Introducing a Genetic Fuzzy Linguistic Combination Method for Bagging Fuzzy Rule-Based Multiclassification Systems

Luciano Sánchez
Computer Science Department
University of Oviedo
Campus de Viesques
33203 - Gijón (Asturias), SPAIN
Email: luciano@uniovi.es

Oscar Cordón, Arnaud Quirin, and Krzysztof Trawinski
European Centre for Soft Computing
Edificio Científico-Tecnológico, planta 3
C. Gonzalo Gutiérrez Quirós, s/n
33600 - Mieres (Asturias), SPAIN
Email: {oscar.cordon, arnaud.quirin, krzysztof.trawinski}@softcomputing.es

***Abstract*—Many different fuzzy aggregation operators have been successfully used to combine the outputs provided by the individual classifiers in a multiclassification system. However, up to our knowledge, the use of fuzzy combination methods composed of a fuzzy system is less extended. By using a fuzzy linguistic rule-based classification system as a combination method, the resulting classifier ensemble would show a hierarchical structure and the operation of the latter component would be transparent to the user. Moreover, for the specific case of fuzzy multiclassification systems, the new approach could also become a smart way to allow fuzzy classifiers to deal with high dimensional problems avoiding the curse of dimensionality. The present contribution establishes the first basis in this direction by introducing a genetic fuzzy system-based framework to build the fuzzy linguistic combination method for a bagging fuzzy multiclassification system.**

I. INTRODUCTION

Multiclassification systems (MCSs), also called classifier ensembles, are machine learning tools capable to obtain better performance than a single classifier when dealing with complex classification problems, especially when the number of dimensions or the size of the data are really large [1]. The most common base classifiers are decision trees [2], neural networks [3], and more recently fuzzy classifiers [4], [5], [6].

MCS design is mainly based on two stages [7]: the learning of the component classifiers and the combination mechanism for the individual decisions provided by them into the global MCS output. The overall accuracy of the MCS relies on the performance and the proper integration of these two tasks.

The research area of combination methods is very active. It considers both the direct combination of the results provided by all the initial set of component classifiers to compute the final output (*fusion-based methods*) and the selection of the best single classifier or classifier subset which will be taken into account to provide a decision for each specific input pattern (static/dynamic classifier selection [8] and overproduce-and-choose strategies [9]). Besides, hybrid strategies between the two groups have also been introduced [1].

While the weighted majority vote could be considered as the most extended fusion-based combination method [10], many other proposals have been developed in the specialized literature, including several successful procedures based on the use of fuzzy set theory and, specifically, of fuzzy aggregation operators [11], [12]. Up to our knowledge, there is only one previous proposal of a MCS combination method considering the use of a fuzzy system to accomplish this task [13]. It is based on a first-order TSK fuzzy system. However, the use of a fuzzy linguistic system can constitute a very interesting alternative due to its higher interpretability.

In this contribution we introduce a framework to derive a fuzzy rule-based classification system (FRBCS) playing the role of the MCS combination method. This fuzzy linguistic combination method presents an interpretable structure as it is based on the use of a single disjunctive fuzzy classification rule per problem class as well as on the classical single-winner fuzzy reasoning method [14], [15]. The antecedent variables correspond to the component FRBCSs (and thus its number is bounded by the existing number) and each of them has a weight associated representing the certainty degree of each ensemble member in the classification of each class. A specific genetic algorithm (GA) to design such FRBCS-based combination method (FRBCS-CM) will be proposed with the ability of selecting features and linguistic terms in the antecedent parts of the rules. In such way, it will perform both classifier fusion and classifier selection at class level. The resulting system is thus a genetic fuzzy system (GFS) [16], [17] (in particular, a genetic fuzzy rule-based classification system (GFRBCS)) dealing with the interpretability-accuracy trade-off in a proper way [18].

