Pattern Recognition Letters 34 (2013) 163–175

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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



# A robust kernelized intuitionistic fuzzy c-means clustering algorithm in segmentation of noisy medical images

Prabhjot Kaur [a,*], A.K. Soni [b,1], Anjana Gosain [c,2]

[a] Department of Information Technology, Maharaja Surajmal Institute of Technology, C-4, Janakpuri, New Delhi 110 058, India
[b] Department of Computer Science, Sharda University, Greater Noida, Uttar Pradesh, India
[c] Department of Information Technology, USIT, Guru Gobind Singh Indraprastha University, New Delhi, India

ARTICLE INFO

*Article history:*
Received 22 March 2011
Available online 4 October 2012
Communicated by S. Sarkar

*Keywords:*
Fuzzy clustering
Intuitionistic fuzzy c-means
Robust image segmentation
RBF kernel based intuitionistic fuzzy c-means
Fuzzy c-means

ABSTRACT

This paper presents an automatic effective intuitionistic fuzzy c-means which is an extension of standard intuitionisitc fuzzy c-means (IFCM). We present a model called RBF Kernel based intuitionistic fuzzy c-means (KIFCM) where IFCM is extended by adopting a kernel induced metric in the data space to replace the original Euclidean norm metric. By using kernel function it becomes possible to cluster data, which is linearly non-separable in the original space, into homogeneous groups by transforming the data into high dimensional space. Proposed clustering method is applied on synthetic data-sets referred from various papers, real data-sets from Public Library UCI, simulated and Real MR brain images. Experimental results are given to show the effectiveness of proposed method in contrast to conventional fuzzy c-means, possibilistic c-means, possibilistic fuzzy c-means, noise clustering, kernelized fuzzy c-means, type-2 fuzzy c-means, kernelized type-2 fuzzy c-means, and intuitionistic fuzzy c-means.



## 1. Introduction

Many neurological conditions alter the shape, volume, and distribution of brain tissue; magnetic resonance imaging (MRI) is the preferred imaging modality for examining these conditions. MRI is an important diagnostic imaging technique for the early detection of abnormal changes in tissues and organs. MRI possesses good contrast resolution for different tissues and has advantages over CT for brain studies due to its superior contrast properties. Therefore, the majority of research in medical image segmentation concerns MR images.

Image segmentation plays an important role in image analysis and computer vision. The goal of image segmentation is partitioning of an image into a set of disjoint regions with uniform and homogeneous attributes such as intensity, color, tone, etc. In images, the boundaries between objects are blurred and distorted due to the imaging acquisition process. Furthermore, object definitions are not always crisp and knowledge about the objects in a scene may be vague. Fuzzy set theory and fuzzy logic are ideally suited to deal with such uncertainties. Fuzzy sets were introduced in 1965 by Lofti Zadeh with a view to reconcile mathematical modeling and human knowledge in the engineering sciences. Medical images generally have limited spatial resolution, poor contrast, noise, and non-uniform intensity variation. The fuzzy c-means (FCM) (Bezdek, 1981), algorithm, proposed by Bezdek, is the most widely used algorithm in image segmentation. FCM is the extension of the fuzzy ISODATA algorithm proposed by DUNN (Dunn, 1974). 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. An image can be represented in various feature spaces and the FCM algorithm classifies the image by grouping similar data points in the feature space into clusters. In case the image is noisy or distorted then FCM technique wrongly classify noisy pixels because of its abnormal feature data which is the major limitation of FCM. Various approaches are proposed by researchers to compensate this drawback of FCM.

Dave proposed the idea of a noise cluster to deal with noisy data using the technique, known as noise clustering (Dave and Krishnapuram, 1997). Another similar technique, PCM, proposed by Krishnapuram and Keller (1993) interpreted clustering as a possibilistic partition. However, it caused clustering being stuck in one or two clusters. Yang and Chung (2009) developed a robust clustering technique by deriving a novel objective function for FCM. Kang et al. (2009) proposed another technique which modified FCM objective function by incorporating the spatial neighborhood information into the standard FCM algorithm. Rhee and Hwang (2001) proposed type 2 fuzzy clustering. Type 2 fuzzy set is the fuzziness

* Corresponding author. Tel.: +91 9810665064/9810165064.
*E-mail addresses:* thisisprabhjot@gmail.com (P. Kaur), ak.soni@sharda.ac.in (A.K. Soni), anjana_gosain@hotmail.com (A. Gosain).
[1] Tel.: +91 9990021800.
[2] Tel.: +91 9811055716.

0167-8655/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2012.09.015

in a fuzzy set. In this algorithm, the membership value of each pattern in the image is extended as type 2 fuzzy memberships by assigning membership grades (triangular membership function) to type 1 fuzzy membership.

While discussing the uncertainty, another uncertainty arises, which is the hesitation in defining the membership function of the pixels of an image. Since the membership degrees are imprecise and it varies on person's choice, hence 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 which was introduced by Atanassov's in 1983. It took into account the membership degree as well as non-membership degree. Few works on clustering is reported in the literature on intuitionistic fuzzy sets. Zhang and Chen (2009) suggested a clustering approach where an intuitionistic fuzzy similarity matrix is transformed to interval valued fuzzy matrix. Recently, Chaira (2011) proposed a novel intuitionistic fuzzy c-means algorithm using intuitionistic fuzzy set theory. This algorithm incorporated another uncertainty factor which is the hesitation degree that aroused while defining the membership function.

This paper proposes a new model called RBF kernel based intuitionistic fuzzy c-means (KIFCM), which is an extension of intuitionistic fuzzy c-means (IFCM) by adopting a kernel induced metric in the data space to replace the original Euclidean norm metric. By replacing the inner product with an appropriate 'kernel' function, one can implicitly perform a non-linear mapping to a high dimensional feature space in which the data is more clearly separable.

The organization of the paper is as follows: Section 2, briefly review fuzzy c-means (FCM), type-2 fcm (T2FCM) (Rhee and Hwang, 2001) and intutionistic fuzzy c-means (IFCM) (Chaira, 2011), possibilistic c-means (PCM), possibilistic fuzzy c-means (PFCM), and noice clustering (NC). Section 3 describes the proposed algorithm; RBF kernel based intuitionistic fuzzy c means (KIFCM). Section 4 evaluates the performance of the propose algorithm using synthetic data-sets, simulated and real medical images followed by concluding remarks in Section 5.

## 2. Background information

This section briefly discusses the fuzzy c-means (FCM), type-2 fuzzy c-means, and intuitionistic fuzzy c means (IFCM), PCM, PFCM, and NC algorithms. In this paper, the data-set is denoted by '$X$', where $X = \{x_1, x_2, x_3, \ldots, x_n\}$ specifying an image with '$n$' pixels in $M$-dimensional space to be partioned 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)

FCM 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:

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

Here '$m$' is a constant, known as the fuzzifier (or fuzziness index), which controls the fuzziness of the resulting partition. Any norm $\|\cdot\|$ 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 $1 \leqslant i \leqslant c$; $1 \leqslant k \leqslant n$ and

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

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

### 2.2. The type-2 fuzzy c-means algorithm (T2FCM)

Rhee and Hwang (2001) extended the type-1 membership values (i.e. membership values of FCM) to type-2 by assigning a membership function to each membership value of type-1. Their idea is based on the fact that higher membership values should contribute more than smaller memberships values, when updating the cluster centers. Type-2 memberships can be obtained as per following equation:

$$a_{ik} = u_{ik} - \frac{1 - u_{ik}}{2} \tag{5}$$

where $a_{ik}$ and $u_{ik}$ are the type-2 and type-1 fuzzy membership respectively. From (5), the type-2 membership function area can be considered as the uncertainty of the type-1 membership contribution when the center is updated. Substituting (5) for the memberships in the center update equation of the conventional FCM method gives the following equation for updating centers.

$$v_i = \frac{\sum_{k=1}^{n}(a_{ik})^{m} x_k}{\sum_{k=1}^{n}(a_{ik})^{m}} \tag{6}$$

During the cluster center updates, the contribution of a pattern that has low memberships to a given cluster is relatively smaller when using type-2 memberships and the memberships may represent better typicality. Cluster centers that are estimated by type-2 memberships tend to have more desirable locations than cluster centers obtained by type-1 FCM method in the presence of noise. T2FCM algorithm is identical to the type-1 FCM algorithm except Eq. (6). 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.

