

# Performance evaluation of FMIG clustering using fuzzy validity indexes

Monia Tlili · Thouraya Ayadi · Tarek M. Hamdani · Adel M. Alimi

Published online: 21 October 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** The clustering of high-dimensional data presents a critical computational problem. Therefore, it is convenient to arrange the cluster centres on a grid with a small dimensional space that reduces computational cost and can be easily visualized. Moreover, in real applications there is no sharp boundary between classes, real datasets are naturally defined in a fuzzy context. Therefore, fuzzy clustering fits better for complex real datasets to determine the best distribution. Self-organizing map (SOM) technique is appropriate for clustering and vector quantization of high-dimensional data. In this paper we present a new fuzzy learning approach called FMIG (fuzzy multilevel interior growing self-organizing maps). The proposed algorithm is a fuzzy version of MIGSOM (multilevel interior growing self-organizing maps). The main contribution of FMIG is to define a fuzzy process of data mapping and to take into account the fuzzy aspect of high-dimensional real datasets. This new algorithm is able to auto-organize the map accordingly to the fuzzy training property of the nodes. In the second step, the introduced scheme for FMIG is clus-

tered by means of fuzzy C-means (FCM) to discover the interior class distribution of the learned data. The validation of FCM partitions is directed by applying six validity indexes. Superiority of the new method is demonstrated by comparing it with crisp MIGSOM, GSOM (growing SOM) and FKCN (fuzzy Kohonen clustering network) techniques. Thus, our new method shows improvement in term of quantization error, topology preservation and clustering ability.

**Keywords** Fuzzy training · Multilevel interior growing self-organizing maps · Quantization · Topology · Fuzzy validity indexes

## 1 Introduction

Real datasets are generally defined in a fuzzy context. In fact, there is no sharp boundary between classes so that fuzzy clustering is better suited for the data classification. Therefore, to classify such data is a very challenging task as in [Ayadi et al. \(2007, 2010, 2011\)](#), [Alimi \(2000, 2003\)](#), [Hamdani et al. \(2011a\)](#), [Dhahri and Alimi \(2005\)](#), [Kohonen \(1998\)](#), [Denaïa et al. \(2007\)](#) and [Chang and Liao \(2006\)](#). The clustering is an unsupervised classification mechanism where a set of entities are classified into a number of homogeneous clusters, with respect to a given similarity measure. Due to the fuzzy nature of many classification problems, a number of fuzzy clustering methods have been developed such as in [Ayadi et al. \(2007, 2010, 2011\)](#), [Alimi \(2003\)](#), [Hamdani et al. \(2008, 2011a,b\)](#), [El Malek et al. \(2002\)](#), [Pascual et al. \(2000\)](#), [Kohonen \(1998, 2001\)](#), [Denaïa et al. \(2007\)](#), [Fritzke \(1994\)](#), [Pettersson et al. \(2007\)](#), [Abraham and Nath \(2001\)](#), [Aguilera and Frenich \(2001\)](#) and [Chang and Liao \(2006\)](#). Fuzzy clustering can be applied as an unsupervised learning

---

Communicated by E. Lughofer.

---

M. Tlili (✉) · T. Ayadi · T. M. Hamdani · A. M. Alimi  
REGIM Lab: Research Groups on Intelligent Machines, National Engineering School of Sfax (ENIS), University of Sfax, BP 1173, Sfax 3038, Tunisia  
e-mail: monia.tlili@gmail.com; monia.tlili@ieee.org

T. Ayadi  
e-mail: thourya.ayadi@ieee.org

T. M. Hamdani  
e-mail: tarek.hamdani@ieee.org

A. M. Alimi  
e-mail: adel.alimi@ieee.org

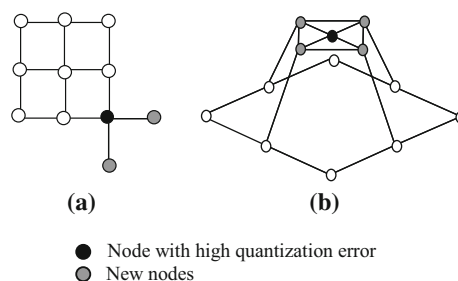
T. M. Hamdani  
College of Science and Arts at Al-Ula, Taibah University, Al-Madinah al-Munawwarah, Saudi Arabia

technique to classify patterns (Alimi 2000, 2003; Alimi et al. 2003).

In the unsupervised training, the network learns to form its own classes of the training data without external adjustment. Unsupervised systems of neural networks use a competitive learning process which is based on similarity comparisons in a continuous space. The resulting system associates similar inputs close to each other in one- or two-dimensional grid. Among the architectures suggested for unsupervised neural networks, the self-organizing map (SOM) (Kohonen 1982) is the most efficient in creating specific organized internal representations of input vectors and their abstractions. SOM is an unsupervised learning model used to visualise and explore linear and nonlinear relationships in high-dimensional data. SOM was first used in speech recognition by Kohonen in (1998, 2001). In literature researches, SOM is successful in visualisation of rich and complex phenomena such as pattern recognition tasks involving different criteria especially in clustering mechanism (Denaña et al. 2007; Fritzke 1994) and (Hsu and Halgarmuge 2001). Then, SOM algorithms have been implemented in various aspects of researches in classification Kohonen (1998), managing ecosystems Chang and Liao (2006) and GIS for addressing remediation Peterssona et al. (2007). SOM approaches were also applied in pattern recognition of big data (El Malek et al. 2002; Tlili et al. 2012; Amarasiri et al. 2004).

Generally, in their training process SOM approaches try to preserve the geometric form of the data called topology with the minimum error of quantization (Alimi et al. 2003; Hamdani et al. 2008; Pascual et al. 2000; Ayadi et al. 2012). However, the topology is not perfectly preserved with SOM algorithms due to their predefined structure. Therefore, dynamic variants of SOM are proposed to define the structure dynamically during the training process.

In this paper we present a new fuzzy dynamic learning method called FMIG (Fuzzy Multilevel Interior Growing Self-Organizing Maps). This algorithm is a hybrid approach in which both the MIGSOM technique (Multilevel Interior Growing Self-Organizing Maps) (Ayadi et al. 2010, 2012) and the FKCN algorithm (Fuzzy Kohonen Clustering Network) (Pascual et al. 2000) have been integrated. Our approach introduces a new form of fuzzy growing process and attempts to overcome some of the clustering difficulties by taking advantage of the best features of the multilevel structure given by MIGSOM and the fuzzy clustering model presented by FKCN. In fact, MIGSOM develops its growing process from the node that accumulates the highest Quantization Error (Fig. 1) (Amarasiri et al. 2004; Ayadi et al. 2012). On the other hand, FKCN presents a fuzzy model of clustering based on the Kohonen's Self-Organizing Feature Maps (SOFM) (Kohonen 1997) and the Fuzzy C-Means (FCM) (Bezdek 1981, 1987). Both of SOFM and FCM are successfully applied in classification.



**Fig. 1** New nodes insertion with MIGSOM growing process. **a** From the boundary and **b** from the interior

One of the advantages of our approach is to introduce a new Growing Threshold (GT) used to control the growth of the map. In fact, we remarked that the GT structure used in other algorithms such as MIGSOM becomes complicated and takes big values as the training process grows (Ayadi et al. 2010; Alimi 2000, 2003; Alimi et al. 2003). In addition, when the dataset is large the computing process becomes more complicated at next iterations and takes more time with the augmentation of GT values. For these reasons, we created a new growing threshold applied in our powerful methodology FMIG for big data clustering. The result is a new method of fuzzy dynamic SOM, which maintains the topology of patterns and considerably decreases the quantization error.