In particular, in the current work the novel FRBCS-CM will be applied on fuzzy rule-based multiclassification systems (FRBMCSs) generated from the bagging and feature selection methodology we proposed in [4], [5], [6]. Therefore, the

This work has been supported by the Spanish Ministerio de Ciencia e Innovación under projects TIN2009-07727 and TIN2008-06681-C06-04, both including EDRF fundings.

resulting FRBMCS will show a clear hierarchical structure composed of two levels of FRBCSs allowing it to deal with high dimensional problems. A preliminary study will be conducted on nine datasets of different sizes from the UCI machine learning repository to test the accuracy and complexity of the derived FRBMCSs in comparison to the original FRBMCS.

This paper is set up as follows. In the next section, the preliminaries required for a good understanding of our work (MCS combination methods, fuzzy MCS combination methods, and our approach for designing FRBMCSs considering bagging and feature selection) are reviewed. Section III describes the proposed FRBCS-CM framework, structure, and the GA considered to design it. The experiments developed and their analysis are shown in Sec. IV. Finally, Sec. V collects some concluding remarks and future research lines.

II. PRELIMINARIES

This section explores the current literature related to classifier ensemble combination methods and reviews our generation method for FRBMCSs.

A. Multiclassification System Combination Methods

Two main approaches arise in the literature for the combination of the outputs provided by a previously generated set of individual classifiers into a single MCS output [19]: *classifier fusion* and *classifier selection*.

Classifier fusion relies on the assumption that all ensemble members make independent errors. Thus, combining the decisions of the ensemble members may lead to increasing the overall performance of the system. Majority voting, sum, product, maximum and minimum are examples of functions used to combine their decisions [20]. The most extended one is the weighted majority voting, which allows to weight the contribution of each individual classifier to the final decision according to its "classification competence" using coefficients of importance [10].

Alternatively, classifier selection is based on the fact that not all the individual classifiers but only a subset of them will influence on the final decision for each input pattern. Two categories of classifier selection techniques exist: static and dynamic [8], [19]. In the first case, regions of competence are defined during the training phase, while in the second case, they are defined during the classification phase taking into account the characteristics of the sample to be classified. There is also another family of static classifier selection methods based on the assumption that the candidate classifiers in the ensemble could be redundant. These methods are grouped under the name of overproduce-and-choose strategy (OCS) [9] and they are based on the fact that a large set of candidate classifiers is generated and then selected to extract the best performing subset (removing duplicates and poor-performing candidate classifiers), which composes the final MCS used to classify the whole test set. In addition, hybrid methods between the latter families have been proposed, such as the GA-based dynamic OCS procedure introduced in [21].

The FRBCS-CM proposed in the current contribution will belong to the static OCS group and it will be able to either completely remove a whole candidate classifier or to reduce its contribution to only some specific classes with a specific weight measuring our confidence in the individual classifier for that specific class (as done in other existing classifier selection methods such as [22], [23]). All the latter will be performed using a human-interpretable structure generated by means of a GFRBCS.

B. Multiclassification System Fuzzy Combination Methods

Fuzzy set theory has been extensively and successfuly considered for classifier fusion. The use of fuzzy connectives to combine the outputs of the component classifiers of an ensemble was first proposed in [24]. Since then, many different fuzzy aggregation operators have been considered in the specialized literature [11], [12], [25]. In [12] the accuracy of some of them was compared to that of seven of the usual crisp (i.e., non-fuzzy) aggregation operators when considered as combination operators for Boosting classifier ensembles. The conclusions drawn from that experimentation were that fuzzy combination methods clearly outperformed non-fuzzy ones.

