A Review on Fuzzy Classifier Design

Arvind Kumar[1] Amandeep Kaur[2]
[1,2]Department of Computer Science & Engineering
[1,2]Central University of Punjab Bathinda (Punjab), India

Abstract— Image classification is an important stage in many image processing applications. Fuzzy classifiers perform well in real world situations where input vs. Output class membership overlapped and no crisp output is easily available. In this report, a survey of existing design techniques for fuzzy rule based classifiers is done. In addition to this various approaches for classifier performance enhancement and application of fuzzy classifiers to various domains have been discussed. Paper discusses with the development of genetic algorithm and boosting etc. methods new dimensions are being opened in field of fuzzy classification.

Key words: Fuzzy Logic, Fuzzy Classifier, Classification, CBIR

I. Introduction

The relatively young Fuzzy logic theory's[1] main benefit is that it allows the natural description, in linguistic terms, of problems that better to be solved in terms of relationships between variables, their precise numerical values. So, the idea of using fuzzy rule-based systems over non-linear systems are their easy interpretation to user. Usually fuzzy rule-based systems are readily understood via linguistic accessment of fuzzy rule base in a human gettable manner. For designing fuzzy rule-based systems, the emphasis has often been placed on both their accuracy and interpretability[2]. Hence, one has to find crucial accuracy vrs. Interpretablity tradeoff.There are a number of approaches [2]–[9]to improve both the accuracy and the interpretability of fuzzy rule-based systems. In these approaches, fuzzy rule-based system design can be viewed as a single-objective problem for optimization. In the field of image processing, one can use these classifiers to classify image data(any other type of digital data), in such a way that resultant classes are clearly represented with the resulting image. A classifier is an algorithm that assigns a class label to input, based its description. One can say that classifier *predicts* the class label. The input image can be represented as a vector having features such as LBP, HOG,SIFT and SURF (attributes) considered good for the classification task. Typically, the classifier learns to predict class labels using a training dataset and a training algorithm. When a training dataset is not available, a classifier can be designed from prior knowledge and experts knowledge. Once trained, the classifier is ready to operate on unseen digital data but data should be first modeled according to classification model.

Here questions arise that can we use the fuzzy logic technique for classification and reasoning to decrease the efforts of a person dealing with supervised classification? Can we eliminate the prejudice?These are the questions that lead to designing a fuzzy rule based classification system. Fuzzy classifiers based on if-then rules might be "transparent" or "interpretable", the end user (expert) is able to verify the classification paradigm. For example, such verification may be done by an expert judging the plausibility, consistency or completeness of the rule-base in fuzzy if-then classifiers. This verification is appropriate for small-scale systems, i.e., systems which do not use a large number of input features and significantantly large rule bases. Many other classification techniques are also available such as a hierarchical quantizer in the form of a decision tree. Such tree is a kind of an approximate nearest neighbor algorithm and constitutes a visual dictionary. Recently, the bag-of-features (BoF) approach has gained in popularity. In the BoF method,clustered vectors of image features are collected and sorted by their (histograms)count of occurrence.The main problem with the afore mentioned techiques is comparision with all approximations of sets of descriptors or individual descriptors presented in the histogram form. Such calculations are very expensive and tedious computationally. Moreover, the BoF approach needs redesign of the classifier structure whenever some new visual classes are added to the system.

A fuzzy rule based classification system consists of input dataset of crisp, linguistic values and knowledge base that consitutes the membership functions,image feautres and rule base. It has fuzzification,inference engine and defuzzification phases. There are several issues that needs to be addressed when constructing a classifier. Let D = {x|x = (x1,...,xn)}∈n be a data set, and L = {1,...,m} a set of labels. To construct a classifier is to obtain a rule (a mapping) f : n → L, which assigns to each data point x ∈ D a label l ∈ L. A special case, referred to as a 2-class classifier, is when the number of labels is m = 2. We base our discussion on the 2-class classifier. Often the general, m-class classifier, is obtained from the simpler case. Deriving the rule f is known as the training of the classifier, and that is done based on a training set T ⊂ D × L. A training tuple is of the form (x,l). Some important choices that are made at this phase like deriving the Fuzzy Sets, selection of the Decision Rule (Class Boundary), best Attribute Selection that helps in dimensionality reduction, Aggregation across Several Attributes when weak classifiers are ensemble via boosting and finally measures for classifier evaluation where some of these are merely checking accuracy by help of classification results can be faulty but confusion matrix and precision-recall curve provide better evaluation results[8].