The rationale underlying the presented FMIG algorithm is that both advantages of fuzzy clustering and MIGSOM approaches are combined. It is worth to mention that FMIG serves two purposes: first, it can be used to produce the best fuzzy distribution of data by finding out the topology corresponding to the fuzzy form of inputs. Second, it gives the minimum value of error quantization providing the best cluster structure. In fact, FMIG is used to produce the best approximation to the real distribution of a given dataset. This algorithm finds a set of codebook vectors to code compactly all the input patterns. The topological FMIG maps introduce an additional relation between the codebook vectors. This topological relationship is a constraint which has been proved to produce robust results with the FMIG algorithm (Alimi 2000; Alimi et al. 2003; Hamdani et al. 2011a).

The classification of high-dimensional datasets presents a computational problem and it increases depending on the data size. Thus, to reduce the computational cost, two-steps have been proposed. First, the datasets are trained using a neural map to find the quantization prototypes. Then, the generated prototypes are clustered using a clustering algorithm such as FCM algorithm (Bezdek 1981, 1987). The choice of FCM is based on its best clustering compared to SOM and hierarchical algorithms (Mingoti and Lima 2006). However, the challenge of big data clustering is how to evaluate the performance of the resulting partition and finding the optimal number of clusters. Currently, some validity indexes are developed focusing on the geometric structure of the data subject

of unsupervised classification. The validity function Dunn (1974) is based on inter-cluster distance and the diameter of a hyperspheric cluster. Inspired from the Xie-Beni validity function, Kim and Ramakrishna developed in 2005 the Xie-Beni Index (XBI) based on a fuzzy compactness criterion (Mingoti and Lima 2006). In 2004, Tsekouras and Sarimveis proposed an index based on a fuzzy measure of separation called Separation Validity Index (SVI) (Wu and Yang 2005). Lung and Yang proposed in 2004 a fuzzy index evaluating noisy points in the dataset (Hu et al. 2004), called Partition Coefficient And Exponential Separation (PCAES) that was normalized by Tlili et al. (PCAESN) (2014). Wang and Lee also presented a method which calculates the fuzzy degree of overlapping clusters called Validity Overlap Separation (VOS) (Pal and Bezdek 1992). In 2009, M. Tlili et al. presented a new cluster validity function called Fuzzy Validity Index with Noise-Overlap Separation (FVINOS) for evaluating fuzzy and crisp clustering (Tlili et al. 2014).

Our analysis is based on cluster validity indexes FVINOS, PCAESN, VOS, SVI, Dunn and XBI. In fact, we have studied these indices in our previous research (Tlili et al. 2014) (Appendix).

In a first step, we have undertaken extensive performance comparisons with MIGSOM, GSOM and FKCN to establish the robustness of FMIG in mapping several synthetic as well as real-world datasets. Then, we have proceeded by the classification of the generated prototypes with the FCM algorithm. Finally, we have applied some validity indexes on the resulting partitions to evaluate FMIG clustering performance.

This paper is organized as following: Sect. 2 reviews the MIGSOM and FKCN functions. Section 3 introduces our new approach FMIG and the details of its steps. Section 4 evaluates FMIG method compared with MIGSOM, GSOM and FKCN algorithms in term of quantization and topologic errors with a variety of synthetic and real datasets. In Sect. 5, we focus our experimentations on the robustness of our approach in data clustering by validating the obtained partitions with the validity indexes described above. Finally, Sect. 6 summarises and concludes our remarks.

## 2 Review of crisp and fuzzy SOM algorithms

### 2.1 Multilevel interior growing self-organizing maps algorithm (MIGSOM)

Proposed by Ayadi et al. (2010) in 2012, this SOM heuristic presents a new method of growing process. The MIGSOM algorithm is based on three phases:

1. The initialization phase giving the values of the Growing Threshold (GT) to control the growth of the map and weight vectors.

2. The growing process is developed in the second phase that increases the map size by adding new nodes from the unit that presents the highest quantization error.
3. The smoothing phase is used to refine the resulting map of the previous phase.

MIGSOM is given by

1. Initialize the map grid with  $(2 \times 2)$  or  $(3 \times 3)$  nodes defined randomly and calculate the growth threshold (GT) as shown:

$$GT = -\ln(n) \times \ln(SF) \tag{1}$$

The GT is used to decide when to initiate new node growth. It specifies the spread of the feature map to generate. The GT value depends on the requirement for map growth. Thus if we require only a very abstract picture of the data, a large GT will result in a map with fewer nodes. Similarly, a smaller GT will result in the map spreading out more. For these reasons, we have integrated the size of the database in the new form NGT to represent faithfully and correctly Big Data.

SF is the Factor of Spread used to control the growth of the map taken in  $[0, 1]$ .  $n$  is the size of the given dataset.

2. In the growing phase, six steps are developed:
  - Step 1-Present the input patterns to the network.
  - Step 2-Calculate the weight of the best vectors (winners) that are close to the input data using Euclidean distance.
  - Step 3-Update weight vectors as following formula:

$$w_i(t + 1) = \frac{\sum_{j=1}^n h_{ci}(t) \times x_j}{\sum_{j=1}^n h_{ci}(t)} \tag{2}$$

with  $w_i(t)$  the weight vector of unit  $i$ ,  $x_j$  is the input pattern ( $j = 1 \dots n$ ).

$h_{ci}(t)$  represents the neighbourhood kernel centred on the BMU (Best Match Unit) given by

$$h_{ci}(t) = \exp\left(-\frac{Ud^2}{2\sigma^2(t)}\right) \times l(\sigma(t) - Ud), \tag{3}$$

where

$Ud$  defines the distance between neuron  $i$  and the winning node  $c$  on the map,

$\sigma(t)$  represents the degree to which excited neurons in the vicinity of the winning neuron.  $l(\sigma_i - Ud)$  is an evaluating function taking two values:

$$l(\sigma_i - Ud) = \begin{cases} 1 & \text{if } (\sigma_i \geq Ud), \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Step 4-Compute the error of each node. Calculate the node presenting the highest quantization error called node  $q$  with  $k$  units mapped in as follows:

$$E_{rr} = \sum_j^k \|x_j - w_q\| \quad (5)$$

Step 5-If ( $Err_q > GT$ ), new nodes will be generated from  $q$ .

Then initialize the new node weight vectors to match the neighbouring node weights.

Step 6-Repeat steps 1–5 until the number of iterations is reached.

### 3. Smoothing Phase:

Step 1-Fix the neighborhood radius to one.

Step 2-Training the map as the growing phase without adding new nodes.

## 2.2 The fuzzy Kohonen clustering network: FKCN

Proposed by Pascual et al. (2000), FKCN presents a fuzzy SOM heuristic. It consists of two layers: The input layer is formed by  $n$  nodes, with  $n$  the number of input patterns. The output layer is composed of  $c$  nodes, where  $c$  is the number of the output map vectors. For each input node, it is assigned a connection to all output units with a weight vector  $v_i$  ( $i = 1 \dots c$ ). Based on a pre-defined learning rate, the neurons in the output layer update their weights for an input vector  $x_k$  ( $k = 1 \dots n$ ). FKCN approach integrates the fuzzy degree of membership  $u_{ik,t}$  introduced by the fuzzy c-means algorithm (Xie and Beni 1991). The weight vector  $v_{i,t}$  at iteration  $t$  is given by

$$v_{i,t} = v_{i,t-1} + \alpha_{ik,t} \times (x_k - v_{i,t-1}) \quad (6)$$

The learning rate  $\alpha$  is defined as

$$\alpha_{ik,t} = (u_{ik,t})^{m_t} \quad (7)$$

$$m_t = m_0 - t \times \Delta m \quad (8)$$

$$\Delta m = \frac{(m_0 - 1)}{t_{\max}}, \quad (9)$$

where  $t_{\max}$  is the value of the iteration limit and  $m_0$  is a positive constant greater than 1.

