Biological image classification using rough-fuzzy artificial neural network

Carlos Affonso [a,*], Renato Jose Sassi [b], Ricardo Marques Barreiros [a]

[a] Department of EIM, Universidade Julio de Mesquita Filho, UNESP, 18400-140 SP, Itapeva, Brazil
[b] Industrial Engineering Post Graduation, Universidade Nove de Julho, UNINOVE, 5001-100 SP, São Paulo, Brazil

ARTICLE INFO

ABSTRACT

This paper presents a methodology to biological image classification through a Rough-Fuzzy Artificial Neural Network (RFANN). This approach is used in order to improve the learning process by Rough Sets Theory (RS) focusing on the feature selection, considering that the RS feature selection allows the use of low dimension features from the image database. This result could be achieved, once the image features are characterized using membership functions and reduced it by Fuzzy Sets rules. The RS identifies the attributes relevance and the Fuzzy relations influence on the Artificial Neural Network (ANN) surface response. Thus, the features filtered by Rough Sets are used to train a Multilayer Perceptron Neuro Fuzzy Network. The reduction of feature sets reduces the complexity of the neural network structure therefore improves its runtime. To measure the performance of the proposed RFANN the runtime and training error were compared to the unreduced features.

1. Introduction

In complex problems as biological cells image classification, the capture of the essential features must be carried out without *a priori* knowledge of the image. The increased amount of attributes requires computational complexity and runtime even bigger. Moreover, due to noise in the database caused by excessive image features can cause a reduction in capacity of representation. According to Shang and Qiang (2008), the employment of Rough-Fuzzy features selection mechanism allows the reduction for a low dimensionality features sets from samples descriptions.

For these complex cases from the real life the use of Rough Sets (RS) in the pre-processing of the database has been efficient, since only the most relevant features are used as input parameters for the neural network. The RS has recently emerged as another major mathematical approach for managing uncertainty that arises from inexact, noisy, or incomplete information. It is found to be particularly effective in the area of knowledge reduction (Petrosino & Salvi, 2006).

In these cases, Fuzzy Set theory (FS) and RS represent two different approaches to vagueness. FS addresses gradualness of knowledge, expressed by the fuzzy membership, whereas rough set theory addresses granularity of knowledge, expressed by the indiscernibility relation (Affonso & Sassi, 2010).

An option to simplify the structure of the Artificial Neural Network (ANN) and reduce the noise caused by non-significant features is to use the Rough Set (RS) approach in order to select the most important features. The present paper proposes a new algorithm to realize the feature selection, with the intention to use RS as a tool for structuring the ANN. The methodology consisted of generating rules from training examples by rough-set learning, and mapping the dependency factors of the rules into the connection weights of a four-layered neural network.

The advantage of the Rough-Fuzzy Artificial Neural Network (RFANN) approach consists in the synergy achieved by combining two or more technical capabilities to achieve a more powerful system regarding to learning and generalization (Gomide, Figueiredo, & Pedrycz, 1998). A sequential architecture is used in this work, in which RS and the FS have distinct functions: RS identifies the most critical features, while the FS generates the surface response (input, output) since the Neuro Fuzzy Network (NFN) has Learnability and can adapt itself to the real world.

The paper is organized as follows: Section 2 presents the Literature review, Section 3 presents the Experimental Methodology and Section 4 presents the Conduct of Experiments. The Conclusion is presented in Section 5.

2. Literature review

Techniques can be combined to obtain a more powerful system in terms of interpretation, learning, parameter estimation, generalization, and less disability as well. Thus, various combinations have

---

* Corresponding author. Tel.: +551535249158.
E-mail address: affonso.unesp@gmail.com, affonso@itapeva.unesp.br (C. Affonso).

been applied in different papers generating systems based on: Fuzzy Min–Max Neural Network, Regression Tree and the Random Forest model as a decision support tool for medical data classification (Seera & Lim, 2014), an hybrid evolutionary dynamic Neural Network for stock market trend analysis and prediction using unscented Kalman filter (Bisoi & Dash, 2014), credit risk evaluation using Multi-Criteria Optimization classifier with Kernel, Fuzzification and Penalty Factors (Zhang, Gao, & Shi, 2014), a novel Support Vector Machine model combining Kernel Principal Component Analysis with Genetic Algorithm is proposed for intrusion detection (Kuang, Xu, & Zhang, 2014), two independent hybrid mining algorithms to improve the classification accuracy rates of Decision Tree and Naïve Bayes classifiers for the classification of multi-class problems (Farid, Zhang, Rahman, Hossain, & Strachan, 2014), a novel Fuzzy Hybrid Quantum Artificial Immune Clustering algorithm based on cloud model (Zhang, Shan, Liu, & Zhang, 2014), an optimization approach based on the Ordinal Optimization Philosophy and Particle Swarm Optimization is used to search in the continuous space of the operational variables (Zhang, Chiang, & Wu, 2014) and a Local Least-Squares Support Vector Machines-Based Neuro-Fuzzy Model for Nonlinear and Chaotic Time Series Prediction (Miranian, & Abdollahzade, 2013).

