Integrated Hybrid Segmentation to Overcome Problems in Brain Tumor Segmentation

A. C. Prabu* and S. Narayanaamoorthy
Department of Mathematics, Bharathiar University, Coimbatore – 641046, Tamil Nadu, India; prabhucphd@gmail.com, snm_phd@yahoo.co.in

Abstract

Background/Objective: To improve the performance of the segmentation of Brain tumor images by introducing integrated hybrid segmentation approach. Methods/Statistical Analysis: An innovative integrated approach combining a set of features of some efficient algorithms is proposed in this paper. The proposed algorithm known as integrated hybrid segmentation approach uses the features that classify the MRI images based on the local independent projections. The hybrid PAM and enhanced possibilistic fuzzy C-means approach is used in the partition and determination of the cluster centers. Findings: The problem of false segmentation and segmentation with low accuracy that prevails in the brain image segmentation can be overcome by using this approach. The classified images are used as training data using which different dictionaries are constructed with all classes. The testing samples are projected to the dictionaries and reconstructed using local anchor embedding approach. Thus the proposed integrated algorithm improves the segmentation process of brain tumor images considerably. Applications/Improvements: Improved segmentation performance is required for accurate detection of brain tumor. The integrated hybrid segmentation approach has better segmentation accuracy than other segmentation approaches with 6% improvement which is very significant.

Keywords: Brain Tumor Segmentation, Local Independent Projections, Integrated Hybrid Segmentation Spatial Contextual Information

1. Introduction

The main objective of image segmentation is to simplify and change the complex image into meaningful and easier format to analyze. In the recent past many methods were presented in which various features are utilized for segmenting the tumor images efficiently in their own way. The LIPC method uses the concept of segmenting the images into smaller segments in order to construct new dictionaries that are utilized for testing data. Though the method seems to be efficient it takes more time to perform and hence it has only limited usage. The enhanced fuzzy C-means method is used to obtain high resolution images at a quick rate but the system does not provide the clear images as the boundaries are not intensified. Thus the system lags behind the efficient metrics. The most better method used currently is the hybrid clustering method based on the integrated approach of Fuzzy C-means and the K-means. This method provides high accuracy and low computation time. Many technologies other than the above methods are presented in All these techniques suffer from either possible problem including Tumor mass effect.

In suggested a new technique to solve the problems in the edge detection of brain tumor images. The contour method is presented to solve the brain
tumor segmentation based on the gradient magnitude information. In this method, the Gradient magnitude data is generated from the brain slice image intensity. Contour map of the brain tumor is created from this gradient information which is normalized to produce edge profiles of the brain tumor region contours. Using this data the brain tumor images are segmented into smaller regions which are eliminated to obtain clear view of lesions. In⁢³ proposed an algorithm based on the edge detection method to segment the brain images. An algorithm based on the symmetry nature of the MRI brain images is prescribed for the segmentation process. This method reduces the over segmentation problem by creating boundaries with markers around the tumor regions. Gaussian filter is used to smooth the images to make them free from noises and other irregularities before marking the segmentation. This method provides much beneficial prototype of the segmentation.

In⁢⁴ presented an algorithm for segmentation based on the watershed and thresholding approaches to detect the brain tumor in MRI brain images. The images that are received from the MRI systems are preprocessed to enhance and sharpen the edges of the images. The noise components are removed from the original images to provide efficient segmentation. Then the images are segmented and the tumor regions are enhanced with outer boundaries by using watershed marking and the thresholding. It provides better segmentation than the other edge technologies. In⁢⁵ also a similar work is presented. In their work an enhanced morphological operator is additionally used to enhance the image segmentation especially the tumor region.

In order to overcome the problems of using edge detection in MRI brain image, in⁢⁶ suggested the graph based approach to detect and segment the tumor areas. In this method the image is smoothened and the noises are removed after which the images are made into elements of the sparse graph. The method uses two layered graph containing low-grade characterization in one layer and voxel-based decisions of tumor presence in the next layer. This approach enhances the optimal segmentation. In⁢⁷ also a similar approach is used but in the form of hybrid algorithm combining the features of genetic algorithm and artificial neural network fuzzy inference system to provide an efficient segmentation of tumor images. In⁢⁸ presented one of the efficient methods known as local independent projection-based classification method to segment the tumor images. This method uses the approach of classifying the segments into smaller elements to make tumor libraries. The MRI tumor images are projected on these dictionaries to obtain sharpened segmentation.

