumor Segmentation and Classification Using Modified FCM and SVM Classifier

Aparna M. Nichat[1], S. A. Ladhake[2]

Computer Science and Engineering Department, Sipna's COET SGBAU, Amravati, India[1]

Sipna's COET SGBAU, Amravati, India[2]

Abstract: The field of medical imaging gains its importance with increase in the need of automated and efficient diagnosis in a short period of time. MRI is the most important technique, in detecting the brain tumor. In this paper data mining methods are used for classification of MRI images. A new hybrid technique based on the support vector machine (SVM) and modified fuzzy c-means for brain tumor classification is used. The purposed algorithm is a combination of support vector machine (SVM) and modified fuzzy c-means, a hybrid technique for prediction of brain tumor. In this algorithm the first stage is noise reduction using Median Filtering. Modified fuzzy c-means (FCM) clustering is used for the segmentation of the image to detect the suspicious region in brain MRI image. Texture based features such as GLCM(Gray Level Co-occurrence Matrix)features is used for extraction of feature from the brain image, after which SVM technique is applied to classify the brain MRI images, which provide accurate and more effective result for classification of brain MRI images.

Keywords: Data Mining, MRI, Modified Fuzzy C-means clustering, Gray level co-occurrence matrix (GLCM), Support Vector Machine (SVM).

I. INTRODUCTION

Data mining is a simple and robust tool to extract the information from large dataset [1]. Classification is a branch of data mining field. In this field many classification techniques are available for medical images such as artificial neural network (ANN), fuzzy c-means (FCM), support vector machine (SVM), decision tree and Bayesian classification. A number of researchers have been implemented the classification techniques for medical images classification. Presently many medical imaging techniques such as positron emission tomography (PET), x-ray, computed tomography (CT), magnetic resonance imaging (MRI), for tumor detection but MRI imaging technique is the good because of its higher resolution and most researchers have used MRI imaging for diagnosing tumor. In this paper, we have to preprocess the given test image for reducing noise and to enhance the contrast. Then, we have to segment the brain image using modified fuzzy c-means algorithm to visually show the abnormalities such as tumor. Afterwards, texture features (GLCM) will be extracted from it. In feature extraction stage, statistical measurements are calculated from the gray level co-occurrence matrix for different directions and distances.

Among the various features extracted. We have to select the distinct features that will be utilized for classification purpose. For the selection of features SFS (Sequential Forward Selection) is used. Support vector machine (SVM) is used to classify whether the test image comes under normal, benign and malignant.In brain tumor detection and classification method, we have to preprocess the given test image for reducing noise and to enhance the contrast. Then, we have to segment the brain image using modified fuzzy c-means algorithm to visually show the abnormalities such as tumor. Afterwards, texture features (GLCM) will be extracted from it. In feature extraction stage, statistical measurements are calculated from the gray level co-occurrence matrix for different directions and distances. Among the various features extracted. We have to select the distinct features that will be utilized for classification purpose. For the selection of features SFS (Sequential Forward Selection) is used. Support vector machine (SVM) is used to classify whether the test image comes under normal, benign and malignant.

II. RELATED WORK

Support vector machines were applied in many researches which are given in [2-4]. H. B. Nandpuru, Dr. S. S. Salankar and Prof. V. R. Bora, worked on MRI brain cancer classification using support vector machine. Support Vector Machines (SVM) was applied to brain image classification. In this paper feature extraction from brain MRI Images were carried out by gray scale, symmetrical and texture features. They achieved good result [2]. A. Padma and R. Sukanesh, their study on SVM Based Classification of Soft Tissues in Brain CT Images using Wavelet Based Dominant Gray Level Run Length Texture Features. They have emphasized on the technique of medical CT imaging as one of the widely applied and reliable technique used for the detection and location of pathological changes efficiently, using SVM. They obtained 98% accuracy [3]. S.H.S.A. Ubaidillah, R. Sallehuddin and N.A. Ali, worked on cancer detection using artificial neural network and support vector machine: A Comparative study. In this paper, they compared the performance on four different cancer datasets using SVM and ANN classifiers. In this study, the

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

ANN classifier obtained good classification performance on the datasets that have bigger amount of input features (prostate and ovarian cancer datasets) SVM also presented good performance on datasets with smaller amount of input features (breast cancer and liver cancer), but finally SVM classifier provided better result for tumor [4].

