Kernelised Rough Sets Based Clustering Algorithms Fused With Firefly Algorithm for Image Segmentation

Srujan Sai Chinta, Vellore Institute of Technology, Vellore, India

ABSTRACT

Data clustering methods have been used extensively for image segmentation in the past decade. In one of the author’s previous works, this paper has established that combining the traditional clustering algorithms with a meta-heuristic like the Firefly Algorithm improves the stability of the output as well as the speed of convergence. It is well known now that the Euclidean distance as a measure of similarity has certain drawbacks and so in this paper we replace it with kernel functions for the study. In fact, the authors combined Rough Fuzzy C-Means (RFCM) and Rough Intuitionistic Fuzzy C-Means (RIFCM) with Firefly algorithm and replaced Euclidean distance with either Gaussian or Hyper-tangent or Radial basis Kernels. This paper terms these algorithms as Gaussian Kernel based rough Fuzzy C-Means with Firefly Algorithm (GKRFCMFA), Hyper-tangent Kernel based rough Fuzzy C-Means with Firefly Algorithm (HKRFCMFA), Gaussian Kernel based rough Intuitionistic Fuzzy C-Means with Firefly Algorithm (GKRIFCMFA) and Hyper-tangent Kernel based rough Intuitionistic Fuzzy C-Means with Firefly Algorithm (HKRIFCMFA). In order to establish that these algorithms perform better than the corresponding Euclidean distance-based algorithms, this paper uses measures such as DB and Dunn indices. The input data comprises of three different types of images. Also, this experimentation varies over different number of clusters.

KEYWORDS

DB Index, Dunn Index, Gaussian Kernel, Hypertangent Kernel, L. Zadeh, RFCM, RIFCM, T. Atanassov

1. INTRODUCTION

A process of splitting the image into pixel bands is called image segmentation. As images contain uncertainties, there are difficulties in classification of images into homogeneous regions. With respect to position of elements into different clusters, clustering methods are classified into two types: Hard clustering methods and Soft clustering methods. In the context of Image segmentation, in hard clustering every pixel belongs to only one cluster. In soft clustering, every pixel may belong to more than one cluster to a certain degree. Fuzzy clustering methods use the model of Fuzzy Sets introduced in (Zadeh, 1965). The most common Fuzzy Sets based clustering algorithm is Fuzzy C-Means (FCM) (Bezdek et al. 1984). In the fuzzy set model, the non-membership value of an element in a collection is one’s complement of its membership value. Generalising this concept by restricting the sum of membership value and non-membership value to be less than or equal to 1, the intuitionistic

DOI: 10.4018/IJFSA.2019100102

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
fuzzy set model was introduced in (Atanassov, 1986). Using this model, a clustering algorithm called Intuitionistic Fuzzy C-Means (IFCM) was developed (Chaira, 2011). The other uncertainty-based model, rough set (Pawlak, 1982) was used to develop the rough C-means algorithm in (Lingras and West, 2004). The proposed Rough K-Means algorithm classifies pixels into the lower approximation and boundary of each cluster depending upon the distance between the pixel and cluster centroids. Hybrid models are shown to be more efficient than the individual components in general. The hybrid models of rough fuzzy sets and fuzzy rough sets were proposed by (Dubois and Prade, 1990). (Mitra et al, 2006) used the rough fuzzy set model to develop a Rough Fuzzy C-Means algorithm (RFCM). This was further generalised in (Bhargav et al, 2013) to propose the rough intuitionistic fuzzy C-Means (RIFCM) algorithm. In all the above algorithms, the Euclidean distance between two data items is used to measure their similarity. However, the Euclidean distance-based clustering algorithms do have the problem of linearly separable datasets. However, using the kernel function by projecting the feature space into a higher dimension using an appropriate non-linear mapping function ensures linear separability of the complex clusters which are otherwise not linearly separable in its original feature space. Thus, in an attempt to avail this generality several kernel function-based algorithms have been developed. Some of these algorithms are the Kernel based K-means clustering using rough set (Tripathy et al, 2012). Kernel based rough fuzzy c-means algorithm (Tripathy and Bhargav, 2013), Kernel based rough intuitionistic fuzzy c-means algorithm (Tripathy et al, 2014). A comparative analysis of kernel functions in uncertainty-based c-means algorithms is made in (Tripathy and Mittal, 2015). Also, we find Kernelised clustering algorithms for decision theoretic rough sets in the form of decision theoretic Kernelised rough c-means (Ryan et al, 2016) and Decision theoretic Kernelised rough intuitionistic fuzzy c-means (Ryan et al, 2017). In all the above algorithms the initial solutions are selected randomly. This technique sometimes does not converge and if at all convergent takes more time. So, in order to optimize the process of convergence one needs efficient selection of the initial solutions. This can be achieved by using any meta-heuristic algorithm to get such solutions. In our previous works we used the firefly algorithm to get such solutions, for rough set-based algorithms (Abhay et al, 2017) and for intuitionistic fuzzy c-means algorithm (Srujan et al, 2017), we have clearly shown that RFCM and RIFCM perform better when they are fused with Firefly algorithm. In this paper, we replace the Euclidean distance formula with Gaussian Kernel, Hyper-tangent Kernel and radial basis Kernel. The aim of this paper is to study the behaviour of these kernels and compare the performance of Kernel-based RFCM and RIFCM with their Euclidean counter-parts. Also, we get a comparison of the three kernels in the context. The remainder of the paper is organised as follows: Section 2 contains detailed information about the various clustering algorithms and their Hybrids, Section 3 contains information about Firefly algorithm. Information about the various distance measures such as Euclidean distance, Gaussian kernel and Hyper-tangent kernel is given in Section 4, Section 5 contains the methodology of our proposed algorithms, our experimental results are shown and discussed in Section 6 and the conclusion is given in Section 7.

2. CLUSTERING ALGORITHMS

In this section we present some basic uncertainty-based clustering algorithms. The mathematical formulations of each of these clustering algorithms are discussed.

2.1. Fuzzy C-Means

In FCM, random cluster centroids are initiated. The distance \(d_{ik}\) between every cluster center \(i\) and every pixel of the image \(k\) is computed using some distance measure such as the Euclidean distance. The Membership Matrix is computed as follows:
Application of Fuzzy Logic for Adaptive Food Recommendation
[www.igi-global.com/article/application-of-fuzzy-logic-for-adaptive-food-recommendation/179323?camid=4v1a](www.igi-global.com/article/application-of-fuzzy-logic-for-adaptive-food-recommendation/179323?camid=4v1a)

Predicting Uncertain Behavior and Performance Analysis of the Pulping System in a Paper Industry Using PSO and Fuzzy Methodology