Instead of LIPC, in⁢⁹ a method is suggested using Enhanced Possibilistic Fuzzy C-Means and Gradient vector flow snake model. It provides accurate and robust results than other methods but it is unable to tackle the low operation speed. Different approaches used in the current medical technologies are listed in¹⁰. The author explains about all the techniques used and also predicts the inability to tackle the tumor mass effect which affects the healthier tissues to resemble as tumor cells. In¹¹ the author came out with a solution to the slow processing by introducing the hybrid clustering technique that includes features of Fuzzy C-means and K-means together. This approach provided with accurate as well as fast processing in segmentation of the MRI brain tumor images which is also obtained in¹².

In¹³ a novel extraction technique is presented for the segmentation of medical images that are considered abnormal for normal processing. The segmentation technique reduces the noise level along with the feature enhancement so that the segmentation is more efficient. In¹⁴ the author presented an Efficient Sequential Pattern Matching Algorithm for processing the classified brain images.

In this paper we propose a highly integrated approach combining the features of the current techniques which are individually insane. The remainder of the paper is organized as follows: Section 2 describes the methodologies proposed in this research. The numerical results and analysis are presented in the section 3. Section 4 represents the conclusion and future work.

2. Methods and Materials

2.1 Brain Tumor Segmentation

In this section the techniques that can be used for MRI brain tumor segmentation are discussed. The efficient techniques LIPC and hybrid of PAM and enhanced possibilistic Fuzzy C-means methods are presented and discussed to develop a more efficient integrated hybrid segmentation approach.
2.1.1 Local Independent Projection-based Classification

In this method, tumor image segmentation is a multiclass problem. This can be solved by assuming the sample data can be represented as linear combinational inputs in any order. Samples from different classes are represented as a linear combination of several nearest neighbors from its corresponding sub-manifold. This procedure helps in achieving better segmentation. The images to be classified are taken as N-class classification with manifolds as $M_i$ and the dictionary as $D$. Thus $N_i$ is the N-class of $i$th sub-manifold, $M_i$ is the manifold of $i$th and $D_i$ is the dictionary corresponding to the $i$th sub-manifold.

The method uses local anchor embedding approach to project the $N_i$ samples into the $D_i$ dictionary. This approach is very unique in the sense that it responsible projecting the samples in the exact class $C$ of the dictionary $D$. The method is based on three main processes namely constructing the dictionary, representing in locally linear combination and classification score computing.

The first process is constructing a dictionary $D$ with all classes of the sampling segments. The samples of training data are used in the construction of a compact $D$. but the samples of a few patients may be enough to build a complex $D$. thus it becomes very difficult to select training samples in a traditional way. Hence a dictionary learning method like the K-means is implemented to obtain the training structures of the original sample. This method simplifies the search process by providing a compact representation of the training samples.

In the next process, the methods like sparse coding are used to represent the sample linearly based on training samples $x_i$={x1, x2, x3… xn}. Sparse coding enables the representation of the samples in a dictionary with minimal error. But some other methods like Locality linear coding and Local anchor embedding approaches limits the sparsity and focuses on the locality alone. The sample training images are represented in a linear fashion with the $k$ nearest neighbors of the testing sample $x$ was selected from $D$ and $N(x)$ ($k$) was constructed. While the data $d_j$ not belonging to $N(x)$ ($k$) are set to 0 and $djs$ belongs to $N(x)$ ($k$) are calculated using the projected gradient method.

The reconstruction error is calculated to analyze the failure in data storage. In the final process the classification scores of $x$ are represented within a multi-resolution framework. The classification scores are calculated by the probability formula. Let $y=\{y_i^T\}_{i=1}^{N}$, then

$$y_i = \frac{\exp\left(\langle w_i^T, e\rangle\right)}{\sum_{k=1}^{N} \exp\left(\langle w_i^T, e\rangle\right)}$$

Where $e = ||e||_2$

Here $T$ is the training samples.

2.1.2 Hybridization of PAM and Enhanced Possibilistic Fuzzy C-Means

In\textsuperscript{11} for the segmentation of brain tumor images a hybrid clustering technique has been proposed by integrating the k-means and fuzzy c-means approaches. The hybrid technique by using the k-means clustering reduces the computation time while the fuzzy c-means improves the accuracy of segmentation. Initially the cluster centers are calculated and the image points are assigned to the nearest cluster center based on minimum distance by checking the distance between the points and the cluster centers. Then the points are clustered and the new cluster centers are computed until the convergence criterion. Then the new cluster centers, the clustered points, and the scattered points can be assigned to a loop to calculate the new distances and the points are clustered due to membership value and mean value.