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

Although T2FCM has proven effective for spherical data, it fails when the data structure of input patterns is non-spherical and complex.

### 2.3. The intuitionistic fuzzy c-means algorithm (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 $\nu(x)$. An intuitionistic fuzzy set $A$ in $X$, is written as:

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

where $\mu_A(x) \to [0,1]$, $\nu_A(x) \to [0,1]$ are the membership and non-membership degrees of an element in the set $A$ with the condition: $0 \leqslant \mu(x) + \nu_A(x) \leqslant 1$.

When $\nu_A(x) = 1 - \mu_A(x)$ 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:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x); \quad 0 \leqslant \pi_A(x) \leqslant 1.$$

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

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

Intuitionistic fuzzy c-means (Chaira, 2011) 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:

$$J_{IFCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m} d_{ik}^2 + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*} \tag{7}$$

$u_{ik}^* = u_{ik} + \pi_{ik}$, 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.

$\pi_{ik}$ is hesitation degree, which is defined as:

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

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:

$$A_{\lambda}^{IFS} = \{x, \mu_A(x), (1 - \mu_A(x)^{\alpha})^{1/\alpha} | x \in X\} \tag{8}$$

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 $\mu_A(x_i)$, $\nu_A(x_i)$, $\pi_A(x_i)$ 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 $\pi_A(x_i) = 1 - \mu_A(x_i) - \nu_A(x_i)$.

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:

$$v_i^* = \frac{\sum_{k=1}^{n} u_{ik}^* x_k}{\sum_{k=1}^{n} u_{ik}^*} \tag{9}$$

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.

*2.4. Possibilistic c-means clustering (PCM)*

To avoid the noise sensitivity problem of FCM, Krishnapuram and Keller (1993) relaxed the column sum constraint

$$\sum_{k=1}^{c} u_{ki} = 1; \quad i = 1, 2, \ldots, n$$

in case of FCM and proposed a possibilistic approach to clustering by minimizing objective function as:

$$J_{PCM}(U, V) = \sum_{k=1}^{c}\sum_{i=1}^{n} u_{ki}^m d_{ki}^2 + \sum_{k=1}^{c} \eta_k \sum_{i=1}^{n} (1 - u_{ki})^m$$

where $\eta_k$ are suitable positive numbers. The first term tries to reduce the distance from data points to the centroids as low as possible and second term forces $u_{ki}$ to be as large as possible, thus avoiding the trivial solution. The updating of centroids is same as that in FCM but the membership matrix of PCM is updated as:

$$u_{ki} = \frac{1}{1 + \left(\frac{\|x_i - v_k\|^2}{\eta_k}\right)^{\frac{1}{m-1}}}$$

PCM sometimes helps when data is noisy. It is very much sensitive to initializations and sometimes results into overlapping or identical clusters.

*2.5. Possibilistic fuzzy c-means clustering (PFCM)*

Pal et al. (2005) integrates the fuzzy approach with the possibilistic approach and hence, it has two types of memberships, viz. a possibilistic ($t_{ki}$) membership that measures the absolute degree of typicality of a point in any particular cluster and a fuzzy membership ($u_{ki}$) that measures the relative degree of sharing of point among the clusters. PFCM minimizes the objective function as:

$$J_{PFCM}(U, V, T) = \sum_{k=1}^{c}\sum_{i=1}^{n} (a u_{ki}^m + b t_{ki}^{\eta}) d_{ki}^2 + \sum_{k=1}^{c} \gamma_k \sum_{i=1}^{n} (1 - t_{ki})^{\eta}$$

subject to the constraint that

$$\sum_{k=1}^{c} u_{ki} = 1 \quad \forall i$$

Here, $a > 0$, $b > 0$, $m > 1$, and $\eta > 1$. The constants '$a$' and '$b$' define the relative importance of fuzzy membership and typicality values in the objective function. The minimization of objective function gives the following conditions:

$$u_{ki} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ki}}{d_{ji}}\right)^{\frac{2}{m-1}}} \quad \forall\ k, i$$

and

$$t_{ki} = \frac{1}{1 + \left(\frac{b}{\gamma_k} d_{ki}^2\right)^{\frac{1}{\eta-1}}} \quad \forall k$$

and

$$v_k = \frac{\sum_{i=1}^{n} (a u_{ki}^m + b t_{ki}^{\eta}) x_i}{\sum_{i=1}^{n} (a u_{ki}^m + b t_{ki}^{\eta})}$$

Though PFCM is found to perform better than FCM and PCM but when two highly unequal sized clusters with outliers are given, it fails to give desired results.

*2.6. Noise clustering (NC)*

Noise clustering was introduced by Dave and Krishnapuram (1997) to overcome the major deficiency of the FCM algorithm i.e. its noise sensitivity. He gave the concept of "noise prototype", which is a universal entity such that it is always at the same distance from every point in the data-set. Let '$v_k$' be the noise prototype and '$x_i$' be any point in the data-set such that $v_k, x_i \,\bar{\subset}\, Rp$. Then noise prototype is the distance, $d_{ki}$, given by:

$$d_{ki} = \delta, \quad \forall i$$

The NC algorithm considers noise as a separate class. The membership $u_{*i}$ of $x_i$ in a noise cluster is defined as

$$u_{*i} = 1 - \sum_{k=1}^{c} u_{ki}$$

NC reformulates FCM objective function as:

$$J_{NC(U,V)} = \sum_{k=1}^{c+1}\sum_{i=1}^{N} u_{ki}^{m} d_{ki}^{2}$$

where '$c + 1$' consists of '$c$' good clusters and one noise cluster and for $k = n = c + 1$.

$$\delta^2 = \lambda\left[\frac{\sum_{k=1}^{c}\sum_{i=1}^{N}(d_{ki})^2}{Nc}\right]$$

and membership equation is:

$$u_{ji} = \left(\sum_{k=1}^{k=c}\left(\frac{d_{ji}^2}{d_{ki}^2}\right)^{\frac{1}{m-1}} + \left(\frac{d_{ji}^2}{\delta^2}\right)^{\frac{1}{m-1}}\right)^{-1}$$

Noise clustering is a better approach than FCM, PCM, and PFCM. Although, it identifies outliers in a separate cluster but does not result into efficient clusters because it fails to identify those outliers which are located in between the regular clusters. Its main objective is to reduce the influence of outliers on the clusters rather than identifying it. Real-life data-sets usually contain cluster structures that differ from our assumptions so a clustering technique should be independent of the number of clusters for the same data-set. In NC, noise distance is given as:

$$\delta^2 = \lambda\left[\frac{\sum_{i=1}^{c}\sum_{k=1}^{N}(d_{ik})^2}{Nc}\right]$$

Here, noise distance depends upon distance measure, number of assumed clusters, and $\lambda$, which is the value of multiplier used to obtain '$\delta$', from the average of distances. From the equation, it is interpreted that if the number of clusters is increased, $\delta$ assumes high values. NC assigns only those points to noise cluster whose distance from regular clusters is less than the distance from noise distance, $\delta$. If the number of clusters is increased for the same data-set, NC does not detect outliers, because in that scenario the average distance between points and regular clusters decreases and the noise distance remains almost constant and assumes relatively high values.

## 3. The proposed algorithm, radial basis kernel based intuitionistic fuzzy c-means (KIFCM)

### 3.1. Kernel based approach

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.

The kernel function can be applied to any algorithm that solely depends on the dot product between two vectors. Wherever a dot product is used, it is replaced by a kernel function. When done, linear algorithms are transformed into non-linear algorithms. Those non-linear algorithms are equivalent to their linear originals operating in the range space of a feature space $\varphi$. However, because kernels are used, the $\varphi$ function does not need to be ever explicitly computed. This is highly desirable, as sometimes our higher-dimensional feature space could even be infinite-dimensional and thus infeasible to compute.

