Short Notes on Unsupervised Learning Method with Clustering Approach for Tumor Identification and Tissue Segmentation in Magnetic Resonance Brain Images

Vishnuvarthan Govindaraj*, Anitha Vishnuvarthan, Arunprasath Thiagarajan, Kannan M, and Pallikkonda Rajasekaran Murugan
Kalasalingam University, Virudhunagar, Tamilnadu, India

Abstract

Malignant and benign types of tumor, infiltrated in human brain are diagnosed with the help of an MRI scanner. Using the slice images obtained using an MRI scanner; certain image processing techniques are implemented to have a clear anatomy of brain tissues. One such image processing technique is hybrid Self Organizing Map (SOM) with Fuzzy K Means (FKM) algorithm, which offers a possible identification of tumor region penetration in the tissues of brain. The proposed algorithm is efficient in terms of Jaccard Index, Dice Overlap Index (DOI), Sensitivity, Specificity, Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE) and Computational time and memory requirement for processing the Magnetic Resonance (MR) brain images. Automatic detection of tumor region in MR (Magnetic Resonance) brain images has a good impact in helping the radio surgeons to identify the exact topographical location of tumor region. In this paper, the proposed hybrid SOM – FKM algorithm supports the radio surgeon by providing tissue segmentation and an automated tumor identification.

Keywords: SOM (Self Organizing Map); FKM (Fuzzy K – Means algorithm); Tumor segmentation; Peak Signal to Noise Ratio (PSNR); Jaccard Index; Dice Overlap Index (DOI); Sensitivity; Specificity; Mean Square Error (MSE); Brain image segmentation and Computational time

Introduction

The effects created by tumor are extremely dangerous, leading to the formation of lesions and tumor within the human body. The most complicated organ of our human body is the brain and the cancer cells affecting the tissues of brain induce tumor formation [1]. Instead of manual segmentation of brain tissues and identification of tumor region, an automated algorithm capable of performing tumor identification and tissue segmentation processes is proposed through this paper.

MR (Magnetic Resonance) brain image sequences are classified as T1, T2, FLAIR (Fluid Attenuated Inversion Recovery), MPR (Multi Planar Reconstruction) and MRS (Magnetic Resonance Spectroscopy) images. T1, T2 and FLAIR image sequences are segmented by the proposed SOM – FKM algorithm.

Dr. N. Nandha Gopal [2] recommended an automatic detection of brain tumor, which is capable of producing a segmentation accuracy of 95.16%.

Govindaraj Vishnuvarthan and Murugan Pallikkonda Rajasekaran [3] recommended an automated tumor extraction and tissue segmentation using Fuzzy. It suffers in terms of high computational time and MSE along with low PSNR values.


Related Work


*Corresponding author: Govindaraj V, Kalasalingam University, Virudhunagar, Tamilnadu, INDIA, Tel: +91 9360654171; E-mail: gviswathan@gmail.com

Received December 05, 2015; Accepted January 04, 2016; Published January 07, 2016


Copyright: © 2016 Govindaraj V, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Scope of Our Work

The main aim of our methodology is to offer an entirely automated hybrid algorithm to segment the tumor region present in the brain images. An automated algorithm greatly assists the radio surgeon in the process of non-invasive diagnostic examination of the patients affected with tumor. Further, multimodal image sequences of an MRI scanner with all three axes namely, coronal, sagittal and axial are segmented with the support of the proposed algorithm. A radio surgeon can efficiently plan a pre and post radio therapeutic procedures using the hybrid SOM – FKM algorithm. Comparison parameters validate the performance of the algorithm and the results are compared with Fuzzy C Means (FCM) algorithm. Since, FCM and FKM are the two competitive clustering algorithms, we choose FCM as an algorithm to compare with SOM combined FKM with strategy.

Materials and Methods

A simple flowchart explaining the sequential steps of the proposed method is given below for easy understanding (Figure 1).