The membership degree  $u_{ik,t}$  is given by

$$u_{ik,t} = \left[ \sum_{j=1}^c \left[ \frac{(x_k - v_{i,t})}{(x_k - v_j)} \right]^{2/(m_t-1)} \right]^{-1} \quad (10)$$

## 3 FMIG: fuzzy multilevel interior GSOMs

We present a new form of fuzzy self-organizing algorithm called FMIG (Fuzzy Multilevel Interior GSOMs). Our approach is based on the MIGSOM algorithm (Ayadi et al. 2010, 2012) which adapts its structure according to the data form. The fuzzy aspect of our method is inspired from the FKCN algorithm (Pascual et al. 2000).

As MIGSOM, the training process of FMIG is composed of three steps: initialization phase, growing phase and smoothing phase. The main idea of FMIG is to present a new parameter of Growing GT to ameliorate the control of map growth and improve the training process. In fact, the GT form used in MIGSOM algorithms becomes complicated as the growing increases. In addition, with big datasets, the computing process becomes more complex and takes more time with the augmentation of GT values.

Our new approach FMIG shows two interesting properties:

- It produces the best distribution of data by finding out the topology corresponding to the form of inputs.
- It gives the minimum value of error quantization providing the best cluster structure.

We define our FMIG algorithm as the following processes:

### 3.1 Initialisation

- Initialize the pre-defined values of the Spread Factor (SF) used to control the growth of the map taken in  $[0, 1]$ , and  $m_0$  the fuzzifier factor. The spread factor has to be specified to allow the data analyst to control the growth of the training process. SF is independent of the dimensionality of the given dataset. This factor is used by the system to calculate the Growing Threshold (GT). Such first step of the training process will act as a threshold value for initiating node generation. A high GT value will result in less distribution (spread) out map, and a low GT will produce a well-distributed map.
- Taking into account the dimension of the training dataset ( $n$ ) and the number of its characteristics ( $D$ ), the growing threshold gives an identical image of the specific data aspects. Hence, mapping such data produces the best distribution and topology. The formula of our New Growing Threshold (NGT) is defined as follows:

$$NGT = \frac{(-\ln(D) \times \ln(SF))}{n} \quad (11)$$

- Initialize the map grid with  $(2 \times 2)$  or  $(3 \times 3)$   $c_0$  nodes randomly and the weight vectors  $v_0 = (v_{1,0}, v_{2,0}, v_{c,0,0})$ .

Initialize randomly the membership matrix, with  $n$  the number of data inputs:

$$U = |u_{ij}|_{(i=1\dots c_0, j=1\dots n)} \tag{12}$$

Fix  $t_{\max}$  the limit of iterations.

### 3.2 Growing process

For  $t=1 \dots t_{\max}$ :

- Calculate the fuzzifier parameter  $m_t$  as shown in (8).
- Determine the membership degree  $u_{ik,t}$  for the iteration  $t$  using the Euclidean distance between the input pattern  $x_k$  and the weight vector  $v_i$ , with  $c_t$  being the number of weight vectors at the iteration  $t$  as defined in (10).
- Compute the learning rate  $\alpha_{ik,t}$  as defined in (7).
- The weight vectors are updated by the following formula Pascual et al. (2000):

$$v_{i,t} = v_{i,t-1} + \frac{\sum_{k=1}^n \alpha_{ik,t} (x_k - v_{i,t-1})}{\sum_{s=1}^n \alpha_{is,t}} \tag{13}$$

- Calculate the quantization error of each node  $i$  as Ayadi et al. (2010):

$$QE_i = \sum_j^{nbu} \|x_j - v_i\| \tag{14}$$

with  $nbu$ , the number of units mapped by the node  $i$ . The Quantization Error (QE) is a measure that completely disregards map topology and alignment. The quantization error is computed by determining the average distance of the sample vectors to the prototype vectors by which they are represented.

- Compute the error of each node. Calculate the node presenting the highest quantization error  $QE_{\max}$  called node  $q$  with  $k$  units mapped in as shown in (5).
- If  $QE_{\max} > NGT$  then generate new nodes from  $q$  as (Ayadi et al. 2010).

### 3.3 Smoothing process

- Initialise the limit of smoothing iterations  $t_{\max}$ .
- Train the map as the growing process without adding new nodes as (Ayadi et al. 2010).

## 4 FMIG map quality preservation

In order to present the performance of our technique, comparative studies are made in terms of quantization and topologic

errors. These comparisons take place between the proposed FMIG algorithm, MIGSOM, FKCN and an inspired version of GSOM methods (Denaña et al. 2007).

We first executed FMIG considering two forms of growing threshold NGT Eq. (11) and GT Eq. (1) to demonstrate the performance of the new structure of NGT.

### 4.1 Experimental settings

The experimentations are realized by executing FMIG, MIGSOM [...], FKCN and GSOM algorithms in the test context parameters of the Euclidean distance function, the fuzzy factor value  $m_0=2$  and the Spread factor  $SF=0.05$ . The limit of iterations in the growing process  $t_{\max}$  is taken at the values 30, 50 and 100. The smoothing process takes its maximum of iterations  $t_{s\max}$  at 10.

The comparison of the algorithms is made in term of quantization and topologic errors which are presented as follows:

- Topologic Error: the Final Topologic Error FTE is computed identically to Ayadi et al. (2010) as follows:

$$FTE = \frac{1}{c_t} \sum_{i=1}^{c_t} mu(v_i) \tag{15}$$

$mu(v_i)$  takes the values 1 if the first and the second Best Matching Units (BMUs) of  $v_i$  are adjacent, 0 otherwise.

- The Final Quantization Error: FQE is computed as the normalized average distance between each input data and its BMU.

$$FQE = \begin{cases} 1 & \text{if no data point matches the unit} \\ \frac{1}{c_t} \sum_{i=1}^{c_t} \left[ \frac{\frac{1}{nbu} \sum_{j=1}^{nbu} \|x_j - v_{BMU_t}\|}{norm(v_i)} \right] & \end{cases} \tag{16}$$

$v_{BMU_t}$  is the weight vector of the BMU at the iteration  $t$ .