Combined techniques can also be applied to the identification, treatment and processing of images, in generating systems based on: extreme Learning Machine and Sparse Representation based classification method, have attracted significant attention due to their respective performance characteristics in computer vision and pattern recognition (Luo & Zhang, 2014), a Neural-AdaBoost based facial expression recognition system (Owusu, Zhan, & Mao, 2014), Artificial Bee Colony approach to information granulation-based Fuzzy Radial Basis Function Neural Networks for image fusion (Yu & Duan, 2013), a novel Multi-Instance Learning algorithm based on Multiple-Kernels Framework has been proposed for image classification (Li, Wang, Zhao, Liu, & Wang, 2014), Fuzzy-Rough feature selection aided Support Vector Machines for Mars image classification (Shang & Barnes, 2013), Rough Sets and Near Sets in Medical Imaging (Hassanien, Abraham, Peters, Schaefer, & Henry, 2009), Implementation and comparative analysis of Rough Set, Artificial Neural Network and Fuzzy-Rough classifiers for Satellite image classification (Juneja, Walia, Sandhu, & Mohana, 2009) and an Analysis of Clustering Algorithms for MR Image Segmentation using IQI (Patel & Patnaik, 2012).

Hybrid techniques have been applied to biological images, generating systems based on: Expert System Approach to the Identification and Clustering of Features of Biological Images (Jordan & Perkins, 1988), Artificial Neural Networks for Classification and Identification of Data of Biological Tissue Obtained by Mass-Spectrometry Imaging (Xiong et al., 2012), Multi-objective Nature-Inspired Clustering and Classification Techniques for image segmentation (Bong & Rajeswari, 2011), Evolutionary Artificial Neural Network Design and Training for wood veneer classification (Castellani & Rowlands, 2009), Image Segmentation Algorithms applied to wood defect detection (Funck, Zhong, Butler, Brunner, & Forrer, 2003), a new Neuro-Fuzzy method to investigate the characteristics of the facial images (Diago, Kitaoka, Hagiwara, & Kambayashi, 2011), Rough Sets combined with various other methodologies such as Neural Networks, Wavelets, Mathematical Morphology, Fuzzy Sets, Genetic Algorithms, Bayesian Approaches, Swarm Optimization and Support Vector Machines in the image processing domain (Hassanien, Abraham, Peters, & Schaefer, 2008); Rough Set frameworks hybridized with other Computational Intelligence Technologies that include Neural Networks, Particle Swarm Optimization, Support Vector Machines and Fuzzy Sets (Hassanien et al., 2009).

2.1. Image identification

It takes a long time to train a person to be competent in wood identification. Furthermore, manual examination of the wood sample can be very subjective. In addition to the macroscopic features of wood, physical features such as weight (different moisture content), color (variation), odour, hardness, texture, and surface appearances are also considered. For unknown specimen, usually dichotomous keys are used on a systematic analytical procedure for the examination of the wood structure.

The identity of the tree in the forest can be easily known by examining their flowers, fruits and leaves. However, once the tree is felled, the identification of the tree becomes very difficult and has to rely on their physical, macroscopic and microscopic features for identification. In this research, an intelligent recognition system using low cost equipment for the identification of wood species based on the macroscopic features of wood has been designed (Pham, Soroka, Ghanbarzadeh, & Koc, 2006).

The image processing techniques are widely used for classification and clustering of plant cells. In most cases, the biological classification is performed by trained operators, but this solution suffers significant disadvantages, so the literature contains several papers in which neural networks are used in image processing plant cells (He, 1997; Khalid, Lee, Yusof, & Nadaraj, 2008; Marzuki, Eileen, Rubiyah, & Miniappan, 2008; Pham et al., 2006; Topalova & Tzokev, 2011), also for prediction of fracture toughness (Dassanayake, 2000; Samarasinghe, Kulasiri, & Jamieson, 2007).

2.2. Rough set theory (RS)

RS was proposed by Zdzislaw Pawlak in 1982 (Pawlak, 1982) as a mathematical model to represent knowledge and to treat uncertainty. An important concept in RS is the reduct.

A reduct is a minimal set of attributes that can represent an object with the same accuracy as the original set of attributes. Elimination of redundant attributes can help in the identification of strong, non-redundant classification rules.

A reduct of B – RED(B) – on information system (IS) is a set of attributes B'⊆ B such that all attributes a ∈ (B – B') are dispensable. Thus, U/INDs(B') =U/INDs(B), where INDs(B) is called the Indiscernibility Relation.

Computing the reduct is an n-p hard problem, and processing the reduct for large databases requires high computational processing. The reduct is generated by discernibility from the Discernibility Matrix.

The Discernibility Matrix of information systems S, denoted DM(B), is a symmetric n × n matrix with: $mD_{(i,j)} = \{a \in B \mid a(Ei) \neq a(Ej)\}$ for $i,j = 1,2,\ldots,n$. with $1 \leq i, j \leq n$ e $n = |U / INDs(B)|$.

Thus, the elements of the Discernibility Matrix mD(i,j) are a set of conditional attributes of B that differentiate the elements of classes in relation to their nominal values.

The reducts of S are generated through the simplification methods of Boolean functions for the Fs(B) function (1).

This simplification is an algebraic approximation of the logical functions, with the goal of reducing the number of attributes.

$$F_S(a_1^*, a_2^*, \ldots, a_m^*)$$
$$= \wedge \left\{ \vee m_D^*(i, j) \mid i, j = 1, 2, \ldots, n, \qquad m_D(i, j) \neq 0 \right\} \qquad (1)$$

With: $m_D^*(i, j) = \{a^* \mid a \in m_D(i, j)\}$

The discernibility function Fs(B) is obtained as follows: for all attributes represented by an element in the Discernibility Matrix MD(B), apply the sum operator ("or" or "∨") and, for each pair of cells in this matrix, apply the "product" element ("and" or "∧"), which results in a Boolean expression of "sum of products".

Fuzzy Sets concern membership among elements from the same class, while RS concerns the relationship between groups of elements in different classes.

