Novel Intuitionistic Fuzzy C-Means Clustering for Linearly and Nonlinearly Separable Data

PRABHJOT KAUR
Department of IT, MSIT
GGSIP University
New Delhi
INDIA
thisisprabhjot@gmail.com

DR. A. K. SONI
Department of Computers
Sharda University
Greater Noida
INDIA
ak.soni@sharda.ac.in

DR. ANJANA GOSAIN
University School of IT
GGSIP University
New Delhi
INDIA
anjana_gosain@hotmail.com

Abstract: - This paper presents a robust Intuitionistic Fuzzy c-means (IFCM-σ) in the data space and a robust kernel Intutitionistic Fuzzy C-means (KIFCM-σ) algorithm in the high-dimensional feature space with a new distance metric to improve the performance of Intuitionistic Fuzzy C-means (IFCM) which is based upon intuitionistic fuzzy set theory. IFCM considered an uncertainty parameter called hesitation degree and incorporated a new objective function which is based upon intutionistic fuzzy entropy in the conventional Fuzzy C-means. It has shown better performance than conventional Fuzzy C-Means. We tried to further improve the performance of IFCM by incorporating a new distance measure which has also considered the distance variation within a cluster to regularize the distance between a data point and the cluster centroid. Experiments are done using two-dimensional synthetic data-sets, Standard data-sets referred from previous papers. Results have shown that proposed algorithms, especially KIFCM-σ is more effective for linear and non-linear separation.

Key-words: -Fuzzy Clustering, Intuitionistic Fuzzy C-Means, Robust Clustering, Kernel Intuitionistic fuzzy c-means, Distance metric, Fuzzy c-means

1. Introduction

Clustering helps in finding natural boundaries in the data whereas fuzzy clustering can be used to handle the problem of vague boundaries of clusters. In fuzzy clustering, the requirement of crisp partition of the data is replaced by a weaker requirement of fuzzy partition, where the associations among data are represented by fuzzy relations. Fuzzy clustering can be applied to a wide variety of applications like image segmentation, pattern recognition, object recognition, and customer segmentation etc. The clustering output depends upon various parameters like distribution of points inside and outside the cluster, shape of the cluster and linear or non-linear separability.

Fuzzy clustering deals with uncertainty, fuzziness and vagueness. While discussing the uncertainty, another uncertainty arises which is the hesitation in defining the membership function of the object. Since the membership degrees are imprecise and it varies on person's choice, there is some kind of hesitation present which arises from the lack of precise knowledge in defining the membership function. This idea lead to another higher order fuzzy set called intuitionistic fuzzy set, introduced by Atanassov's in 1983 [1]. It takes into account the membership degree as well as non-membership degree. The Fuzzy C means (FCM) [2] algorithm, proposed by Bezdek (1981), is the first and most widely used algorithm for clustering because it has robust characteristics for ambiguity and can retain much more information than hard clustering methods. FCM has been successfully applied to feature analysis, clustering, and classifier designs in fields such as astronomy, geology, medical imaging, target recognition, and image segmentation. In case the data is noisy, FCM technique wrongly classifies noisy objects because of its abnormal feature data. FCM has a zero breakdown point i.e. even a single outlier may completely throw off the cluster centroids. One way to reduce the influence of noise points can be to make the memberships associated with them in the cluster smaller. However, due to constraint on membership matrix, noise points and outliers have significantly high membership values and they can severely affect the centroid parameter estimate. This drawback has motivated the researchers to seek alternative formulations. Possibilistic C-means (PCM), proposed by Krishnapuran and Keller [3] interprets clustering as a possibilistic partition. However, it caused clustering being struck in one or two clusters. Few works on clustering is reported in the literature on intuitionistic fuzzy sets. Studies on intuitionistic fuzzy set are done by Atanassov [4] on theory and application. Zhang and Chen [5] suggested a clustering approach where an intuitionistic fuzzy similarity matrix is transformed to interval valued fuzzy matrix. T. Chaira [6] recently proposed a novel intuitionistic fuzzy c-means (IFCM) algorithm using intuitionistic fuzzy set theory. This algorithm incorporates another uncertainty factor which is the hesitation degree that arises while defining the membership function and thus the cluster centers can converge to a desirable location than the cluster centers obtained using FCM. It also incorporates a new objective function which is the intuitionistic fuzzy entropy in the conventional FCM to maximize the good points in the class.

The effectiveness of the clustering method highly relies on the choice of distance metric. FCM and IFCM used Euclidean distance as a distance measure thus can only able to detect hyper spherical clusters. Researchers have proposed various other distance measures like Mahalanobis distance measure, Kernel based distance measure in the data space and in high dimensional feature space so that non-hyper spherical/nonlinear clusters can be detected[7,8]. D.M. Tsai and C. C. Lin [9] proposed FCM-σ and KFCM-σ by incorporating new distance metric that incorporates the distance variation of each individual data group to regularize the distance between a data point and the cluster centroid, and able to detect non-hyperspherical clusters with uneven densities for linear and nonlinear separation. But this method suffers with the same problem as conventional FCM i.e. it is not able to give efficient clustering results in the presence of noise.

This paper presents robust intutitionistic fuzzy c-means (IFCM-σ) and robust kernel intuitionistic fuzzy c-means (KIFCM-σ) algorithms by incorporating the distance measure proposed by D. M. Tsai and C. C. Lin. It is an effort made by authors to enhance the performance of IFCM. Proposed algorithms are the hybridization of IFCM, Kernel function and new distance metric; which are able to avoid the limitations of IFCM and FCM-σ in data space and in feature space. Two dimensional synthetic data-sets, standard data-sets referred from previous papers are used to evaluate the performance of proposed methods.

The organization of the paper is as follows: Section 2, briefly review Fuzzy C-Means (FCM), Intuitionistic FCM (IFCM) [6], FCM-σ, and KFCM-σ[9]. Section 3 describes the proposed algorithms, Robust Intuitionistic Fuzzy C Means (IFCM-σ) and Robust kernel Intuitionistic Fuzzy C Means (KIFCM-σ). Section 4 evaluates the performance of the proposed algorithms using synthetic data-set followed by concluding remarks in Section 5.

