Accurate Detection of Abnormal Tissues of Brain MR Images Using Hybrid Clustering and Deep Learning Based Classifier Methods of Image Segmentation

Shivang M. Patel¹, Jyotindra N. Dharwa²

Research Scholar, Faculty of Computer Applications, Ganpat University, Kherva, Mehsana, Gujarat, India¹
Associate Professor, A. M. Patel Institute of Computer Studies, Ganpat University, Kherva, Mehsana, Gujarat, India²

ABSTRACT: Image analysis by segmentation techniques has wide range of utilities. The research on image analysis includes industrial inspection, object tracking from multiple sequential images, traffic control systems, geographical classification from satellite images, geomorphological mapping at land surfaces, crop analysis in agriculture, treatment planning in medical science and many others. Segmentation of an image is very significant in medical field by improving treatment possibilities like early diagnosis of abnormalities in brain. Manual segmentation for detection of abnormal tissues from large amount of brain MR images are difficult and time-consuming task. In this research paper, brain MR images are taken due to its quality parameters for the purpose of detecting abnormalities compared to other techniques like X-Ray, CT Scan etc. This paper reviews the trade-off between various approaches for analysing the abnormalities from brain Magnetic Resonance Images using different methods of image segmentation. In addition, the proposed study has given a roadmap to improve accuracy and precision to detect the region of interest. From multiple methods of segmentation, segmentation using hybrid clustering approach including K-means and Fuzzy C-means followed by deep learning based classifier method improves the level of precision to detect abnormal tissues from brain MR images. The result by implementing the proposed approach in MATLAB environment leads to plan precise treatment for the medical practitioners.


I. INTRODUCTION

Segmentation of an image is a process of dividing/partitioning/separating a digital image into its several simplified, homogeneous and mutual exclusive regions (set of pixels) as per specific criteria so it is more meaningful to study and easier to analyse. For understanding anatomical structure of the human body using medical imaging applications, segmentation plays a vital role. For effective diagnosis and treatment planning based on the different tissues of a subject, various medical imaging practices are used like MRI, CT Scan, mammography etc. [6] In this study, MR Images are used because to classify them into limited number of groups having comparatively high contrast between different tissues.Different approaches of segmentation like Thresholding, Histogram, Clustering, Classifications, Edge and Region growing, Artificial Neural Networks, Markov Random Field models etc. are suited to various types of images. [1][3][7] The quality of outcome of a particular approach is difficult to measure in terms of accuracy and precision due to the fact that there may be many correct algorithms or methods of segmentation for a single image. In medical imaging field, identification and detection of abnormalities from brain MR images like tumor tissues, region of swelling, infection area etc. are still a challenging and tough because of varieties of characteristics of an image like shapes, locations, intensities etc. [2][4][5]
This paper presents the current status of various methods of segmentations in medical imaging applications. Current segmentation approaches are reviewed with advantages and disadvantages of various methods. It focuses mainly on comparison of various classification and clustering techniques of image segmentation on brain MRI images in terms of their accuracy and performance for detection of abnormalities based on various criteria and parameters. This research paper analyses few methods of segmentation on different brain MRI image modalities like T1 and T2 weighted, FLAIR images of various views such as sagittal, axial etc. and compare results for accuracy. It also proposes a hybrid clustering approach followed by deep learning based classifier that improves the level of precision to detect abnormal tissues from brain MR Images. In this hybrid approach, K-means clustering is better for minimal computation time and Fuzzy C-means clustering gives notable accuracy. K-means technique is simple but suffers from incomplete detection of high damaged tissues. Fuzzy C-means technique recalls more information of original source image. The main idea behind hybrid approach is to reduce the number of iterations for initializing the correct cluster centre. Performance is assessed by comparing gathered results with ground truth normal images in terms of accuracy as well as processing time.

II. RELATED WORK

Segmentation process splits an image into its integral regions till the interested/targeted areas have been detected. Mainly segmentation algorithms depend on similarity or discontinuity in intensity values. Similarity property deals with identification of edges while discontinuity detects regions based on a set of previously defined criteria like region growing, thresholding, region splitting and merging etc. [1][26][27] Different image segmentation approaches are applied in various extents in different fields. Early diagnosis of abnormalities in brain tissues improves treatment planning for the patient by medical practitioners. Let us discuss about medical image segmentation and related methods from the literature survey. The medical image segmentation methods available in the literature can be divided into major following categories. [1][2] Clustering, Thresholding, Classifiers, Region growing, Neural Network, Markov Random Field etc. Manual segmentation involves delineation of the boundaries of tissues manually. It is a tedious task. It requires additional care to mark regions carefully, otherwise it will generate jaggy images and may be result in poor segmentation. In semi-automatic segmentation, human interaction is least as possible. Since it involves both computer and humans’ expertise so the result depends on both as well as subject to variations in between. However, fully automated segmentation is also not without problems to retrieve specific region. In such methods, there is no intervention of human and segmented tissues. Though it involves the human intelligence and is developed with soft computing techniques, which is a difficult task. [26]