Besides, some other works have extended the classifier fusion scope and have proposed some techniques which show some similarities with our new proposal. On the one hand, Cococcioni et al. [13] introduced probably the first and the only proposal where a fuzzy system, specifically a first-order TSK fuzzy system, is considered to combine the outputs of a MCS. In addition, Bulacio et al. [26] proposed a hybrid classifier selection-fusion strategy considering Sugeno's fuzzy integral as combination method and a greedy heuristic for the ensemble member selection. On the other hand, Lu and Yamaoka [27] introduced a fuzzy combination method specifically designed for a hybrid ensemble of three classifiers which shows the novel characteristic of allowing the user to incorporate human expert knowledge on the bias of the component classifiers. This is done by means of an additional refinement module based on a FRBS comprised by Mamdani-type fuzzy rules. In this way, contrary to the FRBCS-CM proposed in the current contribution, Lu and Yamaoka's fuzzy combination method does not make use of fuzzy rules but of a complex fuzzy reasoning process where the following components are considered: a linguistic partition for the ensemble members' outputs, a fuzzy aggregation of their membership degrees and a defuzzification method to modify them, and a new (crisp) aggregation for each class in order to take the final MCS decision corresponding to the largest aggregated class membership value.

As said, the latter procedure can be complemented by expert-defined fuzzy rules to adjust the importance of the decisions taken for each class according to the nature of the component classifiers. Hence, a FRBS is used as a refinement module for the fuzzy combination method decisions. Nevertheless, this strategy shows several problems such as its specificity to the consideration of a simple three-classifier ensemble, its highly complex structure composed of two different nature fuzzy reasoning modules, the need of manually defining the fuzzy rules in the refinement module (which could be feasible when using a very small number of component classifiers – only three– but not with dealing with a more usual larger number. In fact, the FRBSs considered in their experimentation are only composed of a single rule with three inputs as well as the authors mention they were not able to incorporate expert knowledge to the Bayesian component classifier), and the impossibility to perform classifier selection (which of course is not required in the simple ensemble structure considered).

The proposal made in the current contribution is aimed to solve all the latter drawbacks by designing a single fuzzy linguistic combination method in the form of a fully understandable FRBCS, automatically derived by a GFRBCS, which shows the capability of performing both classifier fusion and selection.

C. Bagging Fuzzy Multiclassification Systems

In previous studies [4], [5], we described a methodology based on classical MCS design approaches such as bagging [28], random subspace [2], and mutual information-based feature selection [29] to generate FRBMCSs. The approach uses the basic heuristic fuzzy classification rule generation method proposed by Ishibuchi [15] as well as a GA-based classifier selection technique driven by a multicriteria fitness function either composed of only an error measure [6] or of its combination with a diversity measure [30].

In order to build these FRBMCSs, a normalized dataset is split into two parts, a training set and a test set. The training set is submitted to an instance and feature selection procedure in order to provide the K individual training sets (the so-called *bags*) to train the K simple FRBCSs desired through Ishibuchi's method. In every case, the bags are generated with the same size as the original training set, as commonly done. In this contribution, we consider the use of the random subspace [2], where the feature set of each bag (and thus of each component classifier) is randomly selected from the original dataset.

The component FRBCSs are based on fuzzy classification rules R_j^k with a class C_j^k and a certainty degree CF_j^k in the consequent: If x_1^k is A_{j1}^k and ... and x_n^k is A_{jn}^k then Class C_j^k with CF_j^k, $j = 1, 2, \ldots, N$, $k = 1, 2, \ldots, K$. They take their decisions by means of the single-winner method [14], [15]. To derive the fuzzy knowledge bases, one of the heuristic methods proposed by Ishibuchi et al. in [15] is considered and applied on each of the previous bags. The consequent class C_j^k and certainty degree CF_j^k are statistically computed from all the examples located in a specific fuzzy subspace $D(A_j)^k$. C_j^k is computed as the class h with maximum confidence according to the rule compatible training examples $D(A_j)^k = \{e_1^k, \ldots, e_m^k\}$: $c(A_j^k \Rightarrow Class\ h) = |D(A_j^k) \bigcap D(Class\ h)|/|D(A_j^k)|$. CF_j^k is obtained as the difference between the confidence of the consequent class and the sum of the confidences of the remainder (called CF_j^{IV} in [15]).

After performing the training stage on all the bags, we get an initial whole FRBMCS, which is validated using the training and the test errors as well as a measure of complexity based on the total number of rules in the FRBCSs. The pure voting approach is applied as combination method: the ensemble class prediction will directly be the most voted class in the component classifiers output set. The lowest-order class is taken in the case of a tie.