A kernel function is a generalization of the distance metric that measures the distance between two data points as the data points are mapped into a high dimensional space in which they are more clearly separable. By employing a mapping function $\Phi(x)$, which defines a non-linear transformation: $x \to \Phi(x)$, the non-linearly separable data structure existing in the original data space can possibly be mapped into a linearly separable case in the higher dimensional feature space.

Given an unlabeled data set $X = \{x_1, x_2, \ldots, x_n\}$ in the $p$-dimensional space $R^P$, let $\Phi$ be a non-linear mapping function from this input space to a high dimensional feature space $H$:

$$\Phi : R^p \to H, \quad x \to \Phi(x)$$

The key notion in kernel based learning is that mapping function $\Phi$ need not be explicitly specified. The dot product in the high dimensional feature space can be calculated through the kernel function $K(x_i, x_j)$ in the input space $R^P$.

$$K(x_i, x_j) = \Phi(x_i)\cdot\Phi(x_j)$$

Consider the following example. For $p = 2$ and the mapping function $\Phi$,

$$\Phi : R^2 \to H = R^3 \quad (x_{i1}, x_{i2}) \to \left(x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}\right)$$

Then the dot product in the feature space $H$ is calculated as

$$\Phi(x_i)\cdot\Phi(x_j) = \left(x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}\right)\cdot\left(x_{j1}^2, x_{j2}^2, \sqrt{2}x_{j1}x_{j2}\right)$$
$$= \left(\left(x_{i1}^2, x_{i2}^2\right)\cdot\left(x_{j1}^2, x_{j2}^2\right)\right)^2 = (x_i\cdot x_j)^2 = K(x_i, x_j)$$

where $K$-function is the square of the dot product in the input space. We saw from this example that use of the kernel function makes it possible to calculate the value of dot product in the feature space $H$ without explicitly calculating the mapping function $\Phi$. Some examples of kernel function are:

**Example 1.** *Polynomial kernel*: $K(x_i,x_j) = (x_i \cdot x_j + c)^d$, where $c \geqslant 0$, $d \in N$.

**Example 2.** *Gaussian kernel*: $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$, where $\sigma > 0$.

**Example 3.** *Radial basis kernel*: $K(x_i, x_j) = \exp\left(-\frac{\sum\left|x_i^a - x_j^a\right|^b}{\sigma^2}\right)$, where $\sigma$, $a$, $b > 0$.

RBF function with $a = 1$, $b = 2$ reduces to Gaussian function.

**Example 4.** *Hyper tangent kernel*: $K(x_i, x_j) = 1 - tanh\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right)$, where $\sigma > 0$.

### 3.2. Formulation

Our propose model RBF kernel based Intuitionistic fuzzy c-means (KIFCM) adopts a kernel induced metric which is different from the Euclidean norm in the original intuitionistic fuzzy c-means. KIFCM minimizes the objective function:

$$J_{KIFCM} = 2\sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{*m}\|\Phi(x_k) - \Phi(v_i)\|^2 + \sum_{i=1}^{c}\pi_i^{*}e^{1-\pi_i^{*}} \tag{10}$$

where $\|\Phi(x_k) - \Phi(v_i)^2\|$ 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:

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

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

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

where $a$ and $b$ is greater than 0, and $\sigma$ is defined as kernel width and it is a positive number, then $K(x,x) = 1$. Hence,

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

Thus (10) can be written as:

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

$$J_{KIFCM}=2\sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^{*m}\left(1-\exp\left(-\frac{\sum\left|x_i^a-x_j^a\right|^b}{\sigma^2}\right)\right)+\sum_{i=1}^{c}\pi_i^{*}e^{1-\pi_i^{*}} \tag{11}$$

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

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



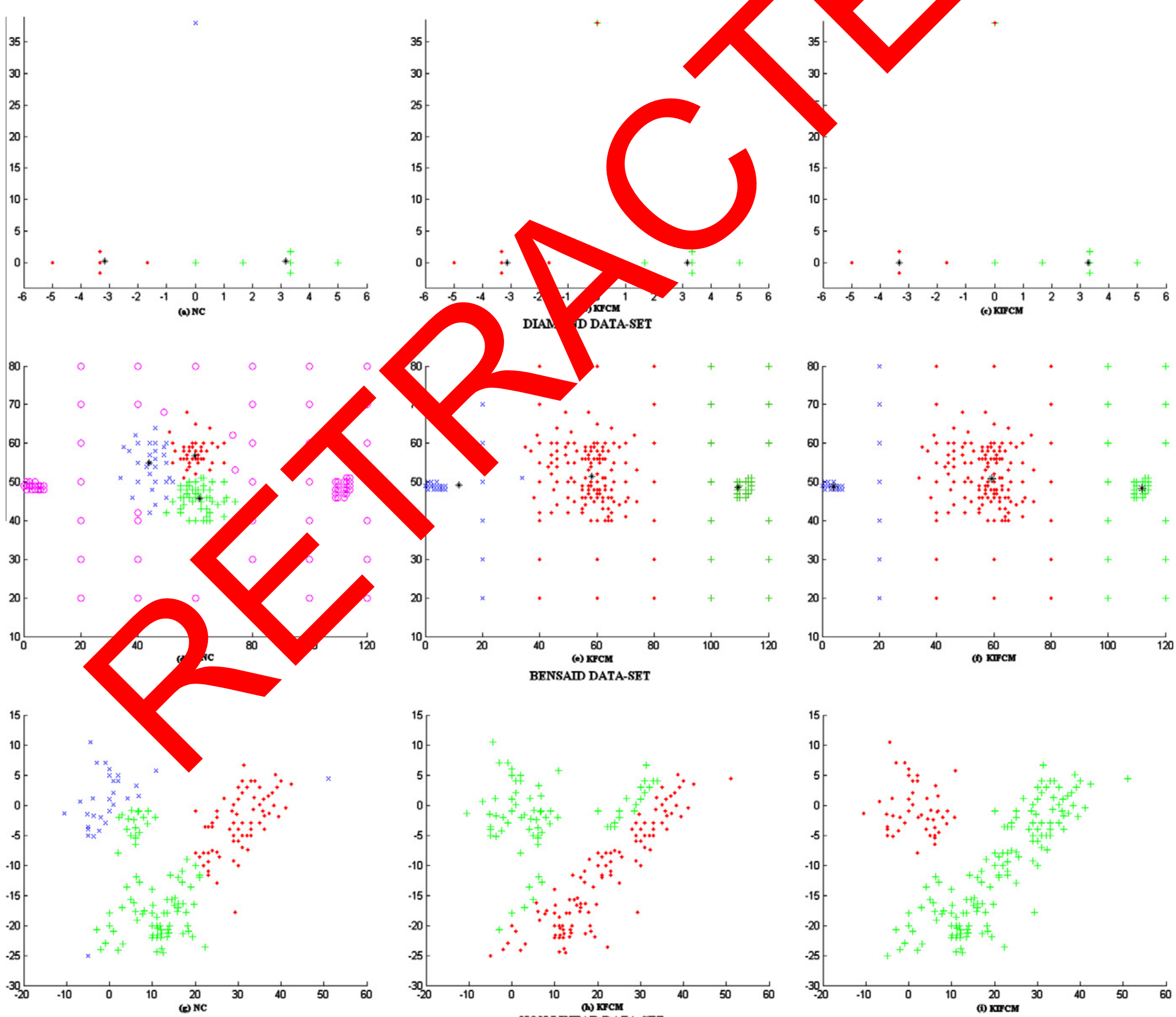


**Fig. 1.** (a–c) shows Clustering result of NC with $\lambda = 0.6$, KFCM with $h$ (kernel width) = 6 and KIFCM with $h = 6$, $a = 2$, $b = 4.2$, $\alpha = 7$ with Diamond data-set (d–f) shows Clustering result of NC with $\lambda = 0.17$, KFCM with $h$ (kernel width) = 100 and KIFCM with $h = 300$, $a = 2$, $b = 2$, $\alpha = 3$ with Bensaid data-set (g–i) shows result of NC with $\lambda = 0.17$, KFCM with $h$ (kernel width) = 55, $a = 2$, $b = 2.6$ and KIFCM with $h = 55$, $a = 2$, $b = 3$, $\alpha = 0.7$ with non-linear data-set. Centroids are shown with '$*$' symbol and the clusters are differentiated with '.' and '+' and '$\times$' symbol. Noise in detected with 'o'.

