Applied Soft Computing xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A study on fuzzy clustering for magnetic resonance brain image segmentation using soft computing approaches

Sanjay Agrawal [a], Rutuparna Panda [a,*], Lingraj Dora [b]

[a] Department of Electronics & Telecommunication Engineering, VSS University of Technology, Burla 768018, India
[b] Department of Electrical &Electronics Engineering, VSS University of Technology, Burla 768018, India

ARTICLE INFO

Article history:
Received 21 February 2013
Received in revised form 2 July 2014
Accepted 6 August 2014
Available online xxx

Keywords:
Fuzzy C-means (FCM) clustering
K-means clustering
Genetic algorithm (GA)
Particle swarm optimization (PSO)
Bacteria foraging optimization (BFO)

ABSTRACT

This paper presents a novel idea of intracranial segmentation of magnetic resonance (MR) brain image using pixel intensity values by optimum boundary point detection (OBPD) method. The newly proposed (OBPD) method consists of three steps. Firstly, the brain only portion is extracted from the whole MR brain image. The brain only portion mainly contains three regions – grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF). We need two boundary points to divide the brain pixels into three regions on the basis of their intensity. Secondly, the optimum boundary points are obtained using the newly proposed hybrid GA–BFO algorithm to compute final cluster centres of FCM method. For a comparison, other soft computing techniques GA, PSO and BFO are also used. Finally, FCM algorithm is executed only once to obtain the membership matrix. The brain image is then segmented using this final membership matrix. The key to our success is that we have proposed a technique where the final cluster centres for FCM are obtained using OBPD method. In addition, reformulated objective function for optimization is used. Initial values of boundary points are constrained to be in a range determined from the brain dataset. The boundary points violating imposed constraints are repaired. This method is validated by using simulated T1-weighted MR brain images from IBSR database with manual segmentation results. Further, we have used MR brain images from the Brainweb database with additional noise levels to validate the robustness of our proposed method. It is observed that our proposed method significantly improves segmentation results as compared to other methods.

Introduction

Image segmentation has been a very critical and important stage in any image processing application. It deals with dividing the pixels in an image into groups or regions having similar features or characteristics for effective object identification. The segmentation of magnetic resonance (MR) brain image has got significant focus in the field of biomedical image processing. Segmentation of MR brain image has got wide application in the field of bio-medical analysis, such as identification of tumours, classification of tissues and blood cells, multi modal registration [1] etc. There are various segmentation techniques proposed for MR brain image like thresholding [2], edge based detection [3] and region growing [4].

Thresholding techniques are effectively used when the histograms of the objects and background are clearly identifiable. But for brain image, these techniques give the inaccurate segmentation result as distribution of pixels in brain image is very complex. Edge based methods rely heavily on detection of boundaries in the image. It is observed in the brain image that grey level distribution of pixels of grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) result in incorrect detection of boundary. Region growing techniques use the homogeneity and connectivity criteria for segmentation. It is not effectively used for brain image segmentation as the brain image does not contain well defined regions. The above methods are found effective for relatively simple images.

So one of the efficient techniques used for complex brain image segmentation is clustering. It classifies the pixels into larger groups depending on certain criteria. Again, several types of clustering methods have been discussed in literature like Expectation–maximization [5], hard C-means, K-means and fuzzy clustering techniques [6]. Among fuzzy clustering techniques, fuzzy C-means (FCM) is the most widely used technique [7,8]. It aims at minimizing an objective function according to some criteria. It permits one data point to belong to more than one cluster defined by a membership matrix. But the random selection of centroids makes the technique fall into local optimum. To overcome this problem, soft computing approaches like genetic algorithm (GA) [9–11], Particle swarm optimization (PSO) [12], ant colony optimization (ACO) [13] etc. have been applied to improve FCM. Castillo et al. [14] presented optimization of the FCM algorithm by using evolutionary methods. They used GA and PSO only. They used it to find the optimal number of clusters and the weight exponent for different types of synthetic datasets. They emphasized on cluster validation. Hruschka et al. [15] presented a survey of evolutionary algorithms for clustering. They emphasized on partition algorithms that focused on hard clustering of data. The survey did not use any particular evolutionary method, but focused on advanced topics like multi-objective and ensemble based evolutionary clustering. Then a taxonomy that highlights on some very important aspects of evolutionary clustering was presented at the end.

