



Type-reduced vague possibilistic fuzzy clustering for medical images

Ankita Bose*, Kalyani Mali

Department of Computer Science & Engineering, University of kalyani, India



ARTICLE INFO

Article history:

Received 17 January 2020

Revised 18 October 2020

Accepted 8 December 2020

Available online 11 December 2020

Keywords:

Fuzzy membership

Typicality

Vague set

Type-reduction

Medical images

ABSTRACT

Soft computing provides the framework for dealing with the uncertainty and imprecision inherent in real-life applications. Soft computing has become a long-standing notable paradigm for medical image processing. A typical fuzzy clustering uses the fuzzy membership function. Nevertheless, there is an alternative membership representation, known as typicality or possibilistic membership. Unlike fuzzy membership that is probabilistic in nature, typicality represents an absolute membership and it is the degree of belonging of an object to a class that does not depend on its distances from the other classes. However, both fuzzy membership and typicality play important role in assigning membership to an object. This study proposes a novel clustering model that creates a vague environment enriched with the concept of fuzzy membership and typicality, while the use of type-reduction plays an essential role in capturing all the vagueness present in the data set. The proposed model is called type-reduced vague possibilistic fuzzy clustering (TVPFC), and we use MRI images to demonstrate its superior robustness over that of FCM (fuzzy c-means), PCM (possibilistic c-means), VCM (vague c-means) and IPFCM (interval-valued possibilistic fuzzy c-means).

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Image segmentation partitions an image into groups of pixels to create a simplified representation of an image that uncovers the intrinsic significance of the underlying objects. Among other useful applications, pattern identification and object detection have been largely adopted to serve many significant purposes. Medical image analysis has gained prominence over the last few years, and the domain of medical image segmentation is of meaningful use in aspects of diagnosis such as in the clinical study of the interior body regions, location of tumors, analysis of anatomical structures, and the measurement of tissue volume.

Because medical image segmentation is application-specific, with imaging modality playing an important role, we chose the segmentation of magnetic resonance imaging (MRI) and computerized tomography (CT) scan images for this study. MR imaging is used to acquire the structural details [1] of different body parts and is also useful for detecting white matter diseases. MR imaging achieves variable image contrast by considering variable parameters of the magnetic field, such as longitudinal relaxation time (T_1) and traverse relaxation time (T_2). In contrast, a CT scan uses the X-ray absorption level to create an image, and the reconstructed image holds structural and functional information on different body

parts. Both these methods suffer from artifacts such as noise, intensity inhomogeneity [2], and partial volume error. Image segmentation retrieves the clusters from a multidimensional space based on the relationships among the image voxels.

Machine learning (ML) is one of the most popular platforms for soft computing. ML provides a large number of tools [3] such as Support Vector Machine (SVM), Neural Network (NN), deep learning algorithms like convolutional neural network (CNN), Recurrent neural Network (RNN), Extreme Learning Model (ELM) gained popularity in medical image analysis, these approaches can identify disease specific key features and discover the relationships among the patterns. Depending on the working methodology, ML methods can be divided into two categories: (i) supervised learning and (ii) unsupervised learning. Some of the well-known supervised approaches widely applied in medical research [4] include, the decision tree (DT), Bayesian networks (BNs), artificial neural networks (ANNs), and the widely popular deep neural network more commonly known as deep learning [5]. All these supervised techniques require a lot of labelled data for training. In contrast, unsupervised approaches do not require a training dataset. Optimal segmentation results can be achieved by minimizing some cost function that takes into account the entire dataset. Our proposed approach is an unsupervised clustering technique.

Furthermore, soft Computing is a computational approach that provides approximate solutions for real world computationally hard problems. On the basis of ML, another significant contribution of soft computing is the fuzzy set. The uncertainty representation

* Corresponding author.

E-mail addresses: boseankita8@gmail.com, ankita@klyuniv.ac.in (A. Bose), kalyanimali@klyuniv.ac.in (K. Mali).

of the fuzzy set enables qualitative evaluation that improves diagnosis via medical images. The journey of fuzzy clustering begins with fuzzy c-means (FCM) [6], proposed by Bezdek in 1981. Fuzzy set theory proposes the membership generation as a probabilistic approach. The idea of possibilistic membership, or typicality, was first proposed under possibilistic c-means (PCM) [7]. Possibilistic membership does not support the probabilistic constraint of fuzzy membership. If we want the membership to be represented as a measure of compatibility, then fuzzy membership is not a suitable choice. Fuzzy membership does not reveal anything about uncertainty but treats the noise points equally irrespective of their distances from the cluster prototype, but like the other fuzzy algorithms, PCM is sensitive to the choice of initial cluster prototypes [8] and to essential parameters [9]. TVPFC deals with the problem of initialization either by obtaining initial prototypes using any suitable fuzzy clustering or by re-estimating the balancing factor β .

We cannot say whether fuzzy membership or typicality would be the best choice for obtaining the membership of an object. This dilemma of sharing versus non-sharing can be appropriately expressed by creating a vague environment. Our proposed approach considers both fuzzy and possibilistic membership and creates the hesitation region of a vague environment. This resembles an interval-valued representation of the memberships. To obtain the crisp prototype from an interval-valued membership representation, type-reduction is required.

The rest of the paper is organized as follows. A brief discussion of the related works is given in Section 2. Section 3 provides the background study, Section 4 outlines the proposed TVPFC algorithm along with a relevant discussion on vague environment. In Section 5, we demonstrate the effect of noise on related algorithms. Section 6 represents a comparative study based on the performance evaluation of the proposed TVPFC against FCM, PCM, VCM, and IPFCM. A large set of images is also compared against IFCM and IPCM. Section 7 is devoted to the conclusion and future scope.

2. Related work

The development of fuzzy logic has evolved through various stages. In this section, we discuss some of these stages, which are relevant to this study.

Pal et al., in 1997 developed the fuzzy possibilistic c-means (FPCM) [10] algorithm, which demonstrates that both probability and possibility are essential. However, the FPCM does not work well with large datasets, and an increase in the amount of noise results in coincidence clusters that are overlapping in nature. An improved FPCM, possibilistic fuzzy c-means (PFCM) [11], was introduced in 2005. However, due to the use of a large number of parameters, superior results cannot always be expected.

The concept of a type-2 fuzzy set was first introduced by Zadeh [12], and then further modified by Mizumoto and Tanaka [13]. The type-2 fuzzy set represents uncertainty by implementing a secondary membership function [14]. Consequently, the membership value of an object in a type-2 fuzzy set is a fuzzy set in [0,1]. A type-2 fuzzy set, \tilde{A} , defined over dataset $X = \{x_1, x_2, x_3 \dots x_N\}$ is characterized by a type-2 membership function $\mu_{\tilde{A}}(x, u)$, where $x \in X$ and $u \in J_x \subseteq [0,1]$. J_x is the primary membership of x defined from the secondary membership of x . For each primary membership, there is a secondary membership, that defines the range of possibilities for that primary membership. Basically, a type-1 fuzzy set is a crisp decision of a type-2 fuzzy set. A general type-2 fuzzy logic increases the computational complexities in decision-making due to numerous embedded type-2 fuzzy sets. To solve this issue, an interval-valued fuzzy set is developed. The interval-valued fuzzy set is a special case of the type-2 fuzzy set, where secondary membership can take the value of either 0 or 1. In this study, we con-

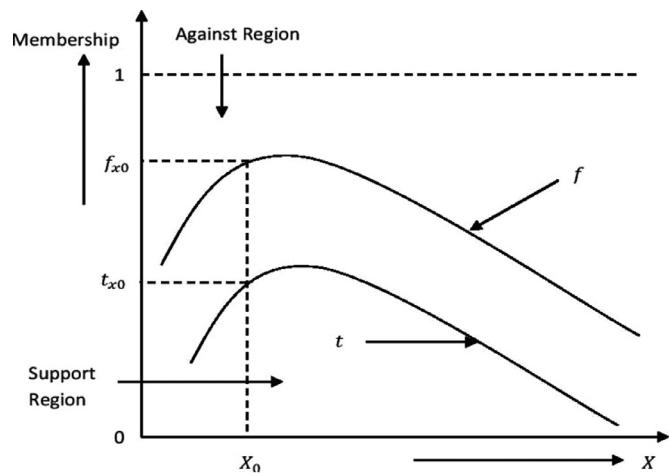


Fig. 1. Vague set.

sider the type-reduction technique proposed by Karnik and Mendel [15].

Interval-valued clustering approaches have shown to have a significant impact on pattern classification. Interval-valued fuzzy c-means (IFCM) [16] and interval-valued fuzzy possibilistic c-means (IPCM) [17] algorithms apply the interval-valued concept to FCM and PCM, respectively. Ji et al., in 2014, proposed interval-valued possibilistic fuzzy c-means clustering (IPFCM) [18] and interval-valued concept of PFCM. However, greater dependence on parameter adjustment makes IPFCM computationally expensive.

Although the concept of possibility solves the shortcomings of fuzzy membership, a single point representation of memberships does not specify the amount of support an object requires to be within a class. Gau, in 1992 [19], solved this problem using the vague set (VS). The VS not only specifies how likely an element is but also specifies its likelihood. A VS is characterized by a truth membership function (t) and a false membership function (f), where t represents the greatest lower bound derived from the evidence in favor of an object for a class, and f represents the greatest lower bound derived from the evidence against that object (Fig. 1). Though the notion of intuitionistic fuzzy set, introduced by Atanassov in 1986 and vague set are same, vague set provides better graphical representation which is more expressive in capturing vagueness of the data. Xu et al., in 2013 [20], proposed a clustering algorithm based on the vague set: vague c-means (VCM). VCM does not consider typicality, and is thus unable to deal with noisy objects using that approach.

Uncertainty is the only reason behind all decision-making. Furthermore, the fuzzy set and vague set were developed to achieve a better understanding of decision-making capability under uncertainty. While performing segmentation, most of the uncertainties lie in the object boundary [21]. The design of fuzzy membership supports the effectiveness of all clusters present in the dataset, while the typicality of an object does not consider the presence of other clusters in the dataset. Our proposed approach ensures that the influence of the core cluster points is maximized, while maintaining the impact of unrepresentative objects at minimal level and also minimizing vagueness with regard to decision-making. However, the use of one membership function is insufficient. Thus, we utilize two memberships, which relaxes the objective function to meet the clustering goal.