However, the theory of RS does not compete with the Fuzzy Sets Theory but rather complements it. In fact, RS theory and Fuzzy
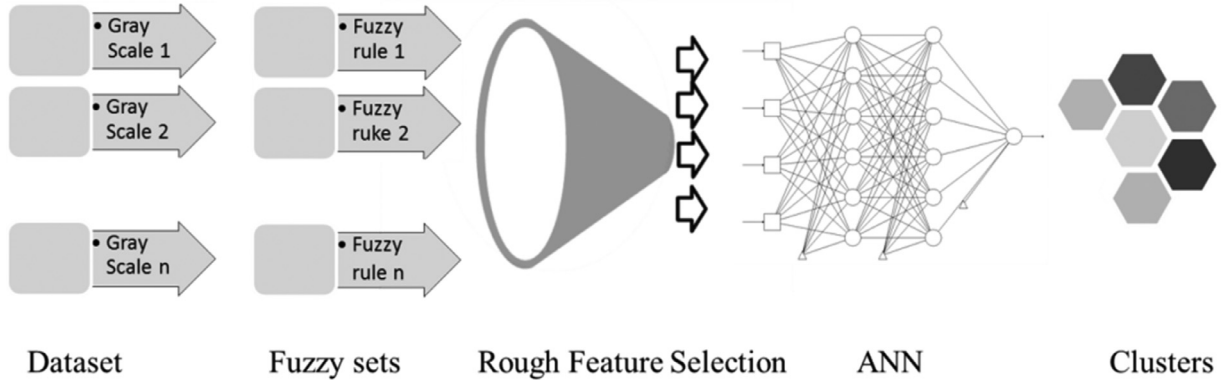
Fig. 1. Full process of the Rough-Fuzzy Artificial Neural Network (RFANN).

Sets theory are two independent approaches for the treatment of imprecise knowledge.

The knowledge acquisition bottleneck is a significant problem that hinders the building of intelligent monitoring systems. The generation of good knowledge bases for this task is notoriously difficult. This problem is particularly prevalent where experts are not readily available.

Machine learning techniques (especially rule induction methods) can be of great benefit to this area by providing strategies to automatically extract useful knowledge, given enough historical data.

2.3. Fuzzy set (FS)

In 1965, Zadeh (1964) assigned a number to every element in the universe, which indicates the degree (grade) to which the element belongs to a Fuzzy set. To formulate this concept of Fuzzy set mathematically, we present the following definition. Let X be the universe. A mapping A: $X \rightarrow [0,1]$ is called a Fuzzy set on X. The value $\mu(x)$ of A, at $x \in X$ stands for the degree of membership of x in A. The set of all Fuzzy sets on $\underline{X}$ will be denoted by F(X). $\mu(x) = 1$ means full membership, $\mu(x) = 0$ means non-membership, and intermediate values between 0 and 1 mean partial membership. $\mu(x)$ is referred to as a membership function as x varies in X.

Based on the database, membership functions for all the variables were defined, and the criterion for defining the deviation and center of each of these functions was developed with support from the injection experts. We considered levels of sensitivity (linguistic labels) α for each membership function with respective centers c_α and standard deviations σ. Considering the components x e X, to model the membership functions, we used the Gaussian function (2).

$$\mu^\alpha(x) = e^{-\frac{1}{\sigma}(x-c_\alpha)^2} \qquad (2)$$

The premises of all rules are compared with controlled entries to determine which rules apply to a situation; the outputs are compared with the established rules that have been determined. In this paper, we used the T-norm applied to $x = (x_1, x_2,...,x_q)$ as suggested by Jensen, 2005, where the value of $\Phi_{j, j=1,...,p}$ rules of inference is calculated (3):

$$\Phi_j(x) = \mu_{x1}^{\alpha 1}(x_1) \vee \mu_{x2}^{\alpha 2}(x_2) \vee ... \vee \mu_{xp}^{\alpha p}(x_p) \qquad (3)$$

2.4. Artificial neural network (ANN)

The ANN architecture Multilayer Perceptron (MLP) typically consists of a specification of the number of layers, the type of activation function of each unit, and the weights of connections between the different units, and should be established for the construction of the neural architecture (Haykin, 2001). This paper presents a supervised feedback ANN architecture in three layers: the input layer represents

the values of the reduced rule base. The neurons of the network hidden layer are trained from the set of inference rules (reduced). The algorithm used in training the MLP is the error back propagation that works as follows: first, a standard is presented. In this work, a standard will be a prototype vector and its label - the input layer of the network. This pattern is processed layer by layer until the output layer provides the response rendered, f_{MLP}, as calculated in Eq. (4):

$$f_{MLP}(x) = \varphi\left(\sum_{i=1}^{n} v_l \varphi\left(\sum_{j=1}^{m} w_{lj} x_l + b_{l0}\right) + b_0\right) \qquad (4)$$

where w_{lj} are synaptic weights; b_l and b_0 are the biases; φ is the activation function, usually specified as the sigmoid function.

ANN has the ability to solve problems in complex systems such as image classification, due to its ability to generalize, however, has difficulty working with redundant information, or with very large data set.

2.5. Rough-fuzzy artificial neural network (RFANN)

In this paper the pre-processing through RS serves to identify which attributes are most relevant for the classification of image. Rough selection provides a means by which discrete or real-valued noisy data (or a mixture of both) can be effectively reduced without the need for user-supplied information. In addition, this technique can be applied to data with continuous or nominal decision attributes, and, as such, can be applied to regression as well as classification datasets. The only additional information required is in the form of Fuzzy partitions for each feature, which can be automatically derived from the data (Jensen, 2005). The attributes selection for problems such as image classification is a complex problem, and its complexity grows exponentially in the number of attributes to be discretized. In this paper a new algorithm was proposed to do the discretization.