* Corresponding author. Tel.: +91 6632431857; fax: +91 6632430204.
E-mail addresses: lingraj02uce157ster@gmail.com, agrawals_72@yahoo.com (S. Agrawal), r_ppanda@yahoo.co.in (R. Panda), lingraj02uce157ster@gmail.com (L. Dora).

http://dx.doi.org/10.1016/j.asoc.2014.08.011
1568-4946/© 2014 Published by Elsevier B.V.

Mukhopadhyay and Maulik [16] proposed a multiobjective real coded genetic fuzzy clustering scheme for the segmentation of multispectral MR images of the human brain. Their technique is able to determine the number of clusters along with clustering results. They emphasized on including the spatial information for improved segmentation result.

All the above mentioned approaches emphasize on selecting a random initial cluster centre for FCM. Then evolutionary computing techniques are used to obtain optimum cluster centroids. FCM is then iteratively applied to obtain a final membership matrix for segmentation. However, in this paper, a new strategy for intracranial (also coined as brain extraction) segmentation of MR brain image using hybridized fuzzy C-means clustering technique is proposed. Instead of randomly selecting centroids of clusters and optimizing them, we have used OBPD method. We first determine the number of boundary points from the dataset to divide the region into required number of clusters. These boundary points are optimized using a new hybrid GA–BFO technique. Other soft computing approaches GA, PSO and BFO are also used for a comparison. We have also used a classical method coined as K-means clustering for a comparison. Final centroids of the clusters are then computed. These final centroids are used to obtain the fuzzy membership matrix by executing FCM once only. To the best of our knowledge, hybrid GA–BFO has not been used so far for MR brain image segmentation. This has motivated us to use the proposed technique.

It has already been reported in the literature that a brain image mainly consists of three regions: grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) [9,17]. The grey level distribution is used to identify these three regions. For accurate identification, ideal clustering is needed.

Two optimum boundary points are obtained from the grey level distribution to divide the brain image into three regions or clusters. Initial values of the boundary points are constrained to be in a range determined from the brain dataset. The proposed study aims at optimizing these boundary points by using hybrid GA–BFO technique to select final cluster centres for FCM algorithm. The objective function used is reformulated in terms of cluster centres only. Using the final cluster centres, the proposed hybrid FCM algorithm is executed only once to obtain the fuzzy membership matrix. Segmentation is then done using this fuzzy membership matrix. Several standard brain images (simulated T1-weighted) from the IBSR database with manual segmentation results are considered in the experiment. The results obtained are compared using various performance parameters. The segmentation performance parameters are also calculated for different noise levels. For the experiment, we consider brain images from Brainweb database with additional noise levels: 1%, 3%, 5%, 7% and 9%. Results are presented in 'Results and discussions' section to validate the robustness of our proposed method.

The rest of the paper is organized as follows. 'FCM and soft computing methods' section presents a brief introduction about FCM technique and soft computing approaches i.e. GA, PSO, BFO and GA–BFO. 'Proposed methodology' section explains the proposed methodology. 'Results and discussions' section presents the results and discussions. The last section is the conclusion.

FCM and soft computing methods

Fuzzy C-means clustering (FCM) algorithm

Fuzzy clustering allows objects to belong to more than one cluster by specifying a membership matrix with different degree for each cluster. It is a local optimum search technique. In this algorithm, a set of n objects $x = \{x_1, x_2, \ldots, x_n\}$ each having d dimensions are divided into c number of clusters of similar features. The features could be the position or intensity of a pixel in an image. The fuzzy clusters of objects are characterized by a fuzzy membership matrix with n rows and c columns. The set of all constrained fuzzy matrices of size $n \times c$ is defined as [8]:

$$M_f = \left\{ \mu \in \Re^{n \times c} | \sum_{j=1}^{c} \mu_{ij} = 1, \quad 0 < \sum_{i=1}^{n} \mu_{ij} < n, \quad \mu_{ij} \in [0, 1] \right\} \tag{1}$$

for $1 \le i \le n$; $1 \le j \le c$.

