RESEARCH ARTICLE OPEN ACCESS

A Novel Kernel Based Fuzzy C Means Clustering With Cluster Validity Measures

D.Vanisri
Department of Computer Technology
Kongu Engineering College, Perundurai
Tamil Nadu - India

ABSTRACT

Clustering algorithms are an integral part of both computational intelligence and pattern recognition. It is unsupervised methods for classifying data into subgroups with similarity based inter cluster and intra cluster. In fuzzy clustering algorithms, mainly used algorithm is Fuzzy c-means (FCM) algorithm. This FCM algorithm is efficient only for spherical data when the input of the data structure is not spherical or complex this method is unsuccessful. For this, modification of the FCM is done by the labeling of a pixel to be partial by the labels in its immediate neighborhood and this modification is called BCFCM (Bias-Corrected FCM). Since it is computationally time taking and lacks enough robustness to noise for that kernel versions of FCM with spatial constraints, such as KFCM, were proposed to solve the drawbacks of BCFCM. In this paper, a novel kernel-based fuzzy C-means clustering algorithm (KFCM) is proposed for clustering. It is recognized by replacing the kernel-induced distance metric over the original Euclidean distance, and the corresponding algorithms are called kernel fuzzy c-means (KFCM) algorithm. The experimental results shows that proposed clustering technique provides better accuracy with less error rate than the BCFCM algorithm.

Keywords:- Clustering, Fuzzy clustering, Bias corrected Fuzzy C Means, Kernel based Fuzzy C Means

I. INTRODUCTION

Clustering method is a process in which a data set or say pixels are replaced by cluster, pixels may belong together because of the same color, texture etc. Among the clustering methods, one of the most popular methods for clustering is fuzzy clustering, which can retain more information than hard clustering in some cases. Fuzzy c-means (FCM) is one of the most promising fuzzy clustering methods. Although this Fuzzy C Means is well accepted clustering method it is unsuccessful for large spherical cluster. Because the FCM only use squared norm for classification similarity between data points, it can only be successful in clustering 'spherical' clusters. But many clustering algorithm can be resultant from the FCM for clustering more general dataset. This derived clustering algorithm can be formed by substitution the squared-norm in the object function of FCM with other similarity measures (metric).

One of the methods proposed from FCM is Bias-corrected FCM (BCFCM). It is very useful for noise and intensity inhomogeneity image segmentation, but it can't estimate accurately the pixels on the boundary especially in the regions with heavy level of intensity inhomogeneous. Thus, it is very sensitive to noise and inhomogeneities in the image moreover, it remains dependent on the initialization of the cluster centers. To rectify the problem of BCFCM algorithm, FCM is extended to KFCM (Kernel based Fuzzy C Means) Algorithm. The main idea of this KFCM is to alter perfectly the input data into a higher dimensional feature space, and then it will increases the possibility of linear seperability of the patterns in the feature space, then perform FCM in the feature space. The KFCM also determines the number of clusters in the dataset which is another good quality of KFCM.

The rest of the paper describes the related work about clustering techniques in Section 2, Methodology about existing technique and drawbacks which leads to proposed in Section 3, experimental proved results for better performance of proposed technique shown in Section 4, and conclusion regarding results in Section 5.

II. LITERATURE SURVEY

Data clustering is one of the important data mining methods. In this paper, a comparative study of these algorithms with different distance measures such as Chebyshev and Chi-square is proposed by Jafar et al (2013). A novel Fuzzy c Means (FCM) algorithm with modified distance computation is proposed by Li et al (2013) in this paper. We modify the distance in FCM with the neighborhood information of cluster centers. In this paper, to overcome the above problems, Chen first propose two variants, FCM_S_1 and FCM_S_2, of FCM_S to aim at simplifying its computation and then extend them, including FCM_S, to corresponding robust kernelized versions KFCM_S, $KFCM_S_1$and $KFCM\_S_2$ by the kernel methods. Cui et al (2013) presents a new fuzzy clustering algorithm for simultaneous segmentation and bias field estimation of medical images.

In this paper, a common misunderstanding of Gaussian-function-based kernel fuzzy clustering is corrected, and a kernel fuzzy c-means clustering-based fuzzy SVM algorithm (KFCM-FSVM) is developed by Yang et al (2011) to deal with the classification problems with outliers or noises. Huang et al (2012) applying kernel tricks, the kernel fuzzy c-means algorithm attempts to address this problem by mapping data with nonlinear relationships to appropriate feature spaces. Chen et al (2011) propose a multiple kernel fuzzy c-means (MKFC) algorithm that extends the fuzzy c-means algorithm with a multiple kernel-learning setting. In addition, we show multiple kernel k-means to be a special case of MKFC In this paper, a novel clustering algorithm using the 'kernel method' based on the classical fuzzy clustering algorithm (FCM) is proposed by Zhang et al (2003) and called as kernel fuzzy c-means algorithm (KFCM).An automatic method is presented by Xiang (2013) for segmentation of MS lesions in multispectral MR images.

