Soft Computing (2020) 24:83–99
https://doi.org/10.1007/s00500-019-04157-2

FOCUS

Toward a development of general type-2 fuzzy classifiers applied in diagnosis problems through embedded type-1 fuzzy classifiers

Emanuel Ontiveros-Robles[1] · Patricia Melin[1]

Published online: 21 June 2019

Abstract
Nowadays, with the emergence of computer-aided systems, diagnosis problems are one of the most important application areas of artificial intelligence. The present paper is focused on a specific kind of computer-aided diagnosis system based on General Type-2 Fuzzy Logic. The main goal is the generation of General Type-2 Fuzzy Classifiers that can handle the data uncertainty. The concept of embedded Type-1 Fuzzy membership functions has been proposed to be used in the design of General Type-2 Fuzzy Classifiers. A methodology for generating the embedded Type-1 fuzzy membership functions is introduced, and the subsequent approach for developing the Footprint of Uncertainty of the General Type-2 Fuzzy Classifier is presented. On the other hand, the proposed approach performance is evaluated by the experimentation with different diagnosis benchmark problems. In addition, a statistical comparison with respect to another existing approach of General Type-2 Fuzzy classifiers is presented.

Keywords General type-2 fuzzy logic · Fuzzy classifier · Footprint of uncertainty

1 Introduction

The emergence of computer-aided diagnosis systems has demonstrated the reliability of artificial intelligence in real-world problems. For example, in Liao et al. (2018) the authors show the efficiency of deep convolutional neural networks for the diagnosis of multiple types of cancer, in Erkaymaz and Ozer (2016) the authors introduce an approach based on feedforward neural networks for the diagnosis of diabetes with interesting results, in Babapour Mofrad et al. (2019) the authors propose to use a decision tree for the interpretation of CSF biomarkers in the diagnosis of Alzheimer's disease, and more cases can be found in the literature, for example Saritas (2012), Subasi (2013), Elyan and Gaber (2016), Davari Dolatabadi et al. (2017), Asl and Zarandi (2017), Rakhmetulayeva et al. (2018), Vogado et al. (2018), Wang et al. (2018), Acharya et al. (2018), Qi et al. (2019), Afifi et al. (2019).

The present paper aims at designing a computer-aided diagnosis system based on General Type-2 Fuzzy Logic and called General Type-2 Fuzzy Classifier (GT2 FC). The methodology for obtaining the parameters of the system and a new approach to estimate the uncertainty of the system is presented.

The main contribution of the present paper is applying the concept of embedded Type-1 Fuzzy memberships for the parameterization of the Footprint of Uncertainty (FOU) of GT2 membership functions in a GT2 Fuzzy Classifier. Remembering that the FOU is modeling the uncertainty in the Type-2 Fuzzy Systems, it is proposed that is possible to find the parameters for modeling the uncertainty based on n subsets resulting from applying a uniform sampling with replacement, and based on multiple Type-1 Fuzzy membership functions (one per each subset) it is possible to generate a single GT2 Fuzzy Classifier. The proposed approach is in focused on Diagnosis problems; however, the methodology for uncertainty modeling can be extended to other kind of problems for example time series. The concept of embedded Type-1 fuzzy membership functions is not new, it was presented for example in Hagras (2008), but the methodology to be applied in classification problems and especially in diagnosis problems is interesting and obtains interesting results.

Communicated by O. Castillo, D. K. Jana.

✉ Patricia Melin
pmelin@tectijuana.mx

[1] Tijuana Institute of Technology, Calzada Tecnologico s/n, Fracc. Tomas Aquino, 22379 Tijuana, Mexico

2 Materials and methods

In this section, a brief introduction of the necessary concepts and definitions for understanding the proposed approach is presented. On the other hand, the methodology for the design of the proposed GT2 Fuzzy Classifier is explained.

2.1 Uniform sampling with replacement

The uniform sampling with replacement is a technique for sampling, where the data has the same probability to be selected even for multiple samplings. This technique considers that every sampling is independent of the others and with the same population. One of the main applications of this technique is the bagging meta-algorithm for example in Hothorn and Lausen (2005), Baraldi et al. (2011), Fernández-Carrobles et al. (2016).

The method consists on generating N new training sets with size M from the D standard training set uniformly and with replacement. Based on these new training subsets, a clustering algorithm can be performed obtaining N sets of equivalent clusters.

Figure 1 illustrates the *Uniform sampling with replacement*.

2.2 Type-2 fuzzy logic

Type-1 fuzzy logic was originally introduced by Liang and Mendel (2000) as an approach to represent vagueness. On the other hand with the emergence of the Interval Type-2 Fuzzy Logic (Liang and Mendel 2000), this approach provides a method to handle uncertainty, and this uncertainty is modeled by an area between two type-1 fuzzy sets and is called Footprint of Uncertainty (FOU) (Mendel et al. 2006).

However, the present paper introduces an approach to generate the FOU for General Type-2 Fuzzy Inference Systems (GT2 FIS). The main difference with respect to an Interval Type-2 Fuzzy Inference System (IT2 FIS) is in handling uncertainty, this is because in IT2 FIS the uncertainty is considered to be uniform, on the other hand, in GT2 FSs the uncertainty is defined by a secondary membership function on the secondary axis, and Eqs. 1 and 2 describe the IT2 FIS and GT2 FIS, respectively (Mendel et al. 2016).

$$J_x = \left\{ ((x,u)) | u \in \left[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x) \right] \right\} \tag{1}$$

$$J_x = \left\{ ((x,u)) \middle| u \in [0,1], \mu_{\tilde{A}}(x) \rangle 0 \right\} \tag{2}$$

where X is the primary axis of the corresponding input and the secondary axis J is related to the uncertainty.

The Footprint of Uncertainty (FOU) (Mendel and John 2002; Ontiveros et al. 2018a) is represented as the area between the boundaries of the Type-2 Fuzzy Sets. For example, in Interval Type-2 Fuzzy Sets the FOU is the area between the lower and the upper membership functions (Fig. 2).

On the other hand, the rules of the Type-2 Fuzzy Systems are very similar to the Type-1 Fuzzy Systems. In this case, there is an antecedent and a consequent, the inference is realized by the introduction of the extension of the T-Norm and S-Norm now called *meet* and *join*, respectively. The structure of the Type-2 Fuzzy rules is expressed in Eq. (3).

$$R^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \text{ and}\ldots\text{and } x_p \text{ is } \tilde{F}_p^l, \text{ THEN } y \text{ is } \tilde{G}^l, \quad \text{where } l = 1, \ldots, M \tag{3}$$

2.2.1 α-plane representation

The α-plane representation is an approach for modeling the GT2 FIS in order to achieve a good approximation of this system (Mendel et al. 2009; Ontiveros et al. 2018b). This representation consists on approximating the GT2 FIS by horizontal slices called α-planes and finally aggregating the

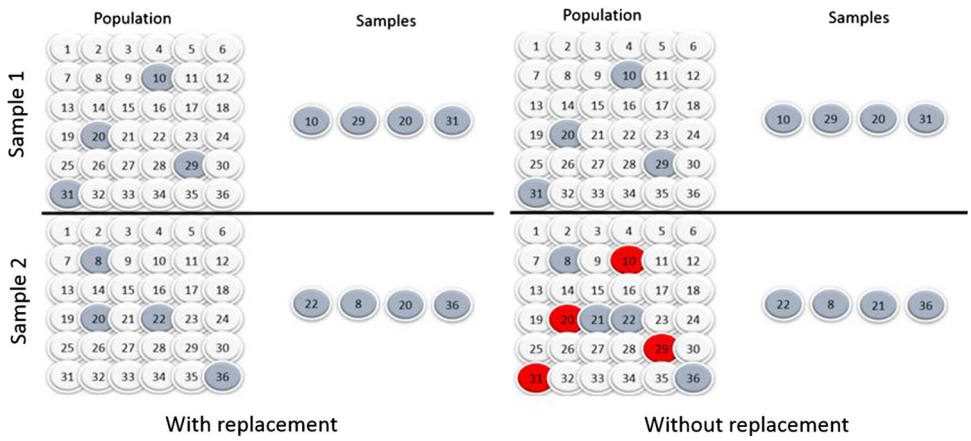

Fig. 1 Uniform sampling with replacement

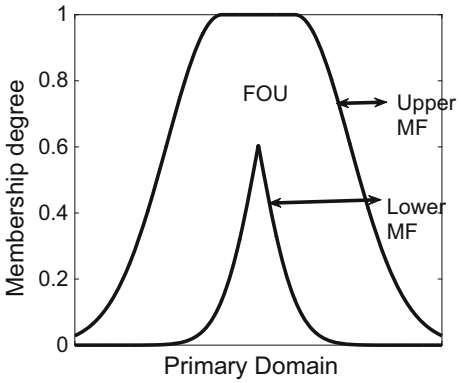

Fig. 2 IT2 MF

results. These slices can be computed as IT2 FIS and have a high computational cost, but are computable for applications that not require real-time execution. Equation (4) represents the expression of the α-planes and Eq. (5) represents the aggregation of the α-planes.

$$\tilde{J}_\alpha = \{((x,u))|u \in [0,1], \mu_{\tilde{A}}(x) = \alpha\} \tag{4}$$

$$\tilde{\tilde{J}} = \cup \tilde{J}_\alpha \tag{5}$$

A method for approximating a GT2 FIS by the α-planes representation is presented in Fig. 3.

