Artificial Intelligence in Medicine (2009) 47, 43—52

ARTIFICIAL INTELLIGENCE IN MEDICINE

http://www.intl.elsevierhealth.com/journals/aiim

Intuitionistic fuzzy set vs. fuzzy set application in medical pattern recognition

Vahid Khatibi, Gholam Ali Montazer *

Information Technology Department, School of Engineering, Tarbiat Modares University, P.O. Box 14115-179, Tehran, Iran

Received 22 July 2008; received in revised form 4 February 2009; accepted 11 March 2009

KEYWORDS
Pattern recognition;
Fuzzy sets;
Intuitionistic fuzzy sets;
Similarity measures;
Bacteria detection

Summary

Objective: One of the toughest challenges in medical diagnosis is uncertainty handling. The detection of intestinal bacteria such as *Salmonella* and *Shigella* which cause typhoid fever and dysentery, respectively, is one such challenging problem for microbiologists. They detect the bacteria by the comparison with predefined classes to find the most similar one. Consequently, we observe uncertainty in determining the similarity degrees, and therefore, in the bacteria classification. In this paper, we take an intelligent approach towards the bacteria classification problem by using five similarity measures of fuzzy sets (FSs) and intuitionistic fuzzy sets (IFSs) to examine their capabilities in encountering uncertainty in the medical pattern recognition.
Methods: FSs and IFSs are two strong frameworks for uncertainty handling. The membership degree in FSs and both membership and non-membership degrees in IFSs are the operators that these frameworks use to represent the degree of which a member of the universe of discourse belongs to a subset of it. In this paper, the similarity measures, which both frameworks provide are used, so as the intestinal bacteria are detected and classified through uncertainty quantification in feature vectors. Also, the experimental results of using the measures are illustrated and compared.
Results: We obtained 263 unknown bacteria from microbiology section of Resalat laboratory in Tehran to examine the similarity measures in practice. Finally, the detection rates of the measures were calculated between which IFS Hausdorf and Mitchel similarity measures scored the best results with 95.27% and 94.48% detection rates, respectively. On the other hand, FS Euclidean distance yielded only 85% detection rate.
Conclusions: Our investigation shows that both frameworks have powerful capabilities to cope with the uncertainty in the medical pattern recognition problems. But, IFSs yield better detection rate as a result of more accurate modeling which is involved with incurring more computational cost. Our research also shows that among different IFS similarity measures, IFS Hausdorf and Mitchel ones score the best results.

* Corresponding author. Tel.: +98 21 82883990; fax: +98 21 82883990.
E-mail addresses: khatibi@modares.ac.ir (V. Khatibi), montazer@modares.ac.ir (G.A. Montazer).

0933-3657/$ — see front matter
doi:10.1016/j.artmed.2009.03.002

1. Introduction

Uncertainty is an inseparable aspect of medical diagnosis problems [1,2]. A symptom is an uncertain indication of a disease as it may or may not occur with the disease. Uncertainty characterizes a relation between symptoms and diseases [1—3]. Hence, coping efficiently with uncertainty leads us to more accurate decision-making and this is considered as a fundamental challenge in medicine.

Fuzzy set (FS), proposed by Zadeh [4], as a framework to encounter uncertainty, vagueness and partial truth, represents a degree of membership for each member of the universe of discourse to a subset of it. Therefore, we have a spectrum of truth values. By adding the degree of non-membership to FS, Atanassov proposed intuitionistic fuzzy set (IFS) [5] which looks more accurately to uncertainty quantification and provides the opportunity to precisely model the problem based on the existing knowledge and observations [6]. Both of these frameworks are considered as soft methods which in turn, lead to soft computing and approximate reasoning [7].

In most of the medical diagnosis problems, there exist some base patterns, and experts make decisions based on the similarity between unknown sample and the base patterns [8,9]. Medical diagnosis problems are inherently involved with uncertainty, so fuzzy approach could be appropriate to envisage these problems [6]. On the other hand, the uncertainty is not just in the judgment, but also in the identification. Hence, there is a need for a kind of fuzzy sets which could support the latter uncertainty. The concept of type II fuzzy sets provide such tools through interval valued fuzzy sets and intuitionistic fuzzy sets as two specific implementations. Each of the mentioned sets has its advantages. The interval valued fuzzy sets demonstrate the uncertainty in the membership degree by representing an interval for it, but the IFS reveals the uncertainty by a non-membership degree. IFS is the appropriate choice when representation of the non-membership degree is simpler than the membership degree [10]. Also, Atanassov and Gargov [11] further generalized IFS in the spirit of ordinary interval-valued fuzzy sets and defined the notion of interval-valued intuitionistic fuzzy set (IVIFS).

The main advantages of using the IFS framework are [12]:

- By using IFSs, we transform a vague pattern classification problem into a precise and well-defined optimization problem.
- IFSs, unlike ordinary FSs, retain a controlled degree of the uncertainty.

The main disadvantage of using the IFS framework is [12]:

- The relatively high computational complexity.