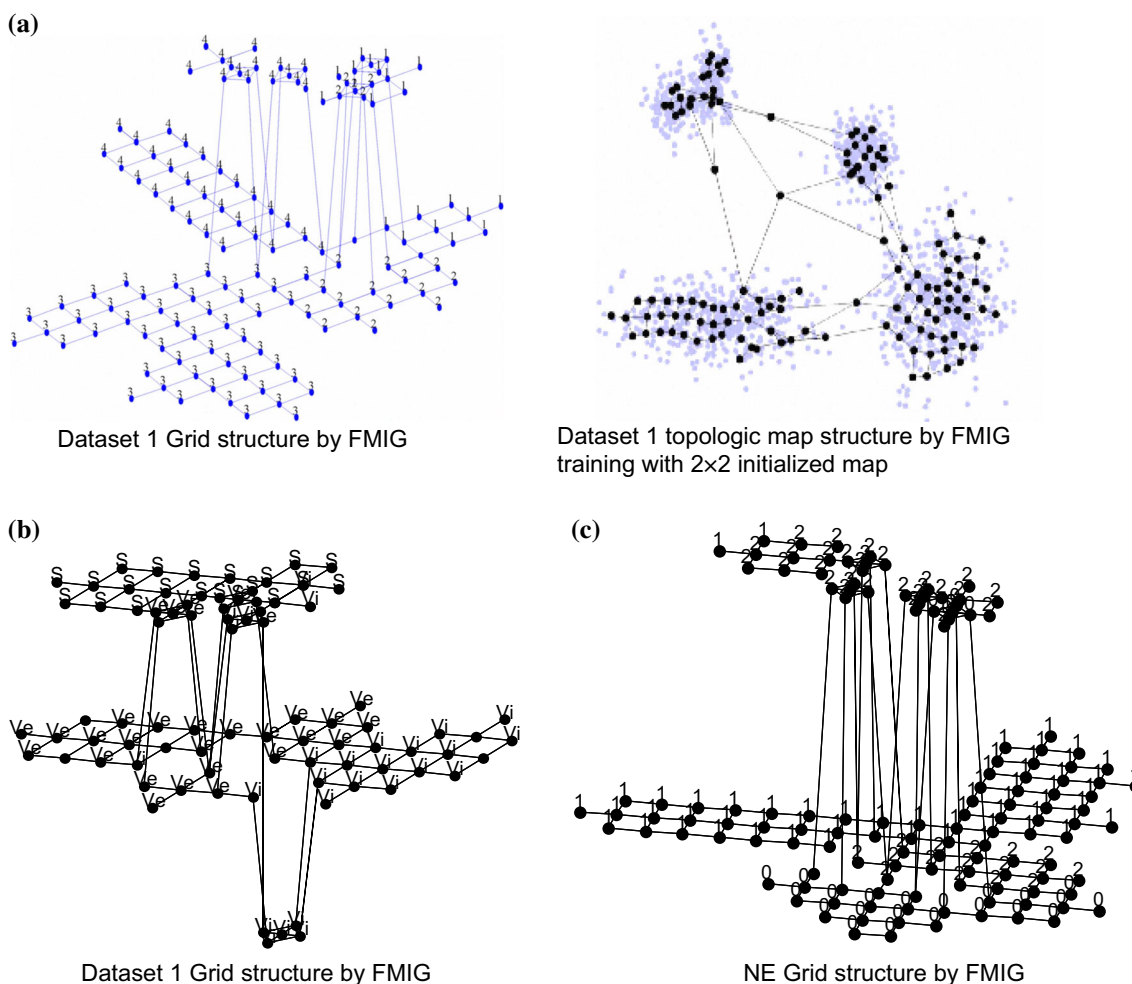
### 4.2 Datasets description

In order to test the effectiveness of the proposed algorithm FMIG, we performed a sensitivity analysis in synthetic datasets with different dimensions (Dataset1, Dataset2 and Dataset3) and real datasets (IRIS, Ionosphere, MNIST, DNA and NE) (Blake and Merz 1998). NE (Theodoridis 1996) consists of postal addresses of three metropolitan cities in US (New York, Philadelphia and Boston). All datasets are described in Table 1.

- Dataset 1 is composed of 1,600 points, 2-Dimension and distributed in 4 overlapping clusters.
- Dataset 2 presents 1,300 points, 2-Dimension and distributed in 13 noisy clusters.

**Table 1** Datasets description

Database	Nb data points ( $n$ )	Dimension	Nb clusters (nbc)	Characteristics
Dataset1	1,600	2	4	Overlapping
Dataset2	1,300	2	13	Noisy
Dataset3	700	2	3	High level of noise
IRIS	150	4	3	Overlapping
Ionosphere	351	34	2	Heterogeneous (noise+overlap)
MNIST	10,000	85	10	Overlapping
DNA	2,000	180	3	Heterogeneous (noise+overlap)
Letter	20,000	16	26	Overlapping
NE	50,000	2	3	Overlapping



**Fig. 2** Grid structure of FMIG algorithm for **a** dataset1, **b** IRIS and **c** NE datasets

4.3 Results and analyses

4.3.1 Multilevel map structure with FMIG

Figure 2 illustrates the labelled FMIG map structures for Dataset1, IRIS and NE. These maps are the result of the execution of FMIG with 50 iterations of growing phase. Thus, as we show FMIG is able to produce grid structure with multi-

levels oriented maps. In addition, the maps present homogeneous clusters.

4.3.2 New growing threshold NGT

In order to prove the amelioration introduced by our new structure of NGT (11) in quantization data and topology preservation, we performed some tests on different datasets.

**Table 2** Comparison of growing thresholds NGT and GT

Algorithm	FMIG (NGT)		FMIG (GT)	
	FQE	FTE	FQE	FTE
Dataset1	0.018	0.131	0.029	0.146
Dataset2	0.021	0.107	0.041	0.125
Dataset3	0.017	0.015	0.024	0.019
IRIS	0.132	0.111	0.209	0.127
Ionosphere	0.100	0.204	0.149	0.311
MNIST	0.497	0.146	0.554	0.261
DNA	0.277	0.110	0.307	0.288
Letter	0.325	0.202	0.433	0.309
NE	0.098	0.153	0.125	0.234

In fact, we executed our FMIG algorithm in two versions, taking into account, respectively, the GT and NGT formula to demonstrate the capacity of integrating the new parameter NGT. Table 2 shows results of FQE and FTE by executing FMIG 100 times, considering the two forms of GT and NGT.

Table 2 proves that the using of the NGT gives the minimum values of quantization error FQE. Therefore, FQE using the NGT form that we propose is clearly smaller than FQE resulting from the classic GT form for all datasets especially when we have a big data (Letter and NE).

Based on results presented in Table 2, the topologic error is improved by the new growing threshold. In fact, FTE with NGT gives minimum values for the tested databases compared to FTE using GT. The minimum values of the topologic error lead to a well-distributed map.

This experimental study confirms that in iterative processes, the number of iterations is important to evaluate the results because it has the probability to give the global solution of the problem. If the stop condition is chosen correctly, the iterative process will have the ability to execute the max of iterations and produce the optimal values of the algorithm.

For these reasons, we generated the new form of growing parameter in FMIG algorithm that corresponds to the stop condition of the growing process:

If  $Q_{E_{max}} < NGT$  then stops growing, else continue generating new nodes.

When we have a small NGT, the probability to continue the process is important. A high GT value will result in less spread out map. However, a low GT will produce a well-distributed map.

Our NGT is smaller than the interior GT that is why the iterative process will continue until the maximum of iterations leading to a well-spread map.

### 4.3.3 Comparison of FMIG with MIGSOM, GSOM and FKCN

In this study, we executed our proposed algorithm FMIG on different databases with different iterations values. The experimental results of comparison with MIGSOM, GSOM and FKCN are presented by Tables (3, 4, and 5), with, respectively, 100, 200 and 500 iterations.

Table 3 shows that from 100 iterations, results of quantization and topographic quality given by the two methods FMIG and MIGSOM are similar for the synthetic databases. In fact, the two methods give the same value of quantization error for Dataset3 which is a small dataset. Whereas, applied on real data that present different geometric aspects, FMIG generates the best values compared to MIGSOM with similar map size. Such improvement could be explained by the combination of MIGSOM advantages in multilevel maps and the fuzzy aspect of FKCN to develop FMIG that produces a well distributed map taking into account the characteristics of the dataset. Results given by GSOM and FKCN are far from the optimal values found by our method for synthetic and real datasets.

**Table 3** Experimental results with 100 iterations

Algorithm	FMIG		MIGSOM		GSOM		FKCN	
	FQE	FTE	FQE	FTE	FQE	FTE	FQE	FTE
Dataset1	0.018	0.131	0.026	0.132	0.092	0.138	0.029	0.167
Dataset2	0.021	0.107	0.030	0.111	0.139	0.209	0.078	0.305
Dataset3	0.017	0.015	0.018	0.026	0.020	0.111	0.020	0.101
IRIS	0.132	0.111	0.260	0.191	0.204	0.289	0.293	0.310
Ionosphere	0.100	0.204	0.128	0.220	0.163	0.251	0.161	0.255
MNIST	0.497	0.146	0.559	0.153	0.664	0.305	0.602	0.344
DNA	0.277	0.110	0.346	0.167	0.378	0.305	0.380	0.311
Letter	0.325	0.202	0.353	0.240	0.477	0.404	0.500	0.402
NE	0.098	0.153	0.104	0.157	0.220	0.179	0.215	0.210