The aggregation of the results of the α-planes is performed by (6).

$$\tilde{\tilde{A}} = \frac{\sum \alpha \tilde{A}_\alpha}{\sum \alpha} \tag{6}$$

The number of α-planes impact the performance of the GT2 FIS; however, it was found that 10 α-planes can be enough for obtaining a good performance in many applications (Melin et al. 2014).

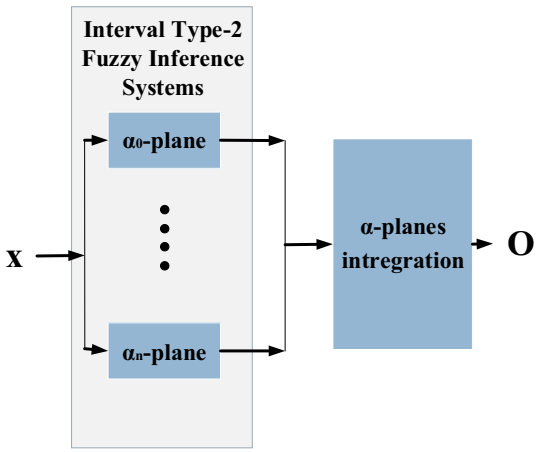

Fig. 3 GT2 FIS based on α-planes

2.2.2 Double Gaussian general type-2 membership function

In order to define the Double Gaussian GT2 Membership Function, it is necessary to model the FOU of the MF, remembering that the FOU is basically the first α-plane of the representation and is equivalent to an IT2 MF. Based on this, the Double Gaussian IT2 Membership Function (DGaussIT2MF) is proposed inspired in the GT2 membership functions presented in Mendel (2017). Figure 4 illustrates this function, and Eq. (7) describes the mathematical expression.

$$\mu_{\tilde{A}}(x) = \begin{cases} \bar{\mu}_t(x) = \begin{cases} \exp\left((-1/2)(x-m_1)^2/\sigma_1^2\right) x < ; m_1 \\ 1 \quad m_1 < x < m_2 \\ \exp\left((-1/2)(x-m_2)^2/\sigma_1^2\right) x > m_2 \end{cases} \\ \underline{\mu}_t(x) \begin{cases} \exp\left((-1/2)(x-m_1)^2/\sigma_2^2\right) x < \dfrac{(m_1+m_2)}{2} \\ \exp\left((-1/2)(x-m_2)^2/\sigma_2^2\right) x > \dfrac{(m_1+m_2)}{2} \end{cases} \end{cases} \tag{7}$$

where $\bar{\mu}_t(x)$ and $\underline{\mu}_t(x)$ are the upper and lower membership functions and are obtained by the evaluation of the $\text{DGaussIT2MF}(x, [\sigma_1, \sigma_2, m_1 m_2])$ function, and based on these membership functions as boundaries, Eq. (8) express the Double Gaussian General Type-2 Membership Function (DGaussGT2MF).

$$\begin{aligned} &\text{DGaussGT2MF}(x, [\sigma_1, \sigma_2, m_1 m_2]) \\ &= \left\{ \left((x,u), \text{trimf}\left(u, \left[\underline{\mu}_t(x), \frac{\bar{\mu}_t(x) + \underline{\mu}_t(x)}{2}, \bar{\mu}_t(x)\right]\right)\right) \middle| \right. \\ &\left. \forall u \in \left[\bar{\mu}_t(x), \underline{\mu}_t(x)\right] \right\} \end{aligned} \tag{8}$$

This General Type-2 Membership Function has a triangular function as secondary membership function. However, this function can be substituted in the future for

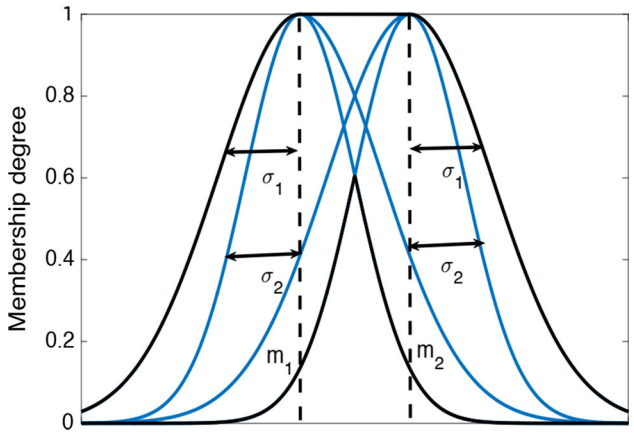

Fig. 4 Double Gaussian IT2 MF

different kinds of membership functions. The triangular function (trimf) is defined as follows by Eq. (9)

$$\text{trimf}(x,[a,b,c]) = \begin{cases} \dfrac{x-a}{b-a} & a<x<b \\ \dfrac{c-x}{c-b} & b<x<c \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

For a GT2 MF represented with α-planes, the number of α-planes depend on the discretization of u, and the approximation is better with the increase in the discretization level of u. On the other hand, the computational cost is proportional to u.

A graphical illustration of the GT2 Double Gaussian Membership Functions can be observed in Fig. 5.

2.3 Proposed approach to generate the FOU

Based on the concepts above presented, we propose the use of uniform sampling with replacement to generate subtraining sets and with this model the FOU of the General Type-2 Fuzzy Sets based on multiple embedded Type-1 Fuzzy Sets.

The T1 Membership Functions are generated based on the subtraining sets, the number of clusters is proportional to the number of membership functions of the system and is proportional to the number of inputs and the number of granules, and the steps to generate the FDS are presented in Fig. 6.

The first step consists on obtaining the centers of the Type-1 Gaussian Membership Functions, and these centers are obtained by the implementation of a clustering algorithm, in this case, the Fuzzy C-Means (FCM) algorithm. For example, consider the variable "years old" of the Mammographic dataset. The output data provided for the FCM algorithm are the centers of the clusters and the membership degree for every cluster, and Fig. 7 illustrates an example of this membership degrees, the data is the years old parameter of the mammographic dataset. The centers of the Gaussian membership functions in this paper

Fig. 6 Steps for FDS generation

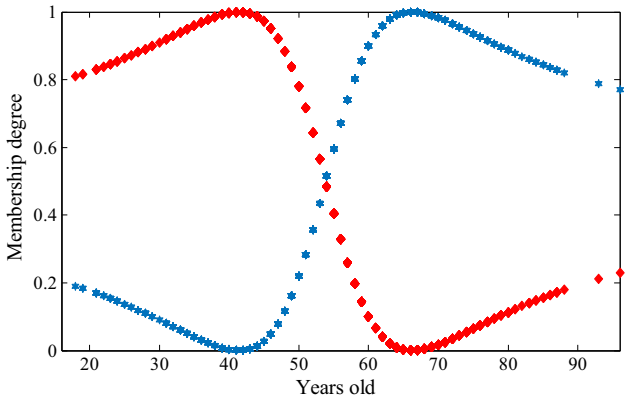

Fig. 7 FCM clusters membership functions for years old feature in mammographic dataset

are proposed to be the centers of the clusters provided by the FCM algorithm.

The second step consists on obtaining the Standard deviations; based on the membership degrees obtained with the FCM algorithm, and this is expressed in Eq. (10), and this equation is the result of a least square regression.

$$\sigma^n = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m^n)^4}{-2(\ln|\mu_i^n|)\sum_{i=1}^{d}(x_i - m^n)^2}} \tag{10}$$

where σ^n is the standard deviation of the nth membership function, m^n is the center of the nth membership function, x_i is the input data and μ_i^n is the membership degree of the ith data to the nth cluster. As we can observe, the σ^n depends on the m^n, for the proposed approach, the m^n is considered the center of the nth cluster obtained with the FCM algorithm.