Though the hybrid clustering using K-means provides faster processing, the accuracy of segmentation is low which is compensated by the use of fuzzy c-means. This problem of low accuracy can be overcome by replacing the k-means with the Partitioning around the Medoids (PAM) approach. PAM provides high accuracy as the boundaries are segmented with fine edges. In PAM, the medoids are found in each cluster and the points are clustered towards the most similar medoids. Thus the hybrid of PAM and fuzzy c-means can considerably improve the accuracy of the segmentation.

Still the segmentation accuracy of the fuzzy c-means of hybrid approach in the imprecise edges is poor. To improve the segmentation in imprecise edges of the image, fuzzy c-means is replaced with Enhanced Possibilistic Fuzzy c-means (EPFCM). In EPFCM, the distance metrics are modified by including the membership, typicality, local and non-local spatial information to reduce the noise effects in the imprecise edges of the MRI brain images. Now the approach is called as Hybrid PAM and EPFCM clustering approach. In this method, the distance
metrics $D_{ij}$ is modified in order to incorporate into objective function. The membership, typicality and cluster centers are needed to be calculated to cluster the image points.

Membership

$$u_{ij} = \left[ \sum_{m=1}^{n} \left( \frac{D_{m}}{D_{ij}} \right)^{\frac{2}{m-1}} \right]^{-1}$$

Typicality

$$t_{ij} = \frac{1}{1 + \left( \frac{b}{D_{ij}} \right)^{\frac{1}{m}} - 1}$$

Where $m$, $n$, $b > 0$ are scalars.

The cluster centers or centroid are calculated as

$$v_{i} = \frac{(1: p) * M}{p + 1}$$

Where $p$ is number of clusters and $M$ is defined as

$$M = \max \text{(MRI image)} + 1$$

To compute the resultant metrics

$$y_{ij} = \frac{K \sum_{j=1}^{n} \mu_{ij}^{m} D_{ij}^{k}}{\sum_{j=1}^{n} \mu_{ij}^{m}}, \quad k > 1$$

The non-local distance metric $D_{nl}$ is calculated to reduce noise effects.

$$d_{nl}^{2}(x_{i}, v_{i}) = \sum_{i=1}^{m} w_{ij}(x_{i}, x_{j}) d^{2}(x_{i}, v_{j})$$

The trade-off parameter is a weighted factor which can be calculated as

$$\lambda_{ij} = \frac{1}{m} \sum_{i=1}^{m} 5(x_{i}, x_{j})$$

Thus efficient segmentation is achieved using hybrid PAM and EPFCM.

### 2.2 Integrated Hybrid Segmentation

In this section, integrated hybrid segmentation approach is developed by integrating the beneficial features of the LIPC and hybrid PAM and EPFCM to improve the segmentation of the MRI brain tumor images. The approach of integrating these techniques provides flexible performance in the segmentation of tumor images.

In the pre-processing phase the ever sensitive brain images are made free from the noise and the image quality are needed to be enhanced. The de-noising process is carried out to remove the noises like Gaussian and Poisson noises. The median filters are used to remove the noises. In general, the image background makes the processing time longer and it does not add to quality. Hence a skull removal technique is used to remove such irregularities which don't improve the quality of the image. Then the images are split into smaller segments to build the dictionary $D$. The features of LIPC method is utilized to provide the training samples dictionary. The training set is partitioned and the cluster centers are calculated using hybrid PAM and EPFCM. The $N$-class images are partitioned around the medoids in order to enhance the tumor region and provide accurate segmentation using EPFCM. The method as said built the dictionaries of different class. The LAE method is used to project the images into the dictionaries $D$ and then reconstruct the training samples. The training samples are extracted from the original images as the edges are fine tuned by the projection based classification. The sample images $X_{i}$ are taken from different patients to assign in the order of the different classes to be analyzed. The images which are segmented by this process are used to test the samples from the current set. The reconstruction error and classification score are calculated to label the segmented images. The concept of hybrid PAM and EPFCM...
has higher utility as it can be used to segment brain tumor images accurately irrespective of the differences in imaging ages so that the detection of tumor can be efficient. The segmented image is labeled as x for diagnosis.