This ensemble is finally selected using a multicriteria GA in order to look for the best cooperating subset of individual classifiers, following a OCS. The final FRBMCS is validated using different accuracy (training error, test error) and complexity (number of classifiers, total number of rules) measures.

For a more detailed description on the methodology, the interested reader is referred to the provided references.

III. A GENETIC FUZZY CLASSIFIER SYSTEM TO DESIGN A FUZZY LINGUISTIC COMBINATION METHOD FOR BAGGING FRBMCS

The next subsections will respectively provide a detailed description of the FRBCS-CM structure and of the composition of the GFS designed to derive its fuzzy knowledge base.

A. Fuzzy linguistic combination

As said in Sec. II-C, the FRBCSs considered in the ensemble will be based on fuzzy classification rules with a class and a certainty degree in the consequent. Let R_j^k be the j-th rule of the k-th member of an ensemble of K components,

$$\text{if } x \text{ is } A_j^k \text{ then Class } C_j^k \text{ with } CF_j^k,$$

where $C_j^k \in \{1, \ldots, n_c\}$ and n_c is the number of classes.

We will use the expression $\mathcal{G}^k = \{R_1^k, \ldots, R_{N_k}^k\}$ to denote the list of fuzzy rules comprising the k-th ensemble member. Let us partition each one of these lists into so many sublists $\mathcal{G}_c^k$ as classes. $\mathcal{G}_c^k$ contains the rules of $\mathcal{G}^k$ whose consequent is the class c.

Let us also define $R^k(x)$ to be the intermediate output of the k-th member of the ensemble, which is the fuzzy subset of the set of class labels computed as follows:

$$R^k(x)(c) = \bigvee_{\{j|C_j^k=c\}} CF_j^k \cdot A_j^k(x). \tag{1}$$

Each component FRBCS maps an input value x to so many degrees of membership as the number of classes in the problem. The highest of these memberships determines the classification of the pattern. That is to say, the k-th FRBCS classifies an object x as being of class $\text{FRBCS}^k(x) = \arg\max_{c\in\{1,\ldots,n_c\}} R^k(x)(c)$. Observe also that $R^k(x)(c)$ is the result of applying the fuzzy reasoning mechanism to the knowledge base defined by the sublist $\mathcal{G}_c^k$.

The simplest linguistic combination of the component FRBCSs consists in stacking a selection of some of the rules R_j^k into a single large rule base. Let us define a binary matrix $[b_{ck}] \in \{0,1\}^{n_c \times K}$, and let us agree that, if b_{ck} is zero, then $\mathcal{G}_c^k$ is removed from the ensemble and $R^k(x)(c) = 0$. This selection is equivalent to the hierarchical FRBCS comprising n_c expressions of the form:

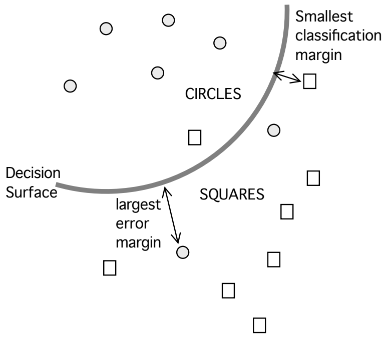

Fig. 1. The fitness of an ensemble has three components: (a) a quantile of the bootstrap estimation of the training error, (b) the largest distance between a misclassified example and the decision surface, and (c) the smallest distance between a correctly classified example and the decision surface.

if (member$_1$ says that class is c) or ... or (member$_K$ says that class is c) then class is c,

where the asserts "(member$_k$ says that class is c)" have a degree of certainty b_{ck} determined by the rules in the sublist $\mathcal{G}_c^k$, and those asserts for which b_{ck} is zero are omitted. The fuzzy output of this selected ensemble is

$$R^I(x)(c) = \bigvee_{\{(j,k)|C_j^k=c\}} b_{ck} \cdot CF_j^k \cdot A_j^k(x). \qquad (2)$$