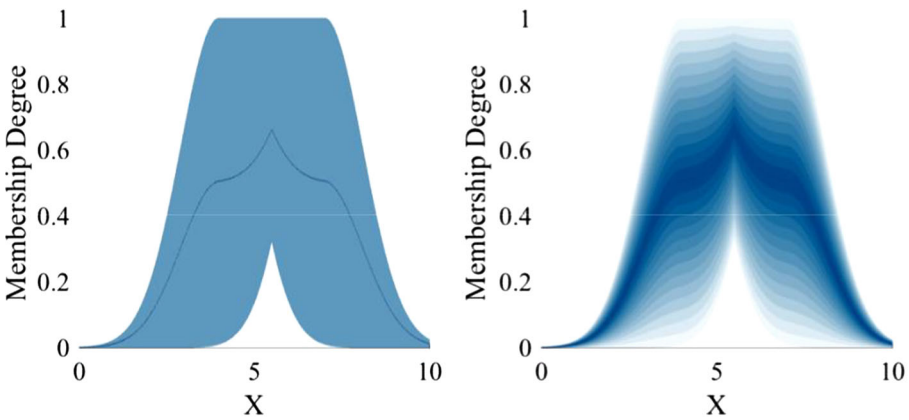

Fig. 5 **a** FOU of GaussG GT2MF **b** GaussG GT2MF

Based on the multiple Type-1 Gaussian MFs generated for every training subset, Eq. (11) describes the General Type-2 Double Gaussian MFs obtained for every cluster.

$$\bar{\bar{\mu}}_i(x,u) = \text{DGaussGT2MF}\left(x, \left[m^i_{\max}, m^i_{\min}, \sigma^i_{\max}, \sigma^i_{\min}\right]\right) \tag{11}$$

where $m^i_{\max}$ and $m^i_{\min}$ are the maximum and minimum m parameters of the ith cluster and $\sigma^i_{\max}$ and $\sigma^i_{\min}$ are the maximum and minimum standard deviations of the ith cluster.

2.4 Type-2 fuzzy inference systems for diagnosis

The architecture of the General Type-2 Fuzzy Classifier (GT2 FC) proposed in the present paper is illustrated in Fig. 8. This architecture is inspired by ANFIS the architecture proposed in Jang (1993) and widely used for complex problems.

The parameters of the input membership functions are estimated as was explained in the previous section. However, in this section, is defined how the output parameters of this architecture are obtained. The output of the architecture is defined in Eq. (12)

$$\tilde{\tilde{Z}} = \frac{\sum \alpha Z_\alpha}{\sum \alpha} \tag{12}$$

where Z represents the different outputs of every α-plane and the output is the α-planes aggregation. However, in order to reduce the computational cost, and considering that the secondary membership function is a triangular membership function, we decide to use only three α-planes and implement the equation proposed in Ontiveros et al. (2018b) that consists on a high-order α-planes integration based on Newton–Cotes integrators. Then, the output with this consideration is expressed in Eq. (13).

$$\tilde{\tilde{Z}} = \frac{\mathbf{Z}_1 + 2\mathbf{Z}_2 + \mathbf{Z}_3}{3} \tag{13}$$

where Z_1 has an $\alpha = 0^+$, in Z_2 $\alpha = 0.5$ and for Z_3 $\alpha = 1^-$.

On the other hand, the solution of the individual α-planes is realized based in the Wu–Mendel type reduction (Wu and Tan 2005), that is one of the fastest methods for this process. Equation (14) expresses the output of the IT2 FIS corresponding to the lth α-plane.

$$\begin{aligned} [\underline{Z}_l, \bar{Z}_l] &= \left[\sum_{i=1}^{N} \underline{\Phi}^l_i \vec{f}_i, \sum_{i=1}^{N} \overline{\Phi^l}_i \vec{f}_i\right] \\ Z_l &= \frac{(\underline{Z}_l + \bar{Z}_l)}{2} \end{aligned} \tag{14}$$

where $\underline{Z}_l$ and $\bar{Z}_l$ are the left and right output of the lth α-plane and Z^l is the output of the lth α-plane, $\underline{\Phi}^l_i$ and $\overline{\Phi^l}_i$ represents the normalized firing force of the ith rule of the lth α-plane, and finally $\vec{f}_i$ is the Sugeno polynomial of the ith rule. It is interesting to observe that this polynomial does not change for the different α-planes and is the same for the left and right outputs, and this is because in this paper we consider that the uncertainty is handled in the input membership functions and not in the consequent. Equation (15) expresses the Sugeno polynomial.

$$\vec{f}_i = a^i_0 + a^i_1 x_1 + \cdots + a^i_m x_m \tag{15}$$

where a^i_n represents the Sugeno coefficient of the ith rule and the mth input.

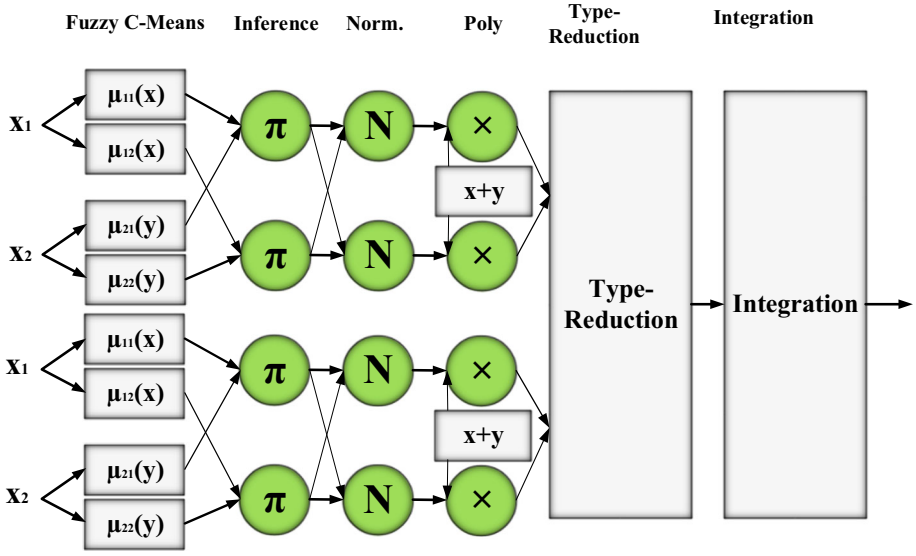

Fig. 8 Proposed architecture of GT2 FC

These Sugeno coefficients in the present paper are proposed to be obtained by minimizing Eq. (16).

$$e = \sum \left(T - \sum_{i=1}^{N} \left(\frac{\overline{\Phi^3}_i + \underline{\Phi}_i^3}{2} \right) \vec{f}_i \right)^2 \tag{16}$$

where T is the target $\overline{\Phi^3}_i$ and $\underline{\Phi}_i^3$ are the normalized firing force of the ith rule of the third α-plane that have an $\alpha = 1^-$, the process consists in a Least Square Error optimization.

3 Experimental results

This section introduces the results obtained by the experimentation with a set of benchmark medical diagnosis datasets, and the structure of this section is described as follows. First, the results by Hold-Out data separation with 60% for training and 40% for testing are presented, and the documented results are the average of thirty experiments for each dataset. On the other hand, it is presented a statistical comparison with respect another approach of GT2 Fuzzy Systems applied in diagnosis that is based on the principle of justifiable granularity presented by Sanchez et al. (2017).

3.1 Benchmark problems

The datasets selected for experimentation have been widely used for evaluating the performance of different kinds of diagnosis systems or classifiers, and a brief description is provided below in Table 1.

Before presenting the results obtained for the presented datasets, the membership functions obtained by uniform sampling with replacement for Type-1 Fuzzy Sets and the proposed approach of General Type-2 Fuzzy Sets are presented. Figures 9 and 10 illustrate the membership functions of the first three inputs (Features) of two of the first datasets presented, and for the Type-1 membership functions used for the estimation of the FOU.

3.2 Hold-out validation

In order to evaluate the performance of the proposed approach, we realize a Hold-Out validation. However, before doing the comparison with other fuzzy approaches we evaluate the performance of the proposed approach by increasing the number of clusters of the systems. The performance results have been reported with accuracy, sensitivity, and specificity. These metrics are illustrated in Fig. 11.

Tables 2, 3 and 4 summarize the results of the different performance measures, accuracy, sensitivity and specificity, respectively.

As can be noted, the accuracy decreases with the increasing of the number of clusters; this can be related to the architecture proposed where the number of rules is defined by the number of clusters and are very simple.

For the case of sensitivity, the behavior is a little chaotic, this performance measure can be affected for the data sampling, because can be obtained different measures for unbalanced samples.

By similar way, the specificity results are very chaotic and also can be related to the sample data, this is the reason to have standard deviation very large in comparison with the accuracy measure.

Figures 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22 illustrate the accuracy of every dataset with the different number of clusters. This kind of graphic illustrates better the point observed in Table 3, the objective is to observe the accuracy of the T2 Fuzzy Classifiers with different number of clustering and conclude over how many clusters are recommendable for this approach.