2. Background Information

This section briefly discusses the Fuzzy C-Means (FCM), Intuitionistic Fuzzy C means (IFCM), and FCM-σ algorithms. In this paper, the data-set is denoted by 'X', where X={x_1, x_2, x_3, …… x_n} specifying an image with 'n' pixels in M-dimensional space to be partitioned into 'c' clusters. Centroids of clusters are denoted by v_i and d_{ik} is the distance between x_k and v_i.

2.1 The Fuzzy C-Means Algorithm

FCM [2] is the most popular fuzzy clustering algorithm. It assumes that number of clusters 'c' is known in priori and minimizes the objective function (J_{FCM}) as:

$$J_{FCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} d_{ik}^{2} \tag{1}$$

Where $d_{ik} = \|x_k - v_i\|$, and u_{ik} is the membership of pixel 'x_k' in cluster 'i', which satisfies the following relationship:

$$\sum_{i=1}^{c} u_{ik} = 1;\;\; i = 1,2,\ldots..n \tag{2}$$

Here 'm' is a constant, known as the fuzzifier (or fuzziness index), which controls the fuzziness of the resulting partition. m=2 is used in this paper. Any norm $\|*\|$ can be used for calculating d_{ik}. Minimization of J_{FCM} is performed by a fixed point iteration scheme known as the alternating optimization technique. The conditions for local extreme for (1) and (2) are derived using Lagrangian multipliers:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \quad \forall\; k,i \tag{3}$$

where $1 \le i \le c;\ 1 \le k \le n$ and

$$v_i = \frac{\sum_{k=1}^{n}(u_{ik}^{m} x_k)}{\sum_{k=1}^{n}(u_{ik}^{m})} \quad \forall\, i \tag{4}$$

The FCM algorithm iteratively optimizes $J_{FCM}(U,V)$ with the continuous update of U and V, until $|U^{(l+1)} - U^{(l)}| <= \varepsilon$, where 'l' is the number of iterations. FCM works fine for the data-sets which are not corrupted with noise but if the data-set is noisy or distorted then it wrongly classifies noisy pixels because of its abnormal feature data and results in an incorrect membership and improper clustering.

2.2 Intuitionistic Fuzzy C Means (IFCM)

Intuitionistic fuzzy c-means clustering algorithm is based upon intuitionistic fuzzy set theory. Fuzzy set generates only membership function $\mu(x), x \in X$, whereas Intutitionistic fuzzy set (IFS) given by Atanassov considers both membership $\mu(x)$ and non-membership $v(x)$. An intuitionistic fuzzy set A in X, is written as:

$$A = \{x, \mu_A(x), v_A(x) | x \in X\}$$

Where $\mu_A(x) \to [0,1], v_A(x) \to [0,1]$ are the membership and non-membership degrees of an element in the set A with the condition:

$$0 \le \mu(x) + v_A(x) \le 1.$$

When $v_A(x) = 1 - \mu_A(x)$ for every x in the set A, then the set A becomes a fuzzy set. For all intuitionistic fuzzy sets, Atanassov also indicated a hesitation degree, $\pi_A(x)$, which arises due to lack of knowledge in defining the membership degree of each element x in the set A and is given by:

$$\pi_A(x) = 1 - \mu_A(x) - v_A(x); \qquad 0 \le \pi_A(x) \le 1$$

Due to hesitation degree, the membership values lie in the interval

$$[\mu_A(x), \mu_A(x) + \pi_A(x)]$$

Intuitionistic fuzzy c-means [6] objective function contains two terms: (i) modified objective function of conventional FCM using Intuitionistic fuzzy set and (ii) intuitionistic fuzzy entropy (IFE). IFCM minimizes the objective function as:

$$J_{IFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m} d_{ik}^{2} + \sum_{i=1}^{c} \pi_i^{*} e^{1-\pi_i^{*}} \tag{5}$$

$u_{ik}^{*} = u_{ik} + \pi_{ik}$, where u_{ik}^{*} denotes the intuitionistic fuzzy membership and u_{ik} denotes the conventional fuzzy membership of the k^{th} data in i^{th} class.

π_{ik} is hesitation degree, which is defined as:

$$\pi_{ik} = 1 - u_{ik} - (1 - u_{ik}{}^{\alpha})^{1/\alpha}, \alpha > 0,$$

and is calculated from Yager's intuitionistic fuzzy complement as under:

$N(x) = (1 - x^{\alpha})^{1/\alpha}, \alpha > 0$, thus with the help of Yager's intuitionistic fuzzy compliment , Intuitionistic Fuzzy Set becomes:

$$A_{\lambda}^{IFS} = \{x, \mu_A(x), (1 - \mu_A(x)^{\alpha})^{1/\alpha} | x \epsilon X\} \tag{6}$$

and

$$\pi_i^{*} = \frac{1}{N}\sum_{k=1}^{n} \pi_{ik}, \quad k \in [1, N]$$

Second term in the objective function is called intuitionistic fuzzy entropy (IFE). Initially the idea of fuzzy entropy was given by Zadeh in 1969. It is the measure of fuzziness in a fuzzy set. Similarly in the case of IFS, intuitionistic fuzzy entropy gives the amount of vagueness or ambiguity in a set. For intuitionistic fuzzy cases, if $\mu_A(x_i), v_A(x_i), \pi_A(x_i)$ are the membership, non-membership, and hesitation degrees of the elements of the set $X=\{x_1, x_2, \ldots .. x_n\}$, then intuitionistic fuzzy entropy, IFE that denotes the degree of intuitionism in fuzzy set, may be given as:

$$IFE(A) = \sum_{i=1}^{n} \pi_A(x_i) e^{[1-\pi_A(x_i)]}$$

Where $\pi_A(x_i) = 1 - \mu_A(x_i) - v_A(x_i)$.

IFE is introduced in the objective function to maximize the good points in the class. The goal is to minimize the entropy of the histogram of an image. Modified cluster centers are:

$$v_i^{*} = \frac{\sum_{k=1}^{n} u_{ik}^{*} x_k}{\sum_{k=1}^{n} u_{ik}^{*}} \tag{7}$$

At each iteration, the cluster center and membership matrix are updated and the algorithm stops when the updated membership and the previous membership i.e.

$max_{ik}|U_{ik}^{*new} - U_{ik}^{*prev}| < \varepsilon, \varepsilon$ is a user defined value.

As IFCM used Euclidian distance measure, hence can only detect hyper-spherical clusters in the data. It cannot be used for non-linearly separable data.