In this work, Variable Kernel Fuzzy C-Means (VKFCM) was used by saikumar et al (2013) to generate an initial contour curve which overcomes leaking at the boundary during the curve propagation. Recently Kernelized Fuzzy C-Means clustering technique where a kernel-induced distance function is used as a similarity measure instead of a Euclidean distance which is used in the conventional Fuzzy C-Means clustering technique, has earned popularity among research community is proposed by Das et al (2014). Ferreira et al (2014) paper presents variable-wise kernel fuzzy *c*-means clustering methods in which dissimilarity measures are obtained as sums of Euclidean distances between patterns and centroids computed individually for each variable by means of kernel functions. In order to improve the segmentation accuracy, Zeng et al (2014) propose an unsupervised method which combines an improved bias correction Fuzzy C-means (BCFCM) and class-adaptive hidden markov random field Modelling (HMRF). The BCFCM segmentation result is used as the initial labeling for class-adaptive HMRF, which is utilized to REFINE the segmentation results. Cui et al (2013) combine the global fuzzy energy with the local fuzzy energy using an adaptive weight function whose value varies with the local contrast of the image. In this paper, a hybridization of rough c-means and spatial fuzzy c-means clustering is presented by Srivastava et al (2013) whose objective function has been modified to accommodate INU field as well.

III. PROPOSED METHODOLOGY

The proposed methodology Kernel based Fuzzy C Means Algorithm is described in this section.

3.1 Fuzzy C Means Algorithm

Fuzzy c-means (FCM) is a technique of clustering which permits data points to cluster based on similarity and mainly used in pattern recognition. The alteration of objective function by minimizing as follows:

$$J_m(U,V) = \sum_{i-1}^{c}\sum_{j-1}^{n} V_{ij}^{m}\left\|x_i - c_j\right\|^2 \qquad (1)$$

Where m is any real number higher than 1, V_{ij}^{m} is the degree of membership of x_i in the cluster j, x_i is the d-dimensional measured data, c_j is the d-dimension center of the cluster. Fuzzy classification is done by objective function iteration as shown above, with the update of membership V_{ij}^{m} and the cluster centers c_j by:
Iteration of this process will be terminated when there is termination condition between 0 and 1, whereas k are the iteration steps and N is the total number of clusters. This procedure converges to a local minimum or a saddle point of J_m. Lastly, in squared error function case, this algorithm aims by minimizing the objective function.

3.2 Bias-Corrected (BC) FCM (BCFCM) Objective Function

When high membership values are used this FCM objective function is minimized whose data points are close to the centroid of its similar class, and low membership values are dispensed when the data are far from the centroid. Then, modification is done to Equation 1 by establishing a term that allows the labeling of a pixel to be inclined by the labels in its immediate neighborhood. As already told, the neighborhood effect proceeds as normalize and biases the solution toward piecewise-homogeneous labeling. The modified objective function is given by

$$J_m = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{p}\|x_k - v_i\|^2 + \frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{p}\left(\sum_{x_r \in N_k}\|x_r - v_i\|^2\right) \quad (2)$$

Where N_k stands for the set of neighbors that exist in a window around x_k and N_R is the cardinality of N_k. The effect of the neighbors' term is controlled by the parameter α. The relative importance of the regularizing term is inversely proportional to the signal-to-noise ratio (SNR). Lower SNR would require a higher value of the parameter α.

3.3 Kernel Fuzzy C-Means Clustering Algorithm

Chen and Zhang pointed out a shortcoming of the BCFCM and then replaced the Euclidean distance $\|x_j - \alpha_i\|$ with a kernel-induced distance $1 - k(x_j, \alpha_i) = 1 - \exp(-\|x_j - \alpha_i\|^2/\sigma^2$. They gave the kernel version of FCM.