$$u_{ik}^{*}=\frac{\left(\frac{1}{(1-K(x_k,v_i))}\right)^{\frac{1}{m-1}}}{\sum_{i=1}^{c}\left(\frac{1}{(1-K(x_k,v_i))}\right)^{\frac{1}{m-1}}} \tag{12}$$

$$v_i=\frac{\sum_{k=1}^{n}u_{ik}^{*m}\times K(x_k,v_i)\times x_k}{\sum_{k=1}^{n}u_{ik}^{*m}\times K(x_k,v_i)} \tag{13}$$

**Proof.** We differentiate $J_{KIFCM}$ with respect to $u_{ik}^{*}$ and $v_i$ and set the derivatives to zero. Thus, we get (12) and (13). The details are given in Appendix A. □

### 3.3. Steps involved in KIFCM are

---

Radial basis kernel based intuitionistic fuzzy c-means clustering

---

***Input parameters:*** Image data ($X$), number of clusters ($K = c + 1$), number of iterations, stopping criteria ($\bar{C}$).
***Output***: Cluster centroids matrix, membership matrix.
Step 1: Get data from image.
Step 2: Select initial prototypes.
Step 3: Obtain the memberships using (12).
Step 4: Update the prototypes using (13).
Step 5: Update the memberships using (12) with updated prototypes.
Step 6: Repeat steps (3)–(5) till the updated membership satisfies the condition: $\left\|u_{ik}^{(t+1)}-u_{ik}^{t}\right\|<\varepsilon\ \forall i,k$ is met for successive iterations $t$ and $t+1$
where $\varepsilon$ is a small number.

---

## 4. Simulations and results

In this section, experimental results are presented to compare the segmentation performance of radial basis kernel based intuitionistic fuzzy c-means (KIFCM) with FCM, PCM, PFCM, NC, KFCM, T2FCM, KT2FCM and IFCM. Experiments are implemented and simulated using MATLAB Version [illegible]0. We considered following common parameters: $m = 2$, which is a common choice for fuzzy clustering, $\varepsilon = 0.03$, $\alpha = 0.8[illegible]$ for IFCM (as referred in (Chaira, 2011)), and maximum iteration = 200. We used RBF kernel for the kernelized methods. Note that the kernel width '$h$' in RBK kernel has a very important effect on the performances of the algorithms. However, how to choose an appropriate value for the kernel width in RBF kernel is still an "open problem". In this paper, we adopt the "trial-and-error" technique to find the kernel width. We used synthetic data-set, standard data-sets, and real medical images for the experiments.

### 4.1. Synthetic data-sets

*Three synthetic data-sets, diamond (D12) data set* (referred from Pal et al. (2005)), *BENSAID Data-set* (Bensaid et al., 1996) *and Non-linear synthetic data-set are considered in this section.*

Diamond Data-set, D11, D12 (referred from Pal et al. (2005)). D11 is a noiseless data-set of points $\{x_i\}_{i=1}^{11}$. D12 is the union of D11 and an outlier x12. BENSAID's two-dimensional data-set consists of one big and two small size clusters. 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 $[0,120]\times[10,80]$. To evaluate the effect of non-linear data structure, a synthetic non-linear data-set consisting of one circular and one elliptic cluster is considered. All the seven algorithms, FCM, IFCM, PCM, PFCM, NC, KFCM, KIFCM are implemented with these data-sets to check the performance.

For diamond data-set, it is observed that FCM, IFCM, PCM, and PFCM could not detect the original clusters and their performance is badly affected with the presence of noise, NC and KFCM detected the clusters but the centroid location is still affected with noise. To show the effectiveness of the proposed algorithm, we also calculated the error for recognizing correct cluster centers in case of Diamond Data-set with the equation $E_{*}=\|V_{ideal}-V_{*}\|^2$, where $*$ is PCM/PFCM/NC/KFCM/KIFCM. The ideal (true) centroids of Diamond data-set are:

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

FCM and IFCM could not detect the clusters. Average error with PCM, PFCM, NC, KFCM and KIFCM is 1[illegible]17, 11[illegible]2, 0.077, 0.0362 and 0.003 respectively. Fig. 1 shows the clustering results with three best performed algorithms (NC, KFCM and KIFCM) for all the three data-sets.

Clearly after observing the results of Fig. 1 and considering the error percentage, it is observed that proposed method can produce more accurate centroids than other methods and is highly robust against noise.

In case of BENSAID data-set PCM, PFCM and NC could not detect the original clusters and the performance of FCM and KFCM is affected with the presence of noise. Best performance is achieved in the case of IFCM and KIFCM and the results are not much affected as in other cases.

In case of non-linear data-set, all the algorithms except KFCM and KIFCM could not detect the elliptical cluster. KFCM although detected the elliptical cluster but misaligned some members of the elliptical cluster into the circular cluster, whereas KIFCM correctly partitioned all the data points.

#### 4.1.1. Performance evaluation based upon misclassifications

To verify the performance of seven algorithm based upon misclassifications, we calculated their scores defined by the following quantitative index (Masulli and Schenone, 1999).

$$r_{ij}=\frac{A_{ij}\cap A_{refj}}{A_{ij}\cup A_{refj}}$$

where $A_{ij}$ represents the set of pixels belonging to the $j$th class found by the $i$th algorithm and $A_{refj}$ represents the set of pixels belonging to the $j$th class in the reference segmented image. $r_{ij}$ is a fuzzy similarity measure, indicating the degree of equality between $A_{ij}$ and $A_{refj}$, and the larger the $r_{ij}$, the better the segmentation is. Table 1 lists the misclassifications and comparison scores using seven methods.

From Fig. 1 and Table 1, we observed that although many algorithms out of seven detected the right clusters in case of diamond and Bensaid data-set but in case of synthetic non-linear data-set no algorithm except KIFCM detected the original clusters. Apart from that, in case of numerical data-sets, location of cluster centers is the major criteria to compare various algorithms. So considering these points into view, we can say that KIFCM outperformed the other six algorithms.

### 4.2. Real data sets

Four real data-sets, Wisconsin Breast cancer, iris, Wine and PIDD (Pima Indians Diabetes database) from public data-bank UCI (UC Irvine Machine Learning Repository) are used to evaluate the performance of these algorithms.

