Magnetic resonance volumetric techniques: a new segmentation method based on interval type-2 Fuzzy Logic and clinical applications

Diego S. Comas*[a,b], Gustavo J. Meschino[a,c], Sebastián Costantino[c,d],
Juan I. Pastore[a,b], Carlos Capiel[c,d], Virginia L. Ballarin[a,c]

[a]Instituto de Investigaciones Científicas y Tecnológicas en Electrónica, ICYTE, Facultad de Ingeniería, UNMDP-CONICET, Juan B. Justo 4302, Mar del Plata, Argentina
[b]CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), Argentina
[c]Universidad Fasta, Avellaneda 3345, Mar del Plata, Argentina
[d]Instituto Radiológico, Catamarca 1542, Mar del Plata, Argentina

ABSTRACT

The analysis of structural changes in the brain through magnetic resonance imaging (MRI) provides useful information for diagnosis and clinical treatment of patients with some pathologies, like Alzheimer disease and dementia. While complexity achieved by the MRI equipment is high, quantification of structures and tissues has not been entirely solved. This paper presents a method for segmentation of magnetic resonance images of the brain, based on a classification method using interval type-2 fuzzy logic called Type-2 Label-based Fuzzy Predicate Classification (T2-LFPC) which enables computing volumes occupied for the different tissues into the intracranial cavity. In the first stage, a random partition of observations is performed. Data contained in data subsets are analyzed, applying clustering to the observations corresponding to each class in order to discover groups of data with similar properties. Then, interval type-2 membership functions and fuzzy predicates are defined. In the final stage, optimization of parameters regarding the classification system is done. A comparison against various known classification methods was performed. A method of measuring the progressive atrophy and possible changes compared to a therapeutic effect should be essentially automatic and therefore independent of the radiologist. Results show that the performance of the proposed method is highly acceptable as a contribution for this goal. Advantages of this approach are presented throughout this paper.

Keywords: Brain volume, magnetic resonance, segmentation, fuzzy logic

1. INTRODUCTION

Magnetic resonance imaging (MRI) is currently the preferred imaging diagnosis method for the assessment of pathologies of the central nervous system (CNS).

The MRI allows obtaining images in multiples views, having high anatomic resolution and the ability of characterizing the tissues of the brain and their pathologies. These features make MRI clearly superior to other medical imaging types.

The determination of volumes constitutes a useful practice for several diagnoses of nosological entities. Knowing alterations of the standard measurements of the brain volume is very interesting for the differential diagnosis of different pathologies, for assessment of progression and treatment efficacy and monitoring for prognostic evaluation[1].

Volume decreasing can be subjectively assessed, but it is common to find inter-observer discrepancies[2]. Brain MRI allows detecting increase in volumes of the intracranial spaces, normally occupied by cerebrospinal fluid (CSF). This increasing is related to the deepening of sulcus of the brain convexity.

By means of image processing techniques, objective volume measurements can be done in MRI, once brain parenchyma could be discriminated from non-parenchyma content. In this work, a voxel classification is performed, trying to assign them to white matter (WM), grey matter (GM) and CSF. This classification allows easily computing volumes related to these tissues.

Some neurological diseases may require MRI-based brain volumetrics support, such as:

Alzheimer disease: volumetric alterations can be observed a few years before clinical synthons coming from typical cognitive alterations appear, like progressive memory loss. Hippocampal atrophy, cingulate and entorhinal cortex atrophy and the later cortical frontoparietal compromise allow recognizing all of the stages of the pathology so relating to them early and advanced neurological affectation[3].

Multiple sclerosis: inflammation and demyelinating associated to neurodegenerative processes are the precursors in the probable brain volume loss related to this pathology. Volume loss is bigger than expected in same-age individuals. This volume reduction globally supports some prediction of the disease progress, what makes volume measurement relevant. It begins prematurely and involves not only WM, but also cortical GM and basal ganglia. A high grade initial atrophy will lead to diminishing motor and cognitive capacities. No differences are seen between atrophy patterns in different sub-groups of multiple sclerosis[1].

Frontotemporal dementia: determination of focal atrophy in the anterior sector of frontal and temporal lobes constitutes a marker for the diagnosis of this disease. Subjective observation does not allow assess the low initial volume loss, but it can be properly characterized by volumetric measurements in RM. Two types of frontotemporal dementia develop different atrophy patterns that can be recognized via segmental volumetric assessment, even in early stages, allowing differential diagnosis of entities like Alzheimer disease[4].

There exist some methods to compute brain volumes. Recently an online available software was released, named volBrain[5], based on multi-atlas label fusion technology. This aims to compute volumes of the intra cranial cavity and those covered by the different tissues. It presents a pipeline including denoising and inhomogeneity correction and the segmentation is based on 50 MRI volumes manually segmented. A discussion of previous approaches for the volume computation can be found in the same work.

Given the previous motivation of brain MRI-based diagnosis, the objective of this work is to present preliminary results of volume estimation of the brain tissues (WM, GM, CSF), using a new automated image segmentation based on interval type-2 fuzzy logic (FL) called Type-2 Label-based Fuzzy Predicate Classification (T2-LFPC). From some examples of the different brain tissues (pixels labeled by medical experts), the T2-LFPC method automatically generates a FL predicate system which is used to perform the whole MRI sequence segmentation. In addition, considerations of noise and inhomogeneity field are covered in the FL properties, since it has been proved its robustness to these undesirable effects, considered as data uncertainty[6]. Gray levels of pixels in each slice represent voxel intensities. Once all voxels are classified by means of the T2-LFPC method, volume of different tissues are computed.

In the next Section, the approach of the proposed method is explained and the image acquisition process is described. Then, in Section 3, results are presented and discussed.

2. MATERIAL AND METHODS

2.1 Fundamentals of interval type-2 FL and classification based on predicates

FL was introduced by Lofti Zadeh in 1965[7,8], extending Boolean (classic) logic in order to deal with linguistic expressions and work with concepts containing vague or imprecise expressions. FL considers truth values between 0 (false) and 1 (true) instead using only 0 and 1 as in classic logic, which is known as "principle of gradualism". Specifically, FL provides an effective conceptual framework for dealing with the problem of knowledge representation in environments of uncertainty and imprecision like human reasoning and pattern recognition[9]. FL has been widely applied in pattern recognition and, specifically, in imagen segmentation[6,10].

The main limitation of the of the traditional FL, called type-1 FL, is that truth values are limited to single values in the [0,1] interval. Due to this, type-1 FL may not be an appropriate model for solving problems with great imprecision or when data is affected by noise[11,9,12].