Fuzzy C Means Algorithm

Fuzzy C - means (FCM) algorithm is a clustering methodology introduced by Dunn, improved by Bezdek and further spruced by Matteo Matteucci; which groups the pixels (data) of the MR brain images as ‘n’ number of clusters. The neighbouring pixel of least mean distance from the centroid pixel are defined with low membership value 0, 0.1, 0.2, 0.4….1), such that,

Here, 

Degree of membership is given as,

During the preliminary processing of the algorithm, \( \hat{\delta}_j = \overline{\delta}_j \overline{\theta}_j \). Here, \( \overline{\theta}_j \) is a random value initiated as, 0 ≤ \( \overline{\theta}_j \) ≤ 1 (Usually expressed as 0, 0.1, 0.2, 0.4…1), such that, \( \sum_j \overline{\theta}_j \overline{\delta}_j = 1 \). In general, FCM algorithm has higher data supervision capacity and has better processing upon diversified data range. On the other end, the convergence rate gets affected if the number of clusters and iterations are amplified. Reducing the number of iterations and clusters to obtain a better convergence rate has a negative impact upon the segmentation accuracy. To overcome these hindrances, a novel segmentation algorithm is proposed through this paper.

Fuzzy K Means algorithm

Fuzzy K means algorithm, also called as FKM, acts an active contestant to FCM algorithm. FKM partitions the given input image into ‘k’ clusters (partitioning the pixels present in the image).

\[
J_k = \sum_{j=1}^{K} \sum_{i=1}^{N} u_{ij}^m d_{ij}
\]

\( m = \) fuzziness co-efficient obtained from the overlapping of clusters in a range of 1.5 ≤ m ≤ 2.5 for optimum segmentation results [15]. Centre vector (Centroid voxel) for the cluster is defined as,

\[
C_j = \frac{\sum_i \delta_{ij} x_i}{\sum_i \delta_{ij}}
\]

During the preliminary processing of the algorithm, \( \hat{\delta}_j = \overline{\delta}_j \overline{\theta}_j \). Here, \( \overline{\theta}_j \) is a random value initiated as, 0 ≤ \( \overline{\theta}_j \) ≤ 1 (Usually expressed as 0, 0.1, 0.2, 0.4…1), such that, \( \sum_j \overline{\theta}_j \overline{\delta}_j = 1 \).

Self-Organizing Map algorithm

The first step of processing a SOM algorithm commences with

Volume 1 • Issue 1 • 1000101

Figure 1: Flowchart of hybrid SOM - FKM algorithm.
initialization, where random variables are chosen for the initial weight vectors $W_j$ derived from the input image. Second step involves sampling process by which an input vector $X$ is drawn from the input image. In third step, finding the winning neuron $I(x)$ with weight vector closest to the input vector is done (Calculation of Euclidean distance). Each point $I$ in the output image will map to a corresponding point $w(I)$ in the input image. In fourth step, the weight equation is updated.

$$w_i = \eta(t)\sum_{j=1}^{n} (I_{i,j} - w_j)$$

(5)

Here, Topographical neighborhood is defined as $\sum_{j=1}^{n} (I_{i,j} - w_j)$. $\eta(t)$ is the dependent learning rate based upon time $t$ for the input vector $X$.

Lastly, the drawing of vector $X$ from the input image is iteratively repeated until the feature map stops changing. SOM itself is an unsupervised neural network based clustering technique, which can be directly used in segmentation process [10,17].

**Proposed Methodology**

The combination of SOM and FKM is the main novelty proposed in this paper. SOM is used for initial clustering and reduction of dimensionality of the input image. The image given as input to the algorithm is of 256x256 dimension, which is reduced with the aid of SOM algorithm. Updated weight function $W_j$ [18] is the resultant value obtained using SOM. The weight function $W_j$ is given as an input to FKM algorithm. Thus the problem aroused due to FKM in handling wide and large data range is minimized with the inclusion of SOM. Next, FKM utilizes the output from SOM and performs the clustering operation [16]. Clinical and live images of patient suffering from SOM. Next, FKM utilizes the output from SOM and performs the handling wide and large data range is minimized with the inclusion of SOM. In fourth step, the weight equation is updated.