3. Background

3.1. Fuzzy c-means (FCM)

FCM is an unsupervised clustering approach that assigns each data point to the classes with some degree of membership. Let $X =$

$\{x_1, x_2, x_3, x_4, x_5 \dots x_N\}$ be a set of N numbers of data points, and C = { $v_1, v_2, v_3, v_4 \dots v_c$ } is a set of c number of prototypes. Initially, we select c cluster prototypes from the dataset randomly then calculate fuzzy membership μ_{ij} , and based on the fuzzy membership we calculate new cluster prototypes (v_i). μ_{ij} and v_i , can be computed as shown in Eq. (1) and Eq. (2).

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}}} \quad (1)$$

$$v_i = \frac{\sum_{j=1}^n \mu_{ij}^m x_j}{\sum_{j=1}^n \mu_{ij}^m} \quad (2)$$

$$J = \sum_{i=1}^c \sum_{j=1}^N \mu_{ij}^m d_{ij}^2 \quad (3)$$

In the equations above, μ_{ij} is the membership of the j_{th} data point to the i_{th} cluster, v_i is the i_{th} cluster prototype, where $i = \{1, 2, \dots, c\}$, and $m > 1$ represents the fuzzifier that controls the fuzziness and d_{ij} is the Euclidian distance between j_{th} data point and the i_{th} cluster.

FCM minimizes the objective function of Eq. (3). In FCM, the sum of the memberships of an object across the classes must be one. Thus, it can avoid the trivial condition of all zeros. Although FCM gives good result for datasets that are overlapping in nature, in some situations it does not work well due to the problem of initialization, Euclidian distance measure, and pre-specification of the number of clusters.

3.2. Possibilistic c-means (PCM)

In 1993, Krishnapuram and Keller [7] introduces a possibilistic approach to clustering. Here is the objective function of PCM.

$$J = \sum_{i=1}^c \sum_{j=1}^N t_{ij}^m d_{ij}^2 + \sum_{i=1}^c \beta_i \sum_{j=1}^N (1 - t_{ij})^m \quad (4)$$

$$t_{ij} = \frac{1}{1 + \left(\frac{d_{ij}}{\beta_i} \right)^{\frac{1}{m-1}}} \quad (5)$$

$$\beta_i = \frac{\sum_{j=1}^N \mu_{ij}^m d_{ij}^2}{\sum_{j=1}^N \mu_{ij}^m} \quad (6)$$

In Eq. (4), N number of data points are distributed among c clusters and t_{ij} is the possibilistic membership of j_{th} object to i_{th} cluster, d_{ij} is the distance measure of j_{th} prototype from i_{th} cluster center, $m > 1$ is the degree of fuzziness and β_i is a positive balance factor that actually measures the average intra-cluster distance of i_{th} cluster. It can be observed that the first section of Eq. (4) is similar with the traditional FCM. However, the membership represents a possibilistic measure that can be obtained from Eq. (5). The possibilistic membership does not consider the idea of sharing of the memberships among the clusters. Therefore, a penalty term $(1-t_{ij})$ is added for the purpose of maintaining a balance with the other clusters.

3.3. Vague c-means (VCM)

In 2013, Xu et al. proposed the VCM [20] clustering algorithm. The VCM uses the concept of the vague set as a framework for FCM clustering. The objective function of a VCM is illustrated using the following equations:

$$J = \sum_{i=1}^c \sum_{j=1}^N t_{ij}^m d_{ij}^2 + \partial \sum_{i=1}^c \sum_{j=1}^N (1 / \max(1 - f_{ij})) \quad (7)$$

$$t_{ij} = \frac{d_{ij}^{-\eta}}{\sum_{i=1}^c d_{ij}^{-\eta}} \quad (8)$$

$$f_{ij} = \frac{d_{ij}^\eta}{\sum_{i=1}^c d_{ij}^\eta} \quad (9)$$

In the equations above, we consider a dataset with N data points distributed among c clusters, truth membership (t_{ij}) and false membership (f_{ij}), where ∂ is a balancing factor between the two parts, η is a positive constant, m is the degree of fuzziness, and d_{ij} is the distance measured between the j_{th} data point and the i_{th} cluster prototype. Quantum-behaved particle swarm optimization (QPSO) [22] is utilized for updating the prototype.

Experimental evidence shows that in the presence of noise, the VCM behaves like an FCM and does not show good clustering results. Both FCM and VCM fail to measure degree of compatibility of an object to a cluster (Section 5 of this paper provides a proper analysis in support of this notion). The VCM also suffers from parameter adjustment, and the original paper on VCM does not provide any useful definition of ∂ . According to the authors of the VCM, when there are two clusters, the degree of vagueness becomes zero, as each object will belong to any of the two clusters. However, this is not a proper representation of vagueness, this concept will not work for noisy objects.

3.4. Interval-valued possibilistic fuzzy c-means (IPFCM)

Zexuan et al., in 2014 [18], proposed the IPFCM, which is an interval-valued concept of PFCM and effectively addresses issues of initialization and sensitivity towards the outlier. Here we show how IPFCM works.

In this regard, m_1 and m_2 represent two different fuzzifiers identifying the upper bound and lower bound of the interval of fuzzy memberships. Similarly, u_1 and u_2 represent two different fuzzifiers identifying the upper bound and lower bound of the interval of possibilistic memberships. Interval representation of the memberships helps with identifying clusters of varying densities.

$$\overline{T}_{ij} = \begin{cases} \frac{1}{1 + \left(\frac{d_{ij}}{\beta_i} \right)^{\frac{1}{m_1-1}}}, & \text{if } u_1 > u_2 \\ \frac{1}{1 + \left(\frac{d_{ij}}{\beta_i} \right)^{\frac{1}{m_2-1}}}, & \text{otherwise.} \end{cases} \quad \underline{T}_{ij} = \begin{cases} \frac{1}{1 + \left(\frac{d_{ij}}{\beta_i} \right)^{\frac{1}{m_1-1}}}, & \text{if } u_1 \leq u_2 \\ \frac{1}{1 + \left(\frac{d_{ij}}{\beta_i} \right)^{\frac{1}{m_2-1}}}, & \text{otherwise.} \end{cases} \quad (10)$$

$$\overline{F}_{kj} = \begin{cases} \frac{1}{\sum_{k=1}^c \left(\frac{d_{kj}}{d_{kj}} \right)^{\frac{2}{m_1-1}}}, & \text{if } m_1 > m_2 \\ \frac{1}{\sum_{k=1}^c \left(\frac{d_{kj}}{d_{kj}} \right)^{\frac{2}{m_2-1}}}, & \text{otherwise.} \end{cases} \quad \underline{F}_{kj} = \begin{cases} \frac{1}{\sum_{k=1}^c \left(\frac{d_{kj}}{d_{kj}} \right)^{\frac{2}{m_1-1}}}, & \text{if } m_1 \leq m_2 \\ \frac{1}{\sum_{k=1}^c \left(\frac{d_{kj}}{d_{kj}} \right)^{\frac{2}{m_2-1}}}, & \text{otherwise.} \end{cases} \quad (11)$$

\overline{T}_{ij} and \underline{T}_{ij} represent the upper and lower membership functions utilizing the fuzzifiers u_1 and u_2 , where $1 < i < c$ and $1 < j < N$ with c number of clusters and N number of data points. \overline{F}_{kj} and \underline{F}_{kj} represent the upper and lower membership functions utilizing the fuzzifiers m_1 and m_2 , where $1 < k < c$ and $1 < j < N$ with c number of clusters and N number of data points. IPFCM suggests a generalized framework that considers FCM, PCM, IFCM, and IPCM, each as a special case. The uncertainty representation of IPFCM is as follows:

$$\begin{cases} \lambda_1 = A\overline{F}_{kj} + B\overline{T}_{ij} \\ \lambda_2 = A\overline{F}_{kj} + B\underline{T}_{ij} \\ \lambda_3 = A\underline{F}_{kj} + B\overline{T}_{ij} \\ \lambda_4 = A\underline{F}_{kj} + B\underline{T}_{ij} \end{cases} \quad \begin{cases} \bar{\lambda} = \max \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} \\ \underline{\lambda} = \min \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} \end{cases} \quad (12)$$

In Eq. (12), λ_1 to λ_4 represent all possible combinations of the upper bound and lower bound of fuzzy membership and typicality while $[\bar{\lambda}, \underline{\lambda}]$ determines the domain of uncertainty for IPFCM. The constants A and B of Eq. (12) determine the relative importance of typicality and fuzzy membership. The important constraints of IPFCM are $0 < A, B \leq 1$ and $A + B = 1$. The domain of uncertainty of IPFCM covers the uncertainty representation of FCM, PCM, IFCM, and IPCM. FCM represents uncertainty by incorporating fuzzy membership with a single fuzzifier ($B = 0, m_1 = m_2$), IFCM represents uncertainty as an interval bounded by upper and lower fuzzy membership functions obtained using a pair of fuzzifiers ($B = 0$), PCM defines uncertainty in terms of a probabilistic membership function utilizing a single fuzzifier ($A = 0, u_1 = u_2$), and IPCM represents the domain of uncertainty as an interval obtained using a pair of fuzzifiers ($A = 0$).

During our experimental study, we noticed that IPFCM increases the search space, considering all possible combinations of the upper bound and lower bound of typicality and fuzzy membership. Consequently, in most cases of MRI image segmentation, it fails to find a local optimum, and the use of a large number of parameters increases the computational complexity.

3.5. Karnik-Mendel (KM) algorithm and type-reduction

Single-valued representation of membership leads to a crisp prototype (Eq. (2)) that is also a single value. According to N.N Karnik and J.M Mendel, the interval-valued membership can be defuzzified differently. This type-reduction approach reduces the type-2 interval-valued fuzzy set to a type-1 set and removes the general complexity of obtaining a crisp prototype of a cluster [23]. KM algorithm, which follows an iterative procedure with an exponential convergence rate. The algorithmic structure of the KM algorithm is shown in Algorithm 1.

Algorithm 1 KM algorithm.