2.3 Fuzzy c-means with new distance metric (FCM-σ)

FCM-σ is proposed by D. M. Tsai and C. C. Lin [9] by incorporating a new distance measure into the conventional FCM. New distance metric is defined as:

$$\hat{d}_{ki}^{2} = \frac{\|x_k - v_i\|^2}{\sigma_i}$$

(8)

Where σ_i is the weighted mean distance of cluster 'i' and is given by

$$\sigma_i = \left\{\frac{\sum_{k=1}^{n} u_{ki}^{m}.\|x_k - v_i\|^2}{\sum_{k=1}^{n} u_{ki}^{m}}\right\}^{1/2} \quad (9)$$

It minimizes the objective function as:

$$J_{FCM-\sigma} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}.\frac{\|x_k - v_i\|^2}{\sigma_i} \quad (10)$$

Membership and modified cluster center equations are:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\hat{d}_{ik}}{\hat{d}_{jk}}\right)^{\frac{2}{m-1}}} \quad \forall\ k, i \quad (11)$$

where $1 \le i \le c;\ 1 \le k \le n$

and

$$v_i = \frac{\sum_{k=1}^{n}(u_{ik}^{m}x_k)}{\sum_{k=1}^{n}(u_{ik}^{m})} \quad \forall\ i \quad (12)$$

2.4 Kernel Fuzzy c-means with new distance metric (KFCM-σ)

Conventional IFCM and FCM-σ can only deal with linearly separable data points in observation space. The observed data-set can be transformed to higher dimensional feature space by applying a non-linear mapping function to achieve nonlinear separation. KFCM-σ[9] incorporates a new distance measure into the mapped feature space. New distance metric is defined as:

$$\widehat{\Phi}_{d_{kc}^2} = \frac{\|\Phi(x_k) - \Phi(v_i)\|^2}{\Phi_{\sigma_i}} = \frac{\Phi_{d_{kc}^2}}{\Phi_{\sigma_i}} \quad (13)$$

Where Φ_{σ_i} is the weighted mean distance of cluster 'i' in the mapped feature space and is given by

$$\Phi_{\sigma_i} = \left\{\frac{\sum_{k=1}^{n} u_{ki}^{m}.\Phi_{d_{kc}^2}}{\sum_{k=1}^{n} u_{ki}^{m}}\right\}^{1/2} \quad (14)$$

KFCM-σ minimizes objective function:

$$J_{KFCM-\sigma} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}\|\Phi(x_k) - \Phi(v_i)\|^2 \quad (15)$$

Here, u_{ik} is the membership of pixel 'x_k' in cluster 'i', which satisfies the following relationship:

$$\sum_{i=1}^{c} u_{ik} = 1;\ \ i = 1,2, \ldots .n$$

and $\|\Phi(x_k) - \Phi(v_i)\|^2$ is the square of distance between $\Phi(x_i)$ and $\Phi(v_k)$. The distance in the feature space is calculated through the kernel in the input space as follows:

$$\Phi_{d_{kc}^2} = \|\Phi(x_k) - \Phi(v_i)\|^2 = (\Phi(x_k) - \Phi(v_i)).(\Phi(x_k) - \Phi(v_i))$$
$$= \Phi(x_k).\Phi(x_k) - 2\Phi(x_k)\Phi(v_i) + \Phi(v_i).\Phi(v_i)$$
$$= K(x_k, x_k) - 2K(x_k, v_i) + K(v_i, v_i)$$

KFCM-σ used radial basis kernel (RBF):

$$K(x, y) = \exp\left(-\frac{\sum\left|x_i^a - x_j^a\right|^b}{h^2}\right) with\ a = 1, b = 2,\ \text{h=kernel width} \quad (16)$$

Minimizing eq(16) w.r.t. U, we get,

$$u_{ik} = 1 \Bigg/ \sum_{i=1}^{c}\left(\frac{\Phi_{d_{kc}^2}}{\Phi_{d_{ki}^2}}\right)^{1/(m-1)}$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m}.x_k}{\sum_{k=1}^{n} u_{ik}^{m}}$$

The KFCM-σ algorithm allows the clustering of non-hyper-spherically shaped data with uneven densities in the mapped feature space and achieved nonlinear separation of the data in the observation space.

Although, the results of FCM-σ and KFCM-σ are efficient for noiseless data, they do not perform well in the presence of noise.

3. The proposed Techniques

3.1 Robust Intuitionistic Fuzzy C-means (IFCM-σ)

Intuitionistic fuzzy c-means only takes into account the distance between objects and centroids. Hence, could not efficiently cluster non-spherically separable data. We are proposing a robust method by incorporating the distance metric which also considers the variation of points within the cluster into conventional IFCM algorithm to regularize the distance variation in each cluster.

IFCM-σ minimizes the objective function as:

$$J_{IFCM-\sigma} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}.\hat{d}_{ki}^2 + \sum_{i=1}^{c}\pi_i^* e^{1-\pi_i^*} \quad (17)$$

Where

$$\hat{d}_{ki}^2 = \frac{\|x_k - v_i\|^2}{\sigma_i}$$

and σ_i is the weighted mean distance of cluster i and is given by

$$\sigma_i = \left\{\frac{\sum_{k=1}^{n} u_{ki}^{*m} . \|x_k - v_i\|^2}{\sum_{k=1}^{n} u_{ki}^{*m}}\right\}^{1/2}$$

Here, $u_{ik}^* = u_{ik} + \pi_{ik}$, where u_{ik}^* denotes the IFCM-σ membership and u_{ik} denotes the FCM-σ membership of the kth data in ith class.

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\hat{d}_{ik}}{\hat{d}_{jk}}\right)^{\frac{2}{m-1}}} \quad \forall\ k, i \tag{18}$$

π_{ik} is hesitation degree, which is defined as:

$$\pi_{ik} = 1 - u_{ik} - (1 - u_{ik}{}^{\alpha})^{1/\alpha}, \ \alpha > 0$$

and

$$\pi_i^* = \frac{1}{N}\sum_{k=1}^{n} \pi_{ik}, \quad k \in [1, N]$$

Modified cluster centers are:

$$v_i^* = \frac{\sum_{k=1}^{n}(u_{ik}^{*m} x_k)}{\sum_{k=1}^{n}(u_{ik}^{*m})} \ \forall\ i \tag{19}$$

Second term in the objective function i.e. IFE is introduced to maximize the good points in the class. The goal is to minimize the entropy of the histogram of a data-set.