III. PROPOSED METHODOLOGY

The proposed method illustrates the basic steps followed by our automation system in clearly diagnosing a given tumor data set into its classes as shown in figure 1. It starts by loading the input MR images, image preprocessing, image segmentation, feature extraction; applying modified fuzzy c means algorithm and support vector machine (SVM) classifier.

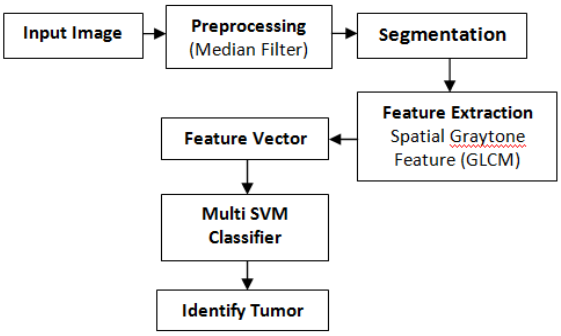

Figure 1: Block Diagram of Proposed System

Image is processed through:

- Preprocessing (Median filter)
- Segmentation (Modified FCM)
- Feature extraction (GLCM technique)
- Feature selection (Sequential forward selection)
- Classification by SVM

A. Preprocessing :

Preprocessing mentions to improvement of tumor intensity, noise, reduction, background removal, contrast manipulation, filtering and edges sharpening etc. Medical images are difficult to interpret the tumor, therefore preprocessing step is required with a specific end goal to make the picture segmentation and improve the quality of image results more precise.

Median filtering operation, the pixel values in the neighborhood window are ranked according to intensity, and the middle value (the median) becomes the output value for the pixel under evaluation. Median filtering does not shift boundaries, as can happen with conventional smoothing filters. Since the median is less sensitive than the mean to extreme values (outliers), those extreme values are more effectively removed. Median filtering preserves the edges.

B. Segmentation :

Segmentation is the technique of separating an image into multiple slices and object region. This provides good result for tumor segmentation. In this work, modified fuzzy c-means algorithm was used in MRI image segmentation. Modified Fuzzy C-Means (FCM) algorithm is used to find out the suspicious region from brain MRI image.

The modified FCM algorithm is based on the concept of data compression where the dimensionality of the input is highly reduced. The data compression includes two steps: quantization and aggregation [5].

The quantization of the feature space is performed by masking the lower 'm' bits of the feature value. The quantized output will result in the common intensity values for more than one feature vector. In the process of aggregation, feature vectors which share common intensity values are grouped together. A representative feature vector is chosen from each group and they are given as input for the conventional FCM algorithm. Once the clustering is complete, the representative feature vector membership values are distributed identically to all members of the quantization level. Since the modified FCM algorithm uses a reduced dataset, the convergence rate is highly improved when compared with the conventional FCM. The modified FCM algorithm uses the same steps of conventional FCM except for the change in the cluster updation and membership value updation criterions. The modified criterions are showed below

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} y_j}{\sum_{j=1}^{n} u_{ij}^{m}} \qquad u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}}$$

$$\text{Where } d_{ij} = y_j - c_i$$

$$\text{y} = \text{Reduced Dataset}$$

C. Feature Extraction

Features are said to be properties that describes the whole image. It can also refer as an important piece of information which is relevant for solving the computational task related to specific application. The purpose of feature extraction is to reduce the original dataset by measuring certain features. The extracted features acts as input to classifier by considering the description of relevant properties of image into feature space.