Also, the intuitionistic fuzzy set theory has independently introduced by Takeuti and Titani [13] as a set theory developed in (a kind of) intuitionistic logic. Takeuti—Titani's intuitionistic fuzzy logic is simply an extension of intuitionistic logic [14]. Even if their interpretive settings and motivation are quite different, Atanassov's construct is isomorphic to interval-valued fuzzy sets and other similar notions, the latter captures the idea of ill-known membership grade, while the former starts from the idea of evaluating degrees of membership and non-membership independently. The main objection to the terminology used by Atanassov is raised from the fact that he calls "intuitionistic fuzzy set theory" something which accepts rules and principles (such as the double negation) that, added to the intuitionistic logic, make it classical, i.e. nothing from intuitionism remains. In this way, calling the Atanassov's theory intuitionistic leads to a misunderstanding. Hence, a solution is proposed to use the terms such as uncertain, imprecise, vague or bipolar fuzzy sets instead of intuitionistic fuzzy sets and replace intuitionistic fuzzy logic with neutrosophic logic which is in line with the terminology in this area and does not borrow its name from another field (intuitionistic logic) [10].

Among diverse medical diagnosis problems, the bacteria detection is considered as a key issue in the microbiology [15]. Detecting intestinal bacteria such as *Shigella*, *Salmonella*, *Klebsiella*, and *Bacillus coli* is an important task of microbiologists, because each of them could originate specific diseases. *Salmonella* and *Shigella* are the causes of typhoid fever and dysentery, respectively [16].

Microbiologists detect the bacteria by the special experiments to investigate their features similarities with the predefined classes' features [17]. But, the detection is involved with uncertainty, so as we could not assign an unknown pattern to a specific class certainly.

In this paper, we applied the similarity measures of FSs and IFSs to the bacteria detection problem to examine their capabilities to cope with uncertainty in the medical pattern recognition.

The organization of this paper is as follows: in Section 2, the bacteria classification and the intestinal bacteria considered in this paper are described. Then, in Section 3, FS, IFS and their respective similarity measures are represented. In Section 4, we use these tools to solve a bacteria detection problem, and then, illustrate their results comparison.

2. Bacteria classification

Microbiologists broadly classify bacteria according to their shapes [17,18]. Most bacteria come in one of three shapes: rod, sphere, or spiral. Rod-shaped bacteria are called bacilli. Spherical bacteria are called cocci, and spiral or cockscrew-shaped ones are called spirilla. Some bacteria come in more complex shapes, e.g. plemorphic bacteria can assume a variety of shapes. It must be noted that the bacterial cell wall generally determines the shape of the bacterial cell. Bacteria may be further classified according to whether they require oxygen and how they react to a test with Gram's stain. Bacteria in which alcohol washes away Gram's stain are called Gram-negative, while bacteria in which alcohol cause the bacteria's wall to absorb the stain, are called Gram-positive. Another system for bacteria classification depends on source of carbon which the bacteria are fed on (autotrophic or heterotrophic) [17,18].

As noted before, bacillus is a rod-shaped bacterium which is active only in the presence of the oxygen (aerobic bacterium). Bacilli occur mainly in chains, produce spores, and include many saprophytes, some parasites, and the bacterium that causes anthrax [19]. In this research, four intestinal bacteria of bacillus category named *Shigella*, *Salmonella*, *Bacillus coli* and *Klebsiella*, shown in Fig. 1, are considered which have some similarity in culture medium and are Gram-negative.

Shigella [19] is a rod-shaped Gram-negative bacterium (bacillus) that lives in the intestinal tracts of human beings and animals and causes bacillary dysentery. There are four species, all causing dysentery, but with varying degrees of severity. *Salmonella* [19] is also a rod-shaped bacterium, found in the intestine that can cause food poisoning, gastroenteritis, and typhoid fever. Other two bacteria considered herein, *Klebsiella* and *Bacillus coli* are also Gram-negative and found in the intestine.

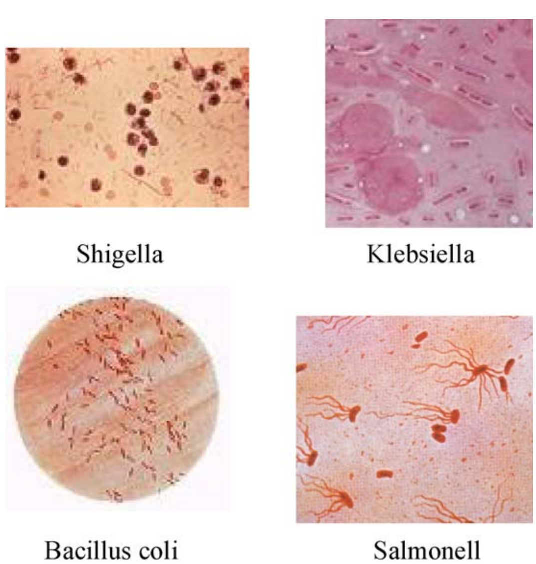

Figure 1 Four intestinal bacteria considered in the detection problem.

The bacteria detection is a microbiologists' task. They use experiments such as the Gram's stain to classify unknown bacteria [17,18]. Hence, they spend much time and energy to detect bacteria, whereas this repetitive task could be done by an expert system. To replace a human expert, the system exploits respective expertise in the pattern recognition. Our aim is to use FS and IFS similarity measures in the bacteria detection to examine their capabilities in encountering uncertainty in the medical pattern recognition.