**Table 1**
Misclassifications and comparison scores of seven methods with different data-sets.

| Type of data-set | Date-set | Name of the algorithm | Number of clusters (Samples in the clusters) | Misclassifications | Comparison score $\mathbf{R}_{ij}$ |
|---|---|---|---|---|---|
| Synthetic and standard data-sets | Diamond data-set with 2 clusters | FCM | 2 (6,5) | 10 | 0.545 |
| | | IFCM | | 10 | 0.545 |
| | | PCM | | 10 | 0.545 |
| | | PFCM | | 0 | 1 |
| | | NC | | 0 | 1 |
| | | KFCM | | 0 | 1 |
| | | KIFCM | | 0 | 1 |
| | Bensaid data-set with 3 clusters | FCM | 3 (18,141,20) | 2 | 0.994 |
| | | IFCM | | 0 | 1 |
| | | PCM | | 40 | 0.899 |
| | | PFCM | | 12[illegible] | 0.625 |
| | | NC | | [illegible]3 | 0.525 |
| | | KFCM | | 2 | 0.994 |
| | | KIFCM | | 0 | 1 |
| | Synthetic non-linear data-set with 2 clusters | FCM | 2 (47,140) | 70 | 0.812 |
| | | IFCM | | [illegible] | 0.791 |
| | | PCM | | 44 | 0.882 |
| | | PFCM | | 70 | 0.812 |
| | | NC | | 70 | 0.737 |
| | | KFCM | | 5[illegible] | 0.866 |
| | | KIFCM | | [illegible]0 | 1 |
| Real data sets from public database UCI | Wisconsin Breast cancer | FCM | 2 (357,[illegible]2) | 162 | 0.8576 |
| | | IFCM | | 158 | 0.8611 |
| | | PCM | | 366 | 0.6783 |
| | | PFCM | | 112 | 0.9051 |
| | | NC | | 138 | 0.8823 |
| | | KFCM | | 88 | 0.9525 |
| | | KIFCM | | 69 | 0.9560 |
| | Iris | FCM | 3 (50,50,50) | 20 | 0.9333 |
| | | IFCM | | 14 | 0.9533 |
| | | PCM | | 50 | 0.667 |
| | | PFCM | | 14 | 0.953 |
| | | NC | | 18 | 0.940 |
| | | KFCM | | 16 | 0.946 |
| | | KIFCM | | 8 | 0.9733 |
| | Wine | FCM | 3 (59, 71, 48) | 4 | 0.9887 |
| | | IFCM | | 4 | 0.9887 |
| | | PCM | | 210 | 0.4101 |
| | | PFCM | | 8 | 0.9775 |
| | | NC | | 27 | 0.8595 |
| | | KFCM | | 14 | 0.9606 |
| | | KIFCM | | 4 | 0.9887 |
| | PIDD | FCM | 2 (500, 268) | 140 | 0.9088 |
| | | IFCM | | 76 | 0.9505 |
| | | PCM | | 528 | 0.6562 |
| | | PFCM | | 82 | 0.9466 |
| | | NC | | 46 | 0.9596 |
| | | KFCM | | 44 | 0.9713 |
| | | KIFCM | | 42 | 0.9726 |

RETRACTED

The Wisconsin Breast cancer data is widely used to test the effectiveness of classification. The aim of the classification is to distinguish between benign and malignant cancers based on the available 30 attributes. The original database contains 569 instances. The class distribution is 357 benign and 212 malignant, respectively. Iris data-set is a four-dimensional data set containing 50 samples each of three types of Iris flowers (Setosa, Versicolor and Vitginica). One of the three clusters (class 1) is well separated from the other two, while classes 2 and 3 have some overlap. The Wine data contains the chemical analysis of 178 wines grown in the same region in Italy but derived from three cultivators. The problem is to distinguish the three different types based on 13 continuous attributes derived from chemical analysis. The three types have values 59, 71, 48. PIDD is a diabetic data-set having 768 instances. It has eight attributes. It has two classes 500, 268. The original feature values are directly used for clustering evaluation without data normalization. Table 1 summarizes the misclassifications and comparison scores for real data-sets. From Table 1 it is observed that the proposed algorithm is far superior to FCM, PCM, PFCM and NC and IFCM. The improvement of the proposed KIFCM with respect to KFCM is comparable. As KIFCM has least misclassifications in three out of four real data-sets so overall performance of KIFCM is high when compare to other algorithms.

### 4.3. Medical image segmentation

This subsection is presented to evaluate the performance of proposed algorithm, KIFCM with medical images.

#### 4.3.1. Simulated MR brain image

Since the ground truth of segmentation for real MR images is not usually available, hence segmentation performance cannot be visually compared and it is impossible to evaluate the segmentation performance qualitatively. However, Brainweb provides a simulated brain database (SBD) including a set of realistic MRI data volumes produced by an MRI simulator. These data enable us to

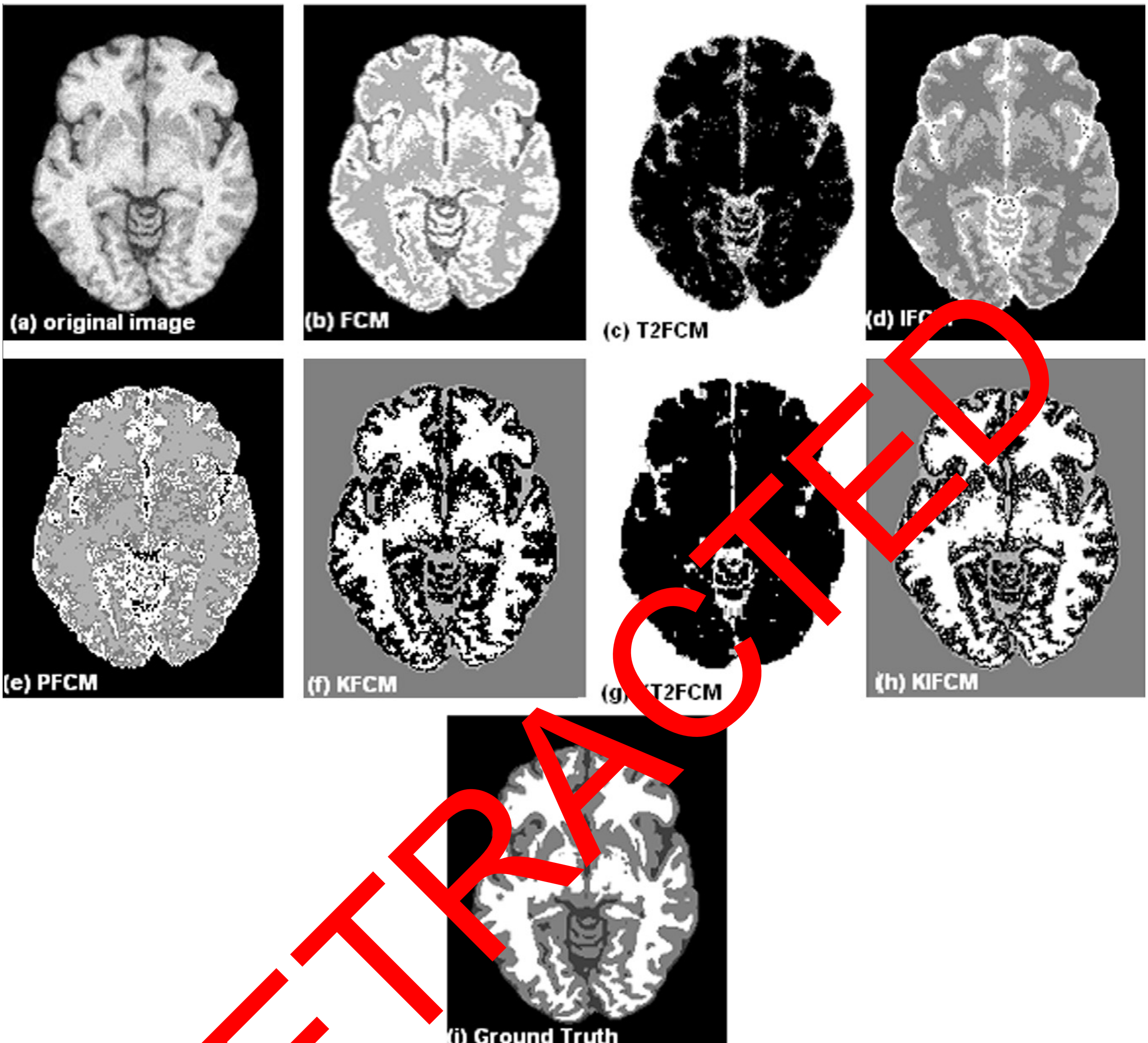


**Fig. 2.** (a) Noisy image, (b) segmentation result of FCM, (c) segmentation result of T2FCM, (d) Segmentation result of IFCM with $\alpha = 0.85$, (e) segmentation result of PFCM, (f) segmentation result of KFCM with $h = 1.$[illegible] and $b =$ [illegible] $a = 1$, (g) segmentation result of KT2FCM with $h = 5$, (h) segmentation result of KIFCM with $\alpha = 1.5$, $h = 0.4$, $a = 1$ and $b = 2.5$, and (i) ground truth.

evaluate the performance of various image analysis methods in a setting where the truth is known (http://www.bic.mni.mcgill.ca/brainweb). A simulated T1-weighted MR image (217 × 181 × 3), 1-mm cubic voxels, including noise, and no intensity inhomogeneity is downloaded from Brainweb. The class number of the image is assumed as four, corresponding to grey matter (GM), white matter (WM), cerebrospinal fluid (CSF), and background (BKD). Fig. 2(a) shows a slice from the simulated dataset. Fig. 2(b–h) shows the segmentation results obtained by applying FCM, T2FCM, IFCM, PFCM, KFCM, KT2FCM and KIFCM respectively. Ground truth is given in Fig. 2(i). As observed from figures, the results of FCM, IFCM and PFCM are not free from noise. T2FCM and KT2FCM are not able to detect all the clusters properly. Compare to FCM, KFCM gave improved results. We tried many combinations of $\alpha$ and $h$ (kernel width) for the kernel methods. It is clearly seen that the segmentation result of KIFCM is much closer to the ground truth than other segmentation techniques. The result of KIFCM is more homogeneous and smoother than other algorithms, especially for WM, which again indicates that our method is effective and robust to noise.