**Table 4** Experimental results with 200 iterations

Algorithm	FMIG		MIGSOM		GSOM		FKCN	
	FQE	FTE	FQE	FTE	FQE	FTE	FQE	FTE
Dataset1	0.013	0.126	0.022	0.116	0.095	0.138	0.029	0.164
Dataset2	0.019	0.102	0.025	0.104	0.136	0.201	0.071	0.301
Dataset3	0.013	0.014	0.015	0.023	0.019	0.107	0.020	0.104
IRIS	0.110	0.106	0.251	0.177	0.148	0.249	0.244	0.306
Ionosphere	0.054	0.195	0.130	0.209	0.162	0.204	0.115	0.242
MNIST	0.349	0.121	0.508	0.110	0.646	0.307	0.599	0.335
DNA	0.205	0.092	0.303	0.144	0.370	0.300	0.311	0.304
Letter	0.304	0.202	0.309	0.208	0.477	0.399	0.479	0.403
NE	0.077	0.139	0.099	0.131	0.224	0.129	0.208	0.207

**Table 5** Experimental results with 500 iterations

Algorithm	FMIG		MIGSOM		GSOM		FKCN	
	FQE	FTE	FQE	FTE	FQE	FTE	FQE	FTE
Dataset1	0.012	0.124	0.021	0.115	0.092	0.136	0.027	0.161
Dataset2	0.017	0.101	0.025	0.104	0.136	0.201	0.071	0.300
Dataset3	0.011	0.013	0.014	0.021	0.017	0.100	0.019	0.100
IRIS	0.108	0.101	0.247	0.175	0.145	0.245	0.241	0.302
Ionosphere	0.053	0.199	0.128	0.207	0.160	0.202	0.113	0.240
MNIST	0.347	0.122	0.505	0.109	0.641	0.301	0.597	0.331
DNA	0.202	0.095	0.301	0.144	0.370	0.300	0.311	0.304
Letter	0.301	0.200	0.302	0.206	0.471	0.398	0.472	0.400
NE	0.075	0.134	0.095	0.129	0.220	0.127	0.206	0.205

The execution of the algorithms 200 and 500 iterations proves that FMIG gives more precision in quantization and topologic errors than MIGSOM, GSOM and FKCN (Tables 4, 5). Our fuzzy algorithm is able to generate the best distribution which produces the minimum values of FTE, especially with big real datasets. With the new structure of NGT taking into account the data size and its characteristics, in addition to the update modifications applied for the predefined parameters, the values of FQE and FTE give their minimum to improve the quantization factor of the map and preserve the topographic aspect of datasets.

To analyse each database separately shown by Tables 4 and 5, we discuss results given by FMIG compared to those generated by MIGSOM, GSOM and FKCN. Dataset1 preserves its distribution significantly with the minimum values of quantization error and topologic error by the map structure given by FMIG. It is the same case with datasets 2 and 3.

The real databases are more preserved by our new algorithm in terms of quantization and topologic errors. In fact, FMIG takes its minimum values with 100 and 200 iterations compared to MIGSOM. We note that GSOM and FKCN become inefficient to train big data and attain their limi-

tation for 200 and 500 iterations leading to unacceptable errors.

As an example, MNIST is considered a big data with a single aspect of data distribution that is why results given by our FMIG are more explicit and fiddle to the reality.

As we shown in Tables 4 and 5, with the evolution of growing iterations, the difference between crisp and fuzzy algorithms is significant. Therefore, 200 and 500 iterations of FMIG growing generate the best maps in parameters FQE and FTE compared with crisp MIGSOM and GSOM. These optimal results are returned due to the new fuzzy form of the weight vectors (13) taking into account the fuzzy membership degree (10). It is worth mentioning that GSOM and FKCN fail to map databases quite there preserving the topology criterion. Furthermore, these algorithms take an important execution time with big data to achieve their learning phase. Considering the GSOM and FKCN training, the resulting maps present important errors compared to FMIG. It could be explained by the drawback of GSOM (Ayadi et al. 2007; Amarasiri et al. 2004) and FKCN to learn big datasets. Therefore, it is clear that the results shown in the tables above are influenced by the number of iterations in



**Table 6** Clustering evaluation with validity indexes

Database	Validity index	FMIG nb. cluster	MIGSOM nb. cluster	GSOM nb. cluster	FKCN nb. cluster
Dataset1 $C = 4$	FVINOS	4	4	4	4
	PCAESN	4	4	4	4
	SVI	4	5	6	5
	VOS	4	4	4	4
	Dunn	4	4	5	6
	XBI	4	5	6	6
Dataset2 $C = 13$	FVINOS	13	13	13	13
	PCAESN	13	13	13	13
	SVI	13	14	15	15
	VOS	13	13	14	14
	Dunn	13	12	11	12
	XBI	13	10	10	10
IRIS $C = 3$	FVINOS	3	3	3	3
	PCAESN	3	2	2	2
	SVI	3	4	4	4
	VOS	3	3	3	3
	Dunn	3	2	4	4
	XBI	4	4	5	5
MNIST $C = 10$	FVINOS	10	10	10	10
	PCAESN	10	9	8	9
	SVI	11	11	12	11
	VOS	10	11	9	9
	Dunn	10	12	9	11
	XBI	11	11	13	12
DNA $C = 3$	FVINOS	3	3	3	3
	PCAESN	3	5	6	6
	SVI	3	7	5	7
	VOS	3	3	3	3
	Dunn	3	4	4	5
	XBI	3	6	5	7
Letter $C = 26$	FVINOS	26	26	26	26
	PCAESN	26	26	25	23
	SVI	28	28	30	28
	VOS	26	26	31	32
	Dunn	27	28	28	30
	XBI	29	30	29	32
NE $C = 3$	FVINOS	3	3	3	3
	PCAESN	3	4	5	6
	SVI	3	4	7	9
	VOS	3	3	4	5
	Dunn	4	4	6	5
	XBI	5	6	8	8

the growing phase of FMIG, MIGSOM, GSOM and FKCN algorithms.

## 5 Evaluation of FMIG clustering using validity indexes

In this study, we applied three stages of the clustering process to evaluate our method FMIG performance in a comparison with MIGSOM, GSOM and FKCN. We first trained synthetic and real datasets I with the four algorithms for 100 iterations of growing. The resulting prototypes were then clustered by FCM algorithm. Finally, we applied five validity indexes: FVINOS (Tlili et al. 2014), PCAESN (Hu et al. 2004), SVI (Wu and Yang 2005), VOS (Pal and Bezdek 1992), Dunn (Dunn 1974) and XBI (Mingoti and Lima 2006). Table 6 presents the clustering evaluation given by FMIG compared with other algorithms.

Results given by Table 6 prove that FCM partitions obtained by FMIG neural maps are optimal and produce the best number of clusters which is evaluated by the validity indexes. In fact, with different databases presenting multiple geometric aspects described by Table 1, FVINOS detects the optimal number of clusters given by FMIG training as well as in MIGSOM, GSOM and FKCN training. In addition, the majority of indexes give their best results with FMIG prototypes that could be explained by the ability of FMIG to produce well-distributed maps preserving the quality of data characteristics. Moreover, with noisy Dataset2, FMIG maps generate the best partition; all the indexes detect outliers and give good estimate of 13 clusters. MIGSOM algorithm produces the best partition with FVINOS, PCAESN and VOS, but the other indices are unable to detect the optimum number of clusters. GSOM and FKCN methods as presented in Table 6 return the best partition only with FVINOS and PCAESN because these indexes are able to detect noise aspect (Tlili et al. 2014).

Overlapping datasets such as Dataset1, IRIS and MNIST are well clustered by FMIG maps. As shown in Table 6, all the indexes discover the optimum number of clusters with Dataset1. For IRIS, just XBI is unable to give good estimate with FMIG; the other indices produce the best partition. MNIST is considered a big data and well trained with FMIG that gives 10 clusters in the resulting partition. FVINOS, VOS, Dunn and PCAESN indices are able to detect the overlap aspect (Tlili et al. 2014) of MNIST. MIGSOM, GSOM and FKCN produce the best partition just with FVINOS and fail in detecting the optimum number of clusters in the other cases.

Heterogeneous DNA is well clustered by FMIG partitioning that produces the optimum number of clusters evaluated by each of the indices. MIGSOM, GSOM and FKCN give good estimate with FVINOS and VOS, but the results of the other indexes fail to detect the optimum result.