In contrast to type-1 FL, interval type-2 FL, or interval-valued FL which is the same, provides additional grades of freedom to define truth values using intervals, called intervals of truth values[9]. Defining truth values through intervals allowing to deal with the effects of noise and to considerate opinions of different experts when expert knowledge about how objects should be classified is needed[9]. In previous works, interval type-2 FL was used for pattern recognition showing better performance that equivalent models based on type-1 FL[6,11,13,14].

As it was previously mentioned, in the present work interval type-2 FL is used in order to model truth values of fuzzy predicates, which are generated by means of the T2-LFPC method, enabling image segmentation and quantifying the

proportion of voxels corresponding to WM, GM, and CSF. In the remainder of this section the main necessary concepts related to interval type-2 FL will be introduced in order to provide the notation and concepts used in the T2-LFPC method, which is presented in the next section.

Definition 1: An interval of truth values is an interval $A = [a_L, a_R]$, with $0 \leq a_L \leq a_R \leq 1$[9]. An interval of truth values defines the truth value of a logic expression when interval type-2 FL is used.

Definition 2: An interval type-2 membership function $\bar{\mu}_A$ on a discourse universe U is a function $\bar{\mu}_A : U \rightarrow \chi$, where χ is the set of all the closed intervals contained in $[0,1]$, i.e. the set of all the possible intervals of truth values, and $\mathbf{A}$ is a property (an attribute)[9]. For a specific value $u \in U$, $\bar{\mu}_A(u)$ is an interval of truth values $A_{\bar{\mu}_A(u)} = \left[a_{\bar{\mu}_A(u),L}, a_{\bar{\mu}_A(u),R} \right]$ which defines the truth value in which u satisfies the property $\mathbf{A}$ described by the membership function.

Another notation used for $\bar{\mu}_A$ is the following:

$$\bar{\mu}_A(u) = \left[\varphi_{\bar{\mu}_A}^-(u), \varphi_{\bar{\mu}_A}^+(u) \right], \tag{1}$$

where $\varphi_{\bar{\mu}_A}^- : U \rightarrow [0,1]$ and $\varphi_{\bar{\mu}_A}^+ : U \rightarrow [0,1]$ are strictly type-1 membership functions called respectively the lower and the upper membership functions of $\bar{\mu}_A$.

A predicate is defined as the part of a sentence that tells something about the subject. In FL this concept refers to a property that a variable or a finite collection of variables can have and it is typically assumed as the equivalent of a membership function. On the other hand, a predicate becomes a proposition when specific values are assigned to its variables. Due to this relation between the terms predicate and preposition, in this paper values taken by fuzzy predicates will be named truth values of the predicates, considering that evaluating a fuzzy predicate in specific values of their arguments is equivalent to analyse the truth value of a preposition and both terms will be used as synonymous unless it appears necessary to distinguish between them.

According to these considerations, in this work a fuzzy predicate $p(x)$ is adopted to be the following:

Definition 3: A fuzzy predicate $p(x)$, where x indicates an object or a variable, is a declarative sentence which assigns one or more properties to the object x. Considering interval type-2 FL, values taken by $p(x)$ for different value of x, which are noted as $v(p(x))$, are interval of truth values as the form $A_{p(x)} = \left[a_{p(x),L}, a_{p(x),R} \right]$, with $0 \leq a_{p(x),L} \leq a_{p(x),R} \leq 1$.

As in general fuzzy predicates can relate one or more objects or variables with properties, fuzzy predicates can be both simple or compound. Simple predicates directly associate a variable with an attribute and their truth values are usually obtained through membership functions. On the other hand, compound fuzzy predicates combine logically two or more simple predicates using conjunctions ($\wedge$), disjunctions ($\vee$) and complements ($\neg$), which in a wide sense are known as fuzzy aggregation operators. In interval type-2 FL, values of truth of compound predicates are computed using fuzzy conjunctions $C : [0,1]^n \rightarrow [0,1]$, disjunctions $D : [0,1]^n \rightarrow [0,1]$, and complements $c : [0,1] \rightarrow [0,1]$, applied on the ends of the interval of truth values[9]. In the literature, a wide set of different fuzzy conjunctions, disjunctions and complement had been proposed. Selecting different fuzzy operators should be made according to the properties of each operator and how predicates are interpreted and evaluated by the experts in each particular application. In the case studied in this paper, considering previous results obtained in pattern recognition, both MIN-MAX triangular norms and compensatory FL operators are considered[15,16].

In pattern recognition applications using fuzzy predicates, one or more fuzzy compound predicates combine logically the properties that a datum must meet to belong to a label. By computing the truth value of each predicate explaining the different labels for a datum to be labeled, the truth value in which this datum belong to each of the labels is obtained[9]. Label assignment is performed by determining the label to which the datum belongs with highest truth value. This process requires comparing truth values.

In the case of interval type-2 FL, the comparison between truth value requires comparing intervals of truth values and defining which one has the highest truth value. Despite of interval arithmetic and interval comparison are active research

fields since the end of the 50's, the problem of comparing two intervals deciding a ranking between these in all the possible cases, i.e. in the cases of non-overlapping, overlapping or an interval included in the other, remains unresolved. In the literature, some solutions had been proposed. The proposals of Moore[17], Ishibuchi and Tanaka[18], Kundu[19] and Sengupta and Pal[20] can be cited among others, but these methods cannot rank intervals with the same mean value.

Based on successfully previous results obtained in pattern recognition applications combining fuzzy predicates and interval type-2 FL, in this paper it is used our previous proposal called measure of interval of truth values[9], also previously applied[13,14], which consists in combining the mean value and the maximum value of an interval in order to obtain a number that globally characterizes the truth value of the interval. This measure is defined as follows:

Definition 4: (Measure of interval of truth values). Let χ be the set of all the closed intervals contained in $[0,1]$, i.e. the set of all the possible intervals of truth values, we call *measure of interval of truth values* to the function $f : \chi \to \mathbb{R}^+$ defined as:

$$f(A) = f([a_L, a_R]) = \frac{a_L + a_R}{2} a_R, \tag{2}$$

where $A = [a_L, a_R]$ is an interval of truth values. The function f describes with a number the truth value of the interval of truth values, mapping from the interval space to $\mathbb{R}^+$. The higher the value of $f(A)$, the higher the degree of truth. The reasons of combining the mean value and the maximum of the interval to its characterization are: a) if two intervals have the same mean value, then those with higher maximum value (closer to 1) represents a higher truth value, and b) in the case of two intervals with the same maximum, those with lower mean value represents a lower truth value.