3.2 Kernel version of Intuitionistic Fuzzy C-means (KIFCM)

The present work proposes a way of increasing the accuracy of the intuitionistic fuzzy c-means by exploiting a kernel function in calculating the distance of data point from the cluster centers i.e. mapping the data points from the input space to a high dimensional space in which the distance is measured using a Radial Basis kernel function.

3.2.1 Formulation

We can construct the kernel version of the IFCM algorithm and modify its objective function with the mapping Φ as follows:

$$J_{KIFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\|\Phi(x_k) - \Phi(v_i)\|^2 + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*} \tag{20}$$

and $\|\Phi(x_k) - \Phi(v_i)\|^2$ is the square of distance between $\Phi(x_i)$ and $\Phi(v_k)$. The distance in the feature space is calculated through the kernel in the input space as follows:

$$\Phi_{d_{ki}^2} = \|\Phi(x_k) - \Phi(v_i)\|^2 = (\Phi(x_k) - \Phi(v_i)).(\Phi(x_k) - \Phi(v_i))$$
$$= \Phi(x_k).\Phi(x_k) - 2\Phi(x_k)\Phi(v_i) + \Phi(v_i).\Phi(v_i)$$
$$= K(x_k, x_k) - 2K(x_k, v_i) + K(v_i, v_i)$$

As we are adopting Radial basis kernel in the propose technique so:

$$K(x, y) = \exp\left(-\frac{\sum\left|x_i^a - x_j^a\right|^b}{h^2}\right) with\ 0 < a, b > 2$$

Where h is defined as kernel width and it is a positive number, then $K(x, x) = 1$. Thus eq.(20) can be written as:

$$J_{KIFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\|1 - K(x_k, v_i)\|^2 + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*}$$

$$J_{KIFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\left\|1 - \exp\left(-\frac{\sum\left|x_i^a - x_j^a\right|^b}{h^2}\right)\right\|^2 + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*} \tag{21}$$

Given a set of points X, we minimize J_{KIFCM} in order to determine u_{ik}^* & v_i. We adopt an alternating optimization approach to minimize J_{KIFCM} and need the following theorem:

3.2.2 Theorem 1: The necessary conditions for minimizing J_{KIFCM} under the constraint of U, we get:

Minimizing (21) under the constraint of U, we get:

$$u_{ik}^* = 1 \Big/ \sum_{i=1}^{c}\left(\frac{\Phi_{d_{kc}^2}}{\Phi_{d_{ki}^2}}\right)^{1/(m-1)} \tag{22}$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{*m} . x_k}{\sum_{k=1}^{n} u_{ik}^{*m}} \tag{23}$$

Proof: We differentiate J_{KIFCM} with respect to u^*_{ik} & v_i and set the derivatives to zero. Thus, we get (22) & (23). The details are given in Appendix A.

3.3 Robust Kernel version of Intuitionistic Fuzzy C-means (KIFCM-σ)

We are now in a position to construct the kernel version of robust intutionistic fuzzy c-means (IFCM-σ). The objective function of the KIFCM-σ will be:

$$J_{KIFCM-\sigma} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\widehat{\Phi}_{d_{kc}^2} + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*} \tag{24}$$

where

$$\widehat{\Phi}_{d_{kc}^2} = \frac{\|\Phi(x_k) - \Phi(v_i)\|^2}{\Phi_{\sigma_i}} = \frac{\Phi_{d_{kc}^2}}{\Phi_{\sigma_i}}$$

Where Φ_{σ_i} is the weighted mean distance of cluster i in the mapped feature space and is given by

$$\Phi_{\sigma_i} = \left\{\frac{\sum_{k=1}^{n} u_{ki}^{*m}.\Phi_{d_{kc}^2}}{\sum_{k=1}^{n} u_{ki}^{*m}}\right\}^{1/2}$$

Here, $u^*_{ik} = u_{ik} + \pi_{ik}$, where $u^*_{ik}(u_{ik})$ denotes the KIFCM-σ (IFCM-σ) fuzzy membership of the k[th] data in i[th] class.

$$u_{ik} = 1 \Bigg/ \sum_{i=1}^{c}\left(\frac{\widehat{\Phi}_{d_{kc}^2}}{\widehat{\Phi}_{d_{ki}^2}}\right)^{1/(m-1)} \tag{25}$$

π_{ik} is hesitation degree, which is defined as:

$$\pi_{ik} = 1 - u_{ik} - (1 - u_{ik}{}^{\alpha})^{1/\alpha}$$

and

$$\pi_i^* = \frac{1}{N}\sum_{k=1}^{n} \pi_{ik}, \quad k \in [1, N]$$

Modified cluster centers are:

$$v_i^* = \frac{\sum_{k=1}^{n}(u_{ik}^{*m} x_k)}{\sum_{k=1}^{n}(u_{ik}^{*m})} \quad \forall\, i \tag{26}$$