Huge databases such as Letter and NE clustering prove that FMIG is able to produce the best distribution of prototypes clustered by FCM. In fact, with FMIG, FVINOS, PCAESN, VOS and SVI indexes detect the best partition. MIGSOM, GSOM and FKCN give the optimum number of clusters with FVINOS and fail to produce good results with the other indexes because they are limited in mapping huge data.

## 6 Conclusion

In this paper, we have presented a new form of fuzzy dynamic self-organizing map algorithm called FMIG. This new method improves the topology and quantization errors and generates its structure according the data form. The fuzzy aspect of our approach adapts its growing process to real problems. In this study, we presented a new parameter of Growing Threshold (GT) to ameliorate the control of the map growth and improve the training process. In fact, we proved that error values of quantization and topology are minimized by applying our new Growing Threshold (NGT). FMIG realizes the purposes of producing the best distribution of data by finding out the topology corresponding to the form of inputs. Also, FMIG gives the minimum value of quantization error providing the best cluster structure. Furthermore, FMIG is able to generate the best representations of big data.

Future research should focus on improving the performance of FMIG algorithm over high-dimensional datasets by incorporating some evolutionary techniques and providing a new form of Fuzzy evolutionary SOM.

**Acknowledgments** The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), REGIM-Lab (LR11ES48) Tunisia, under the ARUB program.

## 7 Appendix

### 7.1 Description of the used validity indexes

Let  $X = \{x_1, x_2, \dots, x_n\}$  be a dataset in an  $m$ -dimensional Euclidean space  $R^m$  with its ordinary Euclidean norm  $\|\cdot\|$  and let  $C$ , be the matrix of cluster centres with  $c_i$ , the centre of the cluster  $C_i$  ( $1 \leq i \leq nbc$ ). In case of fuzzy clustering, a pattern (point) may belong to all the clusters with a certain fuzzy membership grade. Consider the matrix  $U = |u_{ik}|$  where  $u_{ik}$  is the value of membership of the point  $k$  in cluster  $i$ , and  $u_{ik}$  is in the interval  $[0,1]$  ( $1 \leq k \leq n$ ).

We present below the validity indices studied in this work depending on the geometric aspects of clusters (compactness,

separation) and the data properties (noise, overlap) evaluated by each index.

### 7.1.1 SVI index

SVI (Separation Validity Index) (Wu and Yang 2005) is given by

$$SVI = \frac{\pi}{S} \tag{17}$$

$\pi$  is a global compactness of the partition, it makes the sum of  $\pi_i, 1 \leq i \leq nbc$ .

The compactness of the cluster  $c^i$  is defined by

$$\pi_i = \frac{\sigma_i}{n_i}, 1 \leq i \leq nbc \tag{18}$$

The variation  $\sigma_i$  and the cardinality  $n_i$  of the cluster  $c_i$  are given, respectively, as

$$\sigma_i = \sum_{k=1}^n (u_{ik})^m \|x_k - v_i\|^2, n_i = \sum_{k=1}^n u_{ik}, 1 \leq i \leq nbc \tag{19}$$

with  $u_{ik}$ , the membership degree of the vector  $x_k$  to the cluster  $c_i$  and  $v_i$  the centre of  $c_i$ .

$S$  is the global separation between the  $nbc$  clusters defined as

$$S = \sum_{i=1}^{nbc+1} \sum_{\substack{j=1 \\ j \neq i}}^{nbc+1} (dev_{ij})^2 \tag{20}$$

$dev_{ij}$  is the deviation between the centres of  $c_i$  and  $c_j$ .

$$dev_{ij} = (\mu_{ij})^{(2+\omega)/2\omega} \|z_j - z_i\| \tag{21}$$

$[z_1, z_2, \dots, z_{nbc}, z_{nbc+1}] = [v_1, v_2, \dots, v_{nbc}, \bar{x}]^T$ , where  $\bar{x}$  is the mean of  $X$ .

$$\bar{x} = \sum_{k=1}^n x_k/n \tag{22}$$

$\mu_{ij}$  is the degree of membership of  $z_j$  to the centre  $z_i$ .

$$\mu_{ij} = \frac{1}{\sum_{\substack{l=1 \\ l \neq j}}^{nbc+1} \left( \frac{\|z_j - z_l\|}{\|z_j - z_i\|} \right)^\omega}, 1 \leq i \leq nbc + 1; 1 \leq j \leq nbc + 1, j \neq i \tag{23}$$

### 7.1.2 XBI index

This index is normalized and gives the best number of clusters at its minimum value (Mingoti and Lima 2006) as shown:

$$XBI(nbc) = \frac{\max_{k=1, \dots, nbc} \left\{ \sum_{j=1}^n \frac{u_{kj}^2 \|x_j - c_k\|^2}{n_k} \right\} + \max \text{Diff}(nbc)}{\min_{i, j \neq i} \|c_i - c_j\|^2} \tag{24}$$

where

$\max_{k=1, \dots, nbc} \left\{ \sum_{j=1}^n \frac{u_{kj}^2 \|x_j - c_k\|^2}{n_k} \right\}$  gives the max fuzzy distance between points  $x_j$  with membership degree  $u_{kj}$  to the center  $c_k$ .

This index introduced the max Diff(nbc) factor in order to compare the  $nbc$ -partitions obtained by increasing  $nbc$ . Thus, we can find out the best partition.

$$\max \text{Diff}(nbc) = \max_{nbc} \max_{\dots, nbc} \text{diff}_{dw} \tag{25}$$

$$\text{diff}_{dw} = dw(nbc) - dw(nbc + 1) \tag{26}$$

with  $dw$  intra-cluster distance measured for different values of  $nbc$ .

### 7.1.3 PCAESN index

Called Partition Coefficient And Exponential Separation (PCAES) (Hu et al. 2004), this validity index takes into account two factors with a normalized partition coefficient and an exponential separation measure to validate separately each fuzzy cluster  $i$ . PCAES is then defined as:

$$PCAES(nbc) = \sum_{i=1}^{nbc} PCAES_i \tag{27}$$

The validity index of cluster  $i$  is measured by:

$$PCAES_i = \sum_{j=1}^n \mu_{ij}^2 / \mu_M - \exp(-\min\{\|a_i - a_k\|^2\} / \beta_T) \tag{28}$$

where  $\sum_{j=1}^n \mu_{ij}^2 / \mu_M$ : the compactness of the cluster  $i$  compared to the most compact cluster with its value of compactness  $\mu_M$  measured by:

$$\mu_M = \min_{1 \leq i \leq nbc} \left\{ \sum_{j=1}^n \mu_{ij}^2 \right\} \tag{29}$$

$\mu_{ij}$  is the membership degree of vector  $j$  ( $1 \leq j \leq n$ ) to cluster  $i$  ( $1 \leq k \leq nbc$ ),

$\exp(-\min\{\|a_i - a_k\|^2\}/\beta_T)$  is the separation measure relative to the total average separation  $\beta_T$  of the nbc clusters given by:

$$\beta_T = \frac{\sum_{l=1}^{nbc} \|a_l - \bar{a}\|^2}{nbc} \tag{30}$$

$a_i$  is the center of cluster  $i$ , and  $\bar{a} = \sum_{i=1}^{nbc} a_i/nbc$ , the average of center vectors  $a_i$ .

Normalisation of PCAES index:

Basically, each validity index measure PCAES $_i$  ( $i = 1 \dots nbc$ ) is obtained by a subtraction between the compactness and separation values which are defined in the interval [0 1]. Consequently, we present PCAES values as

$$-nbc < \text{PCAES}(nbc) < nbc \tag{31}$$

Thus, to obtain the value of PCAES in [0 1], we have specified PCAESN [Tlili et al. \(2014\)](#) as following:

$$\text{PCAESN}(nbc) = 0.5 + [(\text{PCAES}(nbc)/nbc)/2] \tag{32}$$

### 7.1.4 VOS index

Validity Overlap Separation (VOS) ([Pal and Bezdek 1992](#)) gives its values in the interval [0 1] and reaches its best partition for the minimum value.