*4.3.2. T1-weighted MR brain image corrupted with mixed noise*

In order to further examine the algorithms' effectiveness, we are considering the real brain MR slice corrupted simultaneously by Gaussian white noise N(0,180) and unit dispersion, zero centered symmetric $\alpha$-stable noise.

The class number of the image is assumed as three, corresponding to grey matter (GM), white matter (WM), and background (BKD). Fig. 3(a) shows an original slice (referred from Weiling Cai, Songcan Chen, Daoqiang Zhang and Chen, 2007) and (b) shows corrupted slice with mixed noise. Fig. 3(c–i) shows the segmentation results obtained by applying FCM, T2FCM, IFCM, PFCM, KFCM, KT2FCM, and KIFCM respectively. As observed from figures, the performance of T2FCM and KT2FCM is very bad whereas the per-

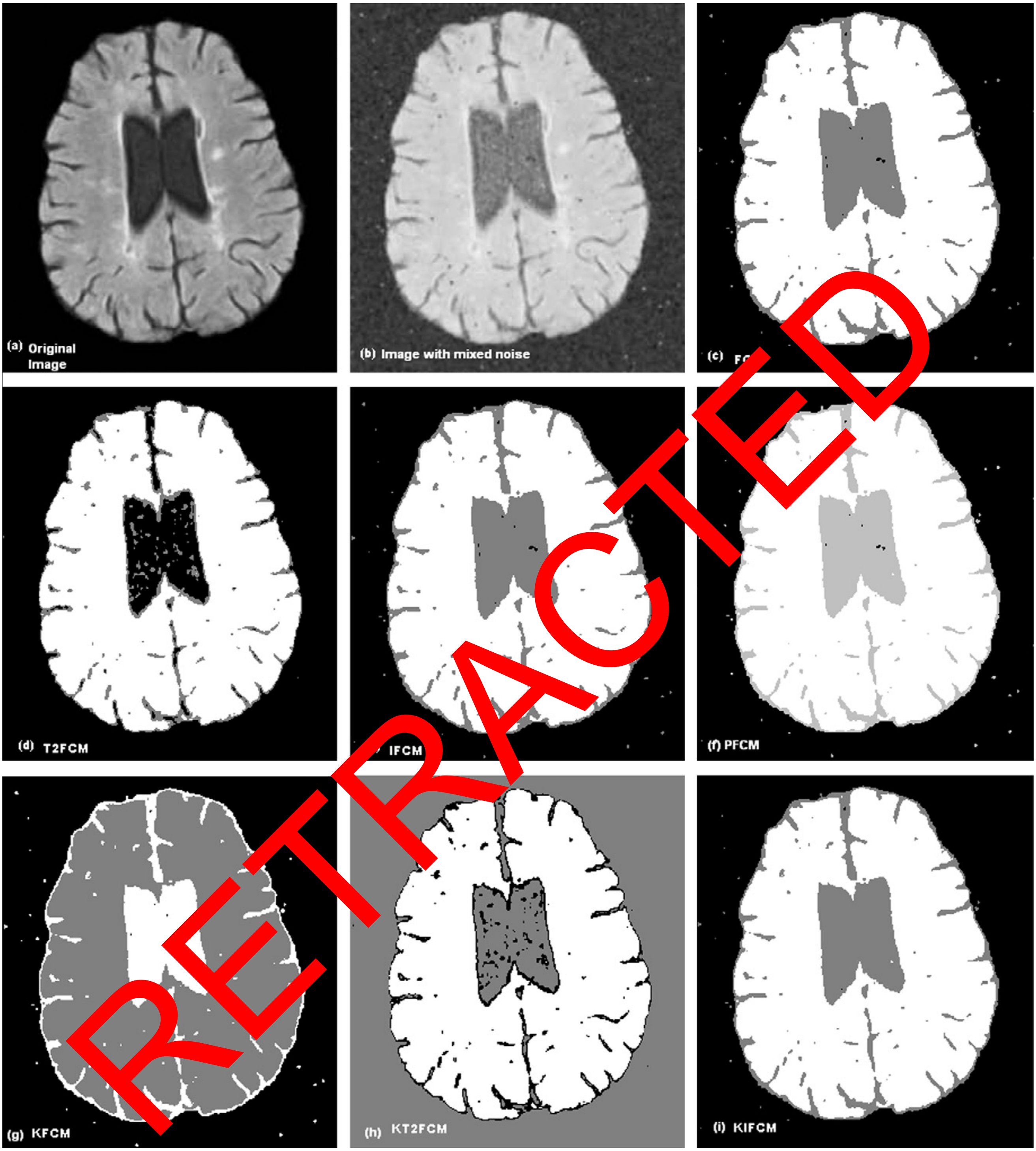




**Fig. 3.** (a) Original image, (b) currupted with mixed noise, (c) segmentation result of FCM, (d) segmentation result of T2FCM, (e) segmentation result of IFCM with $\alpha = 0.85$, (f) segmentation result of PFCM, (g) segmentation result of KFCM with $h = 0.2$, (h) segmentation result of KT2FCM with $h = 5$, (i) segmentation result of KIFCM with $\alpha = 0.85$, $h = 0.3$, and $b = 2.5$.

formances of all other algorithms are almost similar. It is clearly seen that in the segmentation result of KIFCM the noise is suppressed as compared to IFCM and KFCM. The result of KIFCM is more homogeneous and smoother than other methods.

#### 4.3.3. T1-weighted MR brain image corrupted with Gaussian noise

In the second case, we are considering the same image with the Gaussian noise. Note that MR images typically do not suffer from "Guassian" noise. We add such type of noise just to compare various algorithms for robustness against noise.

Fig. 4 shows the segmentation results of all the algorithms. From the results, we observed that segmentation result of FCM and KFCM is better in case of mixed noise than Gaussian noise. T2FCM and PFCM could not detect the clusters in case of Gaussian noise. Results of IFCM are better in case of mixed noise. But the KIFCM has constant results without much variation in the case of both types of noises.