Considering all the previous definition, classifiers based on fuzzy predicates are formed by K fuzzy compound predicates, noted by $\{p_i(x)\}_{i=1,\dots,K}$, each one associated with one of the possible classes $\{1,\dots,K\}$. Given a datum $x \in \mathbb{R}^d$ to classify, where d is the dimension of the space of patterns, the truth values of K predicates for x are computed, defining the set $\{v(p_i(x))\}_{i=1,\dots,K}$, where $v(p_i(x)) = [a_{p_i(x),L}, a_{p_i(x),R}]$ (an interval of truth values). The measure of interval of truth values if applied on each $v(p_i(x))$, $i = 1,\dots,K$, resulting in $\{f(v(p_i(x)))\}_{i=1,\dots,K}$. The label assigned to the datum x is the one whose corresponding fuzzy predicate has the highest measure of interval of truth values, that is, the class k is assigned to datum x if:

$$\lambda(v(p_k(x))) = \max\{f(v(p_i(x)))\}_{i=1,\dots K}. \tag{3}$$

Based on the concepts defined in this Section, in the next one the T2-LFPC is explained in detail.

2.2 Type-2 Fuzzy Predicate Classification method

As it was previously introduced, in the problem addressed in the present paper, first goal is the voxel classification. Here, a voxel acts as a data vector composed by a unique variable (its gray level intensity in the T1 sequence) and the different tissues and the background act as labels. Voxel intensity are represented by the gray level of each pixel in each slice. Voxel classification, i.e. the segmentation of each slice in the entire volume of the MRI is performed using the T2-LFPC method explained in the present Section.

The present explanation is presented considering a general classification problem, i.e. without non restriction to image segmentation. Data to be labeled are represented by $\mathbf{X} \subset [-1,1]^d$, where N is the number of data and d the number of features, and a datum $\mathbf{x} \in \mathbf{X}$ is a d-uple $(x_1, x_2, \dots, x_d)$. In addition, $\mathbf{Y} = \{y_r\}_{r=1,\dots,N}$, $y_r \in \{1,\dots,K\}$ is the vector of labels (gold-standard) for the data in $\mathbf{X}$, i.e. $\mathbf{X}$ and $\mathbf{Y}$ define the training dataset for the classification problem.

The Type-2 Label-based Fuzzy Predicate Classification (T2-LFPC) method was originally proposed and used in by Comas *et al*[13] in order to design nonlinear morphological operators (window morphological operators) for binary images. In the present paper, we propose an improvement of our original method which enables to the general data classification and is applied to MRI segmentation. Considering that the present improvement became the previous method in a general

and complete classification method, including the original application of the morphological operator design, we consider to keep the original name of T2-LFPC as necessary. In the next, the new general T2-LFPC method is explained in detail.

The method consists of four stages: *A) Random data partition*, *B) Class prototype determination*, *C) Fuzzy system generation* and *D) Optimization*. From the training dataset, a classification system based on fuzzy predicates is automatically generated. In the first stage *A*, a random partition of the data in $\mathbf{X}$ is performed. Data contained in each subset are analyzed in stage *B*, applying clustering to the data corresponding to each label in order to discover groups of data with similar properties for each subset and label. For each label, the obtained groups are subsequently combined with the other subsets to determine interval type-2 membership functions and fuzzy predicates in the stage *C*. Finally, in the stage *D* an optimization of the parameters of the fuzzy system is done. Figure 1 shows the processing pipeline for the T2-LFPC method. The steps for each stage are described in detail in the remainder of this Section.
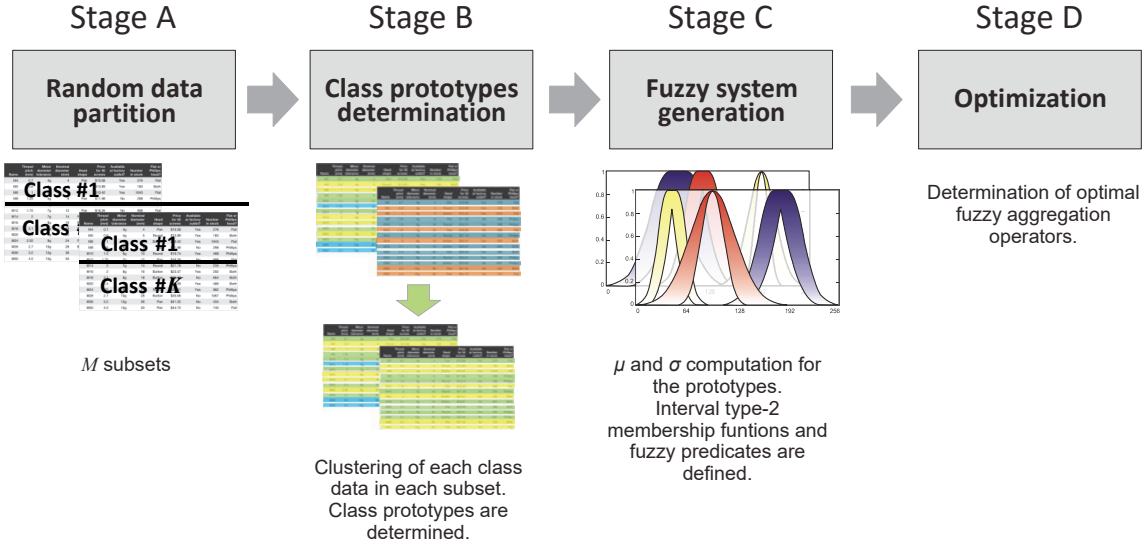

Figure 1. Processing pipeline for the classification method T2-LFPC proposed for the segmentation of the MRI volume.

Stage A: Random data partition

From the training dataset defined by $\mathbf{X}$ and $\mathbf{Y}$, first, a random partition is performed on $\mathbf{X}$, obtaining $M \geq 2$ disjoint data subsets noted as $\left\{ \mathrm{P}_m \right\}_{m=1,...,M}$, where $\mathbf{X} = \mathrm{P}_1 \cup \mathrm{P}_2 ... \cup \mathrm{P}_M$. Such random partition is made simultaneously on the label vector $\mathbf{Y}$, given as a result that each P_m can be written as $\mathrm{P}_m = \Phi_{m,1} \cup \Phi_{m,2} \cup ... \cup \Phi_{m,K}$, where $\Phi_{m,k}$ contains the data in P_m that belong to the label k, $k \in \left\{ 1,..,K \right\}$. The partition is done in such a way that each label has approximately the same quantity of data in the different subsets.

Stage B: Class prototype determination

In order to build fuzzy predicates and membership functions in the next stage, it is required an analysis of the properties that presents the data contained in $\Phi_{m,k}$. Observations corresponding to the same class into a partition do not necessarily share the same properties, i.e. they are not necessarily close in the pattern space.