Given a dataset, $X = \{x_1, \dots, x_n\} \subset R^p$, the original FCM algorithm partitions X into c fuzzy subsets by minimizing the following objective function

$$J_m(U,V) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}\|x_k - v_i\|^2 \quad (3)$$

Where c is the number of clusters and selected as a specified value, n the number of data points, u_{ik} the membership of x_k in class i, satisfying $\sum_{i=1}^{c} u_{ik} = 1$, m the quantity controlling clustering fuzziness, and V the set of cluster centers or prototype $(v_i \in R^P)$. The function J_m is minimized by a famous alternate iterative algorithm.

Now by considering the proposed (KFCM) algorithm, define a nonlinear map as $\Phi: x \longrightarrow \Phi(x) \in F$, where $x \in X$. X denotes the data space, and F the transformed feature space with higher even infinite dimension. KFCM minimizes the following objective function

$$J_m(U,V) = \sum_{i-1}^{c}\sum_{k-1}^{n} u_{ik}^{m}\|\Phi(x_k) - \Phi(v_i)\|^2 \quad (4)$$

Where

$$\|\Phi(x_k) - \Phi(v_i)\|^2 = K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i) \tag{5}$$

where $K(x,y) = \Phi(x)^T \Phi(y)$ is an inner product kernel function. Then by adopting the kernel function, i.e.,$K(x,y) = exp(-\|x-y\|^2/\sigma^2)$, then $K(x, x) = 1$, according to Equations (4) and (5) can be rewritten as

$$J_m(U,V) = 2\sum_{i-1}^{c}\sum_{k-1}^{n} u_{ik}^m(1 - K(x_k, v_i)) \tag{6}$$

Minimizing Equation (6) under the constraint of u_{ik}, we have

$$u_{ik} = \frac{(1/(1 - K(x_k, v_i)))^{1/(m-1)}}{\sum_{j-1}^{c}(1/(1 - K(x_k, v_j)))^{1/(m-1)}} \tag{7}$$

$$v_i = \frac{\sum_{k-1}^{n} u_{ik}^m K(x_k, v_i) x_k}{\sum_{k-1}^{n} u_{ik}^m K(x_k, v_i)} \tag{8}$$

If by using other kernel functions, there will be equivalent modifications in Equation (7) and (8).In fact, Equation (5) can be viewed as kernel-induced new metric in the data space, which is defined as the following

$$d(x,y) \triangleq \|\Phi(x) - \Phi(y)\| = \sqrt{2(1 - K(x,y))} \tag{9}$$

And it can be proven that $d(x, y)$ defined in Equation (9) is a metric in the original space in case that $K(x,y)$ takes as the kernel function. According to Equations (8), the data point x_k is endowed with an additional weight $K(x_k, v_i)$, which measures the similarity between x_k and v_i , and when x_k is an outlier, i.e., x_k is far from the other data points, then $K(x_k, v_i)$ will be very small, so the weighted sum of data points shall be more robust.

The full description of KFCM algorithm is as follows:

KFCM Algorithm

Step 1: Fix c, t_{max} , $m > 1$ and $\varepsilon > 0$ for some positive constant;
Step 2: Initialize the memberships u_{ik}^0;
Step 3: For t =1, 2,…, t_{max} , do:
(a) Update all prototypes v_i^t with Equation (8);
(b) Update all memberships u_{ik}^t with Equation (7);
(c) Compute $E^t = max_{i,k}|u_{ik}^t - u_{ik}^{t-1}|, if\ E^t \leq \varepsilon$
Stop; else t=t+1

3.3.1 Kernelized Cluster Validity Measures

For determining the suitable number of clusters, this kernelized validity measures are proposed. This is classified into two classes: one class only uses membership value. A typical example is the entropy

$$E(U,C) = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ki} \log u_{ki} \quad (10)$$

Whereby the number c that maximizes $E(U, C)$ is selected.

While another classes uses geometrical characteristics. A typical method uses the fuzzy covariance matrix for cluster i:

$$F_i = \frac{\sum_{k=1}^{N}(u_{ki})^m (x_k - v_i)(x_k - v_i)^{\mathrm{T}}}{\sum_{k=1}^{N}(u_{ki})^m} \quad (11)$$

Here, use the sum of the square root of the determinants of F_i :

$$W_{det} = \sum_{i=1}^{c} \sqrt{\det F_i} \quad (12)$$

We also consider the sum of the traces of F_i:

$$W_{tr} = \sum_{i=1}^{c} trF_i \quad (13)$$

When kernel-based clustering is used then kernel based validity measures should be used. Let us consider the kernelized versions of (12) and (13) for this purpose. The kernel-based fuzzy covariance matrix is the following:

$$KF_i = \frac{\sum_{k=1}^{N}(u_{ki})^m (\Phi(x_k) - v_i)(\Phi(x_k) - v_i)^T}{\sum_{K=1}^{N}(u_{ki})^m} \quad (14)$$

Where v_i is not explicitly given.