Step 1: Calculation of left centroid v_{jl} ;
 Step 2: For each pattern $x_j, \forall x_j \in X$ and $1 < j < N$ calculate the primary memberships $[\bar{w}_j, w_j]$; [According to TVPFC, \bar{w}_j and w_j represent the fuzzy membership (f_j) and typicality (t_j) of pattern X_j , respectively]
 Step 3: Sort x_j and assign each x_j along with its corresponding primary memberships.
 Step 4: Initialize w_j by $w_j = \frac{\bar{w}_j + w_j}{2}$
 Step 5: Compute $v_j = \frac{\sum_{j=1}^N x_j w_j}{\sum_{j=1}^N w_j}$
 Step 6: Find a switch point γ ($1 \leq \gamma \leq N-1$) such that $x_\gamma \leq v_j \leq x_{\gamma+1}$
 Step 7: Set

$$w_j = \begin{cases} \bar{w}_j, & \text{if } j \leq \gamma \\ w_j, & \text{if } j > \gamma \end{cases}$$

Step 8: Again compute $v'_j = \frac{\sum_{j=1}^N x_j w_j}{\sum_{j=1}^N w_j}$
 check if $v'_j = v_j$ then stop and set $v_{jl} = v_j$. If not then go to Step 6.
 Step 9: Calculation of the right centroid v_{jr} is the same but the only difference is in Step 7. Here it is.

$$w_j = \begin{cases} w_j, & \text{if } j \leq \gamma \\ \bar{w}_j, & \text{if } j > \gamma \end{cases}$$

Step 10: Finally the crisp H_i can be represented as Crisp $v_j = \frac{v_{jl} + v_{jr}}{2}$

4. The proposed TVPFC algorithm

Gau, in 1992 [19], demonstrated how vagueness can be an efficient way to define the uncertainty present in a dataset. Vagueness can be better understood in terms of the degree of belonging, when the degree of belonging of an object cannot be determined, it gives rise to vagueness. If the degree of belonging is a probabilistic measure, we can represent vagueness as a degree of truth (or

truth membership) and a degree of falsehood (or false membership). A vague set demonstrates that instead of representing the degree of belonging as a fuzzy membership, which is a singleton, it is more meaningful to represent the degree of belonging as a continuous subinterval denoting the vague region. Fig. 1 presents a clear illustration of this concept.

In a vague set, the degree of membership of an object is a continuous subinterval of $[0, 1]$. Therefore, if we consider an object x_0 (Fig. 1) and the degree of membership of x_0 as $[0.4, 0.8]$, then t_{x0} is 0.4 and f_{x0} is 0.8. Thus, we can interpret the meaning of $[0.4, 0.8]$ as x_0 takes 0.4 membership in support and $(1 - 0.8) = 0.2$ membership not in support, and the membership of the interval, i.e., $(0.8 - 0.4) = 0.4$, is abstentious. The interval $(f - t)$ represents a vague region, and this vagueness needs to be reduced to reach the optimum. When we know the exact membership of an object, $t = f$ and the interval is 0. The vague set follows some other useful properties such as (i) $t \leq f$, (ii) $t+f \leq 1$, (iii) $t \leq \mu \leq f$. where μ represents the fuzzy membership.

Both fuzzy membership and typicality have some useful contributions to representing uncertainty. The influence of fuzzy membership and typicality on an object varies depending on the object's distance from a cluster prototype and the position of the other clusters. In some situations, however, both FCM and PCM do not perform as expected. Furthermore, PCM uses only probabilistic membership and FCM uses fuzzy membership. The probabilistic membership of FCM makes it challenging to decide how compatible an object is to a cluster. Moreover, fuzzy membership creates a problem for noisy objects as it assigns an equal degree of membership, regardless of how far or close the object is from the cluster prototype. In contrast, PCM relaxes the probabilistic constraint of FCM but needs good initialization. Otherwise, it may result in coincidence clusters. Thus, careful utilization is required for both FCM and PCM. In the TVPFC, we have proposed a vague region where typicality is upper bounded by fuzzy membership. It prevents an object to be completely independent and doesn't let coincidence clusters to happen.

The vagueness model of TVPFC is presented in Fig. 1. The objective function of TVPFC is expressed in Eq. (13), where truth membership (t_{ij}) of the vagueness model is represented by typicality as it demonstrates an independent nature in determining the objective function and fuzzy membership as false membership (f_{ij}) due to its sharing nature. The region in between denotes the hesitation (vague) region.

In Eq. (13), we consider $C = \{v_1, v_2, \dots, v_c\}$, c number of clusters to be present in the dataset, d_{ij} is the city block distance between the j_{th} object and the i_{th} cluster prototype, μ and m represent the degree of fuzziness for typicality and fuzzy membership, respectively,

$$J_{\mu, m}(C, U, T) = \sum_{i=1}^c \sum_{j=1}^N t_{ij}^\mu d_{ij}^2 + \sum_{i=1}^c \beta_i \sum_{j=1}^N (1 - f_{ij})^m \quad (13)$$

N represents the total number of objects within the dataset, T is a $C \times N$ matrix, represents the probabilistic C -partition matrix, U is a $C \times N$ matrix, represents the fuzzy C -partition matrix, satisfying the conditions that, $\forall t_{ij} \in T, 0 \leq t_{ij} \leq 1, 0 < \sum_{j=1}^N t_{ij} \leq N, \max t_{ij} > 0$ for probabilistic membership t_{ij} and $\forall f_{ij} \in U, 0 \leq f_{ij} \leq 1, 0 \leq \sum_{j=1}^N f_{ij} \leq N$ and $\sum_{i=1}^c f_{ij} = 1$ for fuzzy membership f_{ij} . In these conditions what follows, probabilistic membership constrain relaxes the fuzzy membership constrain $\sum_{j=1}^N f_{ij} = 1$ to $\max_i t_{ij} >$

0. In Eq. (13), β_i represents the average intra-cluster distance that acts as a balancing factor between the upper and lower bound of the vagueness model.

Minimization of the objective function in Eq. (13), demands the distance value d_{ij} of the first term to be low, which eventually causes typicality and fuzzy membership to be high and influences

the second term to be low. This reduces the unrepresentative part of the membership, therefore, the objective function is definitely a minimization function that minimizes the intra-cluster distance.

Theorem. If $X = \{x_1, x_2, x_3 \dots x_N\}$ represents universe of discourse with N number of objects and $C = \{v_1, v_2, v_3 \dots v_c\}$ is a set of c number of clusters with $d_{ij} \geq 0$, denoting the distance between an object and the cluster prototype, then $J_{\mu,m}(C, U, T)$ will be minimum only if $t_{ij} = [1 + (d_{ij}^2/\beta_i)^{\frac{1}{\mu-1}}]^{-1}$ and the difference between f_{ij} and t_{ij} , i.e., $f_{ij} - t_{ij} = 0$. f_{ij} and t_{ij} follow the necessary conditions mentioned before.

Proof. To get the minimum of $J_{\mu,m}$, we need to find the minimum of each additive term of $J_{\mu,m}$, as $J_{\mu,m}$ represents the sum of non-negative terms, thus, a minimum of each term corresponds to the minimum of $J_{\mu,m}$. Another important point is $f_{ij} - t_{ij} = 0$, which suggests that vagueness must be reduced. Therefore, we can find the minimum by solving the following equation and considering Lagrange's multiplier theorem. To simplify the derivation, we consider $\mu = m$ but do not restrict this in application.

$$\min \left\{ J_{\mu,m}(C, U, T) = \sum_{i=1}^c \sum_{j=1}^N t_{ij}^m d_{ij}^2 + \sum_{i=1}^c \beta_i \sum_{j=1}^N (1 - f_{ij})^m \right. \\ \left. - \sum_{i=1}^c \sum_{j=1}^N \lambda_{ij} (f_{ij} - t_{ij}) \right\} \quad (14)$$

$$\frac{\Delta J}{\Delta t_{ij}} = mt_{ij}^{m-1} d_{ij}^2 + \lambda_{ij} = 0 \quad (15)$$

$$\frac{\Delta J}{\Delta f_{ij}} = -m\beta_i(1 - f_{ij})^{m-1} - \lambda_{ij} = 0 \quad (16)$$

$$\frac{\Delta J}{\Delta \lambda_{ij}} = f_{ij} - t_{ij} = 0 \quad (17)$$

$$t_{ij} = f_{ij} = \frac{1}{1 + (\frac{d_{ij}^2}{\beta_i})^{\frac{1}{m-1}}} \quad (18)$$

□

Therefore, we can say that when we reduce vagueness utilizing our proposed model, fuzzy membership shows a possibilistic nature. Increasing typicality allows fuzzy membership to consider only core cluster objects, and consequently, not be influenced by the presence of outliers. Furthermore, β is sensitive to initialization, and fuzzy membership is used for the calculation of β as proposed by Krishnapuram and Keller [7].

Parameter β_i defines the shape and structure of a cluster. In TVPFC, it also works as a balancing factor between the membership and non-membership for each object assigned to a cluster. We estimate (Eq. (6)) the value of β_i that is the average intra-cluster distance. This average weight helps in keeping track of the homogeneity of the surrounding objects as the cluster moves through iterations and update its center continuously. Selection of β_i for a cluster can effect the overall clustering result. While preparing for TVPFC we have found that if we choose to modify β_i with every iteration then clustering becomes very unstable and some times degrades. This is because every time when we update β_i with the updated cluster center we will lost the opportunity to measure how much it gets deviated from the actual weighted average. Actual weighted average provides the idea of expected shape of a cluster. Which we can not get from the randomly chosen initial clusters. Therefore, we can choose β either from random initialization and re-estimate it after convergence or we can start with initial cluster centers obtained by using any fuzzy clustering on that data set.

Algorithm 2 TVPFC.