3. Experimental methodology

This work should be viewed as an application of Artificial Intelligence techniques, in particular, ANN to determine a surface response. The hardware platform used in the experiments was a Core i5-4200 U 1.6 GHz, 8 G RAM and 500GB hard drive.

3.1. Establishment of new algorithm

The images processing and the RFANN routines were written on C Program Language. Fig. 1 illustrates the full process of the Rough-Fuzzy Artificial Neural Network (RFANN). The RS pre-processes the Fuzzy rules and generates the input vectors with most critical attributes. This reduced rule base is used to train the MLP ANN.
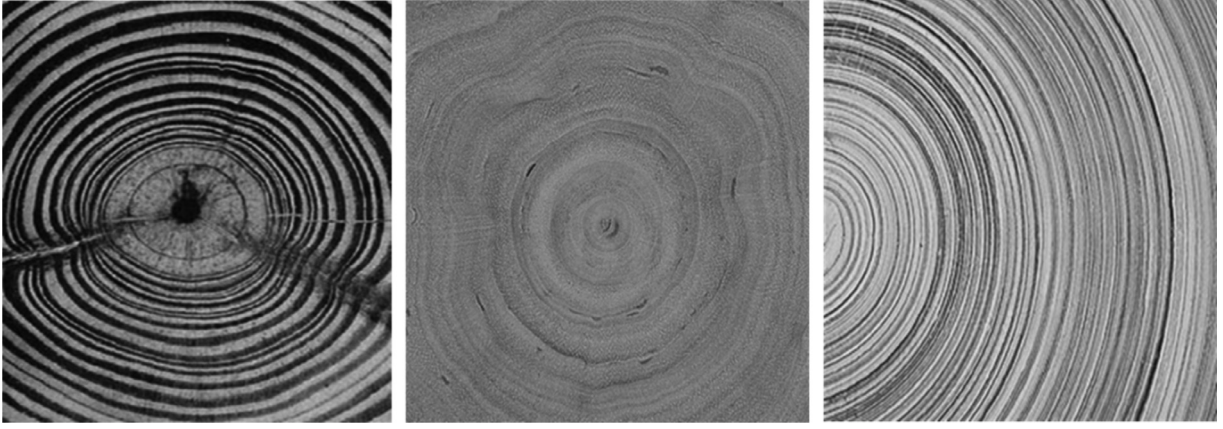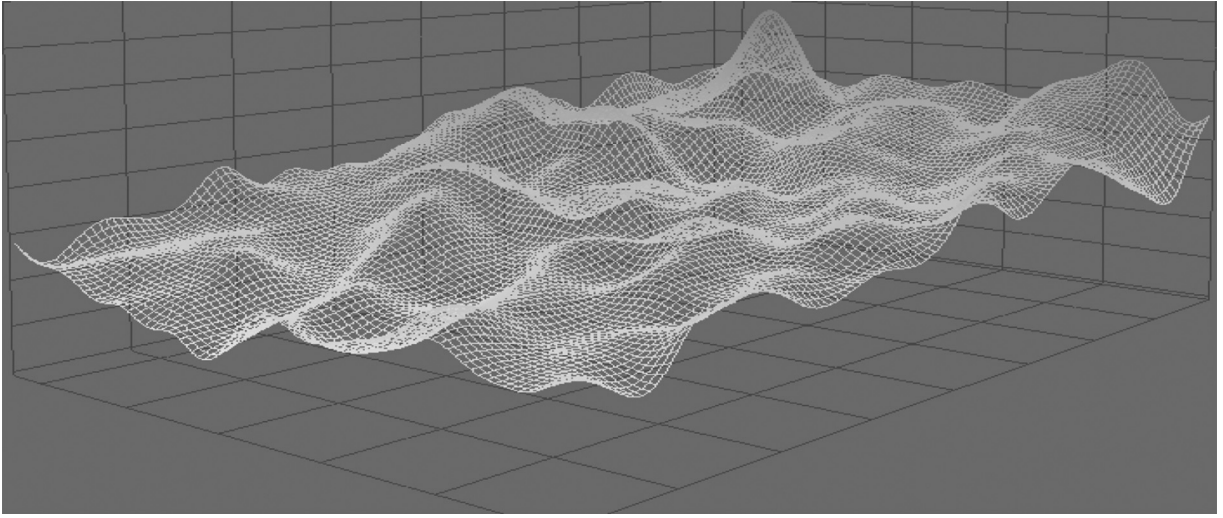
Fig. 2. Image from tree radial cuts.

Fig. 3. Mesh representing the image topography.

The image processing system operates as follows: The features are extracted from colour images by treating each channel of colour image (Red–Yellow–Green) as a monochrome imageand transforming its shape information in pixels surfaces through a C-language program. In the second step, a matrix is created with the numeric values of color intensity corresponding to each pixel (between 0 and 255).

Later it performs normalization into the numerical matrix, where it is assigned a unit value for pixels maximum intensity and zero to minimal intensity.

The images were collected from axial cut of trees, each one with 1756×1326 pixels size. The following is presented in Fig. 2, a dataset of images composed of radial cuts in different samples.

From this normalized basis, vectors are created to identify patterns for each image. These standards will be used as the basis of training data for RFNAN, which after its training; will be able to identify the specimen, even if it exactly feature vector does not belong to the database training.

The computer program converts the images into a 3D mesh through his monochromatic image. At this stage the algorithm creates a mesh representing the topography of the image, where the point is associated to a set of spatial coordinates as shown in Fig. 3.