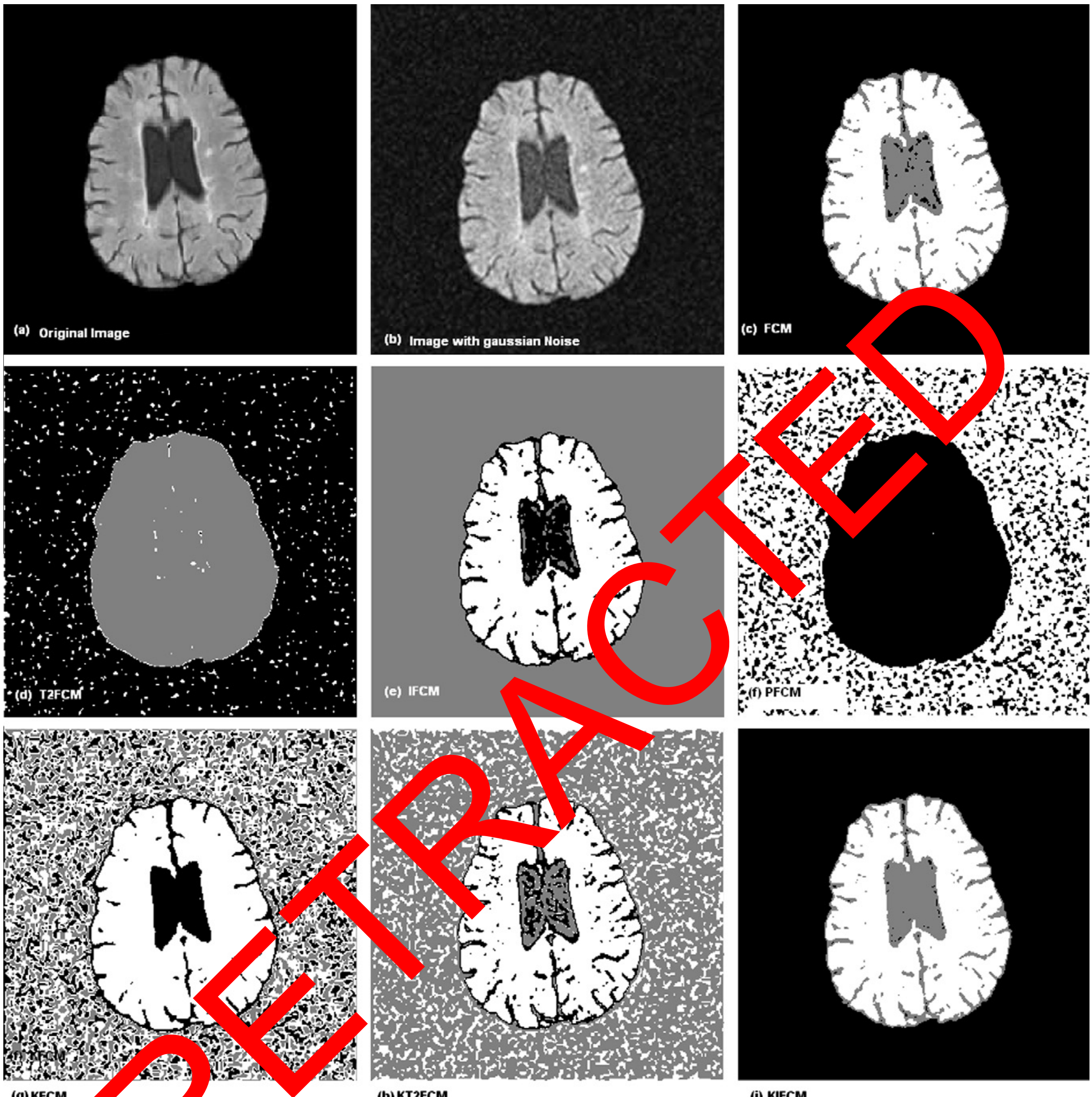




**Fig. 4.** (a) Original image, (b) corrupted with Gaussian noise, (c) segmentation result of FCM, (d) segmentation result of T2FCM, (e) segmentation result of IFCM with $\alpha = 0.85$, (f) segmentation result of PFCM, (g) segmentation result of KFCM with $h = 0.1$, (h) segmentation result of KT2FCM with $h = 0.1$, and (i) segmentation result of KIFCM with $\alpha = 0.85$, $h = 3$, and $b = 2.5$.

#### 4.3.4. CT-scan brain image having tumor region

In this section, we considered a CT-scan brain image. This original image is downloaded from medical image gallery and is segmented into four clusters. For performance comparison, the segmented image is compared with the ground truth image for the tumor region. Fig. 5(a) shows the poorly illuminated CT scan brain image. Fig. 5(b) shows ground truth image. It shows lateral ventricles, third ventricle, and the blood clot (hemorrhagic stroke). All the algorithms, except T2FCM and PFCM, although detected the blood clot region but the size of blood clot region detected is less than the actual size except the blood clot region detected with KIFCM algorithm. After comparing the performance it is observed that KIFCM detected the blood clot region and clot/hemorrhage region almost equal in area as that in the ground truth image (Fig. 5(b)) as compared to other methods.

Table 2 lists the quantitative comparison scores corresponding to Figs. 2–5 using seven methods. In Figs. 2–4, we considered two clusters and for Fig. 5 we considered only the tumor region for calculating scores. From the Figures and Table 2, we observed that KIFCM outperformed consistently the other six algorithms

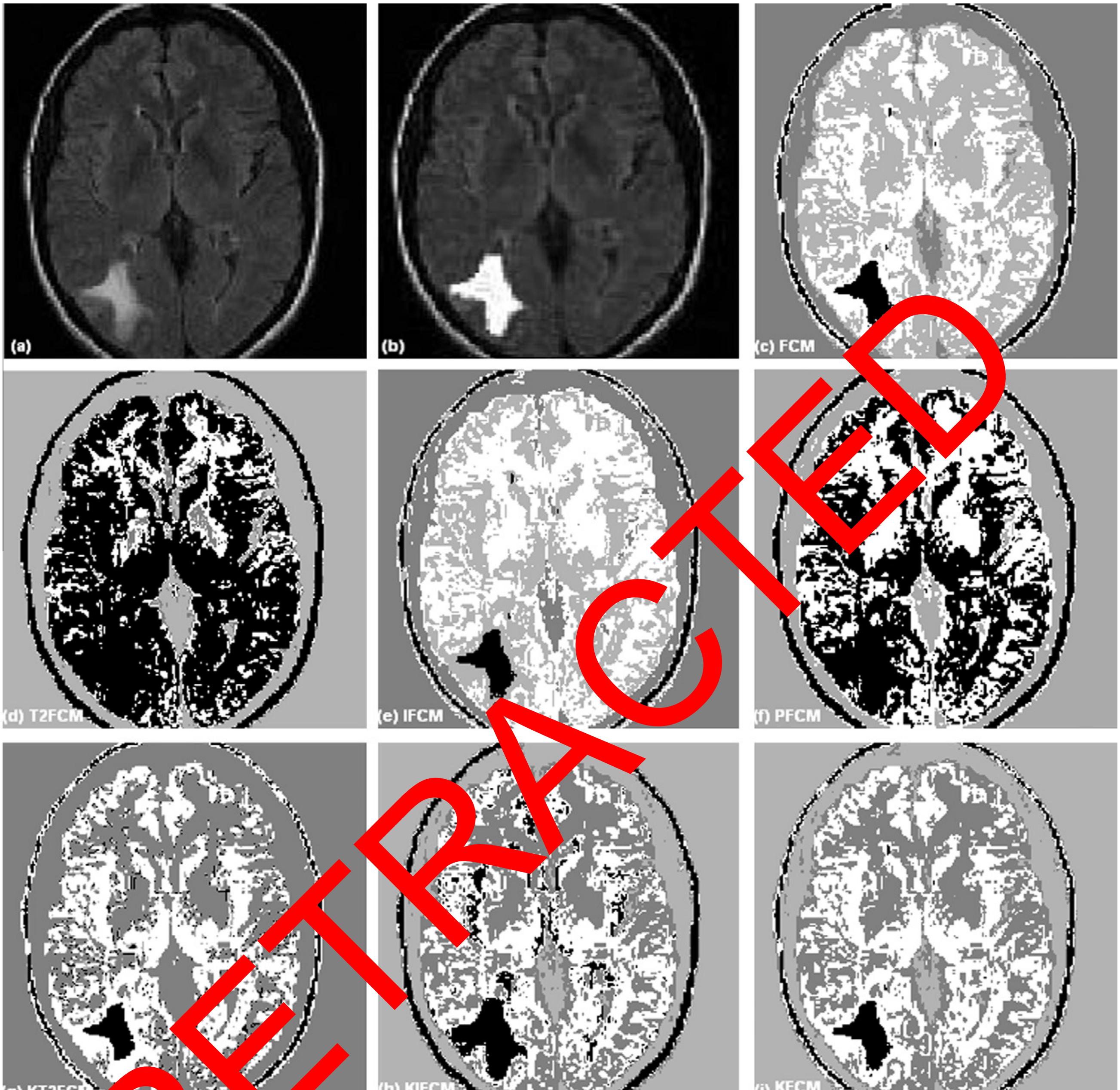




**Fig. 5.** (a) CT scan brain image, (b) ground truth image, (c) segmentation result of FCM, (d) segmentation result of T2FCM, (e) segmentation result of IFCM with $\alpha = 0.85$, (f) segmentation result of PFCM, (g) segmentation result of KT2FCM with $h = 0.1$, (h) segmentation result of KIFCM with $\alpha = 0.85$, $h = 1.11$, and $b = 1.4$, (i) segmentation result of KFCM with $h = 0.5$.