Table 1 Diagnosis datasets

Dataset name	Attributes	Abbreviation
Breast Cancer Wisconsin (original) dataset	9	BCW
Haberman's survival dataset	3	Haber
Fertility dataset	10	Fert
Indian liver dataset	9	Indian
Breast Cancer Wisconsin (diagnostic) dataset	32	BCWD
Pima Indians diabetes dataset	8	Pima
Statlog (heart) dataset	13	Heart
Breast Cancer Coimbra dataset	9	Coimbra
Mammographic mass dataset	5	MMass
Immunotherapy dataset	8	Inmu
Cryotherapy dataset	7	Cryo

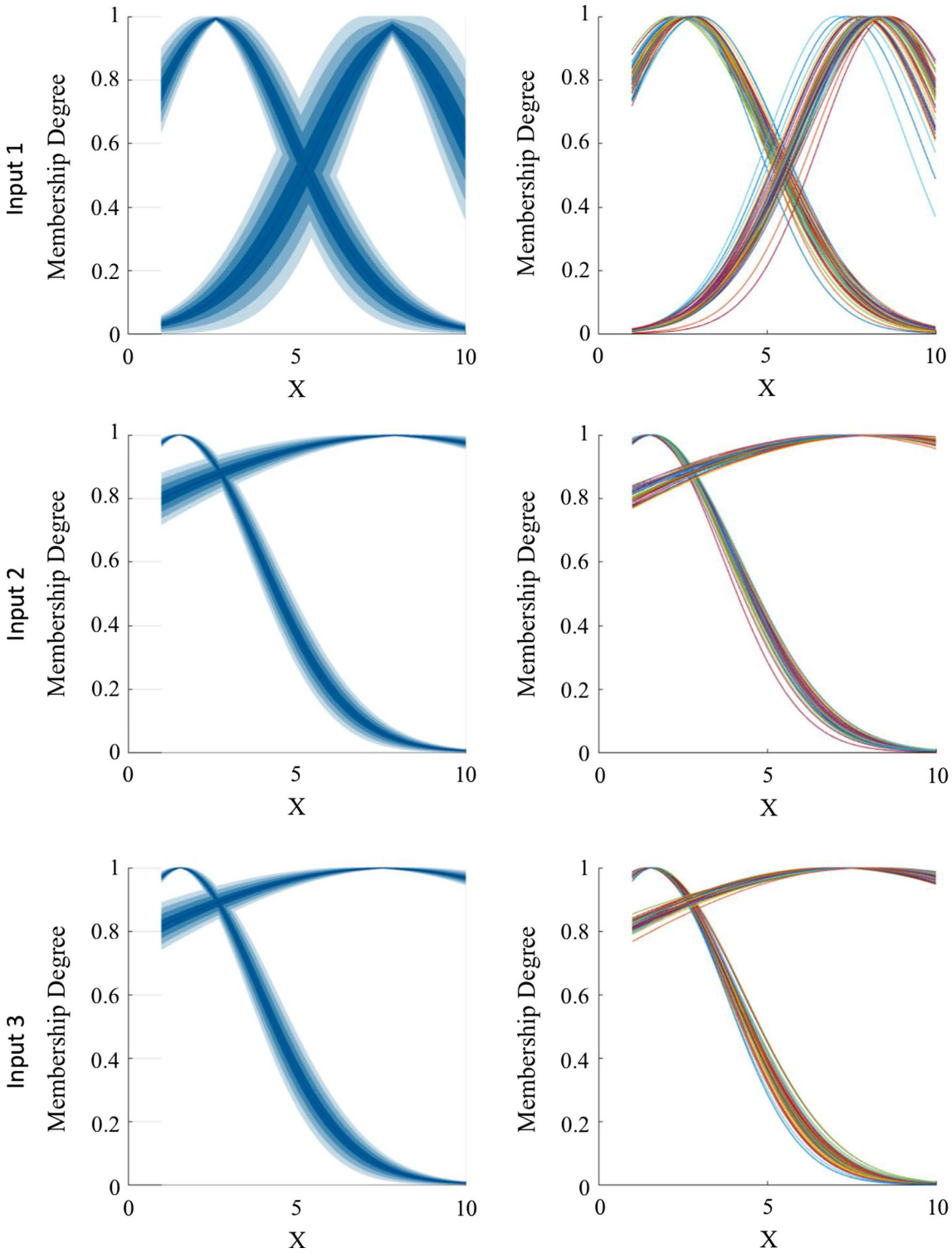

Fig. 9 Breast Cancer Wisconsin (original) dataset

In Fig. 12, we can recommend two clusters, for example, the performance decrease with the increasing of clusters number.

In Fig. 13, we also can recommend to use two clusters, for the Haberman's survival dataset.

For Fertility dataset, the best performance is obtained also for two clusters.

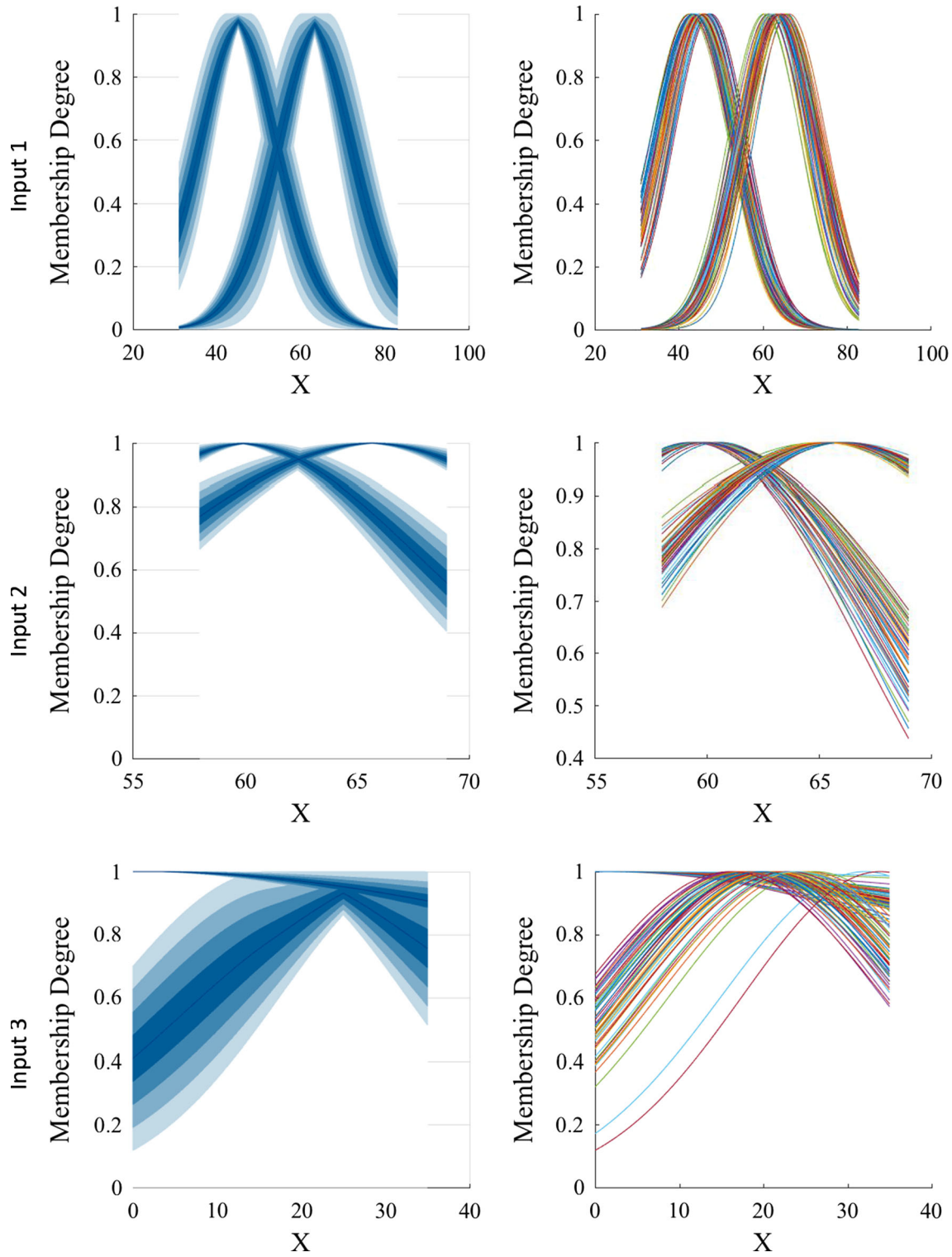

Fig. 10 Haberman's survival dataset

Also for Indian Liver dataset, the best performance is obtained for two clusters.

In Breast Cancer Wisconsin (Diagnostic) dataset, the best performance is obtained for four clusters.

For Pima dataset, the best performance is obtained for three clusters.

For Statlog dataset, the best performance is obtained for three clusters.

