Egyptian Informatics Journal (2012) **13**, 39–58

Cairo University

Egyptian Informatics Journal

www.elsevier.com/locate/eij
www.sciencedirect.com

ORIGINAL ARTICLE

Determining the number of clusters for kernelized fuzzy C-means algorithms for automatic medical image segmentation

E.A. Zanaty

Computer Science Department, Faculty of Science, Sohag University, Sohag, Egypt

Received 12 December 2011; revised 18 January 2012; accepted 29 January 2012
Available online 23 February 2012

KEYWORDS

Medical image segmentation;
Clustering methods;
FCM;
Kernel function;
Validity indexes

Abstract In this paper, we determine the suitable validity criterion of kernelized fuzzy C-means and kernelized fuzzy C-means with spatial constraints for automatic segmentation of magnetic resonance imaging (MRI). For that; the original Euclidean distance in the FCM is replaced by a Gaussian radial basis function classifier (GRBF) and the corresponding algorithms of FCM methods are derived. The derived algorithms are called as the kernelized fuzzy C-means (KFCM) and kernelized fuzzy C-means with spatial constraints (SKFCM). These methods are implemented on eighteen indexes as validation to determine whether indexes are capable to acquire the optimal clusters number. The performance of segmentation is estimated by applying these methods independently on several datasets to prove which method can give good results and with which indexes. Our test spans various indexes covering the classical and the rather more recent indexes that have enjoyed noticeable success in that field. These indexes are evaluated and compared by applying them on various test images, including synthetic images corrupted with noise of varying levels, and simulated volumetric MRI datasets. Comparative analysis is also presented to show whether the validity index indicates the optimal clustering for our datasets.
© 2012 Faculty of Computers and Information, Cairo University.
Production and hosting by Elsevier B.V. All rights reserved.

E-mail address: zanaty22@yahoo.com

1110-8665 © 2012 Faculty of Computers and Information, Cairo University. Production and hosting by Elsevier B.V. All rights reserved.

Peer review under responsibility of Faculty of Computers and Information, Cairo University.
doi:10.1016/j.eij.2012.01.004

Production and hosting by Elsevier

1. Introduction

Clustering is one of the most popular classification methods and has found many applications in pattern classification and image segmentation [1–5]. Clustering algorithms attempt to classify a voxel to a tissue class by using the notion of similarity to the class. Unlike the crisp K-means clustering algorithm [4], the FCM algorithm allows partial membership in different tissue classes. Thus, FCM can be used to model the partial volume averaging artifact, where a pixel may contain multiple tissue classes [2,3]. The kernelized fuzzy C-means

(KFCM) [6–9] used a kernel function as a substitute for the inner product in the original space, which is like mapping the space into higher dimensional feature space. There have been a number of other approaches to incorporating kernels into fuzzy clustering algorithms. These include enhancing clustering algorithms designed to handle different shape clusters [8]. More recent results of fuzzy algorithms have been presented in [9] for improving automatic MRI image segmentation. They used the intra-cluster distance measure to give the ideal number of clusters automatically; more discussion can be found in [9]. Also, possibilistic clustering which is pioneered by the possibilistic c-means (PFCM) algorithm was developed in [10–12]. They had been shown that PFCM is more robust to outliers than FCM. However, the robustness of PFCM comes at the expense of the stability of the algorithm [11]. The PCM-based algorithms suffer from the coincident cluster problem, which makes them too sensitive to initialization [12].

Most fuzzy methods have several advantages including yielding regions more homogeneous than other methods; reducing the spurious blobs; removing noisy spots; reduced sensitivity to noise compared to other techniques. However, they require prior knowledge about the number of clusters in the data, which may not be known for new data [13]. In literature, many studies in dealing with this problem are available in [14–18], and, so, there are many cluster validity indexes in this regard. Compactness and separation are two criteria for the clustering evaluation and selection of an optimal clustering scheme [14]. The variation of data within clusters indicates compactness and isolation between clusters indicates separation, respectively.