The Gray Level Co-occurrence Matrix (GLCM) method is a way of extracting statistical texture features. A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels, G, in the image. The matrix element P (i, j | Δx, Δy) is the relative frequency with which two pixels, separated by a pixel distance (Δx, Δy), occur within a given neighborhood, one with intensity 'i' and the other with intensity 'j'. The matrix element P (i, j | d, ɵ) contains the second order statistical probability values for changes between gray levels 'i' and 'j' at a particular displacement distance d and at a particular angle (ɵ). Using a large number of intensity levels G implies storing a lot of temporary data, i.e. a G × G matrix for

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

each combination of (Δx, Δy) or (d, ɵ). Due to their large dimensionality, the GLCM's are very sensitive to the size of the texture samples on which they are estimated. Thus, the number of gray levels is often reduced.

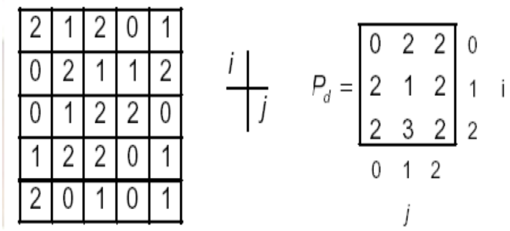

Figure 2: GLCM Matrix

In this matrix Count all pairs of pixels in which the first pixel has a value i, and its matching pair displaced from the first pixel by d has a value of j. This count is entered in the ith row and jth column of the matrix Pd[i,j] Note that Pd[i,j] is not symmetric, since the number of pairs of pixels having gray levels [i,j] does not necessarily equal the number of pixel pairs having gray levels [j,i]. Afterwards various parameters like energy, entropy, mean and contrast has been calculated.

D. Feature Selection

Automatic feature selection is an optimization technique that, given a set of m features, attempts to select a subset of size n that leads to the maximization of some criterion function. Feature selection algorithms are important to recognition and classification systems because, if a feature space with a large dimension is used, the performance of the classifier will decrease with respect to execution time and to recognition rate. The execution time increases with the number of features because of the measurement cost. The recognition rate can decrease because of redundant features and of the fact that small number of features can alleviate the course of dimensionality when the training samples set is limited, leading to overtraining. On the other hand, a reduction in the number of features may lead to a loss in the discrimination power and thereby lower the accuracy of the recognition system.

In order to determine the best feature subset for some criterion, some automatic feature selection algorithm can be applied to the complete feature space, varying the number of selected features from 1 to m. Sequential Forward Selection is the simplest greedy search algorithm. Starting from the empty set, sequentially add the feature x^+ that results in the highest objective function $J(Y_k+x^+)$ when combined with the features Y_k that have already been selected

Algorithm

1. Start with the empty set $Y_0=\{\varnothing\}$
2. Select the next best feature $x^+=\arg\max[J(Y_k+x)]$
3. Update $Y_{k+1}=Y_k+x^+$; k=k+1
4. Go to 2

E. Classification by SVM

Aim of SVM classifier is to group items that have similar feature values into groups. Classifier achieves this by making a classification decision based on the value of the linear combination of the features.

- Data setup: our dataset contains three classes as normal, benign (non-cancerous), malignant (cancerous) each N samples. The data is 2D plot original data for visual inspection.
- SVM with linear kernel (-t 0). We want to find the best parameter value C using 2-fold cross validation (meaning use 1/2 data to train, the other 1/2 to test).
- After finding the best parameter value for C, we train the entire data again using this parameter value.
- Plot support vectors.
- Plot decision area.