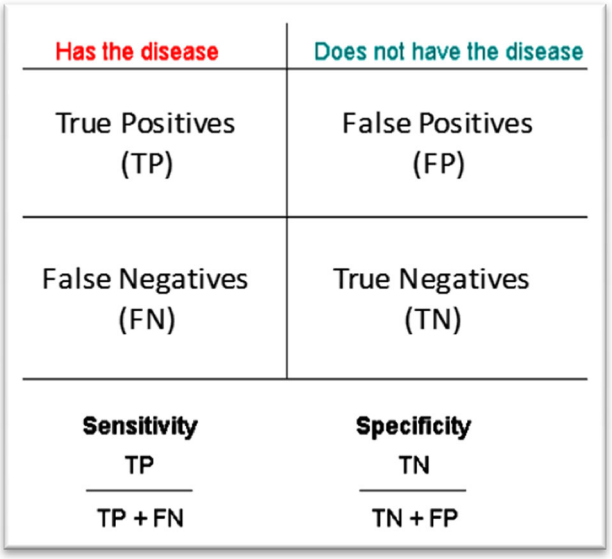

Fig. 11 Performance metrics

For Breast Cancer Coimbra, the best accuracy is obtained with two clusters.

Also for Mammographic Mass dataset, the best performance is obtained with two clusters.

For Immunotherapy dataset, we obtain a perfect performance with two clusters.

Finally, and similar than other cases, for Cryotherapy dataset, the best performance is obtained with two clusters.

As can be observed in the experimental results, the best results are obtained for only two clusters, and this can be related to the rules of the systems that are predefined as the first approach of ANFIS.

3.3 Statistical comparison with another approach of FOU generation

The statistical comparison was made between the proposed method and another approach for selecting the FOU of GT2 FC, which was presented by Sanchez et al. (2017) and is based on granular computing. Below is presented an statistical comparison based in *Z*-Test, this is because the results provided for the authors (Sanchez et al. 2017) are the mean of 30 experiments. Table 5 introduces the parameters of the statistical test.

Table 6 summarizes the results of the statistical test that was realized by a *Z*-test.

As can be observed, the proposed approach has sufficient evidence to be considered better than (Sanchez et al. 2017) in two of the four datasets that were compared. However, for the other two datasets we do not have sufficient evidence to demonstrate a superiority of one of the compared approaches. An explanation for the cases where we do not have enough evidence to show an improvement is because for these cases the architecture proposed for the reference provides a better uncertainty handling, and the fuzzy rules probably describe by a better way the information.

3.4 Cross-validation performances

In this section, the obtained performances for different values of cross-validation are presented and compared with respect to other fuzzy approaches introduced in the literature to solve diagnosis problems. Tables 7, 8 and 9 report the performances obtained for the proposed approach and other approaches of the literature for three, five and ten-folds cross-validation.

The not available results for the cited reference are expressed with a "−," and the reason of these missing values is because the cited papers are not focused in diagnosis problems, as they report results of fuzzy classifiers but we are interested only in the diagnosis datasets for the present paper.

4 Conclusions and future work

The generation of a single GT2 FC based on a set of embedded T1 fuzzy membership functions is interesting because we can consider every T1 membership function as a fuzzy observation and the GT2 FC as the model that aggregates the observations handling the uncertainty.

Regarding the comparison with respect to different fuzzy logic approaches applied in diagnosis problems we conclude that the proposed approach offers competitive performances considering that the proposed approach can be improved in the future with optimization methods, such as metaheuristic algorithms or another kind of algorithms, for example the optimization algorithms presented in Caraveo et al. (2016), Castillo et al. (2016), Olivas et al. (2017), Peraza et al. (2017).

In comparison with respect to the approach proposed in Sanchez et al. (2017), we have enough evidence to demonstrate that the proposed approach is better in two of the four datasets compared and is worst in one of the four, on the other hand, for the other dataset no one shows to be better than the other.

As future work, we have to test the proposed methodology with other kinds of applications, for example, fault diagnosis or time series. On the other hand, is interesting to test different architectures of the system, for example with some methods for selecting the rules of the system or different kinds of membership functions. Maybe some hybridization with other classification methods could be

Table 2 Accuracy results with different number of clusters

Clusters	BCW	Haberman	Fertility
2	96.7279 ± 0.7504	75.1533 ± 2.9331	85.1117 ± 4.7933
3	95.722 ± 1.3409	62.5433 ± 9.1319	76.6749 ± 10.1997
4	95.6527 ± 2.9654	72.9672 ± 4.4053	50.3722 ± 17.498
5	91.7563 ± 7.3005	68.1418 ± 8.8622	48.139 ± 17.5
6	92.0684 ± 7.5924	52.4927 ± 16.9073	48.139 ± 21.1515
7	77.7431 ± 10.7173	55.9851 ± 15.9399	50.5376 ± 19.5073
8	70.5978 ± 9.3902	44.8147 ± 19.6962	49.7932 ± 24.3862
9	67.2563 ± 8.1363	52.9725 ± 17.6949	49.0488 ± 25.7582
10	65.788 ± 6.5418	47.4274 ± 16.3915	NaN ± NaN

Clusters	Indian liver	BCWD	PIMA
2	70.9538 ± 3.1863	95.3247 ± 1.0217	75.7748 ± 2.0387
3	61.1374 ± 9.0914	94.3868 ± 1.2867	75.8486 ± 2.2536
4	48.6513 ± 10.8073	95.2963 ± 1.362	71.5159 ± 3.3146
5	56.8409 ± 11.3326	93.3352 ± 2.2311	68.6591 ± 3.3106
6	66.7964 ± 5.4811	89.3705 ± 5.0571	68.5009 ± 3.934
7	66.129 ± 7.5939	87.2673 ± 8.0873	63.7782 ± 6.4587
8	67.1997 ± 7.3237	80.7162 ± 10.8109	59.2347 ± 9.31
9	58.3843 ± 13.2377	77.6325 ± 11.1665	60.0253 ± 9.6178
10	63.1257 ± 11.9958	82.862 ± 8.1216	56.1459 ± 11.7987

Clusters	Heart	Coimbra	MMass
2	80.4341 ± 3.4545	69.3907 ± 6.8784	83.6847 ± 1.7144
3	81.5496 ± 2.3392	62.2222 ± 8.5683	61.4864 ± 9.1833
4	74.4649 ± 4.5936	57.9211 ± 8.7759	66.9085 ± 11.1064
5	61.3506 ± 10.0499	56.0573 ± 8.1288	58.6038 ± 5.5469
6	55.7431 ± 9.0949	53.1183 ± 8.5559	57.5841 ± 10.8843
7	54.8086 ± 9.4405	49.6057 ± 7.6419	54.8975 ± 9.0803
8	53.06 ± 9.5514	44.5878 ± 8.4315	55.8192 ± 10.8872
9	53.1203 ± 7.5313	52.9749 ± 9.6711	53.7504 ± 6.6734
10	55.0196 ± 7.5916	52.6882 ± 8.8739	51.2991 ± 7.9647

Clusters	Inmu	Cryo
2	100 ± 0	84.977 ± 4.7791
3	97.9724 ± 4.0536	74.8387 ± 11.7669
4	92.0737 ± 7.4764	54.47 ± 9.8681
5	87.2811 ± 9.2975	55.9447 ± 13.1506
6	60 ± 19.1094	53.8249 ± 12.5432
7	65.2535 ± 17.0967	55.576 ± 10.4286
8	53.9171 ± 21.8435	52.4424 ± 11.7467
9	54.8387 ± 17.2917	54.7465 ± 10.769
10	52.5346 ± 14.5999	48.2028 ± 9.7518