Primal features used for the bacteria classification herein comprise of the macroshape, which is domical in the four studied bacteria, the single microscopic shape, the double microscopic shape, the colony size, and existence of the flagellum. In the next section, we represent the fundamental concepts of FS, IFS and their respective similarity measures which are applied to a bacteria detection problem in following.

3. Fuzzy sets, intuitionistic fuzzy sets, and their similarity measures

Fuzzy sets theory, proposed by Zadeh [4] in 1965, has shown successful applications in various fields. In this theory, the membership of an element to a fuzzy set is a single value between zero and one. But in reality, it may not always be certain that the degree of non-membership of an element to a fuzzy set is just equal to 1 minus the degree of membership, i.e. there may be some hesitation degree. So, as a generalization of fuzzy sets, the concept of intuitionistic fuzzy sets was introduced by Atanassov [5] in 1986. Bustince and Burillo [20] showed that this notion coincides with the notion of vague sets (VSs).

The IFS was defined as an extension of the ordinary FS [21]. As opposed to a fuzzy set in X, given by

$$A = \{(x, \mu_A(x)) | x \in X\} \quad (1)$$

where μ_A:$X \rightarrow$ [0,1] is the membership function of the fuzzy set A, an intuitionistic fuzzy set B is given by

$$B = \{(x, \mu_B(x), \nu_B(x)) | x \in X\} \quad (2)$$

where μ_B:$X \rightarrow$ [0,1] and $\nu_B : X \rightarrow [0, 1]$ are such that:

$$0 \leq \mu_B + \nu_B \leq 1 \quad (3)$$

and $\mu_B, \nu_B \in [0, 1]$ denote degrees of membership and non-membership of $x \in A$, respectively. For each intuitionistic fuzzy set B in X, "hesitation margin" (or "intuitionistic fuzzy index") of $x \in B$ is given by

$$\pi_B = 1 - \mu_B - \nu_B \tag{4}$$

which expresses a hesitation degree of whether x belongs to B or not. It is obvious that $0 \leq \pi_B(x) \leq 1$, for each $x \in X$. The illustration of these degrees has been exhibited in Fig. 2.

Therefore, to describe an intuitionistic fuzzy set completely; we need at least two functions from the triplet [21,22]:

- membership function;
- non-membership function;
- hesitation margin.

In other words, using IFS introduces another degree of freedom into a set description (i.e. in addition to μ_B, we also have ν_B or π_B). Since the intuitionistic fuzzy sets are generalized from fuzzy sets and they give us an additional possibility to represent imperfect knowledge, they can make it possible to describe many real problems in a more adequate way [6,23].

Similarity and distance measures as two main pattern recognition mechanisms can be considered as dual concepts [24]. As important subjects in intuitionistic fuzzy mathematics, similarity and distance measures between IFSs have attracted many researchers. Szmidt and Kacprzyk [25] proposed four distance measures between IFSs, which were in some extent based on the geometric interpretation of intuitionistic fuzzy sets, and have some good geometric properties.

Let $A = \{(x, \mu_A(x), \nu_A(x)) | x \in X\}$ and $B = \{(x, \mu_B(x), \nu_B(x)) | x \in X\}$ be two IFSs in $X = \{x_1, x_2, \ldots, x_n\}$. Based on the geometric interpretation of IFS, Szmidt and Kacprzyk proposed the following distance measures between A and B:

- Hamming distance:

$$d^1_{IFS}(A,B) = \frac{1}{2}\sum_{i=1}^{n}(|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|) \tag{5}$$

- Euclidean distance:

$$e^1_{IFS}(A,B) = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + (\pi_A(x_i) - \pi_B(x_i))^2} \tag{6}$$

whereas their respective measures in fuzzy sets are represented with:

- Hamming distance:

$$d^1_{FS}(A,B) = \frac{1}{2}\sum_{i=1}^{n}(|\mu_A(x_i) - \mu_B(x_i)|) \tag{7}$$

- Euclidean distance:

$$e^1_{FS}(A,B) = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(\mu_A(x_i) - \mu_B(x_i))^2} \tag{8}$$

Li and Cheng proposed some IFS similarity measures and applied them to pattern recognition problems [26]. For A and B as IFSs in which $x \in [a,b]$, Li and Cheng [27] proposed the following similarity measure:

$$T^c_L(A,B) = 1 - \sqrt[p]{\frac{1}{(b-a)}\int_a^b \left[\frac{(\mu_A(x) - \mu_B(x)) + (\nu_A(x) - \nu_B(x))}{2}\right]^p dx} \tag{9}$$

which Liu [28] modified it to:

$$T^c(A,B) = 1 - \sqrt[p]{\frac{1}{2(b-a)}\int_a^b \left[|\mu_A(x) - \mu_B(x)|^p + |\nu_A(x) - \nu_B(x)|^p + |\pi_A(x) - \pi_B(x)|^p\right] dx} \tag{10}$$

But Liang and Shi [29] and Mitchell [30] pointed out that Li and Cheng's measures are not always effective, and made some modifications, respectively. Also, Dengfeng and Chantian proposed a similarity measure between the two

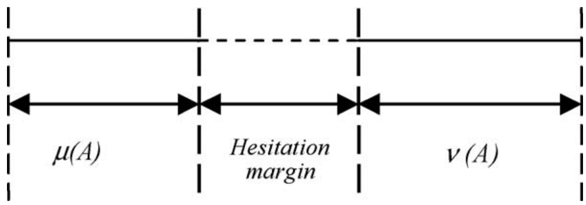

Figure 2 Membership, non-membership, and hesitation degrees.