II. Review of Literature

In past years a lot of research has been done on fuzzy classifiers and their applications in different fields. In [5]authors propose an algorithm for designing a fuzzy classification system using adaboost for meta-learning that takes less learning and testing time as compared to a traditional bag of features image representation model over the support vector machine classification based on chi-square kernel. According to authors, fuzzy classifier gives better classification accuracy and adding a new class of visual objects to the existing system is easier while in the BoF we have to recreate the whole dictionary. The most timeconsuming part of the bag-of-features classification is the SVM learning is replaced by creating fuzzy rules based on most salient image features. In [8] authors have noted various key issues in designing a fuzzy classifier. Some of them are:

A. Selection of the Decision Rule (Class Boundary)

If the training set is viewed as a subset of a high dimensional space (≤ n) the decision rule can be viewed as a surface that divides the training set into two subsets each of which corresponding to one class. The usual approach is to parameterize f and to select its parameters by minimizing the overall misclassification errors. If the training data happens to be linearly separable, f is a linear surface. In the case when this is not the case, linear separability can be obtained by using a kernel to achieve implicit mapping into a higher dimensional space where data are linearly separable. In all but the most simple cases when classes are well separated to begin with, errors of classification of the training data are tolerated in order to improve correct classification of test data.

B. Derivation of Fuzzy Sets for a Fuzzy Classifier

It can be done via initialization-plus-tuning approach, whereby the fuzzy sets in questions are adjusted during the process of training the classifier. The membership functions are selected so as to make the tuning process quite easy. This approach is appealing from an intuitive point of view, and has the advantage that the resulting fuzzy set can be easily expressed by a linguistic label. However, not all classification problems deal with classes whose intuitive meaning is grasped from the beginning. In such cases the fuzzy sets can be obtained directly from the data, either by a clustering procedure [10], or using the mass assignment theory, an approach especially useful for training a classifier from imbalanced data.

C. Attribute Selection and Error modelling

Attribute Selection and Error modelling are some other issues acc. to paper. Often attribute selection for fuzzy classifiers is done in the same way as for traditional classifiers. There is a notable exception in [10], in which a regularity criterion is used in conjunction with a fuzzy model to select the best attributes. Although, the approach does not result in best attributes as it stops at a local minimum of the prediction error. Fuzzy techniques gives us the way to find a fuzzy set of attributes where the membership degree of an attribute is directly proportional its importance for classification. The challenge is to derive a formal (as opposed to some ad-hoc weighting of attributes) technique for obtaining such a set. According to literature models of checking error and classifier performance should be chosen efficiently as precision and recall, and the associate Fα measure can be, in theory, easily generalized using fuzzy sets (using Zadeh's extension principle) providing us with tools to further distinguish between classes and classifier results. However, the challenge here is to develop an approach which is at the same time technically correct and computationally efficient. In [7]an approach for fuzzy feature selection is given when one has to classify high dimensional data. Authors have reviewed a lot of literature on fuzzy feature selection and fuzzy classifier design techniques. According to them, the feature selection methods can be grouped in three categories: filters, wrappers and embedded models. Filters are used to rank all features via a preprocessing stage and then select the best ones. In wrappers, some feature sets are selected and then evaluated on the designed classifiers. The embedded methods, however, are specific to the selected learning machines and the process of feature selection is done in their training step. While designing fuzzy classifier they proposed a class specific approach for selecting relevant features instead of using same global features for all classes. Method combines the interclass distance concept along with the compatibility degree of data in some predefined fuzzy sets on each feature to evaluate that feature. In their experiment they evaluated their classifier's performance with other design techniques such as SGERD, a steady state genetic algorithm for extracting fuzzy classification rules from data. FARC-HD, a fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. In [6]a classifier based on fuzzy if–then rules that uses weighted training patterns to adjust the sensitivity of classification task accordingly for certain classes. The consequent class and the degree of certainty are determined from the compatibility and weights of training patterns while the antecedent part of fuzzy if–then rules are given by partitioning each attributes into fuzzy sets. Paper also introduces a learning method which adjusts the degree of certainty in order to provide improved classification performance and reduced costs. In the end paper evaluate its performance over the Breast cancer diagnosis images and satellite images for classification which shows significantly improved classification rate. In[2] fuzzy rule-based classifier that can visually present its classification results to users i.e. fuzzy rule-based classifier can explain to users why an input pattern is classified as a resultant class in a recognizable manner. The proposed approach consists of a rule selection method and a visualization interface. The basic idea is to use only two antecedent conditions with a genetic algorithm designed to choose only useful rules for classification. For classification, a single winner rule based technique is employed on input pattern and results are shown in two-dimensional space, where winner rule is one with maximum product of compatibility grade and rule weight in a fuzzy set. After generating rules their pre-screening is done on the basis of two measures such as support and confidence of each rule. A minimum threshold value is set for each of this and finally using a single objective function in genetic selection for final candidate rule base .i.e. weighted sum fitness fitness(S) = w1.Accuracy(S) - w2 .Complexity(S),

Where w1,w2 are non-negative weights, Accuracy (S) and complexity(S) are accuracy and complexity measures for fuzzy rule set S.