With the aim of discover groups with "similar" properties within $\Phi_{m,k}$ a clustering algorithm is used, discovering groups of data inside of each $\Phi_{m,k}$, i.e., dividing the data of each subset m in each label k. Due to the appropriate number of groups for each subset is not known *a priori*, a clustering approach combining fuzzy c-means clustering (FCM)[21] and the bayesian information criterion (BIC) is used[22]. This clustering approach is used for automatic determination of the number of clusters in a dataset, following the method developed by Zhao *et al.*[23] and tested by Comas *et al.*[13]. As a result, each subset $\Phi_{m,k}$ in a label k, is divided in $\rho_{m,k}$ clusters.

Formally, the clustering obtained on $\Phi_{m,k}$ can be written as:

$$\Phi_{m,k} = \Delta_{m,k,1} \cup ... \Delta_{m,k,j} ... \cup \Delta_{m,k,\kappa_{m,k}} ,\qquad\qquad(4)$$

where $\Delta_{m,k,j}$ data in the cluster j obtained for the data of the label k, $k = 1,...,K$ in the subset m, $m = 1,...,M$ during the stage A. The set of centroid obtained in the current stage is noted by $\left\{ \Upsilon_{m,k,j} \right\}_{\substack{m=1,...,M \\ k=1,...,K \\ j=1,...,\rho_{m,k}}}$.

It is important to note that $\Delta_{m,k,j}$ represents near data in a label $k \in \{1,..,K\}$ in the subset m, which were automatically grouped by means of the FCM-BIC clustering approach. Specifically, given two subsets Δ_{m,k,j_1} and Δ_{m,k,j_2} with $j_1 \neq j_2$, these subsets contain data of a subset P_m and a label k which are not close in the pattern space.

Finally, the data in $\left\{ \Delta_{m,k,j} \right\}_{\substack{m=1,...,M \\ k=1,...,K \\ j=1,...,\rho_{m,k}}}$ are re-grouped in order to determine the class prototypes. The centroids $\left\{ \Upsilon_{m,k,j} \right\}_{\substack{m=1,...,M \\ k=1,...,K \\ j=1,...,\rho_{m,k}}}$ are used as data and clustered according with their proximity. As a result, the data in $\left\{ \Delta_{m,k,j} \right\}_{\substack{m=1,...,M \\ k=1,...,K \\ j=1,...,\rho_{m,k}}}$ are clustered again, but now considering the group assigned to their respective centroids. That is, considering class by class, the subsets $\left\{ \Delta_{m,k,j} \right\}_{\substack{m=1,...,M \\ k=1,...,K \\ j=1,...,\rho_{m,k}}}$ are re-clustered according to their $\Upsilon_{m,k,j}$, with m varying in $1,...,M$ and j varying in $1,...,\rho_{m,k}$ for a class k, generating κ_k clusters of centroids.

Following the previous procedure, the class prototypes can be determined. The set of class prototypes is noted by $\left\{ \Gamma_{k,s,n} \right\}_{\substack{k=1,...,K \\ s=1,...,\kappa_k \\ n=1,...,\zeta_{s,k}}}$, where k is one of the labels $\{1,..,K\}$, s is an index which represents some of the clusters previously obtained for the centroids and n is another index, which corresponds to one of the data subset $\Delta_{m,k,j}$, $m = 1,...,M$, $j = 1,...,\rho_{m,k}$, whose centroid was re-assigned to the centroid group s in the previous step, with $s \in \{1,...,\kappa_k\}$.

Using the steps described here, the prototypes in $\Gamma_{k,s,n}$ in a fixed k are data of the training dataset $\mathbf{X}$ that belong to the class k and all the data of the class k are in one of the sets $\left\{ \Gamma_{k,s,n} \right\}_{\substack{s=1,...,\kappa_k \\ n=1,...,\zeta_{s,k}}}$. Likewise, $\Gamma_{k,s,n}$ for specific k and s, contains $\zeta_{s,k}$ clusters of data, each near to each other, where each data cluster contains examples of the class k which share same properties, i.e. they are close in the pattern space.

Stage C: Fuzzy system generation

As a result of the procedures described of the stages A and B, the data of each class were clustered in data subsets which have similar properties, resulting in class prototypes. In the present stage, the prototypes in $\left\{ \Gamma_{k,s,n} \right\}_{\substack{k=1,...,K \\ s=1,...,\kappa_k \\ n=1,...,\zeta_{s,k}}}$ are analyzed, defining interval type-2 membership functions and fuzzy predicates generating, in consequence, a FL system able to classify data.

First, for each set of class prototype $\Gamma_{k,s,n}$, with $n = 1,...,\zeta_{s,k}$ and fixed k and s, a Gaussian type-1 membership function is found. Centroids of the membership functions are obtained as the centroid of the data in each $\Gamma_{k,s,n}$, formally noted by $\Omega_{k,s,n}$. As a result, for a label k in a prototype group s, $\zeta_{s,k}$ type-1 membership functions are defined for each feature $i \in \{1,...,d\}$. Widths of the membership functions are computed as the standard deviation of the data in $\Gamma_{k,s,n}$ for each class k in each group $s = 1,...,\kappa_k$. The standard deviation acts in this case as parameter, controlling how the truth

value of the membership functions decreases when the values move away from the centroid found for the class prototypes. The type-1 membership functions are noted by $\left\{\mu_{i,k,s,n}\right\}_{\substack{i=1,\ldots,d\\k=1,\ldots,K\\s=1,\ldots,\kappa_k\\n=1,\ldots,\zeta_{s,k}}}$.

As it was previously mentioned, the class prototypes $\Gamma_{k,s,n}$ share the same properties for fixed values of k and s. After type-1 membership functions are defined, these are aggregated varying the parameter n, as follow:

$$\varphi^-_{\hat{\mu}_{i,k,s}}(x) = \min\left\{\mu_{i,k,s,n}(x)\right\}_{n=1,\ldots,\zeta_{s,k}}, \ \forall x \in [-1,1], \tag{5}$$

$$\varphi^+_{\hat{\mu}_{i,k,s}}(x) = \max\left\{\mu_{i,k,s,n}(x)\right\}_{n=1,\ldots,\zeta_{s,k}}, \ \forall x \in [-1,1], \tag{6}$$

$\forall x \in [-1,1]$, $i \in \{1,\ldots,n\}$, $k \in \{1,\ldots,K\}$, $s = 1,\ldots,\kappa_k$, where $\varphi^-_{\hat{\mu}_{i,k,s}} : [-1,1] \to [0,1]$ and $\varphi^+_{\hat{\mu}_{i,k,s}} : [-1,1] \to [0,1]$ are respectively the lower and the upper membership functions of the interval type-2 membership function $\hat{\mu}_{i,k,s} : [-1,1] \to \xi([0,1])$, which defines with what truth value the feature i of a datum belong to the class k according to the set of prototypes s. As it is easy to note, $\hat{\mu}_{i,k,s}(x)$ is the interval of truth value $\left[\varphi^-_{\hat{\mu}_{i,k,s}}(x), \varphi^+_{\hat{\mu}_{i,k,s}}(x)\right]$.