IFSs, A and B as follows:

$$S_d^p(A,B) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} |m_A(i) - m_B(i)|^p} \tag{11}$$

where $m_A(i) = (\mu_A(x_i) + 1 - \nu_A(x_i))/2, m_B(i) = (\mu_B(x_i) + 1 - \nu_B(x_i))/2$, and $1 \le p < \infty$.

Also, Liang and Shi [29] proposed the following similarity measure between IFSs. Let $\Phi_{tAB}(i) = |\mu_A(x_i) - \mu_B(x_i)|/2$ and $\Phi_{fAB}(i) = |[(1 - \nu_A(x_i))/2] - [(1 - \nu_B(x_i))/2]|$. They used $(\Phi_{tAB}(i) + \Phi_{fAB}(i))^p$ to measure the distance between $[\mu_A(x_i), 1 - \nu_A(x_i)]$ and $[\mu_B(x_i), 1 - \nu_B(x_i)]$. Thus, the distance between IFSs A and B was given by

$$D_e^p(A,B) = \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} (\Phi_{tAB}(i) + \Phi_{fAB}(i))^p} \tag{12}$$

and $S_e^p(A,B) = 1 - D_e^p(A,B)$ was defined as a similarity measure between A and B. Also, Mitchell [30] adopted a statistical approach and interpreted IFSs as ensembles of ordered fuzzy sets to modify Li and Cheng's similarity measure. Let

He then, defined the modified similarity measure between A and B with:

$$S_{\text{mod}}(A,B) = \frac{1}{2}(\rho_\mu(A,B) + \rho_\nu(A,B)) \tag{15}$$

Hung and Yang [31] suggested some similarity measures between IFSs. First, they use the idea of Hausdorf distance to define the distance between IFSs A and B as follows:

$$d_H(A,B) = \frac{1}{n}\sum_{i=1}^{n} \max\{|\mu_A(x_i) - \mu_B(x_i)|, |\nu_A(x_i) - \nu_B(x_i)|\}. \tag{16}$$

Furthermore, they used the distance d_H to generate three similarity measures:

$$S_l(A,B) = 1 - d_H(A,B), \tag{17}$$

$$S_e(A,B) = \frac{\exp(-d_H(A,B)) - \exp(-1)}{1 - \exp(-1)}, \tag{18}$$

$$S_c(A,B) = \frac{1 - d_H(A,B)}{1 + d_H(A,B)}. \tag{19}$$

We can utilize the IFSs correlation measures or association coefficients to investigate the similarity as well. Recently, a new IFS association coefficient has been represented [32]. Let us consider a discrete universe of discourse, $X = \{x_1, x, \ldots, x_n\}$, and let $A, B \in \Phi(X)$, then the association coefficient of A and B can be defined as

$$c(A,B) = \frac{\sum_{i=1}^{n}((\mu_A(x_i) + \mu_B(x_i)) + (\upsilon_A(x_i) + \upsilon_B(x_i)) + (\pi_A(x_i) + \pi_B(x_i)))}{\max\left(\sum_{i=1}^{n}(((\mu_A^2(x_i) + \upsilon_A^2(x_i) + \pi_A^2(x_i)), \sum_{i=1}^{n}((\mu_B^2(x_i) + \upsilon_B^2(x_i) + \pi_B^2(x_i)))\right)} \tag{20}$$

$\rho_\mu(A,B)$ and $\rho_\nu(A,B)$ denote the similarity measures between the "low" membership functions μ_A and μ_B and between the "high" membership functions $1 - \upsilon_A$ and $1 - \upsilon_B$, respectively, as follows:

$$\rho_\mu(A,B) = S(\mu_A, \mu_B) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|^p}, \tag{13}$$

and

$$\rho_\nu(A,B) = S(1 - \nu_A, 1 - \nu_B) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} |\nu_A(x_i) - \nu_B(x_i)|^p}. \tag{14}$$

The larger the value of $c(A,B)$, the more the association between A and B. Clearly, the association coefficient derived from Eq. (20) takes into account all the three parameters (membership, non-membership, and hesitation degrees) describing IFS, and thus contains the most information.

4. FS vs. IFS in bacteria detection

As noted in Section 2, microbiologists generally rely on the experimental results such as Gram's stain to classify the bacteria. We extract the knowledge and expertise of microbiologists regarding the features of the studied intestinal bacteria to propose a soft method for their detection. For this purpose, first we obtained 263 unknown bacteria from microbiology section of Resalat laboratory in Tehran. Then, exis-

Table 1 Some part of the data set.