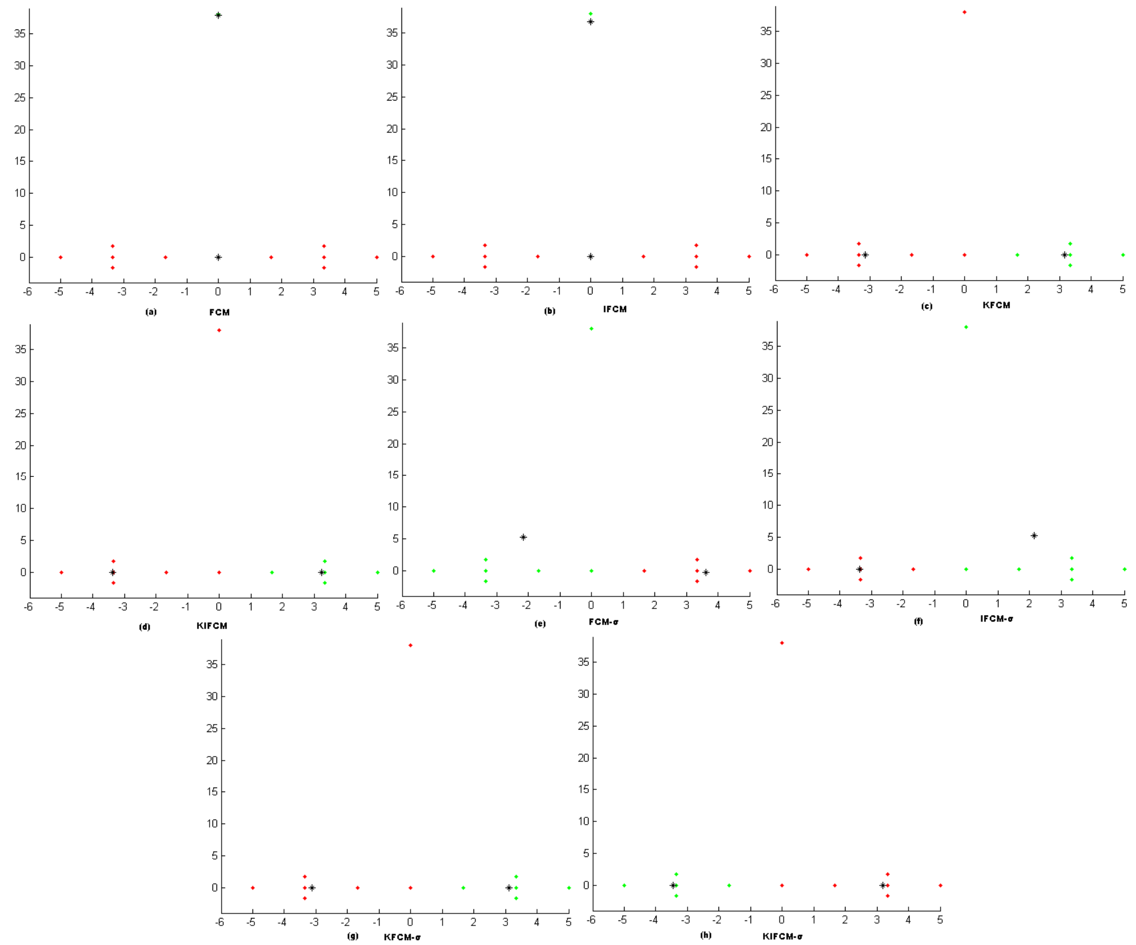

Fig. 1: (a) Clustering result of FCM (b) Clustering result of IFCM with α=0.7, (c) Clustering result of KFCM with h(kernel width)=6 (d) Clustering result of KIFCM with α=1.7, h=11 (e) Clustering result of FCM-σ (f) Clustering result of IFCM-σ with α=0.8 (g) Clustering result of KFCM-σ with h=10 (h) Clustering result of KIFCM-σ with α=1.6, h=10. Centroids are shown with '*' symbol

4. Results and Simulations

We now give some experimental results to show the effectiveness of the eight methods FCM, IFCM, KFCM, KIFCM, FCM-σ, IFCM-σ, KFCM-σ, and KIFCM-σ. Experiments are implemented and simulated using MATLAB Version 7.0. We considered following common parameters: m=2 which is a common choice for fuzzy clustering, ε = 0.03, maximum iterations=100.

Example 1:

Data-set: D11, D12 (referred from [10]). D11 is a noiseless data-set of points $\{x_i\}_{i=1}^{11}$. D12 is the union of D11 and an outlier x12.

Algorithms: FCM, IFCM, KFCM, KIFCM, FCM-σ, IFCM-σ, KFCM-σ, and KIFCM-σ

Number of Clusters: 2 (Identical data with Noise)

Fig. 1 shows clustering results of FCM, IFCM, KFCM, KIFCM, FCM-σ, IFCM-σ, KFCM-σ, and KIFCM-σ. '*' symbol shows centroids. Table 1 and 2 list the centroids generated with the algorithms for D12. It is observed from the figure that FCM and IFCM could not detect the original clusters and their performance is badly affected by the presence of noise whereas kernelized methods detected the clusters but the centroid location is still affected with noise. It is observed from the Fig.1, Table 1, and Table 2 that proposed algorithms especially KIFCM-σ detected almost original centroid locations. The ideal (true) centroids of data-set D11 are:

$$V_{ideal} = \begin{bmatrix} -3.34 & 0 \\ 3.34 & 0 \end{bmatrix}$$

To show the effectiveness of the proposed algorithm, we also calculated the error, $E_* = \|V_{ideal} - V_*\|^2$, where * is FCM/IFCM/ KFCM/KIFCM/FCM-σ/IFCM-σ/KFCM-σ/KIFCM-σ. Table 3 lists the error percentage. Clearly from Fig.1 and Table 3, it is observed that proposed methods can produce more accurate centroids than other methods and are highly robust against noise.

Example 2:

Data-set: DUNN, 2-dimentional Square data [11] (142 points)

Algorithm: FCM, IFCM, KFCM, KIFCM, FCM-σ, IFCM-σ, KFCM-σ, and KIFCM-σ

Number of Clusters: 2

DUNN Square data-set [11] with noise is used in this example. It consisted of one small and one big cluster of square shape. We have increased the quantity of core points to increase the density of data-set. We tried various values of α & h. Fig. 2 shows clustering result of the algorithms and Tables 1 & 2 listed centroids. It is observed that results of FCM in data space and by incorporating new distance are not very good, whereas kernalized FCM methods have improved the performance in comparison to FCM methods in data space. But proposed methods, especially KIFCM-σ detected almost correct centroid locations.

For the performance comparison we also computed error. The ideal (true) centroids of DUNN data-set are:

$$V_{ideal} = \begin{bmatrix} 5.25 & 0.25 \\ 17 & 0 \end{bmatrix}$$

Table 3 lists the error percentage. Clearly from Fig.2, Table 1, Table 2 and Table 3, it is observed that proposed methods can produce more accurate centroids than other methods and are highly robust against noise.

Example 3:

Data-set: BENSAID [12] 2-dimentional data (213 points)

Algorithm: FCM, IFCM, KFCM, KIFCM, FCM-σ, IFCM-σ, KFCM-σ, and KIFCM-σ

Number of Clusters: 3 clusters with 2 small and 1 big size clusters

BENSAID's [12] two-dimensional data-set consisting of one big and two small size clusters is used in first example. We have saved the structure of this set but have increased count of core points and added uniform noise to it, which is distributed over the region [0,120] x [10,80]. Fig. 3 shows the clustering results of all the methods. Centroids of the clusters are plotted with '*.

The clustering results of FCM method, whether in the observed space or in the feature space or including the new distance, in all the cases is highly affected with the presence of noise. Centroids in all the cases are diverted towards noise. But proposed methods gave efficient results in the presence of noise and results are not that much affected as in other cases.