Section 1 Input: Number of cluster prototype c , $m \in [1, \infty]$, $\mu \in [1, \infty]$; Step 2: Iteration counter $k = 1$; Step 3: Initialize the fuzzy partition f randomly; Step 4: Estimate β_i , $i=1 \dots c$ using Eq.~(6); Step 5: Estimate t of an object using Eq.~(5); Step 6: Compute J_k ; Step 7: Update prototype as in KM algorithm [Algorithm 1]; Step 8: Recompute t and f of an object using Eq.~(5) and Eq.~(1) respectively; Step 9: Increment k and compute J_{k+1} ; Step 10: **Repeat** Step 7 to Step 9 Until $|J_{k+1} - J_k| < \epsilon$; [ϵ is a very small number][The following steps are not required if fuzzy initialization is done using any clustering algorithm]

Section 2 Input: Cluster prototypes obtained from Section 1, $m \in [1, \infty]$, $\mu \in [1, \infty]$; Step 11: Re-estimate β_i , $i = 1, \dots, c$ using Eq.~(6); Step 12: Estimate t of an object using Eq.~(5) and f using Eq.~(1); Step 13: Compute J_k ; Step 14: Update prototype as in KM algorithm [Algorithm 1]; Step 15: Recompute t and f of an object using Eq.~(5) and Eq.~(1) respectively; Step 16: Increment k and Compute J_{k+1} ; Step 17: **Repeat** Step 14 to Step 16 Until $|J_{k+1} - J_k| < \epsilon$;

Here we provide the algorithmic structure of TVPFC (Algorithm 2) and illustrate steps for the application of TVPFC on images.

For medical image segmentation, precise detailing of the area specific information is important. From the perspective of local information processing, KM (Karnik-Mendel) algorithm plays a very important role. Unlike pixel by pixel processing of the local information, TVPFC tries to optimize the ambiguity of assigning membership values by optimizing the interval $[t_{ij} - f_{ij}]$ for each pixel. This helps in processing low resolution gray scale MRI and CT scan images. Here we illustrate the major steps of TVPFC for image segmentation. Let us consider an image dataset X with N number of pixels is segmented into c number of homogeneous clusters. We choose initial cluster prototypes and compute the city block distance (d_{ij}) between every pixel x_j and the cluster prototype v_i . Calculate β_i for every cluster, it represents the average intra-cluster distance. With the help of d_{ij} and β_i we estimate the absolute membership t_{ij} (typicality) of every pixel for every cluster and we also estimate relative membership f_{ij} (fuzzy membership) using d_{ij} .

TVPFC and image segmentation

1. Assign initial pixel values that define cluster prototypes v_i , $i=1, \dots, c$. v_i must lie within the maximum and minimum intensity values of image dataset X . Choose values for fuzzifier, m and μ , $m \neq \mu$. Set iteration counter $k=1$.
2. Calculate city block distance $d_{ij} = |v_i - x_j|$, $\forall x_j \in X$, $i=1, \dots, c$ and $j=1, \dots, N$.
3. Compute β_i , $i=1, \dots, c$ using Eq. (6).
4. Compute t_{ij} and f_{ij} for $i=1, \dots, c$ clusters and $j=1, \dots, N$ pixels using, Eq. (5) and Eq. (1).
5. Compute J_k using Eq. (13).
6. Update prototype v_i using KM algorithm.
7. **Repeat** Step 2, 4, 5 and 6 until $|J_{k+1} - J_k| < \epsilon$ [ϵ is a very small number][We need to go for a second run through this algorithm after convergence unless we select the initial cluster prototypes using any fuzzy clustering algorithm]

5. Noise handling

We analyze how TVPFC handles vagueness. We consider three possible cases: Case 1 is on core cluster points that are definite members of a cluster. In this case, the distance of an object from the cluster prototype is very small, and thus, both fuzzy membership and typicality show high values that are almost the same. Therefore, vagueness is minimized as $f \simeq t$. Case 2 is on noisy objects when the distance of an object from the cluster prototype is large.

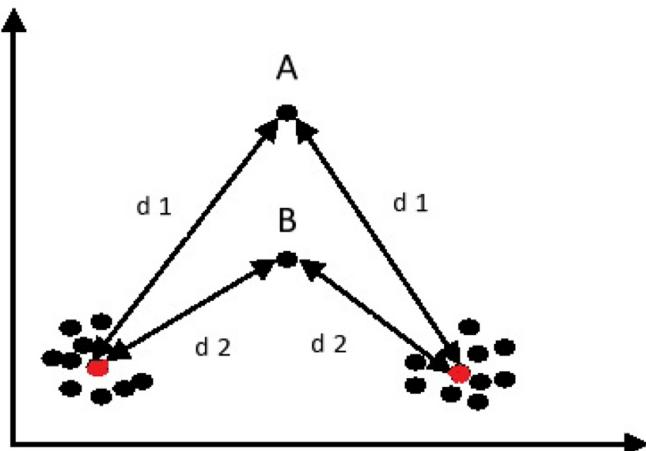


Fig. 2. Noise A and B at a distance d_1 and d_2 respectively from the clusters.

Both fuzzy membership and typicality show low values, but in practice, we have found that fuzzy membership is higher than typicality as fuzzy membership considers the location of the other clusters, where the object may have a shorter distance. Therefore, vagueness is reduced, to some extent, as some objects are excluded from the cluster due to the lower memberships. Case 3 is when objects are on the border of one or more clusters. The sharing nature of fuzzy membership always takes a higher value than the typicality of the border objects. Hence, we chose fuzzy membership as the upper bound of the vagueness model (Fig. 1) and typicality as the lower bound.

Fig. 2 presents a two-cluster dataset with the prototypes denoted by different colors. There are two noise points A and B with distance d_1 and d_2 , respectively. Based on VCM, the truth membership and false membership of A and B for the two clusters are as follows:

$$\begin{cases} t_A = \frac{\frac{1}{d_1}}{\frac{1}{d_1} + \frac{1}{d_1}} = 0.5 \\ f_A = \frac{d_1}{d_1 + d_1} = 0.5 \end{cases} \quad \begin{cases} t_B = \frac{\frac{1}{d_2}}{\frac{1}{d_2} + \frac{1}{d_2}} = 0.5 \\ f_B = \frac{d_2}{d_2 + d_2} = 0.5 \end{cases}$$

From the above expression, it can be seen that, based on Eqs. (8) and (9), the truth membership (t_A) and false membership (f_A) of A are the same because the distance of A from both clusters is equal, and the same is true for B. Furthermore, even with the significant difference between d_1 and d_2 , the numeric values of the truth membership and false membership are the same for both A and B. Thus, it is clear that in terms of typicality VCM does not detect any difference between the noise points.

We have seen that both IPFCM and TVPFC consider typicality. Again, considering Fig. 2, Eqs. (10), and (11), we have:

$$\begin{cases} \bar{F}_A = 0.5, & \text{for } m_1 = 2 \\ F_A = 0.5, & \text{for } m_2 = 5 \end{cases} \quad \begin{cases} \bar{T}_A = \frac{1}{1+d_1}, & \text{for } \mu_1 = 2 \\ T_A = \frac{1}{1+d_1^{\frac{1}{4}}}, & \text{for } \mu_2 = 5 \end{cases}$$

After going through the stages of Eq. (12) we obtain:

$$\begin{cases} \bar{\lambda} = 0.5 + \frac{1}{1+d_1^{\frac{1}{4}}} \\ \lambda = 0.5 + \frac{1}{1+d_1} \end{cases}$$

In Eq. (12), $\bar{\lambda}$ and λ represent the interval for A. From this expression, it is clear that after type-reduction of this interval, A will always get a higher membership value. However, this should not be expected when dealing with noise and outliers. Further, in com-

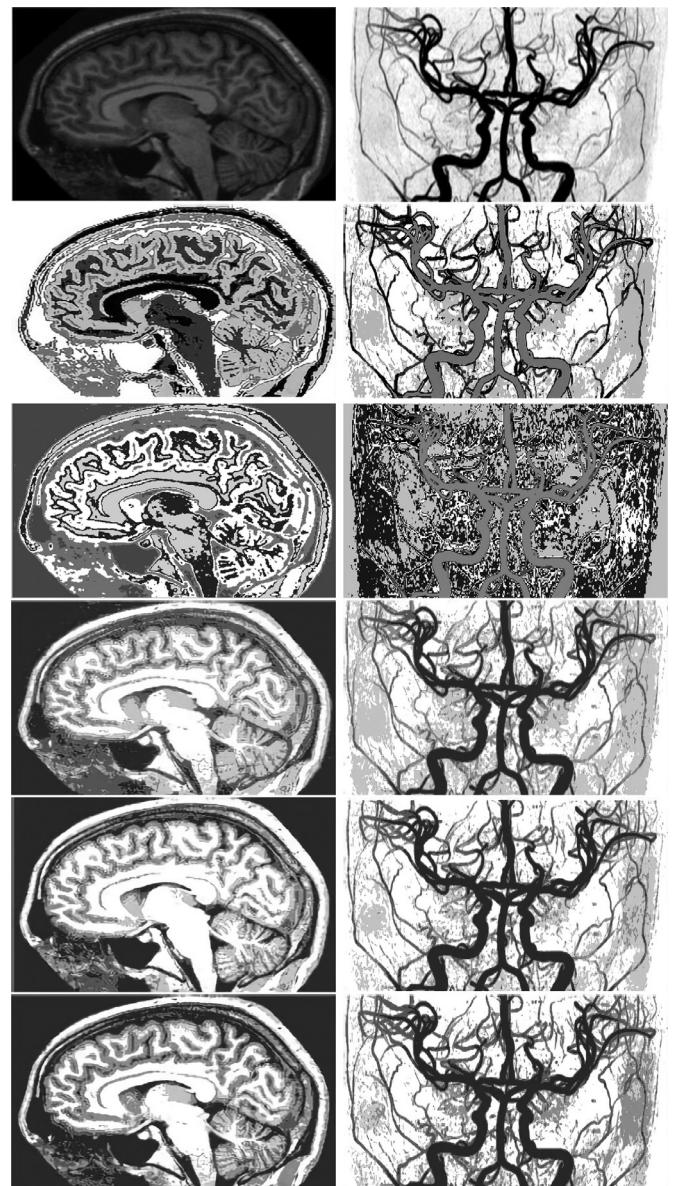


Fig. 3. Illustrates segmentation of T1-weighted brain MRI scan and anatomical structure of brain vessel with different thickness levels (left to right of Row 1). Row1: Original MRI of Brain and Brain vessel, Row2: Segmented images with FCM, Row3: Segmented images with PCM, Row4: Segmented images with VCM, Row5: Segmented images with IPFCM, Row6: Segmented images with TVPFC.

parison with IPFCM, TVPFC will define the interval as follows:

$$\begin{cases} \bar{w}_A = 0.5 \\ w_A = \frac{1}{1+d_1} \end{cases} \quad \text{or} \quad \begin{cases} \bar{w}_A = \frac{1}{1+d_1} \\ w_A = 0.5 \end{cases}$$

In practice, we found that noisy objects generally have a larger d_1 . Consequently, TVPFC will consider the first part of the following expression, and type-reduction, TVPFC will have a membership value of A that is less than 0.5. \bar{w}_A and w_A are the primary memberships of object A in Fig. 2.