Note that the determinant of the kernelized fuzzy covariance is inappropriate, since the next relation holds: $\det KF_i \to o, \quad as\ N \to \infty$

The proof of this relation is simple because the monotone decreasing sequence $\lambda_1, \lambda_2, \ldots$of the Eigen values of KF_i will converge to zero as $N \to \infty$. Hence we have

$$\log\det KF_i = \sum_{i=1}^{N} \log\lambda_1 \to -\infty, \qquad as\ N \to \infty \quad (15)$$

In contrast, the trace of KF_i is useful. After some calculation, we have

$$trKF_i = \frac{1}{\sum_{k=1}^{N}(u_{ki})^m}\sum_{k=1}^{N}(u_{ki})^m\|\Phi(x_k) - v_i\|^2 = \frac{1}{\sum_{k=1}^{N}(u_{ki})^m}\sum_{k=1}^{N}(u_{ki})^m D_{ki} \tag{16}$$

Hence we define

$$KW_{tr} = \sum_{i=1}^{c} trKF_i \tag{17}$$

The discussions on kernel functions are kernelized cluster validity measures and new kernels derived from basic functions of fuzzy *c*-means.

EXPERIMENTAL RESULTS

The proposed approach for clustering unlabeled data is experimented UCI machine learning Repository. All algorithms are implemented under the same initial values and stopping conditions. The experiments are all performed on using MATLAB version 7.5. Table 1 shows the accuracy and execution time for FCM, BCFCM and KFCM.

Techniques	Accuracy	Execution time
FCM	72	35
BCFCM	81	29
KFCM	90	23

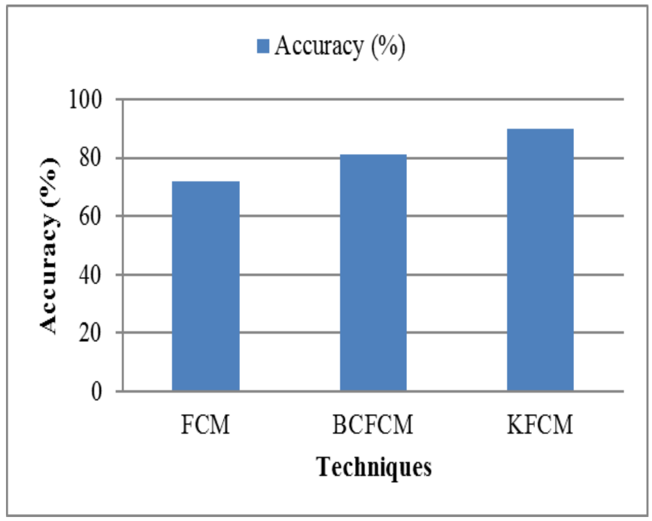

Figure 1: Accuracy

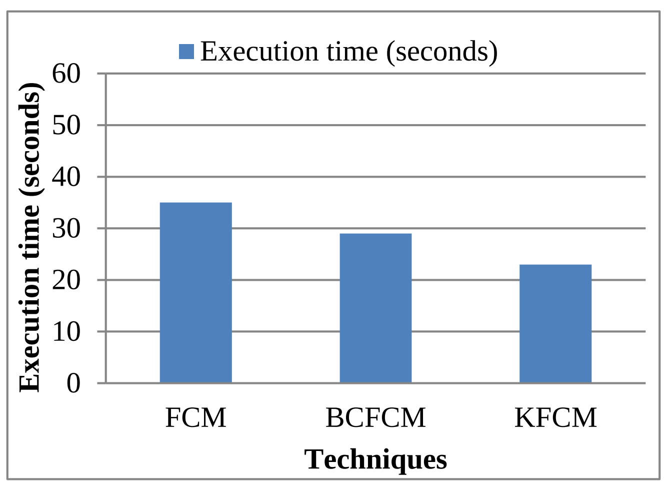

Figure 2: Execution time

From the above table and figure, it is shown that the proposed KFCM provides better accuracy with les execution time in clustering technique.

IV. CONCLUSION

Thus we have succeeded in solving problems addressed in BCFCM algorithm: to obtain a new class of kernel functions, to uncover relations between solutions of fuzzy *c*-means clustering and related methods, and to show usefulness of the derived kernel functions. The results show that the KFCM algorithm outperforms the FCM & BCFCM. The choice of a method is related to several factors and parameters adjustment that governs the algorithm deserves special attention. Finally, it seems interesting to consider the integration of other constraints on the pixels spatial arrangement and to combine several classification algorithms working in cooperation.