Sample	Domical shape		Single microscopic shape		Double microscopic shape		Flagellum	
	$\mu(A_i)$	$\nu(A_i)$	$\mu(B_i)$	$\nu(B_i)$	$\mu(C_i)$	$\nu(C_i)$	$\mu(D_i)$	$\nu(D_i)$
S_1	0.837	0.133	0.718	0.159	0.064	0.897	0.021	0.806
S_2	0.911	0.029	0.831	0.031	0.028	0.894	0.952	0.036
S_3	0.929	0.037	0.812	0.033	0.021	0.926	0.054	0.922
S_4	0.815	0.091	0.949	0.048	0.020	0.880	0.833	0.042
S_5	0.864	0.02	0.610	0.230	0.243	0.624	0.0004	0.964
S_6	0.905	0.016	0.878	0.015	0.072	0.917	0.789	0.114

tence of the aforementioned features was investigated in the unknown bacteria. Consequently, the degrees of membership and non-membership for the features of unknown bacteria were formed. In fact, based on the physical evidence of feature existence in the bacteria by microscopic investigation or experiments, microbiologists determined the membership and non-membership values of the bacteria features. Some part of our data set has been exhibited in Table 1. Based on the acquired knowledge from microbiologists, values of the features in the four studied bacteria were determined which are shown in Table 2. The membership and non-membership functions for each feature indicate the feature's existence and non-existence degrees in a specific class, respectively.

The colony size membership and non-membership function curves have been shown in Figs. 3 and 4, respectively. To detect unknown bacteria, first, the fuzzy Euclidean distance was used to determine the most similar class to an unknown sample. For instance, according to Eq. (8), to compare unknown sample S_1 with *Shigella* class, we have:

$$\mathit{Fuzzy\_Euc}(S_1, \mathit{Shigella}) = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(\mu_{Shigella}(x_i) - \mu_{S_1}(x_i))^2} \tag{23}$$

Then, the IFS Euclidean distance measure was applied to the bacteria detection problem to determine the most similar class for an unknown sample. For instance, based on Eq. (6), we have:

$$\mathit{IFS\_Euc}(S_1, \mathit{Shigella}) = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(\mu_{Shigella}(x_i) - \mu_{S_1}(x_i))^2 + (\nu_{Shigella}(x_i) - \nu_{S_1}(x_i))^2 + (\pi_{Shigella}(x_i) - \pi_{S_1}(x_i))^2} \tag{24}$$

For the colony size, no constant values were observed for membership and non-membership degrees. Instead, according to the colony size, variant behaviors for these functions were obtained. To set equations for these functions, several regression functions were examined to fit the points, and at last, the following polynomial functions resulted in the best fitness:

$$\mu(\text{Colony Size}) = -0.23x^2 + 0.98x - 0.1849 \tag{21}$$

$$\nu(\text{Colony Size}) = 0.24x^2 - 0.97x + 1.1349 \tag{22}$$

Also, three IFS similarity measures of Mitchel, Hausdorf and Xu et al. have been applied to this problem to examine their capabilities to cope with the uncertainty. For instance, according to Eqs. (13)—(15) for Mitchel similarity measure which is itself a modified version of Dengfeng and Chantian measure, we have:

$$\rho_\mu(S_1, \mathit{Shigella}) = S(\mu_{S_1}, \mu_{Shigella}) = 1 - \frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}|\mu_{S_1}(x_i) - \mu_{Shigella}(x_i)|^p}, \tag{25}$$

Table 2 The features' values in the studied bacteria classes (base patterns).

Classes	Domical shape		Single microscopic shape		Double microscopic shape		Flagellum	
	$\mu(A_i)$	$\nu(A_i)$	$\mu(B_i)$	$\nu(B_i)$	$\mu(C_i)$	$\nu(C_i)$	$\mu(D_i)$	$\nu(D_i)$
Bacillus coli	0.85	0.05	0.87	0.01	0.02	0.97	0.92	0.06
Shigella	0.83	0.08	0.92	0.05	0.05	0.92	0.08	0.91
Salmonella	0.79	0.12	0.78	0.11	0.11	0.85	0.87	0.01
Klebsiella	0.82	0.15	0.72	0.15	0.22	0.75	0.12	0.85

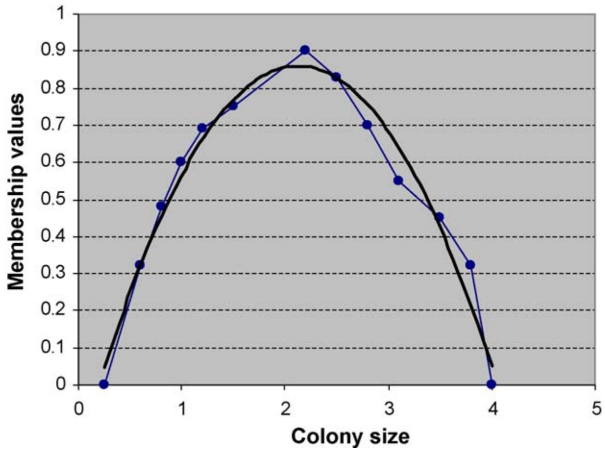

Figure 3 The colony size membership function.