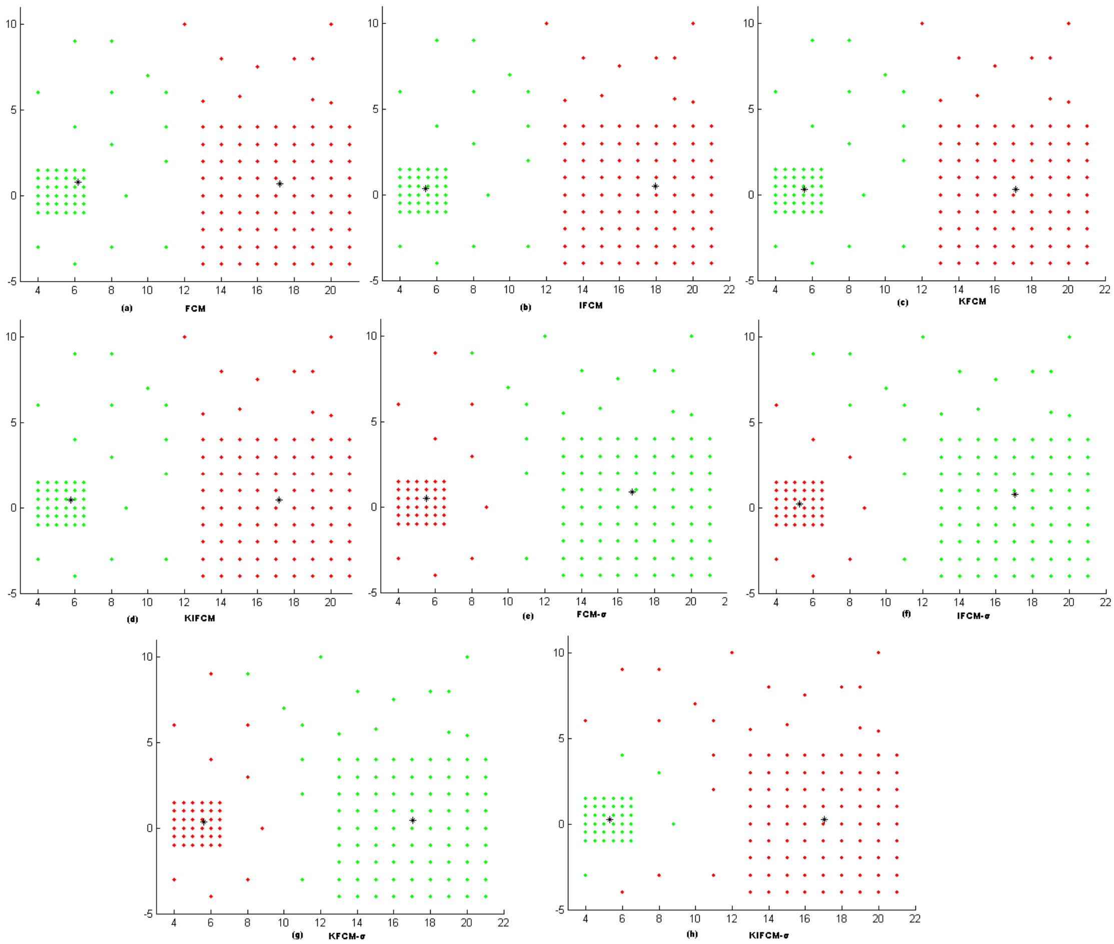

Fig. 2: (a) Clustering result of FCM (b) Clustering result of IFCM with α=2, (c) Clustering result of KFCM with h(kernel width)=6 (d) Clustering result of KIFCM with α=1, h=6 (e) Clustering result of FCM-σ (f) Clustering result of IFCM-σ with α=2 (g) Clustering result of KFCM-σ with h=8 (h) Clustering result of KIFCM-σ with α=1.6, h=6. Centroids are shown with '*' symbol

Table 1 Centroids produced by FCM, IFCM, KFCM, KIFCM for D12, Square and Bensaid Data-sets

Data Sets	Name of the Algorithms							
D12	FCM (m=2)		IFCM (m=2,α=0.7)		KFCM (m=2,h =6)		KIFCM (m=2,h =11, α=1.7)	
	0	0	0	0	3.14	0	3.42	-0.006
	0	37.96	0	36.80	-3.12	0.001	-3.33	0.001
Square Data-set	FCM (m=2)		IFCM (m=2,α=2)		KFCM (m=2,h =6)		KIFCM (m=2,h =6, α=1)	
	6.18	0.78	5.38	0.36	5.54	0.31	5.54	0.31
	17.2	0.70	17.95	0.493	17	0.29	17.1	0.29
Bensaid Data-set	FCM (m=2)		IFCM (m=2,α=2)		KFCM (m=2,h =100)		KIFCM (m=2,h =300, α=3)	
	10.64	49.12	4.39	48.81	11.56	49.25	3.76	48.82
	57.75	51.49	58.35	51.49	58.08	51.44	58.74	51.43
	109.21	48.69	111.57	48.41	109.48	48.63	111.74	48.38