We can define more powerful linguistic selections which extend this basic fuzzy reasoning schema. In this paper we will use a *sparse* matrix of weights $[w_{ck}] \in [0,1]^{n_c \times K}$ and operate as follows:

$$R^{II}(x)(c) = \bigvee_{\{(j,k)|C_j^k=c\}} w_{ck} \cdot CF_j^k \cdot A_j^k(x). \qquad (3)$$

Thus, the selected ensemble can be seen as a hierarchical knowledge base with n_c fuzzy classification rules with weights in the antecedent part

if (member$_1(w_{c1})$ says that class is c) or ...
(member$_K(w_{cK})$ says that class is c) then class is c,

where the asserts "member$_k(w_{ck})$ says that class is c" have a certainty determined by the rules in the sublist $\mathcal{G}_c^k$, after multiplying their confidence degrees by the same factor w_{ck}:

if x is A_j^k then Class C_j^k with $w_{C_j^k k} \cdot CF_j^k$.

Again, those rules where $w_{C_j^k k} = 0$ are omitted.

In this case, any of these hierarchical rule bases we have introduced is univocally determined by a matrix $[w_{ck}]$. Therefore, the genetic search of the best selection involves finding the best matrix $[w_{ck}]$, according to certain criteria that will be defined next. Notice that, this search is a selection, because $[w_{ck}]$ is a sparse matrix. As we will detail later, in this contribution the number of terms w_{ck} different than zero is a design parameter.

B. Fitness function

We propose that the quality of a selected and combined fuzzy ensemble is defined by three components (e, m_1, m_2) (see Figure 1), thus the fitness of a possible FRBCS-CM design is a triplet comprising three real numbers:

1) Training error e: we compute the error of each ensemble for a large number of bootstrapped resamples of the training set, and use a quantile of the distribution of these errors as the first term of the fitness. This is intended to avoid overfitting when there are outliers in the training set, and also to detect the most robust selections, which are expected to generalize better.
2) Error margin m_1 : the second component of the fitness function depends on the distance between the misclassified examples and their nearest decision surface. Given an example x, we have approximated this value by the difference between the highest and the second highest term of $R^{II}(x)(c)$, and defined that the error margin of an ensemble is the worst (i.e. the highest) value of this difference for any example x in the training set.
3) Classification margin m_2: the third component depends on the distance between the correctly classified instances and their nearest decision surface, which is approximated as before, by the difference between the highest and the second highest terms in $R^{II}(x)(c)$. In this case, however, the margin of an ensemble is the lowest value of this difference for all the examples of the training set; we seek a decision surface with the highest margin.

A lexicographical ordering is defined between two triplets: $(e, m_1, m_2) \prec (e', m_1', m_2') \iff$

$$\begin{cases} (e < e') \\ (e = e') \text{ and } m_1 < m_1' \\ (e = e') \text{ and } (m_1 = m_1') \text{ and } (m_2 > m_2') \end{cases} \qquad (4)$$

C. Coding scheme, genetic operators and evolutionary scheme

An individual is an sparse matrix $[w_{ck}]$, which will be stored as two fixed-length ordered lists of indexes (c, k) and their corresponding values w_{ck}, as displayed in Figure 2. The chromosome length is defined according to the maximum percentage of non-zero values in the matrix, which a parameter whose value is set by the user in advance. The initial population is randomly generated. We have decided to apply an arithmetic crossover [31] between the lists of values of both individuals, leaving the lists of indices unchanged. The mutation operator randomly alternates a nonuniform mutation of an element of the list of values [32] and the random generation of a completely new individual.

Lastly, since the fitness function is not scalar, we have decided to implement a tournament-based steady state GA

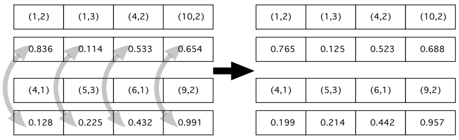

Fig. 2. Coding scheme and crossover operation: an individual is a sparse matrix, which is represented by a list of indices and a list of values.