Among various methods of image segmentation, classifier and clustering based approaches have received an excessive pact of attention in medical imaging research community. Several classifier and clustering strategies have been used by researchers, each of which has its own special characteristics, merits and demerits. Recently, automatic segmentation through deep learning based technique became more effective to achieve desired results. Image classification techniques are mainly divided in two categories: supervised and unsupervised image classification techniques. [10] In supervised approach, first we need to select some pixels having same characteristics from each class to train pixels so later they can be used for identification of issues having similar properties. Some clustering techniques like log based clustering, relevance feedback, hierarchical clustering, k-means clustering, fuzzy based clustering etc. are used according to requirements. [11] In unsupervised approach, it is not required to train pixels. In cases like absence of sufficient basic information, this approach is useful. Various classifier techniques are available like minimum distance, Bayesian, maximum likelihood, convolutional neural network, decision tree etc. [8]

Different clustering and classification methods have their own qualities. So, it is relatively difficult to decide which approach is best suitable for specific study. Results may vary according to the preferred classifier(s). Automated segmentation helps to overcome the necessity of time required in the process of manual segmentation for large datasets. It also reduces the problem occurred by inter and intra observer. Several clustering and classification methods are intended at enhancing the prediction accuracy of diagnosis process in detecting abnormalities from MR images. In recent years, many methods have been recommended for identification and detection of abnormal tissues.
from brain MR Images using different segmentation approaches. Here, few of them are applied to solve different problems by different researchers and related discussion are described briefly.

A fully automated segmentation which uses SVM classification, splits normal and abnormal tissues before they can be classified into: grey matter cell, cerebrospinal fluid, white matter cell, and tumor area. This approach is more robust and spend less time in whole process. [11]Another approach is Bayesian which is considered as supervised as well as parametric. Here, the mean is calculated using training dataset. [12] Threshold based approach uses binary image. In this method, regions are converted into various grey values. Here, intensity values are separated by anticipated classes. [13] Otsu method ensures high intensity level difference between foreground and background images. Then it converts into binary form as needless region. [14] In [15], automated segmentation method was introduced to find tumor region by more accuracy and less diagnosis time using Multi-Layer Perception Network and contrast level. Statistic features like mean, median, variance and correlation were also used. This approach uses neural network for classification. In [16], applications of intelligent computing techniques were discussed for image classification. This study focuses on AI approach to solve different issues related to various diagnostic. Different methods like SVM neural network, fuzzy C means, mammogram, hybrid approach etc. used for feature extraction are discussed in that research study. Segmentation methods for brain tissue analysis, can be separated in two categories: generative and discriminative model. Generative models depend on related prior knowledge about the presence of healthy as well as damaged tissues. These methods depend on structural models of 3D brain MR image tissues. [17][18][19] Techniques such as SVM [20][21], decision forests [22][23] and random forests [24][25] also proved accurate to get targeted regions.

### III. Sample Selection Procedure, Sample Size and Tools Used to Study

This research study considers 16 normal and 56 abnormal standard size (128 X 128 pixels or 512 X 512 pixels) brain MR images like T1, T2, FLAIR (Fluid Attenuation Inversion Recovery) etc. obtained from various imaging centres and few samples from authenticated online sources like brainweb and other databases. Different views of same images are used for better results. Experiments by different techniques are applied on multiple same images, so comparison will become easier and accurate. This research uses MATLAB tool (R2017a trial version 9.2.0.538062) for performing various operations on brain MR images. For more specific processing according to various parameters on images, it is necessary to use special plug-in that is used for clinical data analysis. Here, for pre-processing stage, ITK-SNAP (Insight Segmentation and Registration Toolkit, version 3.6.0) tool is used to open individual images as well as sequences/series stored in different file formats. Medical images obtained from hospital or imaging centres are in .dcm file format and they can be observed here by DICOM (Digital Imaging and Communications in Machine) viewer tool.