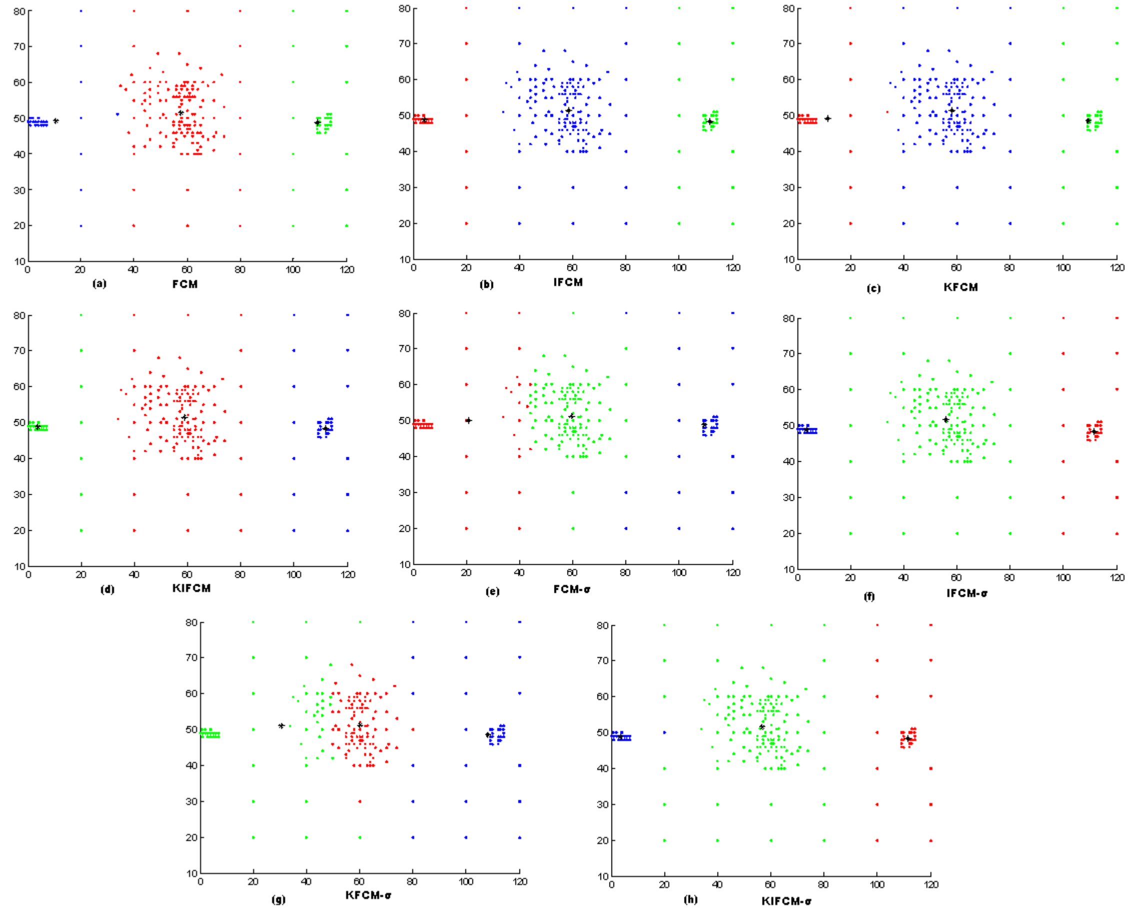

Fig. 3: (a) Clustering result of FCM (b) Clustering result of IFCM with α=2, (c) Clustering result of KFCM with h(kernel width)=100 (d) Clustering result of KIFCM with α=3, h=300 (e) Clustering result of FCM-σ (f) Clustering result of IFCM-σ with α=2.6 (g) Clustering result of KFCM-σ with h=85 (h) Clustering result of KIFCM-σ with α=2, h=300. Centroids are shown with '*' symbol

Table 2 Centroids produced by FCM-σ, IFCM-σ, KFCM-σ, and KIFCM-σ for D12, Square and Bensaid Data-sets

Data Sets	Name of the Algorithms							
D12	FCM-σ (m=2)		IFCM-σ (m=2, α=0.8)		KFCM-σ (m=2,h =10)		KIFCM-σ (m=2,h =10, α=1.6)	
	3.62	0.280	2.1487	5.247	3.1813	0.0003	3.31	0
	-2.14	5.2494	-3.39	-0.088	-3.465	-0.0002	-3.35	0.001
Square Data-set	FCM-σ (m=2)		IFCM-σ (m=2, α=2)		KFCM-σ (m=2,h =8)		KIFCM-σ (m=2,h =6, α=1.6)	
	5.5	0.51	5.26	0.22	5.6	0.36	5.28	0.24
	16.7	0.88	17	0.77	17	0.46	17	0.28
Bensaid Data-set	FCM-σ (m=2)		IFCM-σ (m=2, α=2.6)		KFCM-σ (m=2,h =85)		KIFCM-σ (m=2,h =300, α=2	
	20.75	50.01	3.46	48.83	30.51	51.07	3.52	48.83
	59.67	51.27	55.99	51.64	60.15	51.33	56.93	51.61
	109.41	48.85	111.63	48.36	108.23	48.64	111.65	48.39

Table 3 Error %age

Name of the Algorithm	Error %age					
	D12			Square Data-set		
	Error in Cluster1	Error in Cluster2	Average error	Error in Cluster1	Error in Cluster2	Average error
FCM	Does not recognize clusters			1.14	0.53	0.835
IFCM	Does not recognize clusters			0.029	1.14	0.584
KFCM	0.04	0.0484	0.0442	0.087	0.084	0.085
KIFCM	0.006	0.0001	0.003	0.087	0.084	0.085
FCM-σ	0.1568	28.89	14.52	0.130	1.674	0.902
IFCM-σ	28.89	0.0102	14.45	0.001	0.592	0.296
KFCM-σ	0.0256	0.0144	0.02	0.134	0.211	0.173
KIFCM-σ	0.0009	0.0001	0.0005	0.001	0.078	0.039

5. Conclusion

In this paper, we proposed a Robust approach to Intuitionistic Fuzzy C-Means (IFCM) with a new distance metric that incorporates the distance variation of data-points within each cluster to IFCM in the observed space and KIFCM in the feature space. The proposed IFCM with the new distance in the observed and feature space significantly show improvement over FCM and IFCM. To test the performance of proposed methods, it has been applied to various synthetic and standard data-sets like Bensaid data-set, DUNN data-set. It has been found that proposed algorithms are highly robust to noise and outliers.

Appendix A

Kernel version of Intuitionistic Fuzzy C-means (KIFCM)

A.1 Proof of Kernel based Intuitionistic fuzzy C-means

In this appendix, we give the proof of the kernel based intuitionistic fuzzy c-means clustering which is a kernelized version of intuitionistic Fuzzy c-means clustering algorithm. The problem of minimization of objective function J_{KIFCM} {given in (21)} subjected to the constraint specified by (2) is solved by minimizing a constraint free objective function defined as:

$$J_{KIFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\|\Phi(x_k)-\Phi(v_i)\|^2 + \sum_{i=1}^{c}\pi_i^* e^{1-\pi_i^*} + \sum_{i=1}^{c}\lambda_i\left(\sum_{k=1}^{n}u_{ik}^*-1\right)$$

where $\lambda_i (i = 1,2,3, \ldots . c)$ are Langrangian multipliers

By taking the partial derivatives of J_{KIFCM} with respect to u_{ik}^* & v_i , yields the solution for the problem:

$$\begin{aligned}\|\Phi(x_k)-\Phi(v_i)\|^2 &= \big(\Phi(x_k)-\Phi(v_i)\big).(\Phi(x_k) - \Phi(v_i))\\ &= \Phi(x_k).\Phi(x_k) - 2\Phi(x_k)\Phi(v_i) + \Phi(v_i).\Phi(v_i)\\ &= K(x_k,x_k) - 2K(x_k,v_i) + K(v_i,v_i)\end{aligned}$$