[33], where at each generation the two last elements in each tournament are replaced by the offspring of the two winners. This offspring is the result of the application of the crossover operator mentioned before, followed by a mutation with certain probability.

IV. EXPERIMENTS AND ANALYSIS OF RESULTS

This section is devoted to validate our new fuzzy linguistic combination method proposal. While the first subsection introduces the experimental setup considered, the second one shows the results obtained in the experiments developed and their analysis.

A. Experimental setup

To evaluate the performance of the FRBCS-CM in the ensembles generated, nine popular data sets from the UCI machine learning repository have been selected (see Table I). The number of features ranges from a small value (5) to a large one (60), while the number of examples does so from 208 to 5,477.

TABLE I
DATA SETS CONSIDERED

Data set	#attr.	#examples	#classes
Sonar	60	208	2
Vehicle	18	846	4
P-Blocks	10	5,477	5
Pima	8	768	2
Glass	9	214	7
Breast	9	699	2
Yeast	8	1484	10
Heart	13	270	2
Phoneme	5	5,404	2

In order to compare the accuracy of the considered classifiers, we used Dietterichs 5×2-fold cross-validation (5×2-cv) [34]. The bagging FRBMCS generated are initially comprised by 50 classifiers. The granularity and the number of features used to derive them are both equal to 5. The GA for the FRBCS-CM derivation works with a population of 100 individuals and runs for 2,000 generations (the equivalent to 40 generations in a usual GA with generational replacement and crossover probability equal to 1). The tournament size is 5 and the mutation probability considered is 0.1. Five different values have been considered for the chromosome size: 10, 25, 50, 75, and 90% of the terms of $[w_{ck}]$ matrix were allowed to be non-zero, reporting the best choice in each case. All the experiments have been run in a HP Z600 workstation (Intel eight-core Pentium 2.46 GHz computer with 6 GBytes of memory), under the Mac OS X operating system.

B. Experiments developed

The mean values and standard deviations of the test error are shown in Table II. Observe that the fuzzy ensembles using new fuzzy linguistic combination method have improved the initial ensembles (composed of 50 individual classifiers generated as explained in Sec. II-C without applying any classifier selection mechanism) considering the classical voting combination in both accuracy and complexity in all cases but one. The new FRBCS-CM method outperforms the voting one in Sonar, Vehicle, P-Blocks, Pima, Glass, Yeast, Heart and Phoneme, but it was unable to obtain a linguistically meaningful aggregation in Breast.

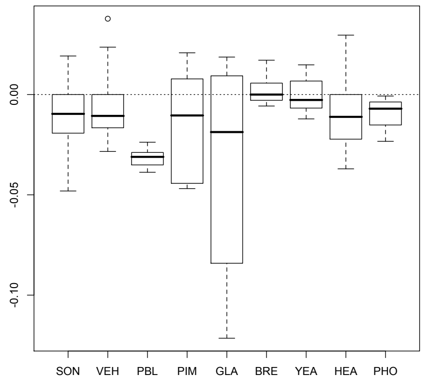

Fig. 3. Results obtained by the fuzzy linguistic combination method.

TABLE II
COMPARED TEST ERROR

Data set	Original FRBMCS	Linguistically selected FRBMCS	%reduction
Sonar	0.2048 ± 0.030	**0.1932 ± 0.028**	75%
Vehicle	0.3570 ± 0.022	**0.3530 ± 0.020**	75%
P-Blocks	0.0896 ± 0.005	**0.0583 ± 0.006**	90%
Pima	0.2628 ± 0.021	**0.2500 ± 0.019**	90%
Glass	0.3879 ± 0.082	**0.3486 ± 0.049**	75%
Breast	**0.0323 ± 0.005**	0.0349 ± 0.006	10%
Yeast	0.4341 ± 0.022	**0.4334 ± 0.021**	75%
Heart	0.1719 ± 0.034	**0.1607 ± 0.031**	50%
Phoneme	0.2137 ± 0.006	**0.2042 ± 0.008**	90%

The statistical relevance of the differences is graphically shown in Figure 3, where we have displayed the 9 boxplots of the paired differences between the results the linguistic selection we have introduced in this paper and those of the bagging-based ensemble. Negative values signal an advantage to our method. The median of the difference has always been negative, but in Breast, where there is a tie between both methods.