An image $= \{y(s)\}$ is assumed to be a Gaussian random field on a $M \times N$ lattice Ω where $y(s)$ denotes the gray level of a pixel at location $s_{(i,j)}$, where $y(s) \in [0; y_{max}]$ The space will be subdivided into q-parts as follows (5) (Stepaniuk, 2008).

$$y^i_{down} = i \cdot \frac{y_{max}}{q}_{i=1,2,...,q} \qquad y^i_{up} = y^i_{down} + i \cdot \frac{y_{max}}{q} \qquad (5)$$

Considering the b samples in the image dataset, given a non-empty set Y, a possible partition Ψ is a collection of non-empty subsets of Y, such that (6):

$$\psi = \left\{ y(s) \in \psi, \ \psi \subset Y/y(s) < y^i_{up} \text{ and } y(s) > y^i_{down} \right\} \qquad (6)$$

For the finite set Ψ_i, cardinality, denoted by $card(\Psi_i)$, is the number of set elements, as shown in Eq. (7):

$$X_i = card(\psi_i) \qquad (7)$$

The feature vectors as defined above will be used as the basis of training data for a RFANN, which after its training; will be able to classify the image.

It is possible to establish a dataset with respective feature vectors for all images pattern, the values obtained are displayed below in Table 1.

3.2. Approximation spaces

To obtain the network output value at the generalization phase, it is necessary to classify the values from inference rules. Approximation spaces can be treated as granules used for concept approximation. They are some special parameterized relational structures.

Table 1
Datasets: All set features for each sample.

Feature	Sample 0	Sample 1	Sample 2	Sample 3	Sample 5	Sample 6
×1	3636	8	29	11	15	119
×2	3284	0	108	0	6	444
×3	1842	1	186	0	25	932
×4	1160	11	286	1	59	1254
×5	796	80	413	0	137	1759
×6	655	433	574	4	248	1937
×7	681	1390	749	67	319	2050
×8	549	3144	998	370	443	2311
×9	677	5431	1278	1276	482	2472
×10	775	7383	1585	3012	610	2509
×11	1051	4516	2176	5006	1039	2409
×12	1484	1110	2653	5916	1806	1950
×13	1999	235	2779	4196	2878	1583
×14	2214	61	2978	2580	3971	871
×15	1757	14	3344	1140	5046	468
×16	967	2	2786	341	4646	149
×17	320	3	1052	106	1950	40
×18	87	5	300	22	421	7
×19	16	0	29	2	56	0

Table 2
Rough inclusion function - $v_{(i,j)}$.

feature	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Σ
1	1,00	0,50	0,33	0,25	0,25	0,17	0,17	0,14	0,14	0,14	0,14	0,14	0,14	0,17	0,17	0,20	0,20	0,25	0,50	**5,01**
2	0,50	1,00	0,50	0,33	0,25	0,17	0,17	0,14	0,14	0,14	0,14	0,14	0,14	0,20	0,20	0,17	0,20	0,20	0,33	**5,07**
3	0,33	0,50	1,00	0,33	0,25	0,17	0,17	0,14	0,14	0,14	0,14	0,17	0,17	0,25	0,25	0,17	0,20	0,20	0,33	**5,05**
4	0,25	0,33	0,33	1,00	0,50	0,25	0,25	0,20	0,20	0,17	0,17	0,14	0,14	0,20	0,20	0,20	0,33	0,25	0,25	**5,37**
5	0,25	0,25	0,25	0,50	1,00	0,33	0,33	0,25	0,25	0,20	0,20	0,17	0,17	0,17	0,17	0,20	0,50	0,25	0,25	**5,68**
6	0,17	0,17	0,17	0,25	0,33	1,00	0,50	0,33	0,33	0,25	0,25	0,20	0,20	0,14	0,14	0,17	0,33	0,25	0,17	**5,35**
7	0,17	0,17	0,17	0,25	0,33	0,50	1,00	0,33	0,33	0,25	0,25	0,17	0,17	0,14	0,14	0,17	0,33	0,25	0,17	**5,29**
8	0,14	0,14	0,14	0,20	0,25	0,33	0,33	1,00	0,50	0,25	0,33	0,17	0,17	0,14	0,17	0,20	0,25	0,20	0,14	**5,06**
9	0,14	0,14	0,14	0,20	0,25	0,33	0,33	0,50	1,00	0,25	0,25	0,17	0,17	0,14	0,17	0,20	0,25	0,20	0,14	**4,98**
10	0,14	0,14	0,14	0,17	0,20	0,25	0,25	0,25	0,25	1,00	0,33	0,20	0,25	0,17	0,14	0,20	0,20	0,17	0,14	**4,60**
11	0,14	0,14	0,14	0,17	0,20	0,25	0,25	0,33	0,25	0,33	1,00	0,25	0,20	0,17	0,14	0,20	0,20	0,17	0,14	**4,68**
12	0,14	0,14	0,17	0,14	0,17	0,20	0,17	0,17	0,17	0,20	0,25	1,00	0,50	0,20	0,17	0,17	0,20	0,14	0,14	**4,43**
13	0,14	0,14	0,17	0,14	0,17	0,20	0,17	0,17	0,17	0,25	0,20	0,50	1,00	0,20	0,17	0,17	0,20	0,14	0,14	**4,43**
14	0,17	0,20	0,25	0,20	0,17	0,14	0,14	0,14	0,14	0,17	0,17	0,20	0,20	1,00	0,25	0,25	0,17	0,17	0,17	**4,29**
15	0,17	0,20	0,25	0,20	0,17	0,14	0,14	0,17	0,17	0,14	0,14	0,17	0,17	0,25	1,00	0,20	0,17	0,17	0,17	**4,17**
16	0,20	0,17	0,17	0,20	0,20	0,17	0,17	0,20	0,20	0,20	0,20	0,17	0,17	0,25	0,20	1,00	0,20	0,20	0,20	**4,45**
17	0,20	0,20	0,20	0,33	0,50	0,33	0,33	0,25	0,25	0,20	0,20	0,20	0,20	0,17	0,17	0,20	1,00	0,25	0,20	**5,38**
18	0,25	0,20	0,20	0,25	0,25	0,25	0,25	0,20	0,20	0,17	0,17	0,14	0,14	0,17	0,17	0,20	0,25	1,00	0,33	**4,79**
19	0,50	0,33	0,33	0,25	0,25	0,17	0,17	0,14	0,14	0,14	0,14	0,14	0,14	0,17	0,17	0,20	0,20	0,33	1,00	**4,92**