This index is given by

$$\text{VOS}(nbc, U) = \frac{\text{Overlap}^N(nbc, U)}{\text{Sep}^N(nbc, U)}, \tag{33}$$

where  $\text{Overlap}^N(nbc, U)$  gives an inter-cluster overlap for different values of nbc, normalized by the max overlap for  $nbc = 2 \dots nbcmax$ .

$$\text{Overlap}^N(nbc, U) = \frac{\text{Overlap}(nbc, U)}{\text{Overlap}_{\max}} \tag{34}$$

$\text{Sep}^N(nbc, U)$  calculates the separation between clusters for different values of nbc, normalized by the max separation for  $nbc = 2 \dots nbcmax$ .

$$\text{Sep}^N(nbc, U) = \frac{\text{Sep}(nbc, U)}{\text{Sep}_{\max}} \tag{35}$$

### 7.2 Overlap function

$$\text{Overlap}(nbc, U) = \frac{2}{nbc(nbc - 1)} \sum_{p=1}^{nbc-1} \sum_{q=p+1}^{nbc} P(\bar{F}_p, \bar{F}_q) \tag{36}$$

$P(\bar{F}_p, \bar{F}_q)$  defines the total overlap between two fuzzy clusters  $\bar{F}_p$  and  $\bar{F}_q$ :

$$P(\bar{F}_p, \bar{F}_q) = \sum_{\mu} f(\mu : \bar{F}_p, \bar{F}_q) \tag{37}$$

The function  $f(\mu : \bar{F}_p, \bar{F}_q)$  calculates the overlap degree between two fuzzy clusters  $\bar{F}_p$  and  $\bar{F}_q$  at a membership degree  $\mu$  given by

$$f(\mu : \bar{F}_p, \bar{F}_q) = \sum_{j=1}^{nbc} \delta(x_j, \mu : \bar{F}_p, \bar{F}_q) \tag{38}$$

$\delta(x_j, \mu : \bar{F}_p, \bar{F}_q)$  indicates whether two clusters are overlapping at the membership degree  $\mu$  for data point  $x_j$ . It returns an overlap value of  $\omega(x_j)$  when the membership degrees of two clusters are both greater than  $\mu$ ; otherwise it returns 0.0:

$$\delta(x_j, \mu : \bar{F}_p, \bar{F}_q) = \begin{cases} \omega(x_j) & \text{if } (\mu_{\bar{F}_p}(x_j) \geq \mu) \text{ and } (\mu_{\bar{F}_q}(x_j) \geq \mu), \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

$\omega(x_j)$  is a weight factor for each point  $x_j$  between clusters.  $\omega(x_j)$  is a value in [0 1].

### 7.3 Separation function

The separation measure is obtained using the similarity distance  $S(\bar{F}_p, \bar{F}_q)$  between two fuzzy clusters  $\bar{F}_p$  and  $\bar{F}_q$ . It is defined as

$$\text{Sep}(nbc, U) = 1 - \min_{p \neq q} S(\bar{F}_p, \bar{F}_q) \tag{40}$$

$S(\bar{F}_p, \bar{F}_q)$  is the maximum membership degree between two clusters  $\bar{F}_p$  and  $\bar{F}_q$  in the interval [0 1]:

$$S(\bar{F}_p, \bar{F}_q) = \max_{x \in X} \min(\mu_{\bar{F}_p}(x), \mu_{\bar{F}_q}(x)), \tag{41}$$

where  $\mu_{\bar{F}_p}(x), \mu_{\bar{F}_q}(x)$  are the membership degrees of vector  $x$ , respectively, in  $\bar{F}_p$  and  $\bar{F}_q$ .

#### 7.3.1 Dunn index

The validity indice [Dunn \(1974\)](#) is based on inter-cluster distance and the diameter of a hyperspheric cluster. Dunn index is given as

$$\text{Dunn}(nbc) = \min_{i=1, \dots, nbc} \left\{ \min_{j=i+1, \dots, nbc, j \neq i} \left\{ \frac{d(c_i, c_j)}{\max_{k=1, \dots, nbc} \{\text{diam}(c_k)\}} \right\} \right\}$$

$$d(c_i, c_j) = \min_{x \in C_i, y \in C_j} \{d(x, y)\}$$

$\text{diam}(c_k)$  is the diameter of the cluster  $c_k$ .

### 7.3.2 FVINOS index

Fuzzy Validity Index with Noise-Overlap Separation (Tlili et al. 2014) (FVINOS) is inspired from the Davies–Bouldin validity index Mingoti and Lima (2006). FVINOS is defined as

$$\text{FVINOS}(\text{NBC}) = \frac{1}{\text{NBC}} \sum_{i=1}^{\text{NBC}} \left( \frac{\max_{k=1, \dots, \text{NBC}, k \neq i} \{S_i + S_k\} + \max \text{Diff}_i(\text{NBC})}{\min_{l=1, \dots, \text{NBC}, l \neq i} \{d_{i,l}\}} \right) \quad (42)$$

$\min_{l=1, \dots, \text{NBC}, l \neq i} \{d_{i,l}\}$  calculates the minimum separation between two clusters  $i$  and  $l$ .

$\max \text{Diff}(\text{NBC})$  factor is given to compare each two successive obtained partitions. In consequence, the best partition of the dataset is found.

$$\max \text{Diff}_i(\text{NBC}) = \max_{\text{NBC}} \max_{i, \dots, \text{NBC}} \text{diff}_i(\text{NBC}) \quad (43)$$

$$\text{diff}_i(\text{NBC}) = \max_{k=1, \dots, \text{NBC}, k \neq i} \{S_i(\text{NBC}) + S_k(\text{NBC})\} - \max_{k=1, \dots, \text{NBC}+1, k \neq i} \{S_i(\text{NBC}+1) + S_k(\text{NBC}+1)\} \quad (44)$$

$\text{diff}_i(\text{NBC})$  calculates the difference between max sums of compactness in a pair of clusters  $i$  and  $k$ ; this is calculated for the obtained partition at  $\text{NBC}$  and  $\text{NBC}+1$ .

We define the average of fuzzy compactness relative to cluster  $i$  as

$$S_i = \frac{1}{n_i} \sum_{x \in C_i} (u_i(x)^m) \times d(x, c_i), \quad (45)$$

where  $u_i(x)$  is the membership degree of point  $x$  to cluster  $i$ ,  $m$  is the fuzzifier factor.

$d(x, c_i)$  is the Euclidean distance separating  $x$  from the cluster centre  $c_i$ .