The highest statistical significances are in Phoneme (p-value of 0.01 with a Wilcoxon test) and P-Blocks (p-value $< 10^{-5}$). Coincidentally, these are the hardest problems of our benchmark. This may indicate that our FRBCS-CM will be able to improve the accuracy of the original fuzzy ensemble in those problems where bagging a random selection of fuzzy classifiers does not produce optimal results. Our bootstrap-based estimation of the training error has prevented an excessive overfitting and has also proven effective detecting those individual FRBCSs that should be removed, in those cases there is actually room for an improvement. Not being able to improve the results of Breast points in the same direction: in this dataset, the GFRBCS converged to ensembles with a null training error, thus all the resamples of the training sets will also have null error and the quantile of the bootstrap distribution does not give more information than a central moment of the training data. The error margin is null, either, so further improvements in the learning are guided by the third term of our fitness function, the classification margin. In this case, the decreased performance attributable to the linguistic interpretability of our ensemble (i.e., the use of a t-norm instead of a sum for combining the classifiers) cannot be compensated by removing the worst ensemble members.

V. CONCLUDING REMARKS AND FUTURE WORKS

We have proposed a novel MCS fuzzy combination method based on the use of a FRBCS automatically derived by means of a GA. The new fuzzy linguistic combination method shows very interesting characteristics, especially its transparency and its capability to jointly perform classifier fusion and selection. In addition, when considered with a fuzzy classifier ensemble, the overall system shows a hierarchical structure, thus making FRBCSs able to deal with high dimensional problems avoiding the curse of dimensionality.

Although the preliminary experiments developed clearly showed the new proposal is very promising, our next steps will be headed to perform a wider experimentation comparing the introduced fuzzy linguistic combination method with some other crisp and fuzzy combination methods, with other OCS classifier selection methods, and with hybrid methods considering both classifier selection and fuzzy classifier fusion (such as [26]). Those results will allow us to validate the actual performance of this novel ensemble fuzzy combination framework.

REFERENCES

[1] L. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. Wiley, 2004.

[2] T. Ho, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[3] D. Optiz and R. Maclin, "Popular ensemble methods: An empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 1999.

[4] O. Cordón, A. Quirin, and L. Sánchez, "A first study on bagging fuzzy rule-based classification systems with multicriteria genetic selection of the component classifiers," in *Third IEEE International Workshop on Genetic and Evolving Fuzzy Systems (GEFS)*, Witten-Bommerholz, 2008, pp. 11–16.

[5] ——, "On the use of bagging, mutual information-based feature selection and multicriteria genetic algorithms to design fuzzy rule-based classification ensembles," in *Eighth International Conference on Hybrid Intelligent Systems (HIS)*, Barcelona, 2008, pp. 549–554.

[6] O. Cordón and A. Quirin, "Comparing two genetic overproduce-and-choose strategies for fuzzy rule-based multiclassification systems generated by bagging and mutual information-based feature selection," *International Journal of Hybrid Intelligent Systems*, 2010, in press.

[7] B. Dasarathy and B. Sheela, "A composite classifier system design: Concepts and methodology," *Proceedings of IEEE*, vol. 67, no. 5, pp. 708–713, 1979.

[8] G. Giacinto and F. Roli, "Dynamic classifier selection based on multiple classifier behaviour," *Pattern Recognition*, vol. 34, no. 9, pp. 1879–1881, 2001.

[9] D. Partridge and W. Yates, "Engineering multiversion neural-net systems," *Neural Computation*, vol. 8, no. 4, pp. 869–893, 1996.

[10] L. Lam and C. Suen, "Application of majority voting to pattern recognition: An analysis of its behavior and performance," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 27, pp. 553–568, 1997.