The condition used to define good clusters for FCM is to minimize the FCM function [8]:

$$J_m(\mu, z) = \sum_{j=1}^{c} \sum_{i=1}^{n} (\mu_{ij})^m d_{ij}^2(z_j, x_i), \tag{2}$$

where μ is the fuzzy membership matrix, $1 \le m \le \infty$ is a scalar weighting exponent which controls the fuzziness. The larger is its value, fuzzier is the partition. An analysis on the weighting exponent is found in Ref. [18]. When the value of m is close to 1, FCM approaches hard c-means algorithm. When m approaches infinity, the mass centre of the data set is the only solution of FCM [18]. Here the value of m is set to 2. It is observed that this value of m is suitable for most MR brain images, as it yield best results [19]. Note that $z = [z_1, z_2, \ldots, z_c]$ is a matrix of cluster centres, and $d_{ij}(z_j, x_i)$ is a measure of Euclidean distance from x_i to jth cluster centre z_j. The algorithm used in this paper is presented below:

Table 1
Parameter setting for the different methods.

The parameter setting for FCM is:
• Scalar weighting exponent $m = 2$,
• Number of iterations = 20,
• Number of clusters = 3
The parameter setting for GA–FCM is:
• Number of iterations = 20,
• Population number = 20
• Crossover probability = 0.8,
• Mutation probability = 0.05
• Selection function is the Roulette wheel selection
The parameter setting for PSO–FCM is:
• Number of iterations = 20,
• Number of particles = 20,
• Acceleration coefficients $C_1 = C_2 = 2$
• Weight factor $w = 0.9$
The parameter setting for GA–BFO–FCM (proposed method) is:
• Number of bacteria = 20,
• Number of chemotactic steps = 4,
• Swimming length = 10,
• Number of reproduction steps = 4,
• Number of elimination and dispersal event = 2
• Probability of elimination and dispersal = 0.02
• Probability of crossover = 0.7
• Mutation probability = 0.01

Algorithm. *Step 1*: Generate brain portion only data set $x = \{x_1, x_2, \ldots, x_n\}$ of MR brain images.

Step 2: Set various parameters (like the scalar weighting exponent m) and the termination condition i.e. the maximum number of iterations.

Step 3: Select the number of clusters c.

Step 4: Get initial set of random cluster centres $z = [z_1, z_2, \ldots, z_c]$.

Step 5: Calculate Euclidean distance $d_{ij}(z_j, x_i)$ for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, c$.

Step 6: Calculate membership matrix μ_{ij} using Eq. (3) as:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (d_{ij}/d_{ik})^{2/m-1}} \quad \text{for} \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, c \tag{3}$$

Step 7: Update the cluster centres z_j using the membership matrix μ_{ij} by using Eq. (4) as:

$$z_j = \frac{\sum_{i=1}^{n} \mu_{ij}^m x_i}{\sum_{i=1}^{n} \mu_{ij}^m} \tag{4}$$

Step 8: If the termination condition is not met, go to step 5.

In this paper, the parameters for FCM are set as given in Table 1. The algorithm is implemented using MATLAB. The pixels of the brain only portion are clustered using the cluster centres z_j obtained after the termination condition is met. Segmentation of the brain image is done using the final membership matrix μ_{ij}.

Soft computing methods

It may be reiterated the fact that the brain portion mainly contains three regions WM, GM and CSF. The pixels in these regions

Please cite this article in press as: S. Agrawal, et al., A study on fuzzy clustering for magnetic resonance brain image segmentation using soft computing approaches, Appl. Soft Comput. J. (2014), http://dx.doi.org/10.1016/j.asoc.2014.08.011