Tuning of parameters makes it possible to search for relevant approximation spaces relative to given concepts (Komorowski, Polkowski, & Skowron, 2002). The Rough inclusion function $P(U) \times P(U) \rightarrow [0, 1]$ defines the degree of inclusion of Φ^I in Φ^j, where Φ^i, $\Phi^j \subseteq U$. In the simplest case the standard rough inclusion function can be defined by (8):

$$v^{ij} = \begin{cases} \dfrac{card(\Phi^j \cap \Phi^i)}{card(\Phi^j)} & if\ \Phi^j \neq 0 \\ 1 & if\ \Phi^j = 0 \end{cases} \quad (8)$$

The Rough inclusion function $v_{(i,j)}$ algorithm is given as follows:

$v[]$: Rough inclusion function
$\Phi[]$: Fuzzy rules of inference

(1) $k,i,j,v_{[]}: \leftarrow 0$
(2) do
(3)   if $(\Phi_{[i,k]} \neq \Phi_{[j,k]})\ v_{[i,j]} = v_{[i,j]} + 1$
(4)   until $k < p$
(5)   until $i < q$ and $j < q$
(6)   return $v_{[]} \leftarrow 1/(v_{[]} + 1)$

This measure is widely used by the data mining and RS communities. It is possible to use this idea to estimate the probability of implications.

The Table 2 presents the values of the Rough inclusion function.

4. Conduct of experiments

4.1. RFANN applied to image identification

The experiment compares the knowledge generated by the RFANN with the data obtained from an Image dataset in two cases: the unreduced dataset and the reduced one by the RS.

Since the processing times were near to zero, the criterion for selection was the simplicity of the architecture of the network.

The choice of MLP parameters of the reduced dataset were: number of input neurons: 4, initial learning rate $\eta = 0.7$; Initial momentum $\mu = 0.7$; the stop criterion was maximum 1000 epochs.

The configuration of the number of neurons and layers are accomplished by choosing the architecture that has the lowest permissible error.

The choice of parameters for the MLP of the full dataset were: input neurons 19, number of hidden neurons 12, and learning constant $\eta = 0.7$, momentum $\mu = 0.7$; the stop criterion was maximum 1000 epochs.

The rule bases reduced by the RFANN exhibit good behavior during the generalization phase, indicating a promising way for to associate RS with FS to replace the human expert in the construction of inference rules.

The RFANN returns four features as the most significant ones, x_{10}, x_{11}, x_{12}, x_{13}, out of the original full set of features. We can check that
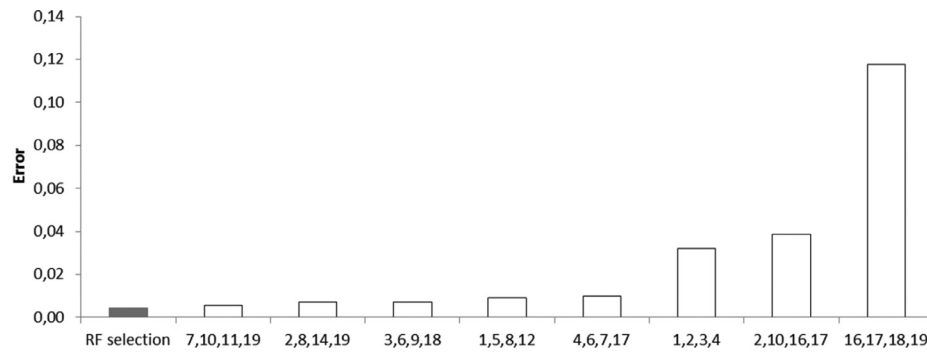
Fig. 4. Error comparison: RFANN vs. randomly selected features.

Table 3
RFANN error: comparing fuzzy-rough selection versus full features.

Features	Dim.	Topology	Error	
			Training	Testing
RF selection	4	4-12-1	0.00307	0.00435
x_1-x_7	7	7-12-1	0.01786	0.01299
x_1-x_{10}	10	10-12-1	0.00519	0.00724
x_1-x_{13}	13	13-12-1	0.00212	0.00301
x_1-x_{16}	16	16-12-1	0.00176	0.00250
Full features	19	19-12-1	0.00208	0.00295

Table 4
RFANN versus randomly selected features.

Features	Dim.	Topology	Error	
			Training	Testing
RF selection	4	4-12-1	0.00307	0.00435
7,10,11,19	4	4-12-1	0.00394	0.00557
2,8,14,19	4	4-12-1	0.00501	0.00706
3,6,9,18	4	4-12-1	0.00556	0.00715
1,5,8,12	4	4-12-1	0.00670	0.00899
4,6,7,17	4	4-12-1	0.00685	0.00978
1,2,3,4	4	4-12-1	0.09206	0.03213
2,10,16,17	4	4-12-1	0.02783	0.03885
16,17,18,19	4	4-12-1	0.11604	0.11777

using the features selection does not significantly reduce the accuracy of the classification compared to the use of the full set of features. Table 3 shows the error produced for the RFANN.