If we adopt Radial Basis kernel in the propose technique then:

$$K(x,y) = \exp\left(-\frac{\sum\left|x_i^a - x_j^a\right|^b}{h^2}\right);\ 0 < a,b > 2$$

Let a=1 and b=2, then

$$K(x,y) = \exp\left(-\frac{\|x_k - v_i\|^2}{h^2}\right)$$

Where h is defined as kernel width and it is a positive number, then $K(x,x) = 1$. Hence,

$$\|\Phi(x_k)-\Phi(v_i)\|^2 = (1 - K(x_k,v_i))$$

So,

$$J_{KIFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}(1-K(x_k,v_i)) + \sum_{i=1}^{c}\pi_i^* e^{1-\pi_i^*} + \sum_{i=1}^{c}\lambda_i\left(\sum_{k=1}^{n}u_{ik}^*-1\right)$$

$$J_{KIFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\left(1-\exp\left(-\frac{\|x_k - v_i\|^2}{h^2}\right)\right) + \sum\nolimits_{i=1}^{c} \pi_i^* e^{1-\pi_i^*} + \sum_{i=1}^{c}\lambda_i\left(\sum_{k=1}^{n} u_{ik}^* - 1\right) \quad (26)$$

A.1.1 Partial derivative of J_{KIFCM} with respect to v_i

The partial derivative of J_{KIFCM} with respect to v_i is:

$$\frac{\partial J}{\partial v_i} = \sum_{k=1}^{n} u_{ik}^{*m} \times \left(-\exp\left(-\frac{\|x_k - v_i\|^2}{h^2}\right)\right) \times \frac{(x_k - v_i)}{h^2} \quad (27)$$

Equating (27) to zero leads to:

$$\frac{\partial J}{\partial v_i} = 0$$

$$\Rightarrow \sum_{k=1}^{n} u_{ik}^{*m} K(x_k, v_i)\,(x_k - v_i) = 0$$

$$\Rightarrow \sum_{k=1}^{n} u_{ik}^{*m} K(x_k, v_i)\, x_k = \sum_{k=1}^{n} u_{ik}^{*m} K(x_k, v_i)\; v_i$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{*m} \times x_k}{\sum_{k=1}^{n} u_{ik}^{*m}}$$

A.1.2 Partial derivative of J_{KIFCM} with respect to u_{ik}^*

$$\frac{\partial J}{\partial u_{ik}^*} = m u_{ik}^{*(m-1)}\left(1 - K(x_k, v_i)\right) + \lambda_i \quad (28)$$

Equating (28) to zero leads to the following:

$$\frac{\partial J}{\partial u_{ik}^*} = 0$$

$$\Rightarrow m u_{ik}^{*(m-1)} \times \left(1 - K(x_k, v_i)\right) + \lambda_i = 0$$

$$\Rightarrow u_{ik}^* = \left(-\frac{\lambda_i}{m}\right)^{1/m-1} \times \left(\frac{1}{1 - K(x_k, v_i)}\right)^{1/m-1}$$

To fulfill the constraint (2),

$$\Rightarrow u_{ik}^* = \frac{u_{ik}^*}{\sum_{i=1}^{c} u_{ik}^*} \quad (29)$$

In view of (29), equation (28) can be written as:

$$\Rightarrow u_{ik}^* = \frac{\left(\frac{1}{(1 - K(x_k, v_i))}\right)^{\frac{1}{m-1}}}{\sum_{i=1}^{c}\left(\frac{1}{(1 - K(x_k, v_i))}\right)^{\frac{1}{m-1}}}$$

By using equation,

$$\Phi_{d_{ki}^2} = \|\Phi(x_k) - \Phi(v_i)\|^2,$$

above equation can be re-written as:

$$u_{ik}^* = 1 \Big/ \sum_{i=1}^{c}\left(\frac{\Phi_{d_{kc}^2}}{\Phi_{d_{ki}^2}}\right)^{1/(m-1)}$$

References

[1] K.T. Atanassov's, Intuitionistic fuzzy sets, VII ITKR's Session, Sofia, 983 (Deposed in Central Science – Technology Library of Bulgaria Academy of Science – 1697/84).

[2] J.C. Bezdek (1981), "Pattern Recognition with Fuzzy Objective Function Algorithm", Plenum, NY.

[3] R. Krishnapuram and J. Keller (1993), "A Possibilistic Approach to Clustering", IEEE Trans. on Fuzzy Systems, vol .1. No. 2, pp.98-1 10.

[4] K.T. Atanassov, Intuitionistic Fuzzy Sets Past, Present and Future, www.eusflat.org/publications/proceedings/.2003/./4Atanassov.pdf.

[5] H.M. Zhang, Z.S.Q. Chen, On clustering pproach to intuitionistic fuzzy sets, Control and Decision 22 (2007) 882–888.

[6] T. Chaira, "A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images", Applied Soft computing 11(2011) 1711-1717.

[7] D. Q. Zhang and S. C. Chen, "Clustering incomplete data using kernel based fuzzy c-means algorithm", Neural Processing Letters 18(2003) 155-162.

[8] D. Q. Zhang and S. C. Chen, "A novel kernelized fuzzy c-means algorithm with application in medical image segmentation", Artificial Intelligence in Medicine 32(2004) 37-50.

[9] D. M. Tsai and C. C. Lin, "Fuzzy C-means based clustering for linearly and nonlinearly separable data", Pattern Recognition, 44(2011) 1750-1760.

[10] Pal N.R., K. Pal, J. Keller and J. C. Bezdek ," A Possibilistic Fuzzy c- Means Clustering Algorithm", IEEE Trans. on Fuzzy Systems, vol 13 (4),pp 517-530,2005.

[11] Dunn, J., 1974. A fuzzy relative of the ISODATA process and its use in detecting

compact well separated clusters. J. Cybernet. 3, 32–57.

[12] Bensaid A. M., L.O. hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, R. F. Murtagh, "Validity-guided clustering with applications to image segmentation", IEEE trans. Fuzzy Systems 4 (2) (1996) 112-123.