**Comparison Parameters**

**MSE (Mean Square Error)**

Mean Square Error defines the process of squaring the differentiated quantities [19]. It is expressed in Equation (6). MSE is the cumulative squared error value between the input image $I(i,j)$ and the segmented image $K(i,j)$.

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i,j) - K(i,j))^2$$

(6)

$m$ & $n$ denote the number of rows and columns present in an input image.

$i = m$, to perform incremental operation for ‘rows’ in a loop statement.

$j = n$, to perform incremental operation for ‘columns’ in a loop statement.

MSE value of the resulting segmented image should be minimized. Smaller MSE values denote minimized error occurrence in the segmented MR brain images.

**PSNR (Peak Signal to Noise Ratio)**

Peak Signal to Noise Ratio refers to the immunity of an image to noise signals [20]. If the PSNR value is high, then the impact of noise interference signal upon the MR brain image is quite low. The algorithm producing PSNR values ranging between 40 to 100 dB is considered to be less sensitive to noise signals. PSNR is defined as:

$$PSNR = 10 \log_{10} \left( \frac{MAX^2}{MSE} \right) = 20 \log_{10} \left( \frac{MAX}{\sqrt{MSE}} \right) = 20 \log_{10} \left( \frac{MAX}{MSE} \right) - 10 \log_{10} MSE$$

(7)

**Memory Requirement**

The memory space occupied by the algorithm to perform the entire segmentation process is expressed in terms of memory required. Usually, the memory requirement for a segmentation process should be as low as possible and is expressed as bytes.

**Jaccard (Tanimoto) Index**

Jaccard index is denoted as the ratio of the common voxels between the input image (A) and the segmented output image (B) to the union function of the input image (A) and the segmented output image (B) [16].

$$J(A,B) = \frac{SA \cap SB}{SA \cup SB}$$

(8)

**Dice Overlap Index (DOI)**

DOI value is expressed using Jaccard index $J(A, B)$. DOI specifies the overlapping function of the input image (A) and the segmented output image (B) [13].

$$D(A,B) = 2 \times \frac{J(A,B)}{1+J(A,B)}$$

(9)

**Similarity Index (SI)**

Similarity Index describes the similarity between the input image (A) and the segmented output image (B) [13,21]. The values used to calculate SI are:

True Positive (TP): Correct classification of tissue region and identification of tumor region.

False Positive (FP): Normal tissue region present in the input image identified as tumor region.

False Negative (FN): Tumor region undetected or misclassified.

$$SI = \frac{2 \times TP}{2 \times TP + FP + FN}$$

(10)

**Overlap Fraction (OF)**

Overlap Fraction (OF) value refers to the correct segmentation of input image. Specifically, OF defines the rate of tumor region and other tissue regions being correctly identified [13,21].

$$OF = \frac{TP}{TP+FN}$$

(11)

**Extra Fraction (EF)**

Extra Fraction (EF) refers to the number of pixels being falsely detected as tumor region and tissue regions misclassified. EF should be as low as possible [13,21].

$$EF = \frac{FP}{TP+FN}$$

(12)

**Sensitivity**

The term sensitivity refers to the ability of an algorithm to classify or segment the tissues of non-tumor region [13,21].

$$Sensitivity = \frac{TN}{TN+FP}$$

Here TN represents True Negative value. It describes the effective
Segmentation of non-tumor region or normal brain tissues by the algorithm.

**Segmentation Results from Hybrid SOM – FKM algorithm**

Figure 2 explicits the segmentation process of MR brain images of the patient suffering from Primitive Neuro Ectodermal Tumor, with an age of 5 years. In T2 – Coronal (d), T2 – Axial contrast enhanced and T2 – Axial (e) images, tumor and edema covered region are identified clearly. In T2 – Axial (a) and T2 – Axial (b), unique identification of tumor region is done. The average time duration required for processing the seven images of Figure 2 by SOM – FKM algorithm is 48.64 seconds and the average memory requirement is 5.31E+14 bytes.