The obtained interval type-2 membership functions $\hat{\mu}_{i,k,s}$, $i = 1,\ldots,d$, $k = 1,\ldots,K$, $s = 1,\ldots,\kappa_k$ are then smoothed, using a numerical interpolation, trying to eliminate possible peaks generated by the MIN-MAX aggregation. This procedure defines new interval type-2 membership functions $\left\{\bar{\mu}_{i,k,s}\right\}_{\substack{i=1,\ldots,d\\k=1,\ldots,K\\s=1,\ldots,\kappa_k}}$.

Finally, fuzzy predicates for each class $k \in \{1,\ldots,K\}$ are defined by logically operating with the truth values defined by the membership functions $\left\{\bar{\mu}_{i,k,s}\right\}_{\substack{i=1,\ldots,d\\k=1,\ldots,K\\s=1,\ldots,\kappa_k}}$ according to the index i, which represents the feature. Considering the class k, κ_k compound fuzzy predicates are defined, each one associated with particular properties of the data belonging to k in the training dataset $\mathbf{X}$. These predicates are:

$$p_{k,s}(\mathbf{x}) \equiv \bar{\mu}_{1,k,s}(x_1) \wedge \bar{\mu}_{2,k,s}(x_2) \wedge, \ldots, \wedge \bar{\mu}_{d,k,s}(x_d),$$
$$k = 1, 2, \ldots, K, \tag{7}$$
$$s = 1, \ldots, \kappa_k,$$

where $\mathbf{x}$ is a datum and the value of $p_{k,s}(\mathbf{x})$ defines the truth value in which $\mathbf{x}$ belong to k. The predicates $\left\{p_{k,s}\right\}_{s=1,\ldots,\kappa_k}$ are combined in a fuzzy predicate $p_k(\mathbf{x})$ using the disjunction operator ($\vee$), obtaining a unique fuzzy predicates for each class as following:

$$p_k(\mathbf{x}) \equiv p_{k,1}(\mathbf{x}) \vee p_{k,2}(\mathbf{x}) \vee, \ldots, \vee p_{k,\kappa_k}(\mathbf{x}),$$
$$k = 1, 2, \ldots, K \tag{8}$$

As a result of the current stage, interval type-2 membership functions and fuzzy predicates are generated, respectively noted by $\left\{\bar{\mu}_{i,k,s}\right\}_{\substack{i=1,\ldots,d\\k=1,\ldots,K\\s=1,\ldots,\kappa_k}}$ and $\left\{p_k\right\}_{k=1,\ldots,K}$, which are obtained from the training dataset $\{\mathbf{X},\mathbf{Y}\}$, defining a FL system. This system can be used both to classify the data in $\mathbf{X}$ and others data resulting of the same process that generated the set $\mathbf{X}$ (generalization).

Fuzzy predicates are evaluated using fuzzy aggregations operators and class are assigned following the procedure detailed at the end of the Section 2.1.

Stage D: Optimization

Using the training dataset $\{\mathbf{X}, \mathbf{Y}\}$, it is possible to compute the training error, or resubstitution error[24], once the FL system was generated. Thus, by quantifying the training error it is possible to set properly intern parameters of the fuzzy system such as the aggregation operators used to evaluate the conjunction and disjunction in the predicates.

As final step of the T2-LFPC method, the FL system is used to classify the training data, testing the aggregation operators MIN-MAX and the compensatory ones, both based in the geometric mean[15] and the arithmetic mean[16]. The classification error is computed in each case computing the percentage of incorrected classified data and considering as gold-standard the label in $\mathbf{Y}$.

As a final result, a FL system is obtained, enabling to classify data of the same kind that the data in the training dataset $\{\mathbf{X}, \mathbf{Y}\}$. The system is formed by interval type-2 membership functions and fuzzy predicates, explaining the classes to be obtained and defining the optimal parameters for the predicate evaluation. Only the size of initial partition M must be defined, which must be determined heuristically in accordance with the quantity of available data.

2.3 Proposed method for volume estimation

Volume estimation requires the next steps: intracranial cavity segmentation, tissue classification and volume computation. The performance of each stage influences dramatically the following ones.

There are plenty of methods for the intracranial cavity segmentation. Many of them have been compared in previous works[25]. In this paper, a method based on the application of morphology alternating sequential filters by reconstruction with structuring elements of growing size is used. Apart from enhancing and filtering in this way, this method captures the interior of a closed simple curve employing geodesic distance[26]. This curve represents the external brain boundary.

Once intracranial cavity area is determined in each slice, a trained operator selects small regions in random slices to construct a training dataset. The operator selects regions for the different classes: GM, WM, CSF. Background, i.e., voxels that does not belong to any of the considered tissues, is not selected in this stage due to this is previously removed by the masks generated to segment the intracranial cavity.

As a result, a training dataset is defined $\{\mathbf{X}, \mathbf{Y}\}$, formed by the intensities of the voxels defined by the MRI in T1 sequence and the corresponding labels defined by the operator. Then, the method described in Section 2.2 is applied, in order to assign voxels to one and only one class tissue. Therefore, the method is parametrized to consider 3 classes.

Once FL system is generated from the examples defined in the training dataset $\{\mathbf{X}, \mathbf{Y}\}$, the system is used to classify all the voxels in the entire volume, obtaining the classification of all the voxels between GM, WM and CSF.

In order to compute the volume of each issue, first, the volume of a single voxel was determined, considering the distance inter-slices (slice thickness) and the pixel spacing (in both directions). The product between these three values is the unitary voxel volume. Finally, the volume of each tissue is obtained as the number of voxels belonging to this tissue multiplied by the unitary voxel volume.

2.4 Image databases

The proposed method was initially tested with images coming from Montréal Neurological Institute Simulated Brain Database, McGill University[27]. Only simulated T1-weighed images were considered, in order to make possible extrapolation of the results to real images. Since these simulations have a gold-standard, the performance of the classification method could be assessed by a quality measure. Dimensions for the volume were 181x217x181 voxels, noise levels: 0%, and levels of intensity non-uniformity (INU): 0%.