### IV. Proposed Framework

Here this analytical study uses the facts/information already available and analyse these for identification and detection of abnormal tissues in brain MR images. This research paper uses statistical data obtained by implementing proposed hybrid approach and evaluate the accuracy of the results. After applying hybrid clustering and before post-processing, obtained images are entered for further processing using deep learning based classifier. One can also use pre-trained network model of brain MRI images for experiment and create N number of hidden layers as per the diagnostic requirement. Generally, at initial stages, two hidden layers should be considered. Each layer filters data and give to next subsequent layer for further processing. In this process, the input image or sequences are compared with pre-trained labelled dataset to get more specific results. The resulted outcome will give more statistics about the images that will be more helpful in preliminary diagnoses. The entire procedure is presented by proposed below framework.

1. **Pre-processing**

In pre-processing, an input image is converted into grayscale using rgb2gray function. To improve the image quality for experiment, it is desirable to reduce/remove noise from original brain MRI image. Denoising can be done using Gaussian or Poisson Function and Filtration can be done by standard median filter to preserve edges without reducing...
the image sharpness of observed MR images. As per experimental requirement, the process of removing background like remove skull, scalp, eyes or any other non-interested area is also initiated using brain surface extractor algorithm. It will decrease the processing time.

2. Clustering
After removing the noise, feed that image to proposed technique.
- Initialize number of clusters, maximum iterations and termination parameters. Initial number of clusters are not chosen at random so it saves time and effort by reducing number of iterations.
- Calculate cluster centre by initial mean = (1:k)*m/(k+1) where k=no. of clusters & m=maximum no. of image count + 1
- Find minimum distance between point and cluster centre to allot each point to nearest cluster centre.
- Then recalculate the cluster centres and repeat this step until some convergence criteria is met.
- Some points are scattered and they are far away from any cluster centre. For such points calculate new values of cluster, distance and mean values.
- There is no inference between points exists in clusters by reclustering process.
- Record output as clustering image, execution time and number of iterations as well as degree of membership by putting each point together which belong to closest cluster.

3. Feature Extraction
After applying the proposed approach, the resulting image is continued to process for feature extraction to get interest area. We can use Thresholding and Contouring function to get region of interest with proper boundary tracking. Then image is entered into binarization process to observe accurate results.

4. An intermediate stage before validating image - Deep Learning based Classifier
Classification algorithms enable to model a categorical response and Statistics and Machine Learning Toolbox offers various functions that cover a variety of parametric and nonparametric techniques. For object recognition, one can create a CNN (Convolutional Neural Network) and train a network from scratch. For creation of such a network, it is required to define the number of layers and filters along with various parameters. It also requires massive amount of sample data and a lot of time. A common alternative to training a CNN from scratch is to use a pretrained model to automatically extract features from a new data set. This method, called transfer learning, is a convenient way to apply
deep learning without a huge data set and long computation and training time. Using MATLAB and ‘nntool’ – Neural Network Toolbox, one can train the required convolutional neural network from scratch or use a pretrained model to perform transfer learning. [Source: mathworkhelp][29]

V. PERFORMANCE MEASURES

Two performance measures are used for measuring performance: i) find sensitivity, specificity and accuracy ii) find dice similarity index. For this research study, we have used clustering and classification approach with various parameters. Classification results are not without any error. It may be possible that experimental study may fail to identify the abnormal tissues or may identify the abnormal tissues which are not actually present. This fact may be described by finding the values of true and false positive and true and false negative. [32][33]

Table 1: Condition of classification results and description

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>Classification result is positive i.e. presence of abnormality (abnormal tissues are correctly detected)</td>
</tr>
<tr>
<td>TN</td>
<td>Classification result is negative i.e. absence of abnormality (normal tissues are not correctly detected)</td>
</tr>
<tr>
<td>FP</td>
<td>Classification result is positive in absence of abnormality (normal tissues are incorrectly detected as abnormal tissues)</td>
</tr>
<tr>
<td>FN</td>
<td>Classification result is negative in presence of abnormality (abnormal tissues are not detected)</td>
</tr>
</tbody>
</table>

We can find the performance of applied classifier using sensitivity, specificity and accuracy measures. [35]

Table 2: Performance Matrix

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Normal</th>
<th>Abnormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>TP / (TP + FN) * 100 %………………………………...eq. (1)</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>TN / (TN + FP) * 100 %………………………………...eq. (2)</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>(TP + TN) / (TP + TN + FP + FN) * 100 %…….eq. (3)</td>
<td></td>
</tr>
</tbody>
</table>

Based on above criteria, we can compare various algorithms by finding FPR, TPR and precisions. Better performance is indicated if the values of accuracy and sensitivity are higher and the value of specificity is lower.