Figure 3 demonstrates the segmentation ability of the proposed algorithm in identifying the tumor region in a challenging clinical case of age 37 years. Both T1 – coronal and T1 – Sagittal have meningioma, which is clearly identified using SOM – FKM algorithm. The challenging tumor regions and their identification are denoted by a circle region in the above images. The average time duration required for processing the two images of Fig 3 by SOM – FKM algorithm is 22.49 seconds and the average memory requirement is 4.66E+14 bytes.

Figure 4 shows the detection of Tumor and Edema regions in T1 – Axial (a), T2 – Axial (f) and T2 – Axial (h) images of Figure 4 are clearly detected with the aid of the proposed methodology. Other images of Figure 4 also show the segmentation ability of SOM – FKM algorithm. During diagnosis, the age of the patient was found to be 42 years. The average time duration required for processing the ten images of Figure 4 by SOM – FKM algorithm is 19.43 seconds and the average memory requirement is 2.70E+14 bytes.

Figure 5 has been completely chosen to validate the segmentation efficiency of the hybrid SOM – FKM algorithm. Images of the patients suffering from different tumor diseases were taken from online Harvard brain repository for validating the segmentation algorithm. The average time duration required for processing the three images of Fig 5 by SOM – FKM algorithm is 11.19 seconds and the average memory requirement is 3.48E+14 bytes.
Images T1 – Axial (a), T1 – Axial (b) and T1 – Axial (d) of Figure 6, illustrate the perfect segmentation of brain tissues and clear identification of tumor region by SOM – FKM algorithm. The average time duration required for processing the four images of Figure 6 by SOM – FKM algorithm is 2.9343 seconds and the average memory requirement is 4.51E+13 bytes.

Images T1 – Axial (a), T1 – Axial (b) and T1 – Axial (c) of Figure 7, show the capability of the proposed algorithm in segmenting the tissue regions and exact identification of tumor region. This helps the radio surgeon to plan pre-surgical procedures. The average time duration required for processing the three images of Figure 7 by SOM – FKM algorithm is 4.2267 seconds and the average memory requirement is 4.23E+13 bytes.

Figure 4: Segmentation of tissues and identification of tumor region for the patient suffering High Grade Astrocytoma.

Figure 5: Segmentation of tissues and identification of tumor region for the Images available in Harvard database taken for validation.
Figure 6: Segmentation of tissues and identification of tumor region for the patient suffering from Pituitary adenomas.

Figure 7: Segmentation of tissues and identification of tumor region for the patient suffering from Nerve Sheath Tumors.
In challenging cases of Figure 8, specifically in T1 – Axial (b) and T1 – Axial (d), tissue regions are poorly captured by a MRI scanner of 1.5 Tesla. These tissue regions are clearly viewed and the tumor region is identified with the assistance of the proposed methodology. The average time duration and memory requirement required for processing the five images of Fig 8 by SOM – FKM algorithm are 3.369 seconds and 4.44E+13 bytes.

Figure 9 depicts the segmentation results of all complex T1 axial images, where the tissue regions are clearly distinguished for the better viewing of a radio surgeon. Exact segmentation of tissue region in T1 – Axial (d), proves the effectiveness of the proposed SOM – FKM algorithm. The average time duration and memory requirement required for processing the six images of Figure 9 by SOM – FKM algorithm are 3.237 seconds and 3.56E+13 bytes.

Figure 10 represents the segmentation results of the proposed algorithm. Distorted or noise affected Skull and Cerebro Spinal Fluid regions of T2 – Axial (b) image are clearly identified using SOM – FKM algorithm. The average time duration and memory requirement required for processing the two images of Figure 10 by SOM – FKM algorithm are 3.447 seconds and 4.54E+13 bytes.

