Expert Systems With Applications 52 (2016) 50–63

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Brain image segmentation using semi-supervised clustering

Sriparna Saha, Abhay Kumar Alok*, Asif Ekbal

Department of Computer Science Engineering, Indian Institute of Technology, Patna 801103, Bihar, India

ARTICLE INFO

Keywords:
Brain image segmentation
Multiobjective optimization
Semi-supervised clustering
AMOSA
Cluster validity index
Sym-index
I-index
MS-index

ABSTRACT

The objective of brain image segmentation is to partition the brain images into different non-overlapping homogeneous regions representing the different anatomical structures. Magnetic resonance brain image segmentation has large number of applications in diagnosis of neurological disorders like Alzheimer diseases, Parkinson related syndrome etc. But automatically segmenting the MR brain image is not an easy task. To solve this problem, several unsupervised and supervised based classification techniques have been developed in the literature. But supervised classification techniques are more time consuming and cost-sensitive due to the requirement of sufficient labeled data. In contrast, unsupervised classification techniques work without using any prior information but it suffers from the local trap problems. So, to overcome the problems associated with unsupervised and supervised classification techniques, we have proposed a new semi-supervised clustering technique using the concepts of multiobjective optimization and applied this technique for automatic segmentation of MR brain images in the intensity space. Multiple centers are used to encode a cluster in the form of a string. The proposed clustering technique utilizes intensity values of the brain pixels as the features. Additionally it also assumes that the actual class label information of 10% points of a particular image data set is also known. Three cluster validity indices are utilized as the objective functions, which are simultaneously optimized using AMOSA, a modern multiobjective optimization technique based on the concepts of simulated annealing. First two cluster validity indices are symmetry distance based Sym-index and Euclidean distance based I-index, which are based on unsupervised properties. Last one is a supervised information based cluster validity index, Minkowski Index. The effectiveness of this proposed semi-supervised clustering technique is demonstrated on several simulated MR normal brain images and MR brain images having some multiple sclerosis lesions. The performance of the proposed semi-supervised clustering technique is compared with some other popular image segmentation techniques like Fuzzy C-means, Expectation Maximization and some recent image clustering techniques like multi-objective based MCMOClust technique, and Fuzzy-VGAPS clustering techniques.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The major challenge in analyzing MR brain images is to classify the pixels of images into homogeneous regions. This type of problem is termed as clustering/segmentation problem (Gonzalez & Woods, 2001). The success of medical imaging system depends on proper segmentation of images. MR images have large number of applications in solving several neurodegenerative disorders like Alzheimer diseases, Parkinson related syndrome etc. To solve these problems, several unsupervised and supervised based classification techniques have been developed (Bhandarkar & Zhang, 1999; Saha & Bandyopadhyay, 2009). Most of the clustering techniques rely on some similarity/dissimilarity criteria of data points by using which points are assigned to different clusters. They evolve the partition matrix U(X) of size $K \times n$ in such a way that $U = [u_{ik}]$, $1 \le i \le K$, where, u_{ij} takes value "0" if pattern x_k does not belong to cluster $C_i (i = 1,\ldots,K)$, and can take value "1" if pattern x_k belongs to cluster $C_i (i = 1, \ldots ,K)$. Unsupervised classification techniques do not take into account any kind of supervised information (Gath & Geva, 1989; Kwan, Evans, & Pike, 1999). They are used to partition the pixels based on some internal characteristics (Suckling, Sigmundsson, Greenwood, & Bullmore, 1999). Thus obtained partitioning may not be perfect always. Supervised classification techniques require large amount of labeled information to generate the models which are further utilized for classifying some unknown pixels (Pedrycz & Waletzky, 1997). But it is both time consuming

* Corresponding author. Tel.: +91 612 255 2067.
E-mail addresses: sriparna@iitp.ac.in (S. Saha), abhayalok@iitp.ac.in, abhayalok@gmail.com (A.K. Alok), asif@iitp.ac.in (A. Ekbal).

http://dx.doi.org/10.1016/j.eswa.2016.01.005
0957-4174/© 2016 Elsevier Ltd. All rights reserved.

and costly to generate huge amount of labeled information. In recent years a new classification technique, namely semi-supervised classification (Ebrahimi & Abadeh, 2012; Handl & Knowles, 2006), is developed to solve the difficulties of both unsupervised and supervised classification techniques. It utilizes the advantages of both supervised and unsupervised classification (Saha, Ekbal, & Alok, 2012). Here some small amount of labeled data and a huge collection of unlabeled data are used. The available labeled data is used to fine-tune the obtained partitionings. Several approaches have been developed to solve this type of semi-supervised classification problem. Literature survey shows that semi-supervised classification techniques are more powerful as compared to supervised or unsupervised classification techniques for solving different real-life problems (Saha et al., 2012).