It is very interesting to note that the error rate when using four selected features is very near the error when using the full feature set. Furthermore, this performance improvement is obtained via a network structurally much simpler. This is indicative of the power resource selection to help reduce redundant feature, not only the measures but also the noise associated with this measurement.

4.2. Comparison with the randomly selected features

The above comparison ensured that no information loss is incurred due to RFANN feature reduction. It is possible to compare the performance of different attributes, reduced according to RFANN criteria and randomly selected features, as shown in Table 4 and Fig. 4.

The average error classifiers, using four randomly selected characteristics, higher than that achieved by the classifier that uses the resources selected by RFANN, considering the same dimensionality. This implies that on these features randomly selected occur losses of information during reduction.

The technique proposed in this paper uses a histogram analysis as feature selection criteria, on this way, there is a computational gain considering the model's simplicity.

This approach presents advantages comparing to the traditional ANN ones, since the running time is considerably reduced, and the ANN processed only 21% (4/19) of the original dataset. The figures shown that with a simple neural network topology, namely, one hidden layer containing 12 neurons, and a sensitive layer (4 neurons), the network was capable to classify biological images. However, this method was applied to a restricted dataset, therefore we suggest its enlargement to another databases.

5. Conclusions

In this paper, an automatic visual inspection system for the recognition of tropical Wood species based on artificial intelligence techniques has been proposed. The system was objectively designed to be cost-effective and as a means to replace wood inspectors due to difficulty in recruiting them as the job is rather laborious.

ANN has the ability to solve problems in complex systems such as image classification, due to its ability to generalize. However, has difficulty working with redundant information, or with very large dataset.

Therefore, the technique proposed in this paper uses a histogram analysis as criteria to feature selection, on this way, there is a computation gain considering the model simplicity. The figures shown on Section 4.2 demonstrates this hypothesis, once with a simple neural network topology, was capable to classify biological images. Additionally, one possible weakness of the method, may be the restricted dataset and the fact the method was been used only as classification.

The application of RFANN showed a great ability to generalize, to identify behavior patterns, and to allow the creation of an inference mechanism in high complex systems.

When applied to a real world dataset, the RFANN was able to identify the significant features as defined by a human expert. The main advantage of using the RFANN is the reduction of dependence on a human expert for the choice and construction of the rules of inference mechanism. This gain is important, considering that one of the weaknesses of the approach using the Fuzzy Sets is its dependence on the human expert. If there is a reasonable number of attributes and a structured database, it may be even possible to eliminate the need for support from the human expert for the construction of inference rules, instead using his support only in the construction of membership functions.

We suggest the RFANN in modeling other problems, such as dynamic routing or business datasets, in order to assess the impact of this approach on the dependence of the human expert in building the inference mechanism.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.eswa.2015.07.075.

References

Affonso, C., & Sassi, J. R. (2010). An inference mechanism for polymer processing using rough-neuro fuzzy network. In K. Diamantaras, W. Duch, & L. S. Iliadias (Eds.), *Computational Intelligence* (pp. 441–450). Berlin Heidelberg: Springe-Verlag ICANN 2010, Part III, LNCS 6354.

Bisoi, R., & Dash, P. K. (2014). A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented Kalman filter. *Applied Soft Computing, 19*, 41–56.

Bong, Chin-Wei., & Rajeswari, M. (2011). Multi-objective nature-inspired clustering and classification techniques for image segmentation. *Applied Soft Computing, 11*(4), 3271–3282.

Castellani, M., & Rowlands, H. (2009). Evolutionary artificial neural network design and training for wood veneer classification. *Engineering Applications of Artificial Intelligence, 22*(Issues 4–5), 732–741.

Dassanayake, D. (2000). Software project on fracture dynamics of wood and artificial neural network model for prediction of fracture toughness. A PHD Thesis submitted in partial fulfilment of the requirements for the degree of Master of Applied Science at Lincoln University.

Diago, L., Kitaoka, T., Hagiwara, I., & Kambayashi, T. (2011). Neuro-fuzzy quantification of personal perceptions of facial images based on a limited data set. *IEEE Transactions on Neural Networks, 22*(12), 2422–2434.

Farid, D. Md., Zhang, L., Rahman, C. M., Hossain, M. A., & Strachan, R. (2014). Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. *Expert Systems with Applications, 41*(4), 1937–1946 Part 2.

Funck, J. W., Zhong, Y., Butler, D. A., Brunner, C. C., & Forrer, J. B. (2003). Image segmentation algorithms applied to wood defect detection. *Computers and Electronics in Agriculture, 41*(1–3), 157–179.

Gomide, F., Figueiredo, M., & Pedrycz, W. (1998). A neural fuzzy network: structure and learning, fuzzy logic and its applications. *Information sciences and intelligent systems* (pp. 177–186). Dordrecht, Netherlands: Kluwer Academic Publishers.

Hassanien, A. E., Abraham, A., Peters, J. F., & Schaefer, G. (2008). An overview of rough-hybrid approaches in image processing. *Proceedings of the ieee international conference on fuzzy systems* (pp. 2135–2142). Hong Kong, China.

Hassanien, A. E., Abraham, A., Peters, J. F., Schaefer, G. &, & Henry, C. (2009). Rough sets and near sets in medical imaging: a review. *IEEE Transactions on Information Technology in Biomedicine, 13*(Issue 6), 955–968.

Haykin, S. (2001). *Neural networks: a comprehensive foundation*. New York: Willey & Sons.