Other measures for testing accuracy and precision of acquired result are Dice and Jaccard similarity index. For each abnormal tissue region, P1 represents the segmented area by the proposed method, and T1 is the actual area in the ground truth. Then, dice and Jaccard index functions [30][31] are written as,

Dice Index (P, T) = (P1.T1) / ( (P1+T1) / 2 ) …………………………………………………………………………………eq. (4)

Jaccard Index (P, T) = (P1.T1) / [P1+T1-(P1.T1)] …………………………………………………………………………………eq. (5)

VI. EXPERIMENTAL STUDY

For this experimental study, 16 normal and 56 abnormal brain MR images are considered. Brain MR Image with abnormal tissues of 512 x 512 pixel dimension is taken as input. The next step is preprocessing. In which, operations like removal of noise, skull etc. is done. Then filters like median filter, morphological filter are applied for edge preservation. Following figures demonstrate these steps:

![Fig. 2 various operations on inputted brain MRI image](image)

(a) Abnormal Brain MRI Image (b) RGB2GRAY Image (c) Threshold Image (d) Enhanced Image

The next stage is to perform various morphological operations on obtained images from previous stage.
As we have discussed in proposed framework, after applying hybrid clustering approach the next step is to input acquired images and results to train network model using ‘nntool’ available in MATLAB environment. Here, this is optional as one can use already pretrained network of same domain for further analysis. Following figures show the wizard for how to input data images and phases of processes like training, validating and testing.

VII. RESULT ANALYSIS

Following results in forms of various tables and charts depict and prove that the proposed approach for segmentation is more precise compared to other methods.

Table 3: Various Parametric Statistics for different segmentation approaches for Image Dataset-I

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>K-means</th>
<th>Fuzzy C-means</th>
<th>ExpectationMaximization</th>
<th>Hybrid Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Clusters</td>
<td>11</td>
<td>9</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Number of Iterations</td>
<td>43</td>
<td>38</td>
<td>63</td>
<td>33</td>
</tr>
<tr>
<td>Processing Time (s)</td>
<td>29.38</td>
<td>17.23</td>
<td>36.31</td>
<td>37.56</td>
</tr>
</tbody>
</table>

Chart 1: Execution Time vs. various Segmentation methods
Table 4: Comparison of accuracies (Confusion Matrix) by different approaches for Image Dataset-I

<table>
<thead>
<tr>
<th>Evaluation Parameters</th>
<th>Clustering Approach</th>
<th>Classification Approach</th>
<th>Hybrid Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k-means</td>
<td>Fuzzy C-means</td>
<td>Expectation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Maximization</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>k-NN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decision Tree</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SVM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Proposed</td>
</tr>
<tr>
<td>TP</td>
<td>36</td>
<td>41</td>
<td>24</td>
</tr>
<tr>
<td>TN</td>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>FP</td>
<td>17</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td>FN</td>
<td>9</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>63.89</td>
<td>73.61</td>
<td>52.78</td>
</tr>
<tr>
<td>Precision (%)</td>
<td>67.92</td>
<td>74.55</td>
<td>51.06</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>80.00</td>
<td>89.13</td>
<td>68.57</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>37.04</td>
<td>46.15</td>
<td>37.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>42.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>48.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>69.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>78.26</td>
</tr>
</tbody>
</table>

Table 5: Performance Analysis parameters for segmented tissues using hybrid approach

<table>
<thead>
<tr>
<th>Image Dataset</th>
<th>Dice Similarity Index</th>
<th>Jaccard Similarity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.94</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Chart 2: Accuracy of proposed framework vs. existing systems
From this research study, we conclude that the proposed framework for detecting abnormal tissues from brain MRI images is more effective and efficient compare to other similar methods of clustering and classification of segmentation. Proposed segmentation method has achieved Dice coefficient of 94%, Jaccard coefficient of 92%, Accuracy of 90.28% and Precision of 90.38%. The proposed approach performs feature extraction and classification better than other similar segmentation methods. Moreover, the application of the proposed approach can be extended to a varying range of tumor and other abnormalities. It can also be applied on larger dataset of MR images having different aspects to get more realistic and more clinically bounded outcomes.

REFERENCES


S. Jayraman, S. Esakkirajan, T. Veerakumal DIP, TMH, 2009