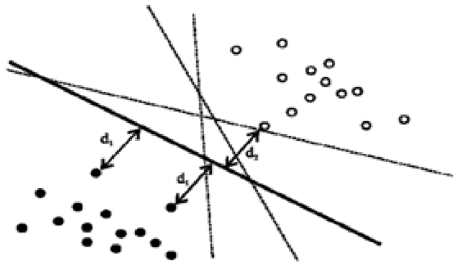
Figure 3: Support Vector Machine classifier

Expression for hyper plane: w.x+b = 0
Margin is d1+d2.
where x – Set of training vectors
w–vectors perpendicular to the separating hyper plane
b–offset parameter which allows the increase of the margin

A grouping of all the classes in two disjoints groups of classes. This grouping is then used to train a SVM classifier in the root node of the decision tree, using the samples of the first group as positive examples and the samples of the second group as negative examples. The classes from the first clustering group are being assigned to the first (left) subtree, while the classes of the second clustering group are being assigned to the (right) second subtree. The process continues recursively until there is only one class per group which defines a leaf in the decision tree.

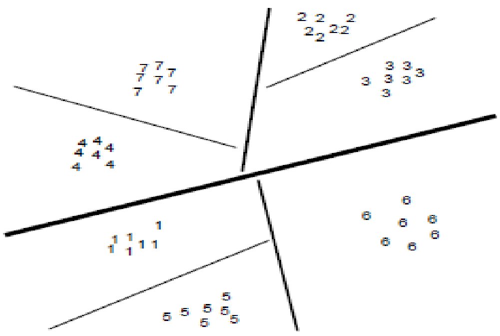
Figure 4: n-class SVM

F. Performance measures

Classification, the sensitivity, specificity and accuracy were calculated using below formulas:

- True Positive (TP): Abnormal brain correctly identified as abnormal.

DOI 10.17148/IJARCCE.2016.5420

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

- True Negative (TN): Normal brain correctly identified as normal.
- False Positive (FP): Normal brain incorrectly identified as abnormal.
- False Negative (FN): Abnormal brain incorrectly identified as normal.

1. Sensitivity = TP/ (TP+FN) *100%
2. Specificity = TN/ (TN+FP) * 100%
3. Accuracy = (TP+ TN)/ (TP+ TN+FP+FN)* 100 %

All these three parameters are used to check the classifiers performance.

IV. CONCLUSION

In this proposed system brain MRI images proved to be a significant way to detect the brain tumor. The hybrid methodology of combining support vector machine and modified fuzzy c-means clustering for classification gives accurate result for identifying the brain tumor. For future work, to get better accuracy rate and less error rate a hybrid SVM algorithm is to be proposed. In future work, different data mining techniques can be used to train in order to improve the performance of the classifiers and the data sets can also be increased.

REFERENCES

[1]. Prakash Mahindrakar and Dr. M. Hanumanthappa, “Data Mining In Healthcare: A Survey of Techniques and Algorithms with Its Limitations and Challenges”, Int. Journal of Engineering Research and Applications, ISSN: 2248-9622, Vol. 3, Issue 6, Nov-Dec 2013, pp.937-941

[2]. H. B. Nandpuru, Dr. S. S. Salankar, Prof. V. R. Bora, MRI brain cancer classification using support vector machine. 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, 978-1-4799-2526-1/14/$31.00 ©2014 IEEE

[3]. A. Padma and R.Sukanesh, “SVM based classification of soft tissues in brain CT images using wavelet based dominant gray level run length texture features”, middle-east journal of scientific research, 2013, 13(7): 883-888.

[4]. S.H.S.A. Ubaidillah, R. Sallehuddin and N.A. Ali, “Cancer detection using artificial neural network and support vector machine: A Comparative study”, journal technology (science & engineering), 2013, 65:1.

[5]. Jude hemanth.D, D.Selvathi and J.Anitha, “Effective Fuzzy Clustering Algorithm for Abnormal MR Brain Image Segmentation”, Page Number 609-614, International/Advance Computing Conference (IACC 2009),IEEE,2009.

[6]. B. Gupta and S. Tiwari, “Brain Tumor Detection using Curvelet Transform and Support Vector Machine”, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.4, April-2014, pg. 1259-1264, ISSN 2320–088X.

[7]. Parveen and Amritpal singh, “Detection of Brain Tumor in MRI Images, using Combination of Fuzzy C-Means and SVM” 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN) ©2015 IEEE.