The second image database came from a diagnostic imaging center, considering real cases, patients with known pathologies and MRI in T1 sequence. In this case, gold-standards are not available and the assessment of the classification results of the slices, but volume results were made for specialists as well as a comparison against volume results obtained by an on-line atlas-based method. Details of classification performance assessment is detailed in the next Section. These real brain images were acquired in a Philips equipment, 1.5 Tesla. Dimensions for the volume were 320x320x160 voxels, resolution 12 bits, in sagittal view. Pixel spacing was 0.8 x 0.8 millimeters and slice thickness was 1 millimeter, defining a unitary voxel volume of 0.64 cubic millimeters (0.64×10^{-3} cubic centimeters). Volumes were processed considering slices in axial view.

2.5 Quality measurements

In order to assess the segmentation results, accuracy was chosen as quality measurement, defined as the percentage of voxels correctly classified to the number of voxels considered. Accuracy was computed taking into account the generalization abilities of the classifiers: once trained the classifiers, accuracy was computed with the output of the classifiers for new data not contained in the training dataset.

In the simulated MRI volume dataset, first, 100 voxels per tissue were randomly selected of the entire volume, defining the training dataset. Then, the classifier was applied to segment the first slice of the volume (considered as test dataset), computing the accuracy for that slice. This procedure was repeated for each slice in the volume, defining in each case a new training dataset. The average of the obtained accuracies is reported, including the standard deviation.

As it was pointed before, for the real MRI volumes gold-standards were not available for the whole slices, but quality measures were computed for the voxels in the training set. From the total of voxels labeled by the operator, 200 voxels per tissue were randomly selected, defining the training dataset with 600 data. A 10-fold was considered in order to estimate the generalization abilities of the classifiers[28]. Then, the entire volume was processed, using all the training dataset.

3. RESULTS

In this section, results for slice segmentation and volume computation are presented, considering both simulated and real MRI volumes.
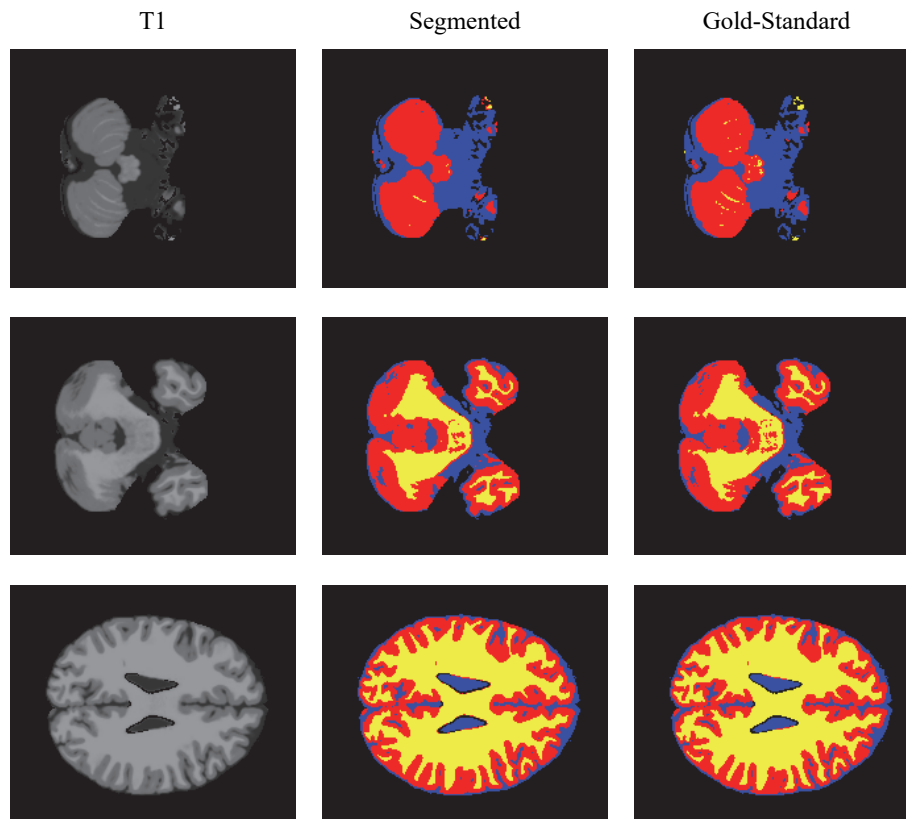

Figure 2. Segmentation obtained for three slices in simulated MRI volumes. The first column shows T1 images after intracranial cavity extraction, second column shows the results obtained by using the T2-LFPC method and the last column shows the gold-standards.

In Figure 2, results obtained using T2-LFPC for three slices (corresponding to #22, #35 and #100) for simulated images are shown. In this case, gold-standards were available and these are included in the third column also for visual comparison. A comparison against other known methods was performed, considering the K-nearest neighbor algorithm (KNN), multilayer feedforward neural networks (MLP), probabilistic neural networks (PNN) and a similar method to the T2-LFPC, but using type-1 FL, called Type-1 Label-based Fuzzy Predicate Classification (T1-LFPC). Unlike the T2-LFPC, type-1 membership functions are not aggregated. Instead, these are using directly generating a type-1 FL system. All of these methods were heuristically parametrized, testing different setting. Reporting results are the best results obtained for each classifier considering the average accuracy and different setting parameters.

Numerical results for segmentation accuracy in simulated MRI volumes are shown in Table 1. Standard deviations were computed along the slices. MLP gave the best in terms of mean accuracy and standard deviation. However, T2-LFPC presents a good trade between accuracy and computation times and differences were not statistically significant after computing P by the non-parametrical Wilcoxon test. Besides, previous deeper studies allowed concluding that this method showed the best performance considering noise and distortion in images. In addition, methods based on fuzzy predicates are able to give a deeper understanding about the classification obtained, which could be helpful for image interpretation. However, these discussions are far beyond the scope of this paper.

Table 1. Accuracy obtained in a simulated MRI volume for the proposed method (T2-LFPC) against other known methods: K-nearest neighbor algorithm (KNN), a multilayer feedforward neural network (MLP), a probabilistic neural network (PNN) and a similar method using type-1 FL (T1-LFPC).

	T2-LFPC	KNN	MLP	PNN	T1-LFPC
Mean	0.975	0.976	0.981	0.977	0.972
Standard deviation	0.009	0.009	0.007	0.008	0.010
Computation time [s]	24	56	88	43	5

In Figure 3, three slices for real MRI images are presented (corresponding to #44, #100 and #200). In the third column of this figure, segmentation was superimposed to the original images. Again, performance of the proposed method T2-LFPC was compared to KNN, MLP, PNN and T1-LFPC, considering different setting parameters in each case. Results presented correspond to the best performance for each method, determined according to the accuracies estimated from 10-fold validation method applied on the voxels in the training set.