**Table 2**
Comparison scores of seven methods.

| Name of the algorithm | $r_{ij}$ (Comparison scores) Fig. 2 (simulated brain image) | | | $r_{ij}$ (comparison scores) Fig. 3 (real MR brain image with mixed noise) | | | $r_{ij}$ (comparison scores) Fig. 4 (real MR brain image with Gaussian noise) | | | Fig. 5 (real CT scan brain image) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Cluster1 (White matter) | Cluster2 (Grey matter) | Average score | Cluster1 (White matter) | Cluster2 (Grey matter) | Average score | Cluster1 (White matter) | Cluster2 (Grey matter) | Average score | Tumor region |
| FCM | 0.888 | 0.904 | 0.896 | 0.966 | 0.915 | 0.940 | 0.9796 | 0.9957 | 0.9876 | 0.8946 |
| T2FCM | No detection | | | 0.897 | 0.459 | 0.678 | No detection | | | No detection |
| IFCM | 0.889 | 0.923 | 0.906 | 0.969 | 0.923 | 0.946 | 0.9907 | 0.9834 | 0.9870 | 0.9264 |
| PFCM | 0.80 | 0.748 | 0.774 | 0.966 | 0.918 | 0.942 | No detection | | | No detection |
| KFCM | 0.903 | 0.916 | 0.909 | 0.963 | 0.910 | 0.936 | 0.4393 | 0.7226 | 0.5809 | 0.9389 |
| KT2FCM | No detection | | | 0.986 | 0.416 | 0.701 | 0.2041 | 0.4130 | 0.3085 | 0.8218 |
| KIFCM | 0.957 | 0.994 | 0.975 | 0.990 | 0.965 | 0.977 | 0.9920 | 0.9980 | 0.9950 | 0.9817 |

which show its insensitivity to different kinds of noises and outliers.

## 5. Conclusions

Medical images generally contain unknown noise and considerable uncertainty, and therefore clinically acceptable segmentation performance is difficult to achieve. In this paper, we proposed a kernel induced metric to replace the Euclidean norm in intuitionistic fuzzy c-means (IFCM) and then derived the RBF kernel based IFCM clustering algorithm. The algorithm is invented by introducing kernel, RBF function, and lagrangian methods with basic objective function of the IFCM algorithm to have effective segmentation in medical images. To test the performance of proposed method it was applied to synthetic data-set, Simulated and Real MR brain images. It has been found that proposed algorithm is robust to noise and outliers compared to FCM, T2FCM, IFCM, PFCM, KFCM, and KT2FCM. Our proposed method selected more desirable cluster centers in case of synthetic data-set and gave better segmentation results in case of simulated and real medical images compared to other methods.


## Acknowledgements

The authors would like to express their sincere appreciation and thanks to the reviewers for their careful and thorough critical appraisement of an earlier draft of the paper which help them to improve their paper.


## Appendix A. Radial basis kernel based intuitionistic fuzzy c-means clustering (KIFCM)

### *A.1. Proof of radial basis kernel based intuitionistic fuzzy c-means*

In this appendix, we give the proof of the radial basis kernel based intuitionistic fuzzy c-means clustering which is kernelized version of intuitionistic Fuzzy c-means clustering algorithm. The problem of minimization of objective function $J_{KIFCM}$ {given in (9)} subjected to the constraint specified by (2) is solved by minimizing a constraint free objective function defined as:

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

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

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

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

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

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

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

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

So,

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

$$J_{KIFCM}=2\sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^{*m}\left(1-\exp\left(-\frac{\sum|x_k^{a}-v_i^{a}|^{b}}{\sigma^2}\right)\right)+\sum_{i=1}^{c}\pi_i^{*}e^{1-\pi_i^{*}}+\sum_{i=1}^{c}\lambda_i\left(\sum_{k=1}^{n}u_{ik}^{*}-1\right)\qquad(15)$$

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

Assuming $a = 1$ and $b = 2$, the partial derivative of $J_{KIFCM}$ with respect to $v_i$ is:

$$\frac{\partial J}{\partial v_i}=\sum_{k=1}^{n}2u_{ik}^{*m}\times\left(-\exp\left(-\frac{(x_k-v_i)^2}{\sigma^2}\right)\right)\times\frac{2(x_k-v_i)}{\sigma^2}\qquad(16)$$

Equating (16) to zero leads to:

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

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

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

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

$$\frac{\partial J}{\partial u_{ik}^{*}}=2mu_{ik}^{*(m-1)}(1-K(x_k,v_i))+\lambda_i\qquad(17)$$

Equating (17) to zero leads to the following:

$$\frac{\partial J}{\partial u_{ik}^{*}}=0\Rightarrow 2mu_{ik}^{*(m-1)}\times(1-K(x_k,v_i))+\lambda_i=0\Rightarrow u_{ik}^{*}=\left(-\frac{\lambda_i}{2m}\right)^{1/m-1}\times\left(\frac{1}{1-K(x_k,v_i)}\right)^{1/m-1}$$

To fulfill the constraint (2),

$$\Rightarrow u_{ik}^{*}=\frac{u_{ik}^{*}}{\sum_{i=1}^{c}u_{ik}^{*}}\qquad(18)$$

In view of (18), Eq. (17) can be written as:

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

## References


Atanassov, K.T., 2003. Intuitionistic Fuzzy Sets Past, Present and Future. <http://www.eusflat.org/publications/proceedings/.2003/./4Atanassov.pdf>.

Atanassov's, K.T. (1983). Intuitionistic fuzzy sets, VII ITKR's Session, Sofia, 983 (Deposed in Central Science–Technology Library of Bulgaria Academy of Science–1697/84).

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

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithm. Plenum, NY.

BrainWeb [Online]. Available from: <http://www.bic.mni.mcgill.ca/brainweb>.

Chaira, T., 2011. A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images. Appl. Soft Comput. 11, 1711–1717.

Dave, R.N., Krishnapuram, R., 1997. Robust clustering methods: a unified view. IEEE Trans. Fuzzy Systems 5 (2).

Dunn, J., 1974. A fuzzy relative of the ISODATA process and its use in detecting compact well separated clusters. J. Cybernet. 3, 32–57.

Kang, Jiayin, Min, Lequan, et al., 2009. Novel modified fuzzy c-means algorithm with applications. Digital Signal Processing 19 (2), 309–319.
Krishnapuram, R., Keller, J., 1993. A possibilistic approach to clustering. IEEE Trans. Fuzzy Systems 1 (2), 98–110.
Masulli, F., Schenone, A., 1999. A fuzzy clustering based segmentation system as support to diagnosis in medical imaging. Artif. Intell. Med. 16 (2), 129–147.
Pal, N.R., Pal, K., Keller, J., Bezdek, J.C., 2005. A possibilistic fuzzy c-means clustering algorithm. IEEE Trans. Fuzzy Systems 13 (4), 517–530.
Rhee, F.C.H., Hwang, C., 2001. A type-2 fuzzy c means clustering algorithm. In: Proc. Joint 9th IFSA World Congress and 20th NAFIPS Internat. Conf., vol. 4, pp. 1926–1929.
Yang, Zhang, Chung, Fu-Lai, et al., 2009. Robust fuzzy clustering-based image segmentation. Appl. Soft Comput. (91), 80–84.
Zadeh, L.A., 1965. Fuzzy sets. Inform. Control 8, 338–353.
Zhang, H.M., Chen, Z.S.Q., 2007. On clustering approach to intuitionistic fuzzy sets. Control and Decision 22, 882–888.