As a future study, for example, various clustering algorithms as well as kernel-based validity measures should be investigated. Moreover, SVM using the non-Gaussian kernels should be studied. Other problems to be solved in near future are how to select appropriate parameters to the three kernels, and various numerical examples should be tested for this purpose. Furthermore, theoretical studies of fuzzy clustering and related kernel functions should be necessary.

REFERENCES

[1] Chen, Long, CL Philip Chen, and Mingzhu Lu. "A multiple-kernel fuzzy c-means algorithm for image segmentation." *Systems, Man, and Cybernetics, Paa B: Cybernetics, IEEE Transactions on* 41, no. 5 (2011): 1263-1274.

[2] Huang, Hsin-Chien, Yung-Yu Chuang, and Chu-Song Chen. "Multiple kernel fuzzy clustering." *Fuzzy Systems, IEEE Transactions on* 20, no. 1 (2012): 120-134.

[3] Xiang, Yan, Jianfeng He, Dangguo Shao, and Lei Ma. "Automatic segmentation of multiple sclerosis lesions in multispectral MR images using kernel fuzzy c-means clustering." In *Medical Imaging Physics and Engineering (ICMIPE), 2013 IEEE International Conference on*, pp. 102-106. IEEE, 2013.

[4] Yang, Xiaowei, Guangquan Zhang, Jie Lu, and Jun Ma. "A kernel fuzzy c-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises." *Fuzzy Systems, IEEE Transactions on* 19, no. 1 (2011): 105-115.

[5] Saikumar, Tara, Khaja FasiUddin, B. Venkata Reddy, and Md Ameen Uddin. "Image Segmentation Using Variable Kernel Fuzzy C Means (VKFCM) Clustering on Modified Level Set Method." In *Computer Networks & Communications (NetCom)*, pp. 265-273. Springer New York, 2013.

[6] Das, Samarjit, Hemanta K. Baruah, Ms Shahin S. Shaikh, Manjusha Deshamukh, Chandrashekhar Kamargaonkar, Linta Ann George, Chandrahas Sahu et al. "A New Kernelized Fuzzy C-Means Clustering Algorithm with Enhanced Performance." (2014).

[7] Ferreira, Marcelo RP, and Francisco De AT De Carvalho. "Kernel fuzzy< i> c</i>-means with automatic variable weighting." *Fuzzy Sets and Systems* 237 (2014): 1-46.

[8] 9.Zeng, Ziming, Chunlei Han, Liping Wang, and Reyer Zwiggelaar. "Unsupervised Brain Tissue Segmentation by Using Bias Correction Fuzzy C-Means and Class-Adaptive Hidden Markov Random Field Modelling." In *Frontier and Future Development of Information Technology in Medicine and Education*, pp. 579-587. Springer Netherlands, 2014.

[9] Cui, Wenchao, Yi Wang, Tao Lei, Yangyu Fan, and Yan Feng. "Brain MR image segmentation based on an adaptive combination of global and local fuzzy energy." *Mathematical Problems in Engineering* 2013 (2013).

[10] Cui, Wenchao, Yi Wang, Yangyu Fan, Yan Feng, and Tao Lei. "Global and local fuzzy clustering with spatial information for medical image segmentation." In *Signal and Information Processing (ChinaSIP), 2013 IEEE China Summit & International Conference on*, pp. 533-537. IEEE, 2013.

[11] Srivastava, Arpit, Jyoti Singhai, and Mahua Bhattacharya. "Collaborative rough-fuzzy clustering: An application to intensity non-uniformity correction in brain MR images." In *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, pp. 1-6. IEEE, 2013.

[12] Jafar, OA Mohamed, and R. Sivakumar. "A Comparative Study of Hard and Fuzzy Data Clustering Algorithms with Cluster Validity Indices." (2013).

[13] Li, Naixiang, and Peng Guo. "A novel Fuzzy C Means algorithm based on distance modification for image segmentation." In *Fuzzy Systems and Knowledge Discovery (FSKD), 2013 10th International Conference on*, pp. 174-178. IEEE, 2013.

[14] Chen, Songcan, and Daoqiang Zhang. "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure."*Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*34, no. 4 (2004): 1907-1916.

[15] Zhang, Dao-Qiang, and Song-Can Chen. "Clustering incomplete data using kernel-based fuzzy c-means algorithm." *Neural Processing Letters* 18, no. 3 (2003): 155-162.