In the continuous processing the image components required to detect tumor are extracted from the original image. Then the images are contoured to a certain level to maximize quality of the segmented images. Then methods like thresholding can be used to extract the tumor segments and detect the presence of tumor.

**Algorithm:** Integrated Hybrid Segmentation

Input: MRI Brain tumor images

Training set \( T = \{T_i\}_{i=1}^N \)

Output: Classification score \( y \) and label \( l \) of \( x \)

1. Initialize set of training samples \( T = \{T_i\}_{i=1}^N \)
2. //Hybrid PAM and EPFCM
3. Partition \( T_i \) into \( N_i \) subsets
4. Compute cluster center \( v_i \) using (4)
5. Calculate membership \( \mu_{ij} \) using (2)
6. Construct sub dictionaries for each class
7. Reconstruct each sample of \( T \) based on dictionary \( D \) using LAE.
8. Calculate locally linear representation coefficients
9. coefficient vectors \( \{a_i\}_{i=1}^N \) for soft regression model
10. Calculate reconstruct error \( e_i \) for all sample training sets

\[
x = D_{a_i} + e_i = \sum_{j=1}^{N} a_i^j d_i^j + e_i
\]

//Where \( \|e_i\| < \tau; \tau \) is the small positive real number for assuring reconstruction accuracy; \( a_i \) is the weight coefficient vector of linear combination
11. Again reconstruct each input sample \( x \) based on \( D \) using LAE
12. Estimate the classification score \( y \) using (1) and \( e_i \) using (11) for \( x \)
13. Achieve the final label \( l \) of \( x \) using (12).

\[
l = \arg\max_{i \in 1,\ldots,N} f_i(x) = \arg\max_{i \in 1,\ldots,N} y_i
\]

//where \( f_i(x) \) is the \( N \) classifiers of sample \( x \) and \( y_i \) is the classification score.

### 3. Results and Discussions

In this section the LIPC, Hybrid PAM and EPFCM, and integrated hybrid segmentation schemes are analyzed by the experimental conclusions. The methods are compared by the parameters such as dice similarity (DS), Jaccard Similarity (JS) and the segmentation accuracy.

#### 3.1 Dice Similarity (DS)

The DS is a statistical similarity index used to compare similarities between the datasets. It can be evaluated by using the following formula. Let’s consider an image of size \( m \times n \). \( P \) is the voxel set of result while \( Q \) is voxel set of ground truth. Then the similarity values can be calculated as,

\[
DS = \frac{2|P \cap Q|}{|P| + |Q|}
\]

Figure 1 shows the comparison of the LIPC, Hybrid PAM and EPFCM, and integrated hybrid segmentation schemes based on the dice similarity parameter. The dice similarity shows the similarity between the datasets used in the methods. As usual, the methods are taken along the x-axis and the dice similarity along y-axis from 0 to 1. The graph shows that the integrated hybrid segmentation has better dice similarity than the LIPC, and Hybrid PAM and EPFCM.

![Figure 1. Dice Similarity.](image)

#### 3.2 Jaccard Similarity (JS)

The JS is a statistical similarity index used to compare similarities and diversities between the datasets. It can be evaluated by using the following formula.

\[
JS = \frac{|P \cap Q|}{|P \cup Q|}
\]
Figure 2 shows the comparison of the LIPC, Hybrid PAM and EPFCM, and integrated hybrid segmentation schemes based on the jaccard similarity parameter. The jaccard similarity shows the similarities and diversities between the datasets used in the methods. As usual, the methods are taken along the x-axis and the jaccard similarity along y-axis from 0 to 100. The graph shows that the Integrated Hybrid segmentation has better Jaccard similarity than LIPC, and Hybrid PAM and EPFCM thus proving that it has better efficiency.

Figure 2. Jaccard Similarity.

### 3.3 Accuracy Rate

Accuracy is defined as the overall accuracy rate or segmentation accuracy which is calculated by the true and false values. In this method the quality of segmentation is considered. The efficiency of the overall system is assigned with the higher rate of accuracy associated with the segmented image.