In the current paper, the automatic MR brain image segmentation problem is posed as a multiobjective optimization (MOO) problem. Here we need to determine the number of clusters and the corresponding partitioning automatically from the given MR images using the search capability of any MOO based technique. During the clustering process, we have also assumed that some labeled information is also available. Thus the clustering problem is treated as a semi-supervised classification problem and a MOO based framework is developed to solve this problem. The proposed technique, namely Semi-MriMOO, utilizes a recently proposed simulated annealing based MOO technique, namely AMOSA (Bandyopadhyay, Saha, Maulik, & Deb, 2008), as the underline optimization strategy. Three objective functions are used to quantify the goodness of the obtained partitionings. These are simultaneously optimized using AMOSA (Bandyopadhyay et al., 2008). First two objective functions are some internal indices for cluster validity based on unsupervised properties of image data sets. These are: symmetry distance based Sym-index (Bandyopadhyay & Saha, 2008), and Euclidean distance based I-index (Maulik & Bandyopadhyay, 2002). Last one is an external index of cluster validity, Minkowski Score or MS-index (Ben-Hur & Guyon, 2003) based on supervised information or prior information of data set (Alok, Saha, & Ekbal, 2012; 2014). This basically checks the compatibility of the obtained partitioning and the available supervised information. A new encoding schema is used to represent clusters in the form of a string. Clusters are divided into multiple small sub-clusters. Centers of these small clusters are then encoded in the form of a string. In the current paper, assignment of points to different clusters is done using the popular Euclidean distance. Different new perturbation operations are defined, and as supervised information, we have assumed that class labels of only 10% data points are available. The segmentation results obtained by the proposed Semi-MriMOO clustering technique for different brain MR images are analyzed quantitatively and visually. Those are further compared with the results obtained by some recent or popularly used clustering techniques like Fuzzy C-means (Bezdek, 1981), Expectation Maximization (Jain, Murty, & Flynn, 1999), MCMOClust (Saha & Bandyopadhyay, 2011) and Fuzzy-VGAPS (Saha & Bandyopadhyay, 2007) clustering techniques.

The contributions of the current paper are as follows:

- To the best of our knowledge this is the first work where a semi-supervised based approach in multiobjective optimization framework is proposed to automatically segment the brain images.
- Semi-supervised clustering is solved using a multiobjective optimization framework.
- A new encoding strategy is proposed to represent the partitions in the form of a solution.
- A set of internal and external cluster validity indices are used as the objective functions.
- Search capability of a simulated annealing based multiobjective optimization technique, AMOSA, is utilized to automatically determine the appropriate partitioning from different brain images.
- Effectiveness of the proposed technique is shown for segmenting several normal brain images and also brain images with multiple sclerosis lesions. Results are compared with several popular and recent image segmentation techniques. Experimental results and a thorough analysis of those results clearly demonstrate the effectiveness of the proposed technique.

2. Literature review

In recent years there have been several attempts to solve the brain image segmentation problem (Klauschen, Goldman, Barra, Meyer-Lindenberg, & Lundervold, 2009; Ortiz, Górriz, Ramírez, Salas-Gonzalez & Llamas-Elvira, 2013). In Portela, Cavalcanti, and Ren (2014), authors have proposed clustering based semi-supervised classification technique to segment the MR brain images. Initially K-means clustering technique is applied on randomly selected MR brain slices. Thereafter, these clusters are labeled by human experts in terms of gray matter(GM), white matter(WM) and cerebrospinal fluid(CSF). Thereafter, some statistical measures of these clusters are computed. Finally, labeled information and statistical measures are given as initial parameters to Gaussian Mixture Model (GMM) to classify the remaining MR brain image slices. In Zhang, Dong, Clapworthy, Zhao, and Jiao (2010), authors have proposed semi-supervised spectral clustering technique to segment the MR brain images. In this approach instance level constraints like must-links and cannot-links are generated based on randomly selected subset of voxels from the same slice. Further, this supervision is also provided in spectral clustering to segment the remaining slices of MR brain images. Nearest neighbor strategy is applied for segmenting images to overcome the computation complexity of spectral clustering. In Ortiz, Gorriz, Ramirez, and Salas-Gonzalez (2014), authors have proposed Self-Organizing Maps(SOMs) and Genetic Algorithm(GA) based unsupervised clustering technique to segment the MR brain images. The whole process is divided into two modules. First module includes the preprocessing, feature extraction, feature selection based on genetic algorithm(GA), and voxel clustering using self-organizing maps(SOMs). Second module basically explores the entropy gradient clustering on extracted features. The obtained clustering solutions are validated using two cluster validity indices like DBI (Davies & Bouldin, 1979) and DUNN (Dunn, 1974). In Ortiz et al. (2013), authors have proposed self-organizing map (SOM) based two clustering techniques to segment the MR brain images. In the first approach, the relevant information is extracted from the whole volume histogram using SOM to classify the voxels. Second approach is divided into four stages including preprocessing, first and second order feature extraction, genetic algorithm based feature selection and finally mapped units are clustered using self-organizing map (SOM). The first approach is a fast procedure and the second approach is more robust under noisy and bad intensity conditions. In Li, Ogunbona, deSilva, and Attikiouzel (2011), authors have proposed a semi-supervised maximum posteriori probability (ssMAP) based segmentation technique for MR brain image data sets. The main intuition behind this algorithm is to exploit the incomplete training or labeled data. Due to the presence of homogeneity in images, performing proper segmentation becomes infeasible. Furthermore, authors have introduced second order polynomial function to handle with in-homogeneity to achieve maximum posterior probability estimation. In Song, Huang, Ma, and Hung (2011), authors have proposed semi-supervised classification technique inspired by Adaboost. Here, nearest neighbor algorithm is applied on unlabeled data to obtain first pseudo class label information.

Download English Version:

https://daneshyari.com/en/article/383880

Download Persian Version:

https://daneshyari.com/article/383880

Daneshyari.com