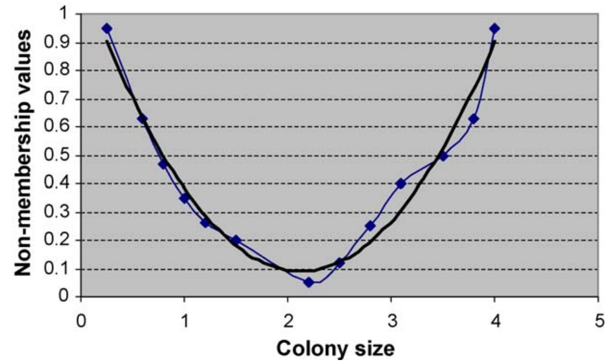

Figure 4 The colony size non-membership function.

and

$$\rho_\nu(S_1, Shigella) = S(1-\nu_{S_1}, 1-\nu_{Shigella}) = 1 - \frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}\left|\nu_{S_1}(x_i) - \nu_{Shigella}(x_i)\right|^p}, \quad (26)$$

where assuming $p = 1$:

$$S_{\text{mod}}(A,B) = \frac{1}{2}(\rho_\mu(A,B) + \rho_\nu(A,B)). \quad (27)$$

Fig. 5 illustrates and compares the detection performance of the fuzzy Euclidean distance, IFS Euclidean distance and IFS Mitchel similarity measures. Also, the comparison of the four IFS measures' results against the microbiologists' opinions has been represented in Fig. 6. Table 3 summarizes the applied approaches' results and compares them with the microbiologists' opinions.

Having applied the five aforementioned measures to all the unknown bacteria, fuzzy Euclidean distance measure could distinguish 85% of the samples properly, whereas, IFS Euclidean distance scored a better result and detected 87% of the samples correctly. Also, the IFS Mitchel and Hausdorf similarity measures even provided us with more accurate classification, so as they could detect 94.48% and 95.27% of the samples properly, respectively. All the three Hausdorf similarity measures Eqs. (15)—(17), were applied to the problem which brought us

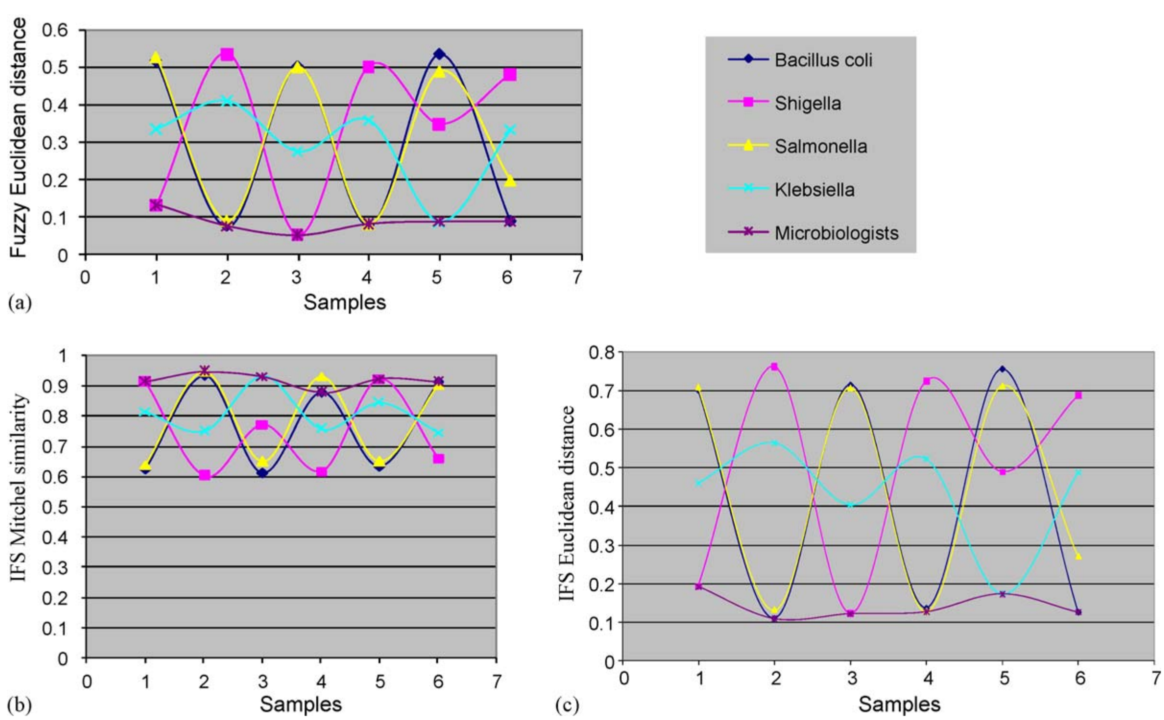

Figure 5 The performance comparison of the three approaches in the bacteria detection: (a) fuzzy Euclidean distance, (b) IFS Mitchel similarity, and (c) IFS Euclidean distance.

Table 3 Classification results based on the FS and IFS measures vs. microbiologists' opinions.

S. No.	Measure	*Bacillus coli*	*Shigella*	*Salmonella*	*Klebsiella*	Microbiologists
S_1	FS Euclidean distance	0.214	1.260	0.256	0.988	*Bacillus coli*
	IFS Euclidean distance	0.267	1.787	0.353	1.360	
	IFS Hausdorf similarity	0.919	0.593	0.889	0.659	
S_2	FS Euclidean distance	1.178	0.190	1.172	0.764	*Shigella*
	IFS Euclidean distance	1.721	0.306	1.734	1.128	
	IFS Hausdorf similarity	0.578	0.901	0.598	0.766	
S_3	FS Euclidean distance	0.212	1.248	0.145	0.923	*Salmonella*
	IFS Euclidean distance	0.288	1.777	0.201	1.273	
	IFS Hausdorf similarity	0.901	0.592	0.939	0.744	
S_4	FS Euclidean distance	1.098	0.625	1.059	0.287	*Klebsiella*
	IFS Euclidean distance	1.588	0.913	1.563	0.424	
	IFS Hausdorf similarity	0.579	0.753	0.629	0.862	
S_5	FS Euclidean distance	0.114	1.191	0.207	0.950	*Bacillus coli*
	IFS Euclidean distance	0.146	1.701	0.307	1.315	
	IFS Hausdorf similarity	0.968	0.622	0.898	0.671	
S_6	FS Euclidean distance	1.203	0.1425	1.212	0.819	*Shigella*
	IFS Euclidean distance	1.737	0.211	1.773	1.158	
	IFS Hausdorf similarity	0.616	0.956	0.584	0.722	