Figure 3 shows the comparison of the LIPC, Hybrid PAM and EPFCM, and integrated hybrid segmentation schemes based on the segmentation accuracy parameter. The accuracy is calculated by the true values and shows the quality of the segmentation of tumor images in the different methods. As usual, the methods are taken along the x-axis and the segmentation accuracy along y-axis from 0 to 100. The graph shows that the integrated hybrid segmentation has better accuracy of tumor image segmentation compared to the LIPC, and Hybrid PAM and EPFCM. This proves that the proposed integrated hybrid segmentation method provides with higher efficient tumor segmentation.

Table 1. Comparison of existing and proposed methods

| Comparison                  | LIPC based Segmentation | Hybrid PAM and EPFCM | Integrated Hybrid Segmentation |
|-----------------------------|-------------------------|----------------------|--------------------------------|
| Dice Similarity             | 0.8500                  | 0.8900               | 0.9200                         |
| Jaccard Similarity          | 83.2000                 | 87.8900              | 93.1000                        |
| Accuracy                    | 84                      | 88                   | 94                             |

Table 1 shows the overall comparison of the LIPC, Hybrid PAM and EPFCM, and integrated hybrid segmentation schemes in terms of the dice similarity, jaccard similarity and the accuracy level. The experimental values of the parameters indicate that the proposed scheme is the efficient of the all methods discussed in this research.

### 4. Conclusion

The proposed system introduces a new approach called integrated hybrid segmentation for detecting the tumor regions in the MRI brain tumor images. The method beneficially avoids the irregularities by removing the noises, complex image background and other irrelevant features thus improving the tumor segmentation process. It also provides a faster processing approach compared to the existing methods. Above all, the proposed method can be used to reduce the effects of tumor mass effect and provides highly accurate tumor segmentation in the future. By utilizing the proposed scheme intensively in the medical analysis of brain tumor, the diagnosis process can be improved.
5. References

1. Priyanka BS. A review on brain tumor detection using segmentation. International Journal of Computer Science and Mobile Computing. 2011; 2(7):48–54.
2. Bandyopadhyay SK. Edge detection in brain images. International Journal of Computer Science and Information Technologies. 2011; 2(2):884–7.
3. Agrawal SS, Gupta SR. Detection of brain tumor using different edge detection algorithm. International Journal of Emerging Research in Management and Technology. 2014; 2(2):85–9.
4. Mustaqeem A, Javed A, Fatima T. An efficient brain tumor detection algorithm using watershed and thresholding based segmentation. International Journal of Image, Graphics and Signal Processing. 2012; 4(10):34–9.
5. Selkar RG, Thakare MN. Brain tumor detection and segmentation by using thresholding and watershed algorithm. International Journal of Advanced Information and Communication Technology. 2014; 1(3):321–4.
6. Parisot S, Duffau H, Chemouny S, Paragios N. Graph-based detection, segmentation and characterization of brain tumors. IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2012. p. 988–95.
7. Sharma M, Mukharjee S. Brain tumor segmentation using hybrid genetic algorithm and Artificial Neural Network Fuzzy Inference System (ANFIS). International Journal of Fuzzy Logic Systems (IJIFLS). 2012; 2(4):31–42.
8. Huang M, Yang W, Wu Y, Jiang J, Chen W, Feng Q. Brain tumor segmentation based on local independent projection-based classification. Proceedings of IEEE Transactions on Biomedical Engineering. 2014 Oct; 61(10):2633–45.
9. Rajendran A, Dhanasekaran R. Brain tumor segmentation on MRI brain images with fuzzy clustering and GVF snake model. International Journal of Computer Communication. 2012; 7(3):530–9.
10. Liu J, Li M, Wang J, Wu F, Liu T, Pan Y. A survey of MRI-based brain tumor segmentation methods. Tsinghua Science and Technology. 2014; 19(6):578–95.
11. Abdel-Maksoud E, Elmogy M, Al-Awadi R. Brain tumor segmentation based on a hybrid clustering technique. Egyptian Informatics Journal. 2015; 16(3):71–81.
12. El-Melegy MT, Mokhtar HM. Tumor segmentation in brain MRI using a fuzzy approach with class center priors. EURASIP Journal on Image and Video Processing. 2014; 2(1):21.
13. Bharathi K, Karthikeyan S. A novel implementation of image segmentation for extracting abnormal images in medical image applications. Indian Journal of Science and Technology. 2015; 8(8):333–40.
14. Harini R, Chandrasekar C. Efficient sequential pattern matching algorithm for classified brain image. Indian Journal of Science and Technology. 2015; 8(14).