Discussion

Table 1 shows the efficiency of the proposed algorithm in terms of MSE, PSNR, Jaccard Tanimoto Coefficient Index (TC) and DOI. The Hybrid SOM – FKM algorithm produces an average MSE value of 0.388095 on comparison with FCM algorithm, which produces an average MSE value of 2.067619. SOM – FKM outperforms FCM in terms of PSNR values. The average PSNR values produced by SOM –
FKM and FCM algorithms are 58.74786 and 45.34333. The proposed algorithm also ranks higher in producing better DOI and TC values. The average DOI and TC values produced by SOM – FKM algorithm are 0.435105 and 0.282381. Whereas, FCM algorithm produces 0.412143 and 0.26881 as DOI and TC values.

Figure 11 shows the computational time required for processing the input images by both FCM and SOM – FKM algorithms. SOM – FKM algorithm has greater efficiency in terms of computational time. The average time required for FCM algorithm to segment 42 input MR brain images is 438.1914 seconds, comparatively, SOM – FKM algorithm requires 16.21499 seconds as an average processing time.

Figure 12 depicts the memory required for the algorithms to process the input images. FCM and the proposed SOM – FKM algorithms require an average memory space of 3.95E+14 MB and 2.08E+14 MB, respectively. Both the algorithms were processed on a 32 Bit operating system with Pentium 4 processor (2.2 GHz) and 2 GB RAM. Lower the memory space better is the segmentation efficiency of the algorithm.

Table 1: Comparison of FCM and proposed SOM – FKM Algorithm.
Table 2 shows a sample of manually selected segmentation (Gold Standard) done by a clinical expert using Region of Interest (ROI) algorithm. The images of table 2 can be compared with T2 Axial (a) of Figure 2. Using ROI, additional tissue region surrounding the tumor part has been detected as tumor, which comes under the account of false positively detected tumor region. The same set back is also observed while manually segmenting T2 – Axial (b), T2 – Coronal (d), T2 – Axial (Contrast Enhanced) and T2 – Axial (e) images of Figure 2.

Table 3 describes the efficiency of the proposed SOM – FKM algorithm using SI, OF, EF and Sensitivity values. The proposed algorithm overrides FCM and expert segmentation methodology by offering better SI, OF and sensitivity values. EF value is considerably reduced using SOM – FKM algorithm. For a good segmentation algorithm SI and OF value should be close to 1 and EF value should be nearer to 0 [21]. All the above said values in table 2 make a concrete statement that the proposed hybrid SOM – FKM algorithm is superior to FCM algorithm and perform better than the expert segmentation method.
Conclusion

The proposed hybrid SOM – FKM algorithm excels FCM algorithm in the process of segmenting MR brain images. The effectiveness of the proposed algorithm is demonstrated using Tables 1, 2 & 3. Varied clinical datasets along with standard database images have been used to prove the efficacy of the proposed algorithm. A novel automated technique to segment the tissues of brain, especially GM, WM and CSF regions, along with identification of different types of tumor at different locations has been proposed through this paper. The ultimate goal of this research work is to assist the clinician or radio surgeon in diagnosing the patients with minimized time duration and is greatly achieved through the combination of SOM and FKM techniques. Future work lies with improving the values of DOI, TC and PSNR values. Time duration for processing the images can also be further reduced. The proposed algorithm can be extensively used in pre and post radio surgical applications.

Acknowledgments

The authors thank Dr. K.G. Srinivasan, MD, RD, Consultant Radiologist and Dr. K.P. Usha Nandhini, DNB, KGS Advanced MR & CT Scan - Madurai, Tamilnadu, India, for supporting the research with the patient information and MR brain images. Also, the authors thank the Department of Instrumentation and Control Engineering of Kalasalingam University, (Kalasalingam Academy of Research and Education), Tamilnadu, India for permitting to use the computational facilities available in biomedical laboratory which was set up with the support of the Department of Science and Technology (DST), New Delhi under FIST program.

References