Table 3 Sensitivity results with different number of clusters

Clusters	BCW	Haberman	Fertility
2	96.7323 ± 1.0647	16.7326 ± 5.667	94.6436 ± 6.0065
3	97.0876 ± 1.0234	51.3794 ± 20.1085	84.5594 ± 13.0527
4	97.3438 ± 1.3688	14.4096 ± 16.0653	51.4878 ± 22.188
5	98.0124 ± 1.5732	16.2187 ± 21.6233	48.7721 ± 22.1029
6	96.261 ± 2.8774	55.0047 ± 39.1578	47.7372 ± 27.0147
7	96.8127 ± 3.9486	41.7275 ± 38.2495	50.9843 ± 25.4955
8	96.1152 ± 6.0354	59.9638 ± 41.858	50.2456 ± 31.4062
9	95.8948 ± 5.5129	41.9813 ± 36.9525	48.4309 ± 33.5734
10	96.7323 ± 1.0647	16.7326 ± 5.667	94.6436 ± 6.0065
Clusters	Indian liver	BCWD	PIMA
2	21.7902 ± 7.1487	93.5037 ± 2.0541	54.4482 ± 4.2487
3	59.333 ± 29.3498	87.2786 ± 3.5871	52.758 ± 6.4167
4	84.434 ± 17.0908	88.6935 ± 4.5419	29.1474 ± 11.7925
5	63.3228 ± 32.4358	93.5507 ± 5.4055	22.572 ± 22.4725
6	31.5907 ± 23.8276	96.4347 ± 2.3789	43.4699 ± 26.4661
7	20.2271 ± 19.3592	88.8419 ± 21.2207	59.8102 ± 24.4454
8	16.4162 ± 19.2009	52.3856 ± 32.0199	64.4169 ± 31.1766
9	28.3632 ± 27.8454	43.9508 ± 32.1674	67.2435 ± 27.7742
10	23.2659 ± 25.4	60.6908 ± 22.2418	66.394 ± 28.3052
Clusters	Heart	Coimbra	MMass
2	74.1086 ± 6.7338	55.7925 ± 13.094	81.2839 ± 2.8103
3	75.5776 ± 7.6232	78.6324 ± 18.4137	18.0943 ± 23.0191
4	58.7025 ± 8.0168	41.0095 ± 31.0406	35.2749 ± 30.7776
5	70.9022 ± 18.885	33.5321 ± 31.0938	18.7556 ± 12.9583
6	82.7613 ± 22.5636	39.2409 ± 32.7047	32.4758 ± 34.5062
7	79.0992 ± 20.4868	48.2062 ± 30.27	46.1433 ± 39.4982
8	56.0071 ± 31.523	49.4557 ± 24.7042	45.007 ± 39.7345
9	53.3443 ± 26.2356	50.9681 ± 21.2602	67.3965 ± 40.0304
10	64.0033 ± 28.8826	52.1163 ± 22.5397	70.717 ± 41.1871
Clusters	Inmu		Cryo
2	100 ± 0		82.0303 ± 8.6147
3	98.4321 ± 3.5007		70.8271 ± 22.5886
4	92.3864 ± 8.2556		53.3215 ± 24.3068
5	88.067 ± 10.287		53.3109 ± 19.7214
6	58.84 ± 22.9477		49.306 ± 26.4932
7	64.6351 ± 23.5999		42.5507 ± 24.1066
8	53.155 ± 29.6901		44.5292 ± 28.9886
9	53.8752 ± 23.7045		54.3408 ± 34.4659
10	55.3543 ± 21.698		41.4895 ± 30.8819

Table 4 Specificity results with different number of clusters

Clusters	BCW	Haberman	Fertility
2	96.6853 ± 1.4436	95.4882 ± 1.9025	11.9892 ± 18.3455
3	93.3013 ± 3.9414	67.1194 ± 18.691	16.0292 ± 22.0994
4	92.358 ± 8.7951	93.0018 ± 9.157	41.4708 ± 25.8425
5	79.4959 ± 23.755	85.7996 ± 18.9664	46.3876 ± 35.1984
6	84.1514 ± 24.5719	50.5677 ± 35.7919	51.0292 ± 37.9906
7	41.5361 ± 35.2751	61.7874 ± 34.7282	48.3641 ± 32.8018
8	21.169 ± 30.1845	39.2678 ± 42.5096	47.3733 ± 36.2437
9	13.2375 ± 24.4566	56.5864 ± 37.8211	61.7896 ± 38.4979
10	10.5391 ± 23.4323	45.8434 ± 35.8848	NaN ± NaN

Clusters	Indian liver	BCWD	PIMA
2	90.5618 ± 5.7225	96.4196 ± 1.3344	87.2717 ± 2.2369
3	61.9929 ± 23.3145	98.6192 ± 1.5559	87.9287 ± 4.1186
4	34.6476 ± 20.4951	99.1861 ± 1.165	93.7474 ± 3.5777
5	54.3175 ± 27.5221	93.2841 ± 4.8932	93.75 ± 9.5608
6	81.85 ± 15.9539	85.2818 ± 8.2849	82.1645 ± 15.095
7	84.3393 ± 16.9811	86.6481 ± 10.4293	66.0435 ± 20.2562
8	87.369 ± 16.8294	97.698 ± 4.4831	56.6773 ± 29.0959
9	70.8339 ± 28.9097	97.9147 ± 6.3665	55.9861 ± 27.338
10	79.5611 ± 25.9269	96.3217 ± 5.5532	50.8056 ± 30.1623

Clusters	Heart	Coimbra	MMass
2	85.5231 ± 4.2524	81.0812 ± 12.186	85.7434 ± 2.9063
3	85.9299 ± 4.8033	49.0157 ± 20.1465	98.4344 ± 3.8843
4	86.9598 ± 5.8732	72.0857 ± 24.6401	93.8375 ± 7.1357
5	53.5289 ± 24.5492	74.6366 ± 24.2804	92.862 ± 4.4807
6	32.714 ± 27.3767	63.9837 ± 27.4555	79.6974 ± 17.637
7	36.0721 ± 22.1854	50.7686 ± 25.1054	61.4254 ± 26.7711
8	51.093 ± 27.4965	41.7275 ± 23.0529	64.7054 ± 27.2057
9	53.6719 ± 25.1569	53.5383 ± 23.5886	42.6002 ± 35.3361
10	48.5687 ± 26.0238	53.5287 ± 23.2557	35.585 ± 38.7614

Clusters	Inmu	Cryo
2	100 ± 0	88.8238 ± 9.7558
3	96.5001 ± 8.172	79.8117 ± 17.1837
4	91.3159 ± 9.7497	56.8222 ± 30.0579
5	86.4437 ± 18.6743	58.8253 ± 21.7298
6	65.4506 ± 23.8011	58.9511 ± 26.7773
7	67.2606 ± 33.1747	70.0746 ± 25.6203
8	57.98 ± 32.8548	61.9066 ± 32.5619
9	61.461 ± 30.2879	52.984 ± 34.6821
10	40.2407 ± 27.0786	57.7868 ± 33.635