6. Experimental results

In this experimental analysis section, we discuss the performance of proposed TVPFC and we also have a comparative study of FCM, PCM, VCM, IPFCM, and TVPFC on various medical images. We precisely chose IPFCM for comparison because it covers IPFCM, IFCM as special case and performs better than PFCM.

Table 1
Datasets.

Database name	No of Images	Colour Space	Ground Truth	Database Description	Database Objective	Data Sources
Database1	2	Gray Scale	Not Available	1. T1-weighted MRI scan of a 74-year old woman. 2. Maximum intensity projection of the coronal plane showing brain vessel of different thickness levels	There is no clinical abnormality, overall image is segmented into multiple regions	www.sciencesource.com
1. LIDC-IDRI 2.TCGA-BRCA, TCGA-PRAD, TCGA-READ 3. NSCLC	82 14, 30, 149 26	Gray Scale	Available	1. lung cancer image database containing computed tomography (CT) scans. 2. Cancer genome atlas for prostate adenocarcinoma. 3. Database on non-small cell lung cancer containing thorax CT scan images.	1. We have tried to separate lung parenchyma nodule. 2. & 3. objective was to segment the lesion that denotes the cancerous region and for comparative study	http://dx.doi.org/10.1111/1JMI.3.4.044504 http://cancergenome.nih.gov/ https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI
Database2	12	Gray Scale	Available	simulator generated T1 - weighted brain MRI images with different noise levels for slice thickness 1mm and 3mm	performance of the algorithms with the increasing noise levels ranging from 1% to 9%(Gaussian noise)	http://www.bic.mni.mcgill.ca/brainweb/
Database3	3	Gray Scale, SPECT imaging	Available	1. Two CT scan images of a 35-year-old man with low glioma (a type of tumor). 2. CT scan image of a 42-year-old woman with a large tumor that is compressing the adjacent midbrain. (For each of these CT scan images, there are corresponding MR images, which indicate signal intensity based on proton density.)	Segmentation is done to identify tumor region.	http://www.med.harvard.edu/AANLIB/home.html

6.1. Dataset description

The Datasets considered for the entire experimental analysis is shown in

Table 1. It is divided into four different sets depending on the type of MRI images available, and the processing approach applied.

6.2. Performance metrics

We considered different measurement techniques for comparative analysis of the proposed TVPFC against FCM, VCM, and IPFCM. For images in Fig. 3, there is no ground truth available. Thus, we used validity indices such as the Beta, Davies-Bouldin, and Xei-Beni indices to evaluate the robustness of the above-mentioned algorithms. Figs. 4 and 6 present ground truth images obtained using an interactive image segmentation approach [24] based on the available clinical data. Although this approach does not provide the exact ground truth, it provides some useful information. For the CT scan images in Fig. 4, we calculated the accuracy (Eq. (19)) of the segmented images. For the images in Fig. 6, we also calculated the L_E of the lesion, as in Eq. (20). On the MRI images for the last segment, Fig. 5 shows one instance with 5% noise, and the segmented results are validated with the index measures considering the available ground truth.

In Eq. (19), TP, TN, FP, and FN represent the true-positive, true-negative, false-positive, and false-negative, respectively. Accuracy defines the misclassification within a pre-specified region (ground truth), while the lesion error (L_E) of Eq. (20) identifies whether the lesion size changes in the segmented image, and L_E can identify this error effectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (19)$$

$$L_E = \frac{2 * abs(L_A - L_S)}{L_A + L_S} \quad (20)$$

In Eq. (20), L_A represents the actual lesion density and L_S is the segmented lesion density. We chose to measure the density of an area by counting the number of pixels or intensity values present in that area. Inspired by the Dice similarity coefficient (DSC) [25], we calculated (L_E). DSC measures the amount of association among the sets. The association indicates the amount of overlap. If a and b are two sets, and h represents the amount of overlap, the association index for a is $\frac{h}{a}$, and for b, it will be $\frac{h}{b}$. Thus, the formula for DSC is $\frac{h}{a} + \frac{h}{b} = \frac{2h}{a+b}$. In this study, it is necessary to measure the change in the size of the lesion after segmentation. To achieve this, we prefer to measure the misclassification error and we termed it lesion error (L_E), where h represents the amount of dissimilarity between the sets.

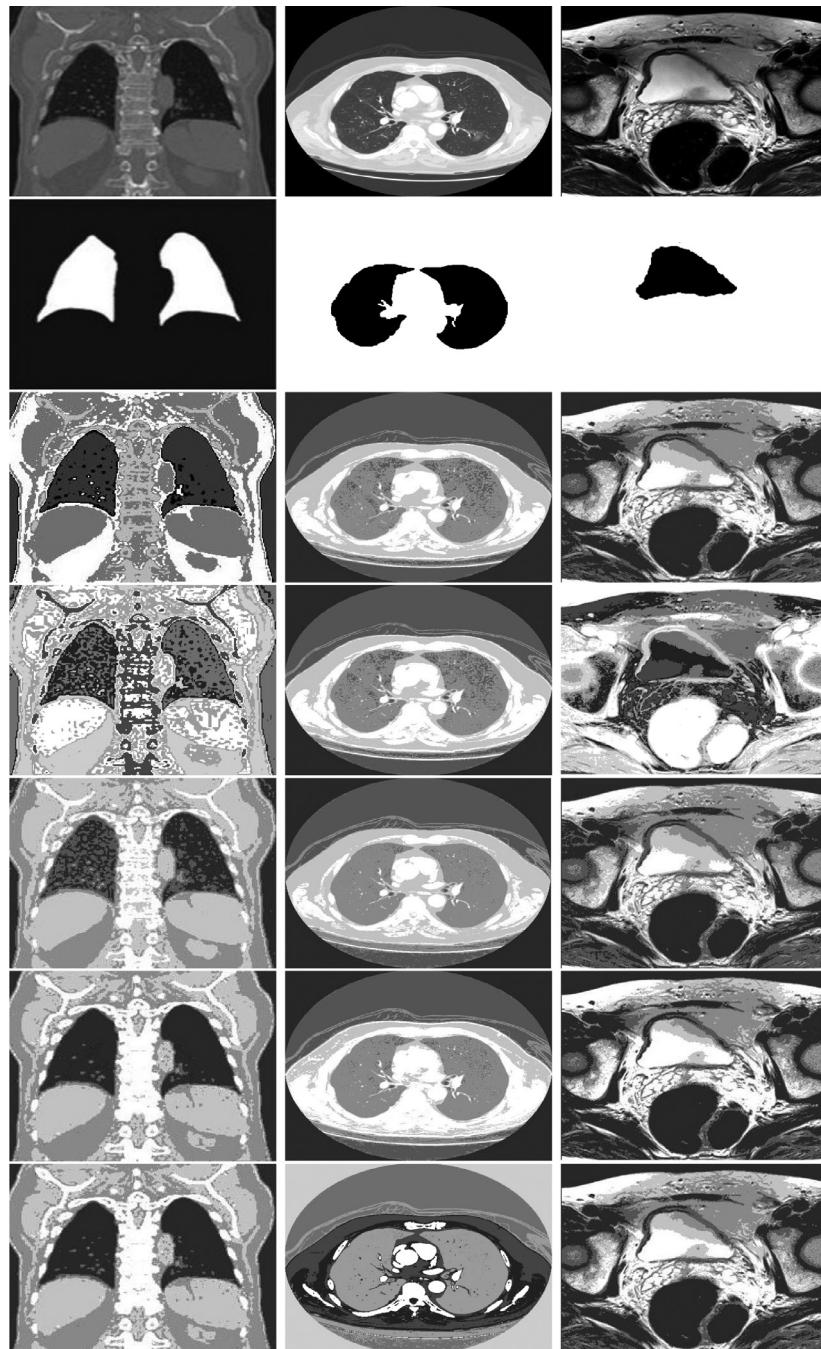


Fig. 4. CT scan of thoracic structure, non small-cell lung cancer and prostate cancer has been shown in the original image (left to right Row 1), ground truth images indicates cancerous lung parenchyma nodule and lesion indicating cancerous region for prostate cancer. Row1: Original CT scan of Thorax, Lung Cancer and Prostate Cancer, Row2: Ground truth images, Row3: Segmented images with FCM, Row4: Segmented images with PCM, Row5: Segmented images with VCM, Row6: Segmented images with IPFCM, Row7: Segmented images with TVPFC.

6.3. Discussion

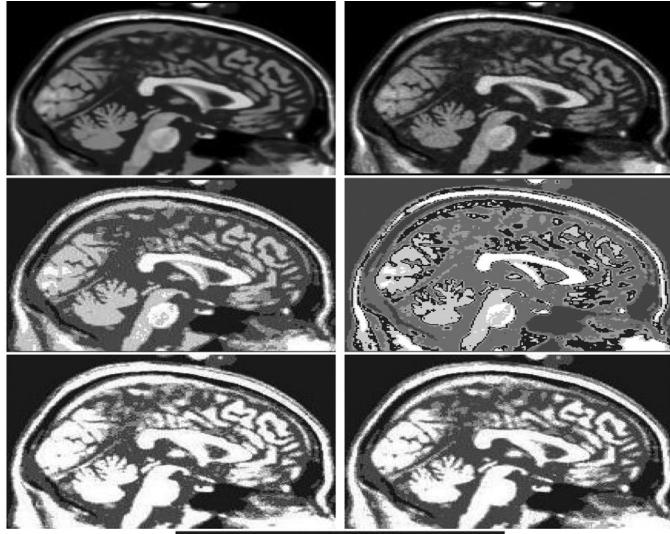
6.3.1. Challenges

This section facilitates the proficiency of TVPFC in dealing with some of the artifacts of MRI and CT scan images.