Though some compatibility or similarity measure can be applied to choose the clusters to be merged, no validity measure is used to guarantee that the clustering result after a merge is better than the one before the merge. Partial results were stated in [19] to answer the questions: "Can the appropriate number of clusters be determined automatically? And if the answer is yes, how?" More existing methods were found in [14–21] to review few validity indexes that can combine with fuzzy c-means algorithms. But, the performance of wide range indexes is not found in any literature before; especially when they applied to kernelized fuzzy c-means (KFCM) or kernelized fuzzy c-means with spatial constraints (SKFCM) methods.

In this paper, we seek the answer to the previous questions for exploring which indexes can achieve high accuracy segmentation whey they performed with KFCM and SKFCM. Our objective is not to improve the segmentation accuracy via enhancing the kernel function, but is to find the indexes with KFCM and SKFCM capable to produce good MRI segmentation. For that; the original Euclidean distance in the FCM algorithm is replaced by the Gaussian radial base function (GRBF)-induced kernel, which is shown to be more robust than FCM (with Euclidean distances). This will make a generalization of the existing FCM methods. The KFCM and SKFCM algorithms based on Gaussian RBF kernel are derived and applied independently on each image. Based on these algorithms, eighteen indexes are implemented to estimate the number of clusters that represents the best structure of a given image. Key existing solutions are evaluated to obtain the cluster validity in the domain of image segmentation. A wide number of various validity indexes from the classical and more recent indexes are examined. As segmentation of medical images is of particular interest in our application, the work here includes the assessment of those indexes on 3D MRI datasets.

The rest of this paper is organized as follows: Section 2 presents the kernel methods. Several criteria to determine the number of clusters are briefly reviewed in Section 3. Experimental comparisons are presented in Section 4. Finally, Section 5 gives our conclusions.

2. Kernel methods

The kernel methods [8,13,22–26] are one of the most researched subjects within machine learning community in the recent few years and have widely been applied to pattern recognition and function approximation. A common philosophy behind these algorithms is based on the following kernel (substitution) trick, that is, firstly with a (implicit) nonlinear map, from the data space to the mapped d feature space, $\Psi: X \to F\,(x \to \Psi(x))$, a dataset $\{x, \ldots, x\} \subseteq X$ (an input data space with low dimension) is mapped into a potentially much higher dimensional feature space or inner product F, which aims at turning the original nonlinear problem in the input space into potentially a linear one in rather high dimensional feature space so as to facilitate problem solving as proved by Girolami [23]. A kernel $K(x, y)$ in the feature space can be represented as:

$$K(x, y) = (\Psi(x), \Psi(y)) \tag{1}$$

where $(\Psi(x), \Psi(y))$ denotes the inner product operation.

An interesting point about kernel function is that the inner product between $\Psi(x)$ and $\Psi(y)$ can be implicitly computed in F, without explicitly using or even knowing the mapping Ψ.

So, kernels allow computing inner products in the space, where one could otherwise not practically perform any computations. Three commonly-used kernel functions in literature [25] are:

(1) Gaussian Radial basis function (GRBF) kernel: $K(x, y) = \exp(-\|x - y\|^2/\sigma^2)$.
(2) Polynomial kernel: $K(x, y) = (\langle x, y\rangle + 1)^d$.
(3) Sigmoid kernel $K(x, y) = \tanh(\alpha\langle x, y\rangle + \beta)$.

where σ, d, α and β are the adjustable parameters of the above kernel functions. The main motives of using the kernel methods consist of: (1) inducing a class of robust non-Euclidean distance measures for the original data space to derive new objective functions and thus clustering the non-Euclidean structures in data; (2) enhancing robustness of the original clustering algorithms to noise and outliers, and (3) still retaining computational simplicity.

Sigmoid kernel is a two-layer neural network kernel and is used as a particular kind of two-layer sigmoid neural network. For this, only a set of parameters satisfying the Mercer theorem can be used to define a kernel function [23–26]. The interested reader may refer to [25] for more details. In this section we only stress on GRBF, which is shown to be more robust than FCM (with Euclidean distances) [7].

2.1. Fuzzy C-means method (FCM)

Fuzzy C-means clustering (FCM), also known as fuzzy ISODATA, is a data clustering algorithm in which each data point belongs to a cluster to determine a degree specified by its

Download English Version:

https://daneshyari.com/en/article/477023

Download Persian Version:

https://daneshyari.com/article/477023

Daneshyari.com