The estimated accuracies and its standard deviation along the 10 folds previously described are shown in Table 2. Again, differences among the different methods were statistically not significant, showing KNN the best performance.

For the MRI volume presented in Table 2 and Figure 3, volume of the intracranial cavity and volumes and percentages covered by the different tissues were computed for the considered methods and a comparison with the online software Volbrain[5] could be made. These results are presented in Table 3.

The determination of the intracranial cavity is crucial to estimate the percentages of each tissue. In the methods considered in this paper, this volume is constant because the same method for intracranial cavity segmentation was applied in all cases.

An objective comparison of these results is a difficult task, because there is no gold-standard. But the observed consistency between results and the subjective expert assessment of the segmented slices are good evidence about the reliability of the results.

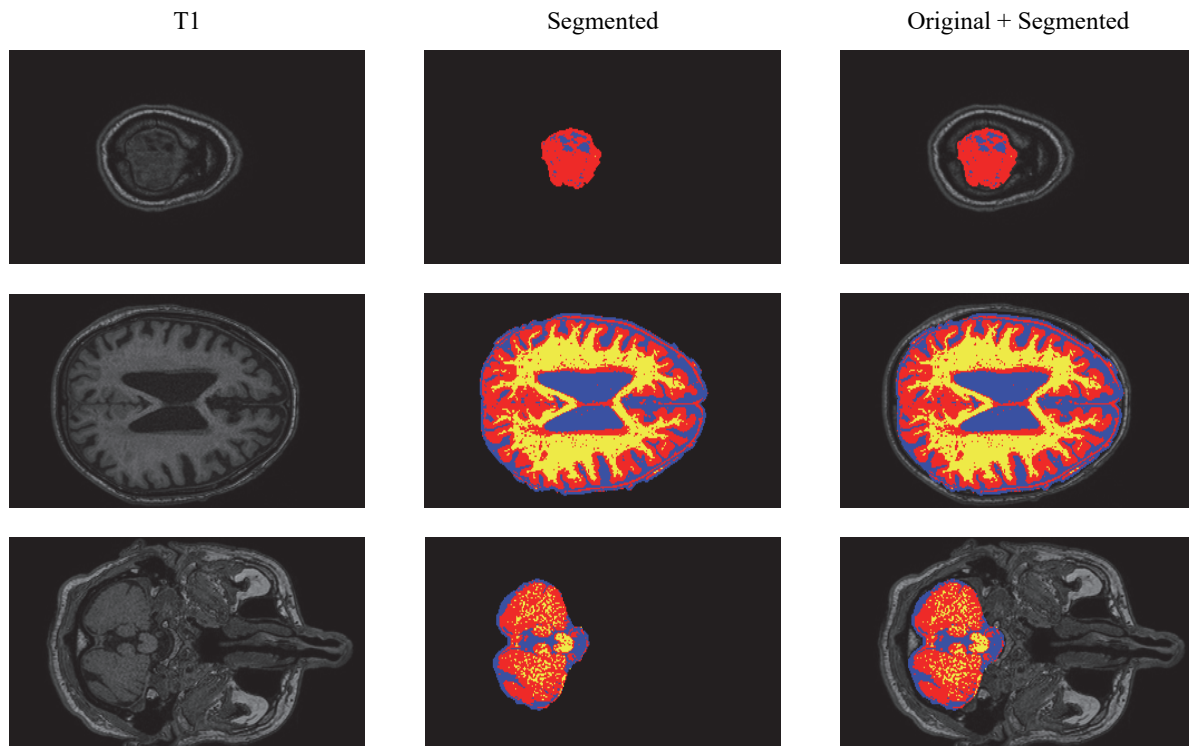|	T1	Segmented	Original + Segmented

Figure 3. Segmentation obtained for three slices in a real sample MRI volume. The first column shows T1 images before intracranial cavity extraction (as they come from the MRI equipment), second column shows the results obtained by using T2- LFPC classification and third column shows results superimposed to original images.

Table 2. Accuracy obtained in a real MRI volume for the proposed method (T2-LFPC) against other known methods: K-nearest neighbor algorithm (KNN), a multilayer feedforward neural network (MLP), a probabilistic neural network (PNN) and a similar method using type-1 FL (T1-LFPC).

	T2-LFPC	KNN	MLP	PNN	T1-LFPC
Mean	**0.943**	0.950	0.948	0.942	0.938
Standard deviation	**0.027**	0.021	0.029	0.029	0.026
Computation time [s]	**2.749**	0.06	10.933	0.78	0.316

Table 3. Comparison of results obtained by the proposed method against those obtained by other classification methods and the online software Volbrain. Volume are presented in cm^3 and percentage.

	T2-LFPC	KNN	MLP	PNN	T1-LFPC	Volbrain
White Matter (WM)	**400.22 (22.93%)**	376.39 (21.57%)	397.29 (22.77%)	413.75 (23.71%)	371.22 (21.27%)	408.42 (23.63%)
Grey Matter (GM)	**760.48 (43.58%)**	786.50 (45.07%)	734.95 (42.11%)	727.52 (41.69%)	816.00 (46.76%)	759.28 (43.92%)
Cerebro Spinal Fluid (CSF)	**584.44 (33.49%)**	582.25 (33.36%)	612.91 (35.12%)	603.87 (34.60%)	557.92 (31.97%)	561.06 (32.45%)
Brain (WM + GM)	**1160.70 (66.51%)**	1162.89 (66.64%)	1132.24 (64.88%)	1141.27 (65.40%)	1187.23 (68.03%)	1167.70 (67.55%)
Intracranial Cavity (IC)	**1745.14 (100.00%)**	1745.14 (100.00%)	1745.14 (100.00%)	1745.14 (100.00%)	1745.14 (100.00%)	1728.74 (100.00%)

4. CONCLUSION

This paper presents a method for segmentation of magnetic resonance images of the brain, based on the Type-2 Label-based Fuzzy Predicate Classification (T2-LFPC) proposed. Results for slice segmentation and volume computation were presented, considering both simulated and real MRI volumes. A comparison against other known methods was performed, considering the K-nearest neighbor algorithm, multilayer feedforward neural networks, probabilistic neural networks and a similar method to the T2-LFPC but using type-1 FL. T2-LFPC presented a good trade between accuracy and computation times and differences with the comparison methods were not statistically significant considering the non-parametrical Wilcoxon test.

The comparison allowed concluding that T2-LFPC method showed good performance considering noise and distortion in images. In addition, methods based on fuzzy predicates are able to give a deeper understanding about the classification obtained, which could be helpful for image interpretation.