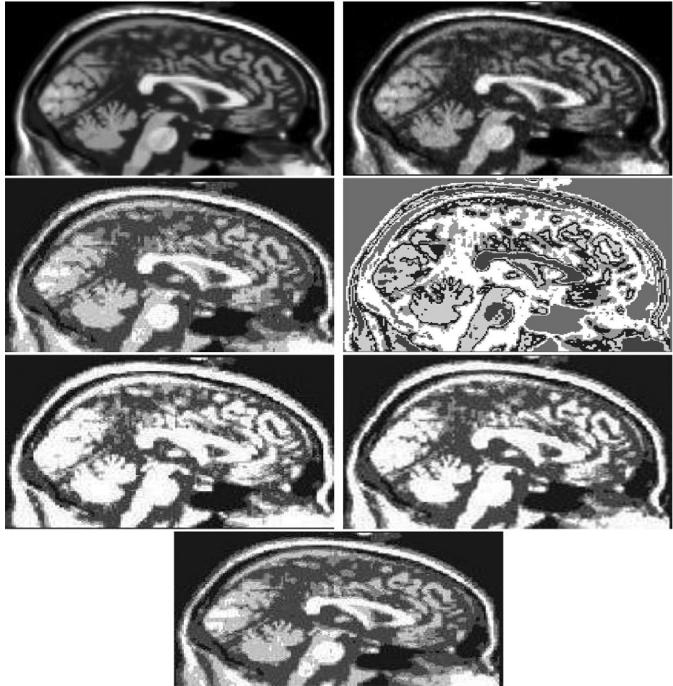
Partial volume effect: MRI images mostly suffer with the partial volume effect because of low resolution, when more tissues occupy the same intensity values and intensity overlap causes a blurring effect across the object boundary. Fig. 7 presents the partial volume effect and intensity inhomogeneity. Commonly used fuzzy clustering algorithms exclusively distribute the fuzzy memberships,

but a single membership does not contain much information regarding uncertainty. The vagueness model of TVPFC supports the overlapping nature of the clusters as it has an interval-valued structure of the membership distribution that uncovers more information regarding uncertainty and iterations through the vague region, thus, effectively deal with ambiguous structural information.

Intensity inhomogeneity Can be the result of the imperfection of the radio frequency coils and gives a shading effect throughout the image. As a result, it gives a shading effect throughout the image. In intensity-based clustering, homogeneity can be attained



(a) Slice Thickness 1mm



(b) Slice Thickness 3mm

Fig. 5. Segmentation of simulator generated brain MR image with 5% noise for (a) and (b), Row1(left to right) noise free image and noisy image, Row2 (left to right) FCM segmented image and PCM segmented image, Row3 (left to right) VCM segmented image and IPFCM segmented image, Row4 TVPFC segmented image.

if we can keep track of the actual average intensity of the surroundings and its deviation after segmentation. In TVPFC, β_i represent the average intra-cluster distance and does not get changed through algorithmic iterations as it retains the effect of initialization. We thus consider it a true intensity distribution if β_i denotes the re-estimation or a fuzzy initialization obtained from any fuzzy clustering algorithm. Furthermore, the multiplicative term

Table 2

Validity indices measure of the segmentation results shown in Fig. 3.

Algorithm	Images	Beta	Davies-Bouldin	Xei-Beni
FCM	MRI	2.9196	1.0457	0.7599
	Brain vessel	2.0341	0.8107	0.9012
PCM	MRI	1.1499	2.0179	1.2210
	Brain vessel	1.5162	1.0213	1.0021
VCM	MRI	3.1892	1.1273	1.0421
	Brain vessel	1.9732	0.8232	0.9142
IPFCM	MRI	3.3991	1.0032	0.6521
	Brain vessel	2.1605	0.7911	0.7456
TVPFC	MRI	3.6724	0.9085	0.5529
	Brain vessel	2.2857	0.7693	0.6831

Table 3

Measurement of Accuracy of the segmentation results shown in Fig. 4.

Algorithms	Accuracy %
FCM	74±6.2
PCM	52±4.2
VCM	57±3.1
IPFCM	85±3.2
TVPFC	92±2.2

$((1 - f)^m)$, acts as an information gain factor because high fuzzy membership (f) not only minimizes the term $((1 - f)^m)$, but also gains more information from the against region. Thus, the information gain in any iteration is influenced by the true intensity distribution and results in homogeneous distribution of the intensities.

The presence of noise in MRI and CT scan images can occur due to hardware adjustment. TVPFC shows compatible attributes for handling noise, as discussed in Section 5, and TVPFC handles noise present in the images more efficiently than other existing algorithms.

6.3.2. Experimental analysis

The entire experimental study considers different types of CT scans and MRI images, and a comparative analysis of TVPFC against FCM, PCM, IPFCM, and VCM was designed. The operational model of TVPFC uses only three parameters: β , the balancing factor for the upper and lower interval of the vague environment, fuzzifier m , and μ . From the experimental evidence, we see that for $m = 2$ and $\mu = 5$, the vagueness model of TVPFC exhibits maximum fuzziness. Thus, throughout the experimental analysis using TVPFC, we chose $m = 2$ and $\mu = 5$. Fuzzy clustering FABC [26] was used for the randomized initialization of the cluster prototypes and we chose fixed number of clusters to maintain consistency in comparative analysis.

From Table 2, it can be seen that the increased fuzziness of the vagueness model of TVPFC can more effectively identify different thickness levels of the brain vessel (Fig. 3) than other algorithms, and the validity indices measures in Table 2 reflects the robustness of TVPFC. Table 3 presents the accuracy measures of the images in Fig. 4, along with the available ground truth. Each measurement of the form $a \pm b$ denotes ' a ' as the mean and ' b ' as the standard deviation. From the segmented images, we evaluated the amount of convergence in the highlighted ground truth region using the segmented output. It can be seen that TVPFC shows superior accuracy over FCM, PCM, VCM, and IPFCM.

The experimental analysis presented in Fig. 5 substantiates that the proposed approach not only provides accurate clustering results but is also only slightly affected by increasing noise. To support this hypothesis, a Chi-square test statistic of the analysis of variance presented in Tables 4 and 5 is provided. In Tables 6 and

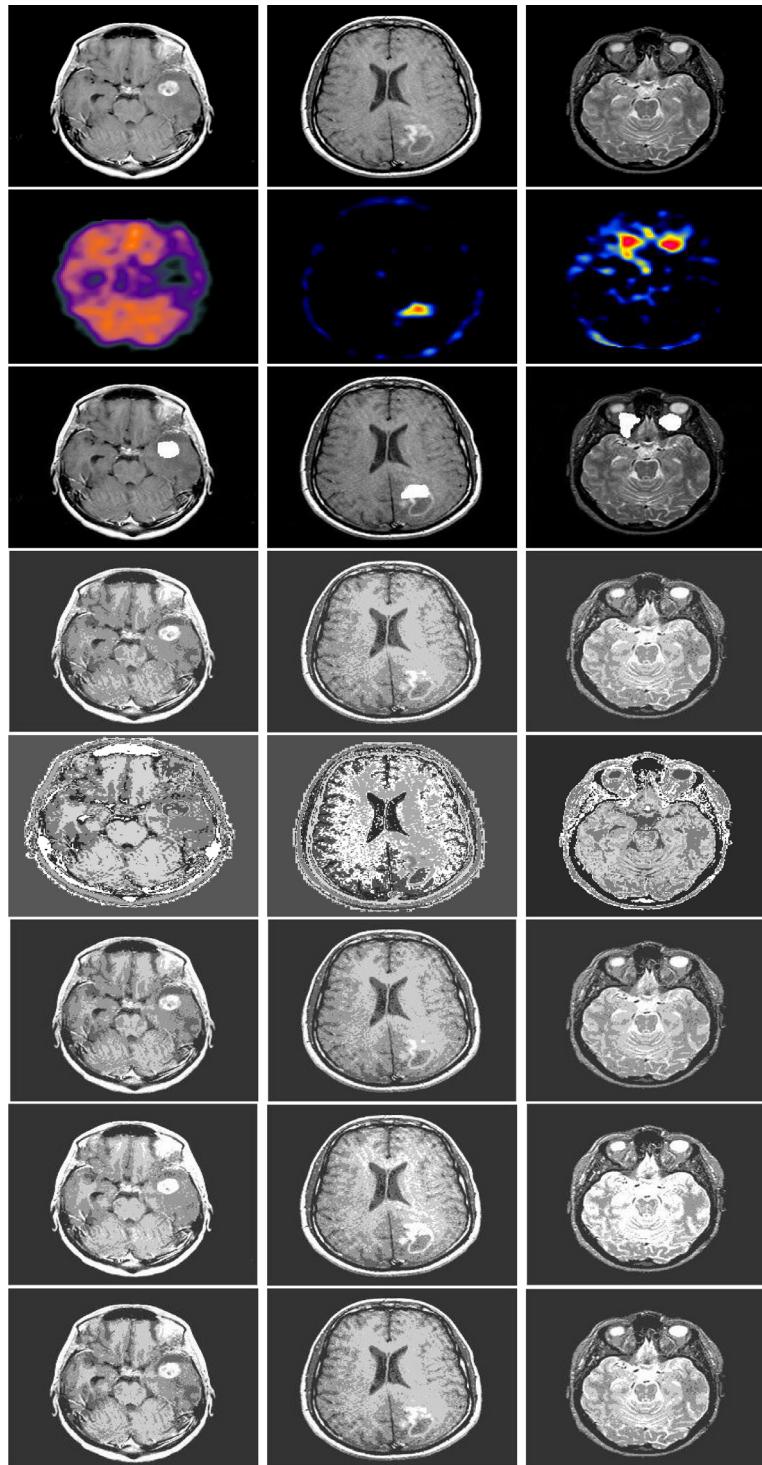


Fig. 6. CT scan images with tumor and their corresponding MR Spectroscopy images, high proton density indicates location of the tumor, Row1: Original CT scan images, Row2: MR Spectroscopy images, Row3: Ground truth images, Row4: Segmented images with FCM, Row5: Segmented images with PCM, Row6: Segmented images with VCM, Row7: Segmented images with IPFCM, Row8: Segmented images with TVPFC.