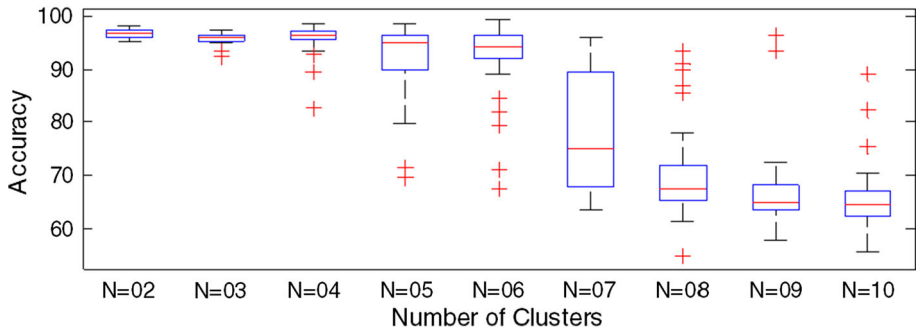

Fig. 12 Breast Cancer Wisconsin (original) dataset

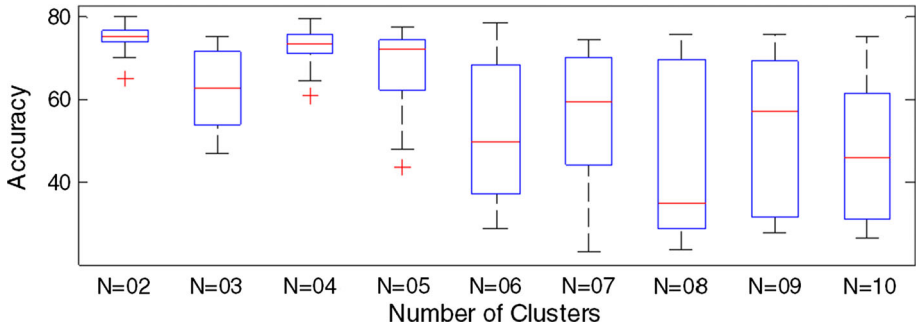

Fig. 13 Haberman's survival dataset

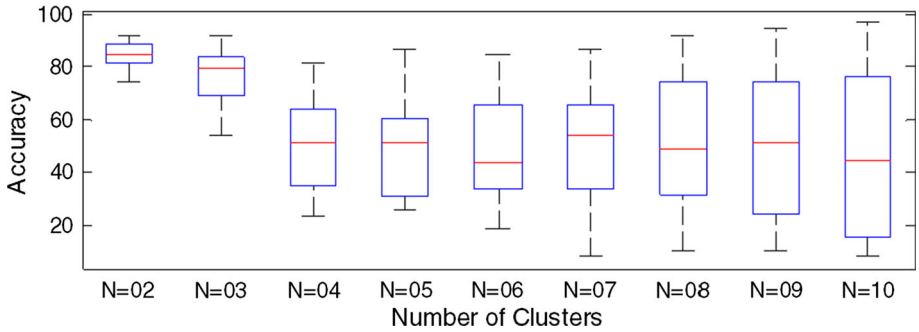

Fig. 14 Fertility dataset

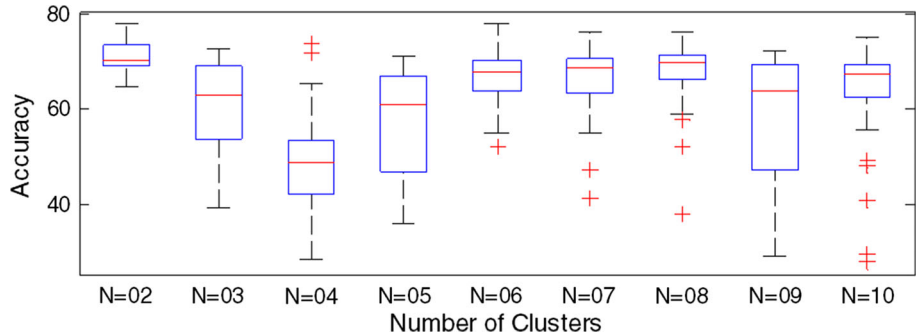

Fig. 15 Indian liver dataset

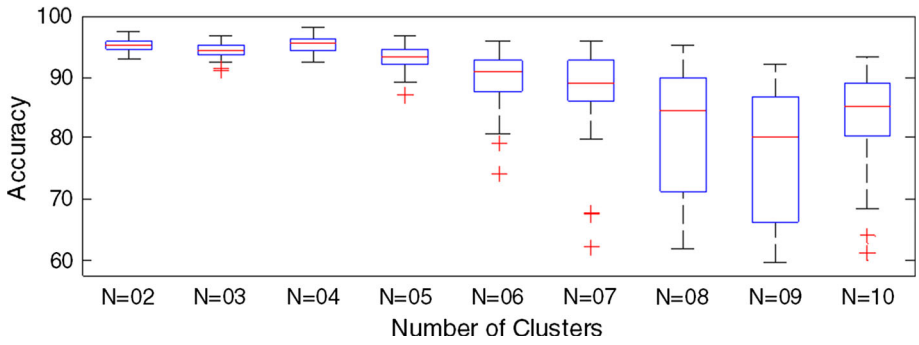

Fig. 16 Breast Cancer Wisconsin (Diagnostic) dataset

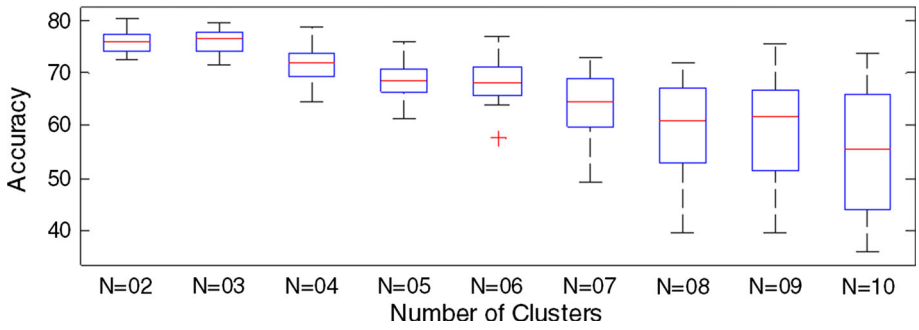

Fig. 17 Pima Indians diabetes dataset

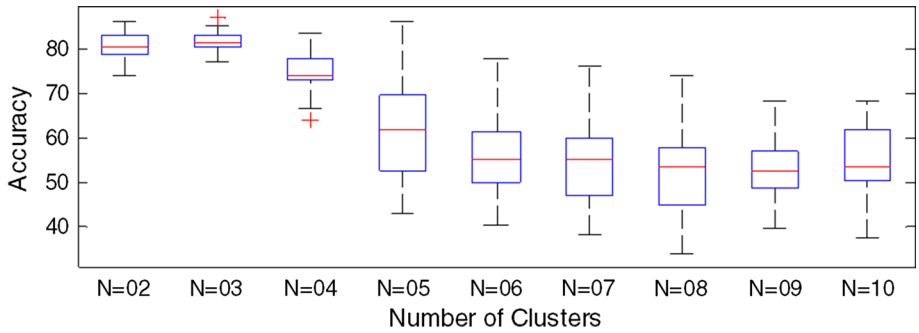

Fig. 18 Statlog (heart) dataset

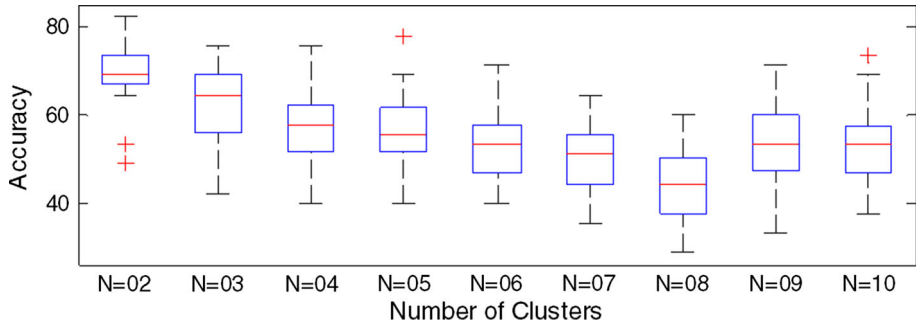

Fig. 19 Breast Cancer Coimbra dataset

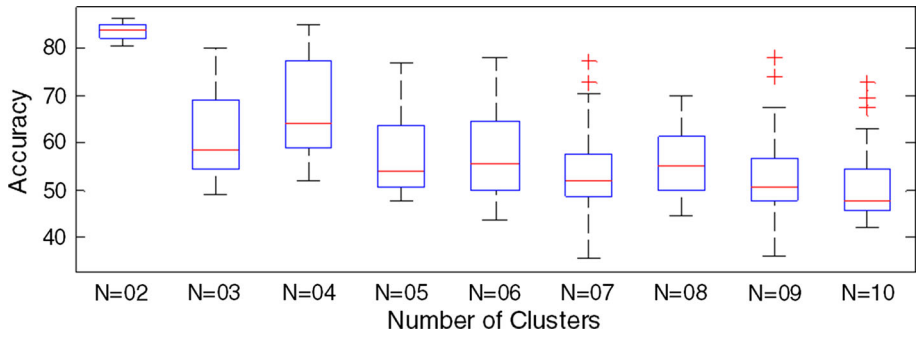

Fig. 20 Mammographic mass dataset

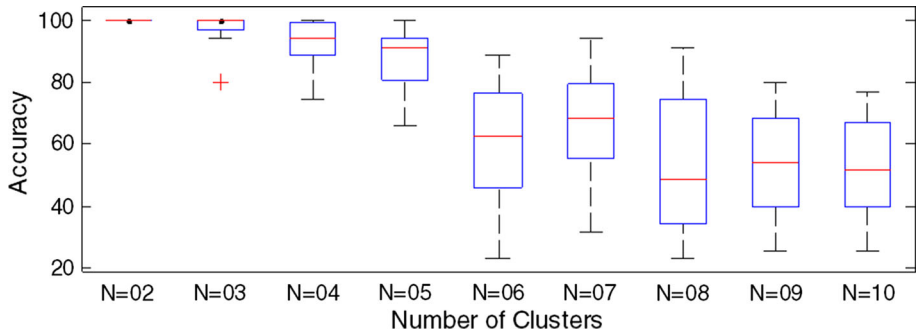

Fig. 21 Immunotherapy dataset

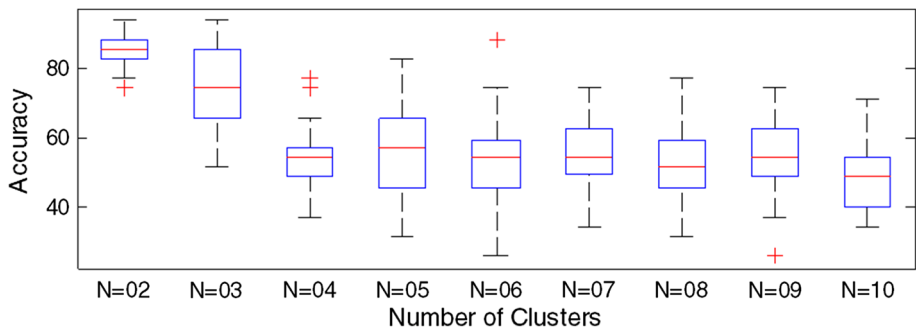

Fig. 22 Cryotherapy dataset

Table 5 Z-test parameters

Parameter	
Significance	95%
α	0.05
Ha	$\mu_1 > \mu_2$
Ho	$\mu_1 \leq \mu_2$
Critical value	1.645

Table 6 Statistical comparison versus Sanchez et al. (2017)

Dataset	Average	SD	Average	SD	Z
BCW	96.7279	0.7504	95.5861	1.1888	8.47184615
Haber	75.3932	2.7219	74.4116	2.1708	2.00790532
Fert	85.1117	4.7933	87.1333	2.6261	− 2.3482345
Indian	70.9538	3.1863	70.5602	3.6878	0.68777957