## References

- Abonyi J, Migaly S, Szeifert F (2002) Fuzzy self-organizing map based on regularized fuzzy c-means clustering. *Advances in soft computing, Engineering Design and Manufacturing*
- Abraham A, Nath B (2001) A neuro-fuzzy approach for modelling electricity demand in Victoria. *Appl Soft Comput* 1(2):127–138
- Aguilera PA, Frenich AG (2001) Application of the Kohonen neural network in coastal water management: methodological development for the assessment and prediction of water quality. *Water Res* 35:4053–4062
- Alimi AM, Hassine R, Selmi M (2003) Beta fuzzy logic systems: approximation properties in the MIMO case. *Int J Appl Math Comput Sci* 13(2):225–238
- Alimi AM (2000) The beta system: toward a change in our use of neuro-fuzzy systems. *Int J Manag, Invited Paper*, no. June, pp 15–19
- Alimi AM (2003) Beta neuro-fuzzy systems. *TASK Q J, Special Issue on Neural Networks* edited by Duch W and Rutkowska D, vol 7, no 1, pp 23–41
- Amarasiri R, Alahakoon D, Smith KA (2004) HDGSOM: a modified growing self-organizing map for high dimensional data clustering. *Fourth international conference on hybrid intelligent systems (HIS'04)*, pp 216–221
- Ayadi T, Ellouze M, Hamdani TM, Alimi AM (2012) Movie scenes detection with MIGSOM based on shots semi-supervised clustering. *Neural Comput Appl*. doi 10.1007/s00521-012-0930-5
- Ayadi T, Hamdani TM, Alimi AM, and Khabou MA (2007) 2IBG-SOM: interior and irregular boundaries growing self-organizing maps. *IEEE sixth international conference on machine learning and applications*, pp 387–392
- Ayadi T, Hamdani TM, Alimi AM (2010) A new data topology matching technique with multilevel interior growing self-organizing maps. *IEEE international conference on systems, man, and cybernetics*, pp 2479–2486
- Ayadi T, Hamdani TM, Alimi AM (2011) On the use of cluster validity for evaluation of MIGSOM clustering. *ISCIH: 5th international symposium on computational intelligence and intelligent informatics*, pp 121–126
- Ayadi T, Hamdani TM, Alimi AM (2012) MIGSOM: multilevel interior growing self-organizing maps for high dimensional data clustering. In: *Neural processing letters*, vol 36, pp 235–256
- Barrea A (2011) Local fuzzy c-means clustering for medical spectroscopy images. *Appl Math Sci* 5(30):1449–1458
- Bezdek JC (1974) Cluster validity with fuzzy sets. *J Cybern* 58–72
- Bezdek JC (1987) Convergence theory for fuzzy c-means. *IEEE Trans SMC*, pp 873–877
- Bezdek JC, Harris JD (1978) Fuzzy partitions and relation: an axiomatic basis for clustering fuzzy sets and systems. Academic Press, Massachusetts
- Bezdek JC, Castelaz PF (1977) Prototype classification and feature selection with fuzzy sets. *IEEE Trans Syst Man Cybern* 7(2): 87–92
- Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York
- Blake CL, Merz CJ (1998) UCI repository of machine learning databases, [<http://www.ics.uci.edu/learn/MLRepository.html>]. University of California, Department of Information and Computer Science, Irvine, CA
- Chang PC, Liao TW (2006) Combining SOM and fuzzy rule base for flow time prediction in semiconductor manufacturing factory. *Appl Soft Comput* 6(2):198–206
- Das S, Abraham A, Konar A (2008) Automatic kernel clustering with a Multi-Elitist Particle Swarm Optimization Algorithm. *Patt Recogn Lett* 29:688–699
- Das S, Abraham A (2008) Automatic clustering using an improved differential evolution algorithm. *IEEE Trans SMC*, vol 38, no 1, pp 218–237
- Denaïa MA, Palisb F, Zeghibb A (2007) Modeling and control of nonlinear systems using soft computing techniques. *Appl Soft Comput* 7(3):728–738
- Dhahri H, Alimi AM (2005) Hierarchical learning algorithm for the beta basis function neural network. In: *Proceedings of third international conference on systems, signals & devices: SSD'2005*, Sousse, Tunisia, March 2005
- Dunn JC (1974) Well separated clusters and optimal fuzzy partitions. *J Cybern* 4:95–104
- El Malek J, Alimi AM, Tourki R (2002) Problems in pattern classification in high dimensional spaces: behavior of a class of combined neuro-fuzzy classifiers. *Fuzzy Sets Syst* 128(1):15–33
- Fritzke B (1994) Growing cell structure: a self organizing network for supervised and un-supervised learning. *Neural Netw* 7:1441–1460

- Hamdani TM, Alimi AM, Khabou MA (2011) An iterative method for deciding SVM and single layer neural network structures. *Neural Process Lett* 33(2):171–186
- Hamdani TM, Alimi AM, Karray F (2008) Enhancing the structure and parameters of the centers for BBF fuzzy neural network classifier construction based on data structure. In: *Proceedings of IEEE international joint conference on neural networks, Hong Kong, IJCNN*, art. no. 4634247, pp 3174–3180
- Hamdani TM, Won JM, Alimi AM, Karray F (2011) Hierarchical genetic algorithm with new evaluation function and bi-coded representation for features selection with confidence rate. *Appl Soft Comput* 11(1):2501–2509, ISSN 1568–4946
- Hsu AL, Halgarmuge SK (2001) Enhanced topology preservation of dynamic self-organising maps for data visualization. *IFSA world congress and 20th NAFIPS international conference*, vol 3, pp 1786–1791
- Hu W, Xie D, Tan T, Maybank S (2004) Learning activity patterns using fuzzy self-organizing neural network. *IEEE Trans Syst Man Cybern B* 34(3):1618–1626
- Kim DW, Lee KH, Lee D (2004) On cluster validity index for estimation of the optimal number of fuzzy clusters. *Patt Recogn* 37(10):2009–2025
- Kim M, Ramakrishna RS (2005) New indices for cluster validity assessment. *Patt Recogn Lett* 26(15):2353–2363
- Kohonen T (1998) Statistical pattern recognition with neural networks: benchmark studies. In: *Proceedings of the second annual IEEE international conference on neural networks*, vol 1
- Kohonen T (1982) Self-organized formation of topologically correct feature maps. *Biol Cybern* 43:59–69
- Kohonen T (1984) *Self-organization and associative memory*. Springer, Berlin
- Kohonen T (1997) *Self-organizing maps*, 2nd edn. Springer, Berlin
- Kohonen T (1998) The self-organizing map. *Neurocomputing* 21:1–6
- Kohonen T (2001) *Self-organizing maps*. Springer, Berlin
- Li Hui-Ya, Hwang Wen-Jyi, Chang Chia-Yen (2011) Efficient fuzzy C-means architecture for image segmentation. *Sensors* 2011(11):6697–6718. doi:10.3390/s110706697
- Mingoti SA, Lima JO (2006) Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms. *Eur J Oper Res* 174:1742–1759
- Pal NR, Bezdek JC (1992) On cluster validity for fuzzy c-means model. *IEEE Trans Fuzzy Syst* 3:370–379
- Pascual A, Barcena M, Merelo JJ, Carazo JM (2000) Mapping and fuzzy classification of macromolecular images using self-organizing neural networks. *Ultramicroscopy* 84:85–99
- Petersson F, Chakraborty N, Saxena H (2007) A genetic algorithms based multi-objective neural net applied to noisy blast furnace data. *Appl Soft Comput* 7(1):387–397
- Ravia V, Kurniawana H, Thalia PN, Kumar PR (2008) Soft computing system for bank performance prediction. *Appl Soft Comput* 8(1):305–315
- Theodoridis Y (1996) *Spatial datasets—an unofficial collection*. Available from: <http://www.dias.cti.gr/~ytheo/research/datasets/spatial.html>
- Tlili M, Ayadi T, Hamdani TM, Alimi AM (2012) FMIG: fuzzy multi-level interior growing self-organizing maps. *IEEE international conference on tools with artificial intelligence (ICTAI)*
- Tlili M, Hamdani TM, Alimi AM, Abraham A (2014) FVINOS: fuzzy validity index with noise-overlap separation for clustering algorithms. *Patt Recogn Lett* (in press)
- Tsekouras GE, Sarimveis H (2004) A new approach for measuring the validity of the fuzzy c-means algorithm. *Adv Eng Softw* 35(8–9):567–575
- Velmurugan T, Santhanam T (2010) Performance evaluation of K-means and fuzzy C-means clustering algorithms for statistical distributions of input data points. *Eur J Sci Res* ISSN 1450–216X, 46(3):320–330
- Wu KL, Yang MS (2005) A cluster validity index for fuzzy clustering. *Patt Recogn Lett* 26(9):1275–1291
- Xie XL, Beni G (1991) A validity measure for fuzzy clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(8):841–847
- Yuksel ME, Besdok E (2004) A simple neuro-fuzzy impulse detector for efficient blur reduction of impulse noise removal operators for digital images. *IEEE Trans Fuzzy Syst* 12(6):854–865