7, h denotes the test decision, which supports the null hypothesis that the value of variance v supports the respective distribution. From the value of v , it can be seen that TVPFC performs better than the other algorithms specified, and is less susceptible to noise and outliers.

Table 8 presents measures of the accuracy as well as the lesion error (L_E) of marked lesion present in the CT scan and segmented images in Fig. 6. This shows the efficiency of TVPFC for

area-specific diagnosis of medical MRI images. The MR images in Fig. 6 mark the tumor as the presence of a tumor increases blood flow and is thus captured in the MR images. To measure the efficiency of the aforementioned algorithms, in terms of CPU time and number of iterations, we considered all the images used in this study. For each image, we performed 20 iterations with randomized initialization of the cluster prototypes. The average CPU time and iteration count are listed in Table 10 and CPU time of some

Table 4
Validity indices measure of the segmentation results shown in Fig. 5a.

Algorithm	Indices	0% noise	1% noise	3% noise	5% noise	7% noise	9% noise
FCM	Beta	7.1030	7.0716	7.0681	6.0982	6.2116	6.0761
	Davies-Bouldin	2.3530	2.5520	5.2043	6.8656	5.4958	5.5396
	Xei-Beni	0.3276	0.1689	4.9292	4.8318	4.4409	3.6813
PCM	Beta	5.7724	5.5710	5.5272	4.7662	4.6100	4.4869
	Davies-Bouldin	2.8825	2.8370	5.0031	5.1652	6.1123	6.3942
	Xei-Beni	1.4111	1.3680	2.8239	4.5842	4.5933	3.6813
VCM	Beta	6.7921	6.8963	6.7667	6.7964	6.7633	6.2471
	Davies-Bouldin	2.5131	2.1082	7.8570	7.6701	4.6450	2.4707
	Xei-Beni	0.8625	0.2039	3.2232	3.3192	1.3610	2.5230
IPFCM	Beta	7.2150	7.6916	7.0610	7.3714	7.2095	5.6642
	Davies-Bouldin	2.3636	2.3517	7.3913	6.3639	7.3736	7.4707
	Xei-Beni	0.3274	0.2444	3.8788	2.9746	3.5462	1.5230
TVPFC	Beta	7.9037	7.8924	7.8733	7.8591	7.4681	6.4650
	Davies-Bouldin	2.0971	1.9684	2.0648	2.1645	3.7257	3.8435
	Xei-Beni	0.0835	0.0786	0.0716	0.1022	0.1580	0.5761

Table 5
Validity indices measure of the segmentation results shown in Fig. 5b.

Algorithm	Indices	0% noise	1% noise	3% noise	5% noise	7% noise	9% noise
FCM	Beta	7.5795	7.5516	7.6279	7.4192	6.6184	6.5976
	Davies-Bouldin	2.9299	3.0774	4.1314	4.2002	4.8852	5.6340
	Xei-Beni	0.3285	1.2577	2.7776	4.6812	4.8920	4.4310
PCM	Beta	6.8755	6.5719	5.8270	6.2602	5.8173	5.1139
	Davies-Bouldin	3.7232	3.6116	3.5275	4.1573	4.8965	5.7843
	Xei-Beni	1.4328	1.3894	2.8567	3.9866	4.9470	4.9872
VCM	Beta	7.1240	7.1260	7.5609	6.9017	6.2738	6.2520
	Davies-Bouldin	2.8384	3.2048	5.9462	5.7243	4.5323	4.4797
	Xei-Beni	1.7537	2.6633	2.1092	4.7360	3.9413	4.1717
IPFCM	Beta	7.7608	7.2628	7.4146	7.8044	7.8844	7.2511
	Davies-Bouldin	2.4416	5.7702	5.7496	3.5569	3.2572	4.6912
	Xei-Beni	0.2405	1.0339	1.5208	2.0620	2.2675	3.2112
TVPFC	Beta	8.5599	8.5218	8.5197	8.3300	8.2365	7.4408
	Davies-Bouldin	1.6946	1.0699	1.7991	1.8112	2.2706	2.2785
	Xei-Beni	0.0814	0.0772	0.1346	0.3160	1.4205	1.4493



Fig. 7. Challenges of MRI and CT scan images.

Table 8
Measurement of Lesion Error and Accuracy of CT scan and MR images of Fig. 6.

Algorithms	Lesion Error (L_E)%	Accuracy %
FCM	51±2.5	66±3.2
PCM	69±3.6	54±7.2
VCM	61±5.3	62±4.3
IPFCM	32±3.2	77±3
TVPFC	24±2.7	94±2.5

Table 6
Variance measure of the validity indices of Table 4.

Algorithm	Beta($h=0$)	Davies-Bouldin($h=0$)	Xei-Beni($h=0$)
FCM	v=0.20	v=1.30	v=0.20
PCM	v=1	v=1	v=0.90
VCM	v=0.30	v=2.80	v=0.40
IPFCM	v=0.20	v=2.50	v=1.10
TVPFC	v=0.20	v=1.00	v=0.02

Table 7
Variance measure of validity indices of Table 5.

Algorithm	Beta($h=0$)	Davies-Bouldin($h=0$)	Xei-Beni($h=0$)
FCM	v=0.10	v=0.50	v=1.5
PCM	v=0.20	v=0.40	v=1.1
VCM	v=0.20	v=0.70	v=0.60
IPFCM	v=0.04	v=0.80	v=0.50
TVPFC	v=0.08	v=0.08	v=0.20

selective images presented in Fig. 8. It shows that TVPFC takes a minimum number of iterations to produce high quality segmented images. Only the result for FCM comes close to that of TVPFC. Although FCM takes less CPU time in some of the cases, it is not able to produce high quality segmented images.

Again, we obtain the inverted Xei-Beni index value of a set of 30 CT scans images taken from TCGA-PRAD. It can be seen in Fig. 9, even with the randomized initialization, that TVPFC shows consistently better performance throughout the dataset than FCM, PCM, VCM, and IPFCM.

Furthermore, Tables 4 and 5 present a comparative study of the algorithms for MRI images with a slice thickness of 1mm and 3mm for different noise levels ranging from 0% to 9%. Tables 6 and 7 indicate that even with the changes in noise levels, in terms of variance, TVPFC shows less dispersion. Again, we can see the consistent performance of TVPFC in Fig. 10. It shows the variance of

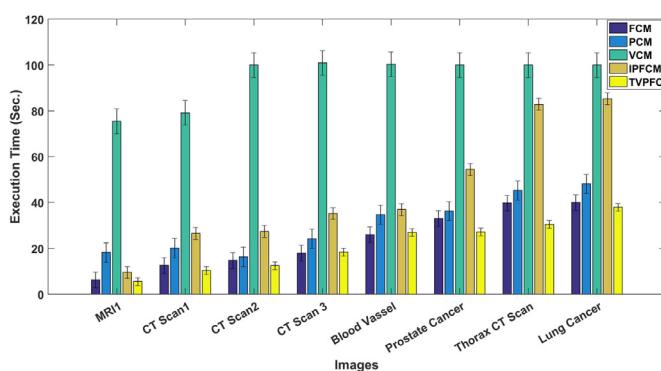


Fig. 8. Average CPU run time of some selective images with randomized initialization for each algorithm and 20 runs for each image. The error bars show the average relaxation of the measured values.

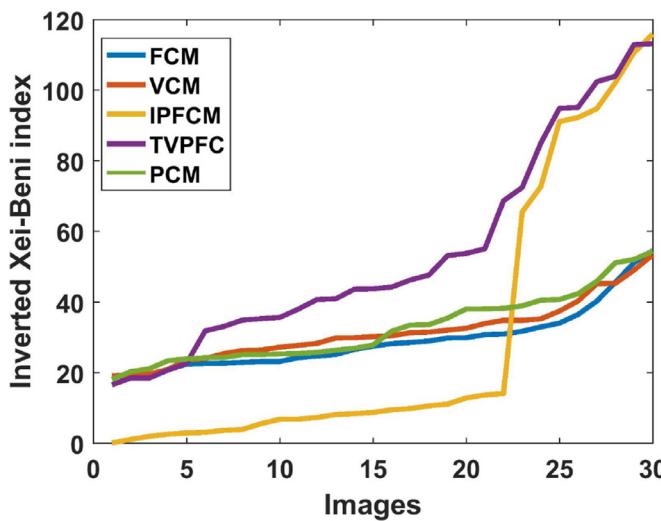


Fig. 9. Inverted Xei-Beni index for a set 30 CT scan images taken from TCGA-PRAD dataset.

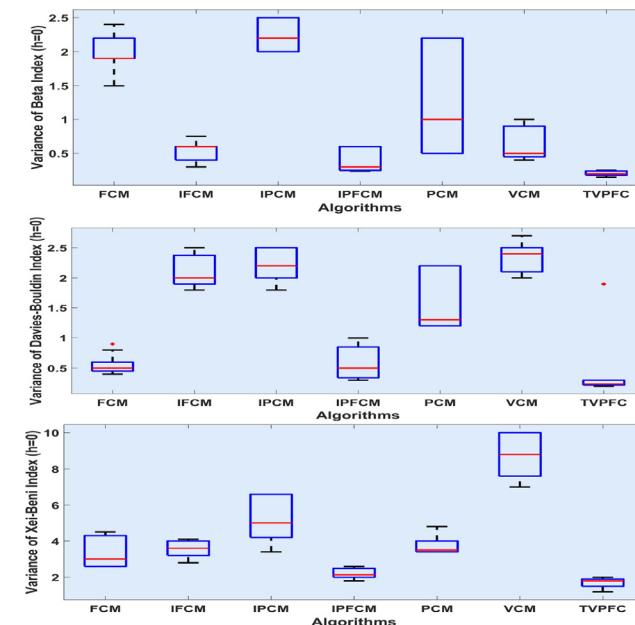


Fig. 10. Variance plot of Beta, Davies-Bouldin and Xei-Beni indices for 231 images with different noise levels taken from LIDC-IDRI and TCGA-READ dataset.

Table 9
Time complexity of different clustering algorithms.