Table 7 Comparison using threefolds CV

Dataset	GT2 FC	Pota et al. (2018)
BCW	97.2844 ± 1.0169	97.28
Haber	75.6435 ± 4.3584	75.49
Fert	87.2727 ± 4.2376	–
Indian	72.0618 ± 1.8947	71.36
BCWD	95.6613 ± 1.4639	–
Pima	77.2156 ± 2.497	78.39
Heart	82.5842 ± 2.4416	85.15
Coimbra	72.8947 ± 9.3699	–
MMass	85.3454 ± 1.9473	–
Inmu	100 ± 0	–
Cryo	87.931 ± 6.1362	–

Table 8 Comparison using fivefolds CV

Dataset	GT2 FC	Froelich (2017)	Fu et al. (2019)
BCW	97.099 ± 1.1219	95.7	96.64
Haber	74.6694 ± 5.7857	–	74.20
Fert	87.6061 ± 6.8688	87.5	–
Indian	73.0812 ± 2.8825	–	–
BCWD	95.4895 ± 1.9165	–	–
Pima	76.6181 ± 2.2935	65.3	75.13
Heart	82.8971 ± 5.0846	–	–
Coimbra	70.4067 ± 8.0245	–	–
MMass	84.7116 ± 2.8841	77.6	–
Inmu	100 ± 0	–	–
Cryo	87.0967 ± 6.8631	–	–

Table 9 Comparison using tenfolds CV

Dataset	GT2 FC	Hu et al. (2018)	Hu et al. (2011)	Lahsasna and Seng (2017)
BCW	97.0079 ± 2.2117	97.53	95.3	96.18
Haber	76.5591 ± 7.7212	–	–	–
Fert	88.5304 ± 8.8349	–	–	–
Indian	72.3581 ± 5.9306	–	–	–
BCWD	95.7373 ± 2.0965	–	96.9	–
Pima	77.3769 ± 3.8047	80.00	75.4	75.44
Heart	84.6153 ± 6.512	88.15	80.4	–
Coimbra	75.953 ± 10.378	–	–	–
MMass	84.6577 ± 4.1268	84.58	–	–
Inmu	100 ± 0	–	–	–
Cryo	89.5161 ± 9.1837	–	–	–

made, for example statistical methods or methods such as Support Vector Machines.

Acknowledgements We would like to express our gratitude to CONACYT, Tijuana Institute of Technology, for the facilities and resources granted for the development of this research.

Compliance with ethical standards

Conflict of interest All the authors in the paper have no conflict of interest

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Acharya UR, Oh SL, Hagiwara Y et al (2018) Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput Biol Med 100:270–278

Afifi S, GholamHosseini H, Sinha R (2019) A system on chip for melanoma detection using FPGA-based SVM classifier. Microprocess Microsyst 65:57–68

Asl AAS, Zarandi MHF (2017) A type-2 fuzzy expert system for diagnosis of leukemia. In: Melin P, Castillo O, Kacprzyk J, Reformat M, Melek W (eds) Fuzzy logic in intelligent system design. Springer, Cham, pp 52–60

Babapour Mofrad R, Schoonenboom NSM, Tijms BM et al (2019) Decision tree supports the interpretation of CSF biomarkers in Alzheimer's disease. Alzheimers Dement Diagn Assess Dis Monit 11:1–9

Baraldi P, Razavi-Far R, Zio E (2011) Bagged ensemble of fuzzy C-means classifiers for nuclear transient identification. Ann Nucl Energy 38:1161–1171

Caraveo C, Valdez F, Castillo O (2016) Optimization of fuzzy controller design using a new bee colony algorithm with fuzzy dynamic parameter adaptation. Appl Soft Comput 43:131–142

Castillo O, Amador-Angulo L, Castro JR, Garcia-Valdez M (2016) A comparative study of type-1 fuzzy logic systems, interval type-2 fuzzy logic systems and generalized type-2 fuzzy logic systems in control problems. Inf Sci 354:257–274

Davari Dolatabadi A, Khadem SEZ, Asl BM (2017) Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM. Comput Methods Programs Biomed 138:117–126

Elyan E, Gaber MM (2016) A fine-grained Random Forests using class decomposition: an application to medical diagnosis. Neural Comput Appl 27:2279–2288

Erkaymaz O, Ozer M (2016) Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes. Chaos, Solitons Fractals 83:178–185

Fernández-Carrobles MM, Serrano I, Bueno G, Déniz O (2016) Bagging tree classifier and texture features for tumor identification in histological images. Procedia Comput Sci 90:99–106

Froelich W (2017) Towards improving the efficiency of the fuzzy cognitive map classifier. Neurocomputing 232:83–93

Fu C, Lu W, Pedrycz W, Yang J (2019) Fuzzy granular classification based on the principle of justifiable granularity. Knowl Based Syst 170:89–101

Hagras H (2008) Developing a type-2 FLC through embedded type-1 FLCs. In: 2008 IEEE international conference on fuzzy systems (IEEE world congress on computational intelligence). IEEE, Hong Kong, China, pp 148–155

Hothorn T, Lausen B (2005) Bundling classifiers by bagging trees. Comput Stat Data Anal 49:1068–1078

Hu Q, An S, Yu X, Yu D (2011) Robust fuzzy rough classifiers. Fuzzy Sets Syst 183:26–43

Hu X, Pedrycz W, Wang X (2018) Fuzzy classifiers with information granules in feature space and logic-based computing. Pattern Recognit 80:156–167

Jang J-SR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685

Lahsasna A, Seng WC (2017) An improved genetic-fuzzy system for classification and data analysis. Expert Syst Appl 83:49–62

Liang Q, Mendel JM (2000) Interval type-2 fuzzy logic systems: theory and design. IEEE Trans Fuzzy Syst 8:535–550

Liao Q, Ding Y, Jiang ZL et al (2018) Multi-task deep convolutional neural network for cancer diagnosis. Neurocomputing 348:66–73

Melin P, Gonzalez CI, Castro JR et al (2014) Edge-detection method for image processing based on generalized type-2 fuzzy logic. IEEE Trans Fuzzy Syst 22:1515–1525

Mendel JM (2017) Uncertain rule-based fuzzy systems. Springer, Cham

Mendel JM, John RIB (2002) Type-2 fuzzy sets made simple. IEEE Trans Fuzzy Syst 10:117–127

Mendel JM, John RI, Liu F (2006) Interval type-2 fuzzy logic systems made simple. IEEE Trans Fuzzy Syst 14:808–821

Mendel JM, Liu F, Zhai D (2009) α-plane representation for type-2 fuzzy sets: theory and applications. IEEE Trans Fuzzy Syst 17:1189–1207

Mendel JM, Rajati MR, Sussner P (2016) On clarifying some definitions and notations used for type-2 fuzzy sets as well as some recommended changes. Inf Sci 340–341:337–345

Olivas F, Valdez F, Castillo O et al (2017) Ant colony optimization with dynamic parameter adaptation based on interval type-2 fuzzy logic systems. Appl Soft Comput 53:74–87. https://doi.org/10.1016/j.asoc.2016.12.015

Ontiveros E, Melin P, Castillo O (2018a) Impact study of the footprint of uncertainty in control applications based on interval type-2 fuzzy logic controllers. In: Castillo O, Melin P, Kacprzyk J (eds) Fuzzy logic augmentation of neural and optimization algorithms: theoretical aspects and real applications. Springer, Cham, pp 181–197

Ontiveros E, Melin P, Castillo O (2018b) High order α-planes integration: a new approach to computational cost reduction of general type-2 fuzzy systems. Eng Appl Artif Intell 74:186–197

Peraza C, Valdez F, Melin P (2017) Optimization of intelligent controllers using a type-1 and interval type-2 fuzzy harmony search algorithm. Algorithms 10:82

Pota M, Esposito M, De Pietro G (2018) Likelihood-fuzzy analysis: from data, through statistics, to interpretable fuzzy classifiers. Int J Approx Reason 93:88–102

Qi X, Zhang L, Chen Y et al (2019) Automated diagnosis of breast ultrasonography images using deep neural networks. Med Image Anal 52:185–198

Rakhmetulayeva SB, Duisebekova KS, Mamyrbekov AM et al (2018) Application of classification algorithm based on SVM for determining the effectiveness of treatment of tuberculosis. Procedia Comput Sci 130:231–238

Sanchez MA, Castro JR, Ocegueda-Miramontes V, Cervantes L (2017) Hybrid learning for general type-2 TSK fuzzy logic systems. Algorithms 10:99

Saritas I (2012) Prediction of breast cancer using artificial neural networks. J Med Syst 36:2901–2907

Subasi A (2013) Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. Comput Biol Med 43:576–586

Vogado LHS, Veras RMS, FlavioHD Araujo et al (2018) Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. Eng Appl Artif Intell 72:415–422

Wang H, Zheng B, Yoon SW, Ko HS (2018) A support vector machine-based ensemble algorithm for breast cancer diagnosis. Eur J Oper Res 267:687–699

Wu D, Tan WW (2005) Computationally efficient type-reduction strategies for a type-2 fuzzy logic controller. In: The 14th IEEE international conference on fuzzy systems, 2005. FUZZ'05, pp 353–358