He, J.A. (1997). Comparison of artificial neural network classifiers for analysis of ct images for the inspection of hardwood logs. PHD Thesis submitted to the Faculty of the Bradley Department of Electrical and Computer Engineering of Virginia.

Jensen, R. (2005). *Combining rough and Fuzzy set for feature selection*. Ph.D. Thesis at University of Edinburgh.

Jordan, M. M., & Perkins, W. J. (1988). An expert system approach to the identification and categorisation of features of biological images. *Machine Intelligence and Pattern Recognition, 7*, 461–470.

Juneja, M., Walia, E., Sandhu, P. S., & Mohana, R. (2009). Implementation and comparative analysis of rough set, artificial neural network (ann) and fuzzy-rough classifiers for satellite image classification. In *Proceedings of the international conference on intelligent agent & multi-agent systems* (pp. 1–6).

Khalid, M., Lee, E. L., Yusof, R., & Nadaraj, M. (2008). Design of an intelligent wood species recognition system. *International Journal of Simulation, Systems, Science and Technology, 9*(3), 9–19.

Komorowski, J., Polkowski, L., & Skowron, A. (2002). *Handbook of data mining and knowledge discovery* (pp. 134–149). NY USA: Oxford University Press.

Kuang, F., Xu, W., & Zhang, S. (2014). A novel hybrid KPCA and SVM with GA model for intrusion detection. *Applied Soft Computing, 18*, 178–184.

Li, D., Wang, J., Zhao, X., Liu, Y., & Wang, D. (2014). Multiple kernel-based multi-instance learning algorithm for image classification. *Journal of Visual Communication and Image Representation, 25*(5), 1112–1117.

Luo, M., & Zhang, K. (2014). A hybrid approach combining extreme learning machine and sparse representation for image classification. *Engineering Applications of Artificial Intelligence, 27*, 228–235.

Marzuki, K., Eileen, L. L., Rubiyah, Y., & Miniappan, N. (2008). Design of an intelligent wood species recognition system. *International Journal of Simulation, Systems, Science and Technology, 9*(3), 9–19.

Miranian, A., & Abdollahzade, M. (2013). Developing a local least-squares support vector machines-based neuro-fuzzy model for nonlinear and chaotic time series prediction. *IEEE Transactions on Neural Networks and Learning Systems, 24*(2), 207–218.

Owusu, E., Zhan, Y., & Mao, Q. R. (2014). A neural-AdaBoost based facial expression recognition system. *Expert Systems with Applications, 41*(7, 1), 3383–3390 1.

Patel, S., & Patnaik, K. S. (2012). Analysis of clustering algorithms for MR image segmentation using IQI. *Procedia Technology, 6*, 387–396.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and information Sciences, 11*, 341–356.

Petrosino, A., & Salvi, G. (2006). Rough fuzzy set based scale space transforms and their use in image analysis. *International Journal of Approximate Reasoning, 41*, 212–228.

Pham, D. T. J., Soroka, A. J., Ghanbarzadeh, A., & Koc, E. (2006). Optimising neural networks for identification of wood defects using the bees algorithm. In *Proceedings of the ieee international conference on industrial informatics* (pp. 1346–1351).

Samarasinghe, S., Kulasiri, D., & Jamieson, T. (2007). Neural networks for predicting fracture toughness of individual wood samples. *Silva Fennica, 41*(1), 105–122.

Seera, M., & Lim, C. P. (2014). A hybrid intelligent system for medical data classification. *Expert Systems with Applications, 41*(5), 2239–2249.

Shang, C., & Barnes, D. (2013). Fuzzy-rough feature selection aided support vector machines for Mars image classification. *Computer Vision and Image Understanding, 117*(3), 202–213.

Shang, C., & Qiang, S. (2008). Aiding zction. In *Proceedings of the ieee international conference on fuzzy systems* (pp. 976–982).

Stepaniuk, J. (2008). *Rough-granular computing in knowledge discovery and data mining*. Berlin Heidelberg: Springer-Verlag.

Topalova, I., & Tzokev, A. (2011). Modular structure of neural networks for classification of wooden surfaces with plc industrial implementation. In *Proceedings of the seventh international conference on autonomic and autonomous systems* (pp. 13–17).

Xiong, Xing-Chuang, Fang, X., Ou, Yang-Zheng, Jiang, Y., Huang, Ze-Jian, & Zhang, Yu-Kui (2012). Artificial neural networks for classification and identification of data of biological tissue obtained by mass-spectrometry imaging. *Chinese Journal of Analytical Chemistry, 40*(1), 43–49.

Yu, J., & Duan, H. (2013). Artificial bee colony approach to information granulation-based fuzzy radial basis function neural networks for image fusion. *Optik - International Journal for Light and Electron Optics, 124*(17), 3103–3111.

Zadeh, L. (1964). A fuzzy sets. *Information and Control, 8*, 338–353.

Zhang, R., Chiang, Wen-Chyuan &, & Wu, C. (2014). Investigating the impact of operational variables on manufacturing cost by simulation optimization. *International Journal of Production Economics, 147*(C), 634–646.

Zhang, Ren-Long, Shan, Ren-Long, Liu, Xiao-Hong, & Zhang, Li-Hong (2014). A novel fuzzy hybrid quantum artificial immune clustering algorithm based on cloud model. *Engineering Applications of Artificial Intelligence, 35*, 1–13.

Zhang, Z., Gao, G., & Shi, Y. (2014). Credit risk evaluation using multi-criteria optimization classifier with kernel, fuzzification and penalty factors. *European Journal of Operational Research, 237*(1, 16), 335–348.