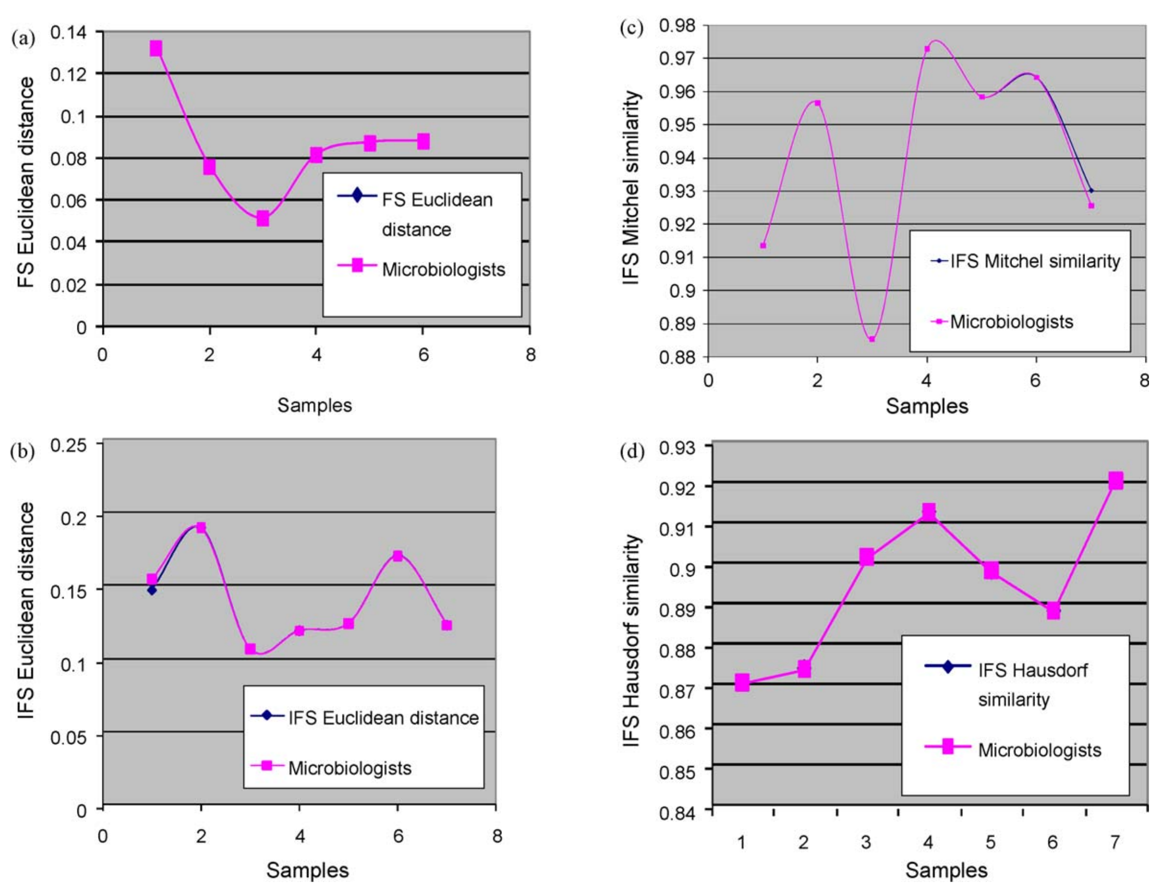

Figure 6 The comparison of the four approaches with microbiologists' opinions: (a) fuzzy Euclidean distance, (b) IFS Euclidean distance, (c) IFS Mitchel similarity, and (d) IFS Hausdorf similarity.

95.27%, 94.4%, and 86.6% detection rates, respectively. Also, the Xu et al. association coefficient approach was examined to classify the unknown bacteria, but it yielded the weak detection rate of 86.61%.

Considering non-membership function and incurring more computational cost in IFS framework led us to better results. Also, the errors in misclassified samples of IFS Hausdorf similarity measure resided in the interval [0.003, 0.07], whereas the same errors were in the intervals [0.007, 0.7] and [0.01, 0.09] for IFS and fuzzy Euclidean distance measures, respectively. This implies utilizing the IFS framework not only provides us with more accurate classification results, but also, the related errors are less than that the fuzzy framework generates.

5. Conclusions

Soft methods which deal with uncertainty, vagueness and partial truth, provide us with better tolerance to cope with uncertainty in the medical diagnosis problems. The bacteria detection by soft computing was undertaken in this paper. For this purpose, different similarity measures of FSs and IFSs were applied to a bacteria classification problem to investigate their capabilities to encounter the uncertainty in the medical pattern recognition.