[11] A. Verikas, A. Lipnickas, K. Malmqvist, M. Bacauskiene, and A. Gelzinis, "Soft combination of neural classifiers: A comparative study," *Pattern Recognition Letters*, vol. 20, no. 4, pp. 429–444, 1999.

[12] L. Kuncheva, ""Fuzzy" versus "nonfuzzy" in combining classifiers designed by boosting," *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 6, pp. 729–741, 2003.

[13] M. Cococcioni, B. Lazzerini, and F. Marcelloni, "A TSK fuzzy model for combining outputs of multiple classifiers," in *2004 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS'04)*, Banff, 2004, pp. 871–875.

[14] O. Cordón, M. del Jesus, and F. Herrera, "A proposal on reasoning methods in fuzzy rule-based classification systems," *International Journal of Approximate Reasoning*, vol. 20, pp. 21–45, 1999.

[15] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling With Linguistic Information Granules*. Springer, 2005.

[16] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena, *Genetic Fuzzy Systems. Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*. World Scientific, 2001.

[17] O. Cordón, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems: Current framework and new trends," *Fuzzy Sets and Systems*, vol. 141, no. 1, pp. 5–31, 2004.

[18] J. Casillas, F. Herrera, R. Pérez, M. del Jesus, and P. Villar, "Special issue on genetic fuzzy systems and the interpretability-accuracy trade-off," *International Journal of Approximate Reasoning*, vol. 44, no. 1, January 2007.

[19] K. Woods, W. Kegelmeyer, and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 405–410, 1997.

[20] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–238, 1998.

[21] E. Dos Santos, R. Sabourin, and P. Maupin, "A dynamic overproduce-and-choose strategy for the selection of classifier ensembles," *Pattern Recognition*, vol. 41, no. 10, pp. 2993–3009, 2008.

[22] B. Gabrys and D. Ruta, "Genetic algorithms in classifier fusion," *Applied Soft Computing*, vol. 6, no. 4, pp. 337–347, 2006.

[23] N. Dimililer, E. Varoglu, and H. Altincay, "Classifier subset selection for biomedical named entity recognition," *Applied Intelligence*, 2009, in press.

[24] S.-B. Cho and J. Kim, "Multiple network fusion using fuzzy logic," *IEEE Trans. on Neural Networks*, vol. 6, no. 2, pp. 497–501, 1995.

[25] M. Abreu and A. Canuto, "An experimental study on the importance of the choice of the ensemble members in fuzzy combination methods," in *Seventh International Conference on Intelligent Systems Design and Applications (ISDA)*, Rio de Janeiro, 2003, pp. 723–728.

[26] P. Bulacio, S. Guillaume, E. Tapia, and L. Magdalena, "A selection approach for scalable fuzzy integral combination," *Information Fusion*, 2010, in press.

[27] Y. Lu and F. Yamaoka, "Fuzzy integration of classification results," *Pattern Recognition*, vol. 30, no. 11, pp. 1877–1891, 1997.

[28] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[29] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Transactions on Neural Networks*, vol. 5, no. 4, pp. 537–550, 1994.

[30] K. Trawinski, O. Cordón, and A. Quirin, "On the combination of accuracy and diversity measures for genetic selection of bagging fuzzy rule-based multiclassification systems," in *Nineth International Conference on Intelligent Systems Design and Applications (ISDA)*, Pisa, 2009.

[31] F. Herrera, M. Lozano, and J. Verdegay, "Tackling real-coded genetic algorithms: Operators and tools for the behaviour analysis," *Artificial Intelligence Review*, vol. 12, pp. 265–319, 1998.

[32] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*. Springer, 1996.

[33] A. Eiben and J. Smith, *Introduction to Evolutionary Computing*, 2nd ed., ser. Natural Computing Series. Springer, 2007.

[34] T. Dietterich, "Approximate statistical test for comparing supervised classification learning algorithms," *Neural Computation*, vol. 10, no. 7, pp. 1895–1923, 1998.