A method of measuring the progressive atrophy and possible changes compared to a therapeutic effect should be essentially automatic and therefore independent of the radiologist. The approach of the methodology followed in this paper is a valuable contribution to achieve this ambitious goal.

REFERENCES

[1] Wattjes, M. P., Steenwijk, M. D.., Stangel, M., "MRI in the Diagnosis and Monitoring of Multiple Sclerosis: An Update.," Clin. Neuroradiol. **25 Suppl 2**, 157–165 (2015).

[2] Kanaly, C. W., Mehta, A. I., Ding, D., Hoang, J. K., Kranz, P. G., Herndon, J. E., Coan, A., Crocker, I., Waller, A. F., et al., "A novel, reproducible, and objective method for volumetric magnetic resonance imaging assessment of enhancing glioblastoma.," J Neuosurg **76**(6), 1–7 (2014).

[3] Johnson, K. A., Fox, N. C., Sperling, R. A.., Klunk, W. E., "Brain imaging in Alzheimer disease.," Cold Spring Harb. Perspect. Med. **2**(4), a006213, Cold Spring Harbor Laboratory Press (2012).

[4] Rohrer, J. D., "Structural brain imaging in frontotemporal dementia," Biochim. Biophys. Acta - Mol. Basis Dis. **1822**(3), 325–332 (2012).

[5] Manjón, J. V.., Coupé, P., "volBrain: An Online MRI Brain Volumetry System," Front. Neuroinform. **10**, 30, Frontiers (2016).

[6] Comas, D. S., Meschino, G. J., Pastore, J. I.., Ballarin, V. L., "A survey of medical images and signal processing problems solved successfully by the application of Type-2 Fuzzy Logic," J. Phys. Conf. Ser. **332**, 12030 (2011).

[7] Zadeh, L. A., "The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I," 199–249 (1975).

[8]   Zadeh, L. A., "Fuzzy sets," Inf. Control **8**(3), 338–353, Academic Press (1965).

[9]   Comas, D. S., Pastore, J. I., Bouchet, A., Ballarin, V. L.., Meschino, G. J., "Type-2 Fuzzy Logic in Decision Support Systems," [Soft Computing for Business Intelligence], R. A. Espin Andrade, R. Bello Pérez, Á. Cobo, J. Marx Gómez, and A. Racet Valdés, Eds., Springer Berlin Heidelberg, Heidelberg, Alemania, 267–280 (2014).

[10]   Meschino, G. J., Espin Andrade, R. A.., Ballarin, V. L., "A framework for tissue discrimination in Magnetic Resonance brain images based on predicates analysis and Compensatory Fuzzy Logic," Int. J. Intell. Comput. Med. Sci. Image Process. **2**(X), 1–16 (2008).

[11]   Melin, P.., Castillo, O., "A review on the applications of type-2 fuzzy logic in classification and pattern recognition," Expert Syst. Appl. **40**(13), 5413–5423 (2013).

[12]   Mendel, J. M., "Type-2 Fuzzy Sets and Systems: An Overview," IEEE Comput. Intell. Mag. **2**(2), 20–29 (2007).

[13]   Comas, D. S., Meschino, G. J., Brun, M.., Ballarin, V. L., "Label-based Type-2 Fuzzy Predicate Classification applied to the design of morphological W-operators for image processing," LA-CCI 2014 - Lat. Am. Congr. Comput. Intell., C. I. S. A. IEEE, Ed., IEEE, CIS Argentina, San Carlos de Bariloche, Argentina (2014).

[14]   Comas, D. S., Meschino, G. J., Nowe, A.., Ballarin, V. L., "Knowledge discovering in data clustering by self-discovered type-2 fuzzy predicates," Fifth Int. Work. Knowl. Discov. Knowl. Manag. Decis. Support (Eureka 2015), J. C. Leyva López, R. A. Espin Andrade, R. Bello Pérez, and P. A. Álvarez Carrillo, Eds., Ciudad México, México (2015).

[15]   Espin Andrade, R. A., MazcorroTéllez, G., Fernández González, E., Marx-Gómez, J.., Lecich, M. I., "Compensatory Logic: a fuzzy normative model for decision making," 178–193 (2006).

[16]   Bouchet, A., Pastore, J. I., Espin Andrade, R. A., Brun, M.., Ballarin, V. L., "Arithmetic Mean Based Compensatory Fuzzy Logic.," Int. J. Comput. Intell. Appl. **10**(2), 231–243 (2011).

[17]   Moore, R.., Lodwick, W., "Interval analysis and fuzzy set theory," Fuzzy Sets Syst. **135**(1 SPEC.), 5–9 (2003).

[18]   Ishibuchi, H.., Tanaka, H., "Multiobjective programming in optimization of the interval objective function," Eur. J. Oper. Res. **48**(2), 219–225 (1990).

[19]   Kundu, S., "Min-transitivity of fuzzy leftness relationship and its application to decision making," Fuzzy Sets Syst. **86**(3), 357–367, North-Holland (1997).

[20]   Sengupta, A.., Pal, T. K., "On comparing interval numbers: A study on existing ideas," Stud. Fuzziness Soft Comput. **238**, 25–37 (2009).

[21]   Bezdek, J. C., Ehrlich, R.., Full, W., "FCM: The fuzzy c-means clustering algorithm," Comput. Geosci. **10**(2-3), 191–203 (1984).

[22]   Fraley, C.., Raftery, A. E., "How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis," Comput. J. **41**(8), 578–588 (1998).

[23]   Zhao, Q., Hautamaki, V.., Fränti, P., "Knee point detection in BIC for detecting the number of clusters," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) **5259 LNCS**, 664–673 (2008).

[24]   Sima, C., Attoor, S., Brag-Neto, U., Lowey, J., Suh, E.., Dougherty, E. R., "Impact of error estimation on feature selection," Pattern Recognit. **38**(12), 2472–2482 (2005).

[25]   Ridgway, G., Barnes, J., Pepple, T.., Fox, N., "Estimation of total intracranial volume; a comparison of methods," Alzheimer's Dement. **7**(4), S62–S63 (2011).

[26]   Pastore, J. I., Moler, E. G.., Ballarin, V. L., "Segmentation of brain magnetic resonance images through morphological operators and geodesic distance," Digit. Signal Process. A Rev. J. **15**(2), 153–160 (2005).

[27]   Aubert-Broche, B., Evans, A. C.., Collins, L., "A new improved version of the realistic digital brain phantom.," Neuroimage **32**(1), 138–145 (2006).

[28]   Doughtery, E. R., Jianping, H.., Bittner, M. L., "Validation of computational methods in genomics.," Curr. Genomics **8**(1), 1–19, Bentham Science Publishers (2007).