IFSs yielded more accurate classification results than FSs, so as 95.27% of unknown samples were classified correctly by IFS Hausdorf similarity measure, while FS Euclidean distance measure detected 85% of the samples correctly. The cause of this phenomenon resides in the nature of the IFS framework. It considers the degree of non-membership and this leads to better detection rate, but it conveys a drawback of more computational cost.

Acknowledgements

The authors would like to thank Ms. Hoda Davarzani, Ph.D. student at the Tarbiat Modares University, for her precious assistance in data analysis of our research. This work was supported by the Iran Telecommunications Research Center (ITRC), Tehran, Iran, under grant no. TMU 87-07-48.

References

[1] De SK, Biswas A, Roy R. An application of intuitionistic fuzzy sets in medical diagnosis. Fuzzy Sets and Systems 2001;117(2): 209—13.
[2] Straszecka E. Combining uncertainty and imprecision in models of medical diagnosis. Information Sciences 2007;176:3026—59.
[3] Han J, Kamber M. Data mining: concepts and techniques. Morgan, San Mateo, CA: Elsevier, Kaufmann; 2006 .
[4] Zadeh LA. Fuzzy sets. Information Control 1965;8: 338—53.
[5] Atanassov K. Intuitionistic fuzzy sets. Fuzzy Sets and Systems 1986;20(1):87—96.
[6] Vlachos IK, Sergiadis GD. Intuitionistic fuzzy information: applications to pattern recognition. Pattern Recognition Letters 2007;28:197—206.
[7] Zadeh LA. Fuzzy logic, neural networks, and soft computing. Communications of ACM 1994;37:77—84.
[8] Meyer-Base A. Neuro-fuzzy classification. Pattern Recognition in Medical Imaging 2004;282—317.
[9] Wolff JG. Medical diagnosis as pattern recognition in a framework of information compression by multiple alignment, unification and search. Decision Support Systems 2006;42:608—25.
[10] Dubois D, Gottwald S, Ha-jek P, Kacprzyk J, Prade H. Terminological difficulties in fuzzy set theory—the case of intuitionistic fuzzy sets. Fuzzy Sets and Systems 2005;156:485—91.
[11] Atanassov K, Gargov G. Interval valued intuitionistic fuzzy sets. Fuzzy Sets and Systems 1989;31:343—9.
[12] Mitchell HB. Pattern recognition using type-II fuzzy sets. Information Sciences 2005;170:409—18.
[13] Takeuti G, Titani S. Intuitionistic fuzzy logic and intuitionistic fuzzy set theory. Journal of Symbolic Logic 1984;49:851—66.
[14] Dalen DV. Intuitionistic logic, handbook of philosophical logic. Dordrecht: Kluwer; 2002.
[15] Singleton P. Bacteria in biology, biotechnology, and medicine. Chichester: Wiley; 1999.
[16] Garrett L. The coming plague: newly emerging diseases in a world out of balance. New York: Penguin Group; 1994 .
[17] Facklam H, Facklam M. Bacteria. Kent: Twenty-First Century; 1995.
[18] Ricciuti ER. Microorganisms: the unseen world. San Diego: Blackbirch; 1993.
[19] Sankaran N. Microbes and people: an A—Z of microorganisms in our lives. Phoneix: Oryx; 2001.
[20] Bustince H, Burillo P. Vague sets are intuitionistic fuzzy sets. Fuzzy Sets and Systems 1996;79(3):403—5.
[21] Atanassov KT. Intuitionistic fuzzy sets. Heidelberg: Physica-Verlag; 1999.
[22] Atanassov K. More on intuitionistic fuzzy sets. Fuzzy Sets and Systems 1989;33(1):37—46.
[23] Abbas SE. On intuitionistic fuzzy compactness. Information Sciences 2005;173:75—91.
[24] Wang W, Xin X. Distance measure between intuitionistic fuzzy sets. Pattern Recognition Letters 2005;26:2063—9.
[25] Szmidt E, Kacprzyk J. Distances between intuitionistic fuzzy sets. Fuzzy Sets and Systems 2000;114(3):505—18.
[26] Li D, Cheng C. New similarity measures of intuitionistic fuzzy sets and application to pattern recognition. Pattern Recognition Letters 2002;23(1-3):221—5.
[27] Li DF, Cheng CT. New similarity measures of intuitionistic fuzzy sets and application to pattern recognition. Pattern Recognition Letters 2002;23:221—5.
[28] Liu H-W. New similarity measures between intuitionistic fuzzy sets and between elements. Mathematical and Computer Modelling 2005;42:61—70.

[29] Liang Z, Shi P. Similarity measures for intuitionistic fuzzy sets. Pattern Recognition Letters 2003;24:2687—93.
[30] Mitchell HB. On the Dengfeng—Chuntian similarity measure and its application to pattern recognition. Pattern Recognition Letters 2003;24:3101—4.
[31] Hung W-L, Yang M-S. Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. Pattern Recognition Letters 2004;25:1603—11.
[32] Xu Z, Chen J, Wuc J. Clustering algorithm for intuitionistic fuzzy sets. Information Sciences 2008;178:3775—90.