Algorithms	Complexity
FCM	$O(nc^2i)$
PCM	$O(nki)$
VCM	$O((mnc^2+e+g) * i)$
IPFCM	$O((nc^2 + knc + a + b) * i)$
TVPFC	$O(nc^2i)$

Table 10
Measurement of CPU time and iteration of different clustering algorithms.

Algorithms	CPU time (s)	Iteration
FCM	28.21 ± 3.4	43 ± 1.7
PCM	37.42 ± 3.4	84 ± 2.5
VCM	107.33 ± 5.4	82 ± 2.4
IPFCM	49.92 ± 2.6	55.66 ± 2.2
TVPFC	19.66 ± 1.7	22.32 ± 2.1

Beta, Davies-Bouldin, and Xei-Beni indices measured on 231 images taken from LIDC-IDRI and TCGA-READ dataset for different noise levels ranging from 0% to 9%. For each image the percent noise represents the ratio of the standard deviation of the white Gaussian noise versus the signal for a reference tissue.

6.4. Complexity analysis

Interval-valued clustering poses high-level complexity when determining a crisp prototype from an interval-valued representation of data. The proposed TVPFC considers two data structures: one for the interval-valued representation of the data, and the other for updating the prototype. For interval-valued representation, TVPFC requires a possibilistic membership matrix and a fuzzy membership matrix in each iteration, which requires nc and nc^2 computations, respectively, where n is the number of objects, and c is the number of clusters. In the next stage, TVPFC updates the prototype matrix, which requires nc^2 computations. Parameter β requires nc computations. Thus, the entire TVPFC algorithm requires $O(nc) + O(nc^2) + O(nc^2) + O(nc) \simeq O(nc^2i)$ computations for i number of iterations. This computational cost is comparable with that of FCM. PCM shows linear time complexity of $O(nki)$ as it only considers possibilistic membership that requires distance measurement of an object from the assigned cluster prototype, while fuzzy membership requires distance measurement from the assigned cluster prototype as well as from the other cluster prototypes. However, possibilistic membership does not provide data mobility, thus, initialization with any fuzzy clustering is essential. Therefore, computational time of PCM is greater than FCM. Furthermore, TVPFC has a complexity of nc^2i , but a reduced search space, as specified by the vague region, and the fast convergence of the KM algorithm help TVPFC to perform better and take less time and iterations than FCM. The computational cost for IPFCM is approximately $O((nc^2 + knc + a + b) * i)$, where k is a constant, which is nearly 10 while a and b represent the time complexity required for the optimization of parameter A and B. VCM requires an $O((mnc^2+e+g) * i)$ computational cost, where m is the size of the swarm and e and g represent the computation time for parameter η and σ . Hence, the parametric complexity and computational cost of IPFCM and VCM are significantly higher than TVPFC. **Table 9** summarized the complexities of all the algorithms considered for experimental analysis.

7. Conclusion and future scope

In this study, we propose a novel clustering algorithm. This algorithm represents uncertainty in a unique way, with the infusion

of fuzzy membership along with typicality providing flexibility to deal with uncertainty. The use of only a few parameters makes TVPFC less complex than other stand-alone techniques. We also demonstrate the robustness of TVPFC via experimental analysis, which shows how TVPFC works efficiently for CT scan and MRI images with fast convergence and superior handling of noise and imprecisions. Proposing a generalized approach for medical image segmentation is always challenging because of variation in appearances, the use of different image acquisition systems, and other imaging modalities. In practice, we have found that the performance of TVPFC is very sensitive to the choice of parameter β , which is responsible for defining the shape of the cluster. A better estimation of β can further improve the quality of segmentation.

Declaration of Competing Interest

The authors, Ankita Bose and Kalyani Mali, certify that they have affiliations with the University of Kalyani, Kalyani, 741235, India and they have non-financial interest in the subject matter discussed in this manuscript.

Acknowledgement

We thank the anonymous referees for their valuable comments and suggestions, which help us to improve the paper and the authors would like to express their cordial thanks to the DST PURSE II programme for providing a grant to carry out this research activity in the University of Kalyani, Kalyani, West Bengal, India.

References

- [1] M. Khatami, T. Schmidt-Wilcke, P.C. Sundgren, A. Abbasloo, B. Schölkopf, T. Schultz, BundleMAP: Anatomically localized classification, regression, and hypothesis testing in diffusion MRI, *Pattern Recognit.* 63 (2017) 593–600.
- [2] H. Min, L. Xia, J. Han, X. Wang, Q. Pan, H. Fu, H. Wang, S.T. Wong, H. Li, A multi-scale level set method based on local features for segmentation of images with intensity inhomogeneity, *Pattern Recognit.* 91 (2019) 69–85.
- [3] Y. Yang, W. Jia, Y. Yang, Multi-atlas segmentation and correction model with level set formulation for 3D brain mr images, *Pattern Recognit.* 90 (2019) 450–463.
- [4] J. Shi, X. Zheng, J. Wu, B. Gong, Q. Zhang, S. Ying, Quaternion Grassmann average network for learning representation of histopathological image, *Pattern Recognit.* 89 (2019) 67–76.
- [5] I. Chaturvedi, R. Satapathy, S. Cavallari, E. Cambria, Fuzzy commonsense reasoning for multimodal sentiment analysis, *Pattern Recognit. Lett.* 125 (2019) 264–270.
- [6] J.C. Bezdek, Objective function clustering, in: *Pattern Recognition with Fuzzy Objective Function Algorithms*, Springer, 1981, pp. 43–93.
- [7] R. Krishnapuram, J.M. Keller, A possibilistic approach to clustering, *IEEE Trans. Fuzzy Syst.* 1 (2) (1993) 98–110.
- [8] M.J. Li, M.K. Ng, Y.-m. Cheung, J.Z. Huang, Agglomerative fuzzy k-means clustering algorithm with selection of number of clusters, *IEEE Trans. Knowl. Data Eng.* 20 (11) (2008) 1519–1534.
- [9] Z.-X. Ji, Q.-S. Sun, D.-S. Xia, A modified possibilistic fuzzy c-means clustering algorithm for bias field estimation and segmentation of brain MR image, *Comput. Med. Imaging Graph.* 35 (5) (2011) 383–397.
- [10] N.R. Pal, K. Pal, J.C. Bezdek, A mixed c-means clustering model, in: *Fuzzy Systems, 1997. Proceedings of the Sixth IEEE International Conference on*, vol. 1, IEEE, 1997, pp. 11–21.
- [11] N.R. Pal, K. Pal, J.M. Keller, J.C. Bezdek, A possibilistic fuzzy c-means clustering algorithm, *IEEE Trans. Fuzzy Syst.* 13 (4) (2005) 517–530.
- [12] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-I, *Inf. Sci.* 8 (3) (1975) 199–249.
- [13] M. Mizumoto, K. Tanaka, Some properties of fuzzy sets of type 2, *Inf. Control* 31 (4) (1976) 312–340.
- [14] H.R. Tizhoosh, Image thresholding using type II fuzzy sets, *Pattern Recognit.* 38 (12) (2005) 2363–2372.
- [15] D. Wu, J.M. Mendel, Enhanced Karnik–Mendel algorithms, *IEEE Trans. Fuzzy Syst.* 17 (4) (2009) 923–934.
- [16] C. Hwang, F.C.-H. Rhee, Uncertain fuzzy clustering: interval type-2 fuzzy approach to c-means, *IEEE Trans. Fuzzy Syst.* 15 (1) (2007) 107–120.
- [17] J.-H. Min, E.-A. Shim, F.C.-H. Rhee, An interval type-2 fuzzy PCM algorithm for pattern recognition, in: *Fuzzy Systems, 2009. FUZZ-IEEE 2009. IEEE International Conference on*, IEEE, 2009, pp. 480–483.
- [18] Z. Ji, Y. Xia, Q. Sun, G. Cao, Interval-valued possibilistic fuzzy c-means clustering algorithm, *Fuzzy Sets Syst.* 253 (2014) 138–156.
- [19] W.-L. Gau, D.J. Buehrer, Vague sets, *IEEE Trans. Syst. Man Cybern.* 23 (2) (1993) 610–614.
- [20] C. Xu, P. Zhang, B. Li, D. Wu, H. Fan, Vague c-means clustering algorithm, *Pattern Recognit. Lett.* 34 (5) (2013) 505–510.
- [21] F. Hoepfner, Fuzzy shell clustering algorithms in image processing: fuzzy c-rectangular and 2-rectangular shells, *IEEE Trans. Fuzzy Syst.* 5 (4) (1997) 599–613.
- [22] J. Sun, X. Wu, V. Palade, W. Fang, C.-H. Lai, W. Xu, Convergence analysis and improvements of quantum-behaved particle swarm optimization, *Inf. Sci.* 193 (2012) 81–103.
- [23] Y. Chen, Study on centroid type-reduction of interval type-2 fuzzy logic systems based on noniterative algorithms, *Complexity* 2019 (2019).
- [24] K. McGuinness, N.E. O’connor, A comparative evaluation of interactive segmentation algorithms, *Pattern Recognit.* 43 (2) (2010) 434–444.
- [25] L.R. Dice, Measures of the amount of ecologic association between species, *Ecology* 26 (3) (1945) 297–302.
- [26] A. Bose, K. Mali, Fuzzy-based artificial bee colony optimization for gray image segmentation, *Signal Image Video Process.* 10 (6) (2016) 1089–1096.

Ankita Bose is presently pursuing Ph.D. (Senior Research Fellow) in the University of Kalyani. She has passed M.Tech as a first-class first rank holder from the University of Kalyani in the year 2014. She has received B.Tech Degree from the West Bengal University of Technology in the year 2010. Her main areas of interest are image processing, soft computing, pattern recognition, and data mining

Kalyani Mali is a Professor at the Department of Computer Science & Engineering, University of Kalyani, West Bengal, India. She was the Head of CSE department 4 times and recently during the period 2014 to 2016. She has more than 25 years of research experience and more than 23 years of teaching experience. She is a life-time member of CSI and a member of IEEE. She has completed her Ph.D. on the topic *Modeling of Heterogeneous Data in Data Mining*. Her research interests include Pattern Recognition, Digital Image processing, data mining, soft computing, and multimedia.