As fuzzy classifiers have significant contribution in the field of remote sensing, in [11] describe a comprehensive approach for designing fuzzy rule based classifiers. In this multistage scheme, first stage is used to develop a set of labelled prototypes representing the distribution of the training data and is generated using a self-organizing feature map (SOFM)[12], [13].A combination of unsupervised and supervised clustering on the training data is used to generate desired set of prototypes representing training data distribution globally as well as locally (i.e. class specific) distribution of the training data. After prototype creation, prototypes are converted into form of a fuzzy rule. Thus, each of the fuzzy rules represents a region, may be overlapped, in the feature space. While in previous work [14] simulated annealing is used for optimizing the performance over the randomly generated initial set of fuzzy rule. Here number of prototypes depends on the complexity of training data. For context tuning, two variant of algorithms are used where conjunction operator used between rules is product and other where softmin is used. According to authors performance may further improve when other T-norms are used for context tuning and if features other than gray levels are used.

In [15] novel method to design an online fuzzy rules repository which can classify multiple type of data. This scheme of on-line classification based on fuzzy rules with an evolving structure is introduced in this paper. The fuzzy classifier can begin either ‘from scratch’ from learning and adapting to the new data samples or from an already existing rule base that can be updated based on the new information extracted from the new data samples. Scheme can be helpful in real-time applications such as, robotic applications, classification streaming data e.g. target and landmark recognition, fault detection and diagnostics, real-time machine health monitoring and prognostics etc. Each processed prototype is a data sample that represents the focal point of a fuzzy rule per class and is selected on the basis of data density. In [16]applications of fuzzy classifiers in field of CBIR have been studied and it states that theoretical method presented by Ishibuchi does not answer the question on how to construct membership functions, especially those corresponding to linguistic values but in [17]gave a suggestion for the construction of membership functions based on the standardized residual analysis but they applied it to continuous data. In [9] problem with usual ways (i.e. sampled input-output and expert human controller) for learning fuzzy rules is that information is incomplete. Here authors suggested to use both linguistic fuzzy rules and neural numerical data pairs combination. The key idea for this new approach is to design fuzzy rules from numerical data pairs. Then combine these fuzzy rules and linguistic fuzzy rules in a common fuzzy rule base.

There are numerous literature [4], [18]–[20]where significant work is done on improving fuzzy classifiers complexity, compatibility, performance and implementing these classifiers for real world problems.

III. FUZZY IF THEN SYSTEMS

A fuzzy if-then system has n inputs ($X = X_1, \ldots X_n$) and c outputs ($Y = Y_1, \ldots\ldots\ldots Y_n$). Here are three popular acronyms for fuzzy and also non fuzzy systems.

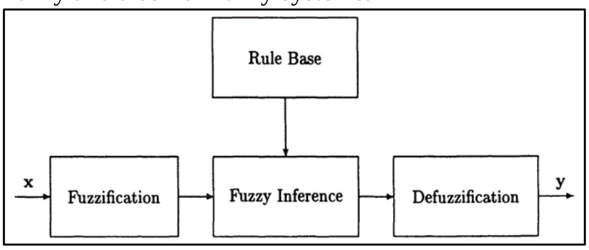

Fig. 1: Operation of a fuzzy if then system

– SISO: Single input-single output systems (n =c = 1).
– MISO: Multiple input - single output systems (n > 1, c = 1)
– MIMO: Multiple input - multiple output systems (n > 1, c > 1).

When a vector x is submitted to the input, the system operates as shown in Figure 1. The input is "fuzzified" and submitted to the fuzzy inference block. A rule base is used to calculate a fuzzy decision which is subsequently "defuzzified" to get the final output.

A. Fuzzification

The first step is the fuzzification process is a compact expression for "finding the degree of match of the input to a set of linguistic terms". Let $A = \{A_{i,1} \ldots ,A_{i,K_i}\}$ be a set of linguistic terms for the i-th component Xi of the input x. For example, the feature age can be described by {very young, young, middle-aged, old}, the feature profit, by {small, reasonable, very good, appreciable, huge}, etc. Each of the linguistic terms is represented as a fuzzy set on the set of possible values of Xi. Fuzzification replaces Xi by a set of Ki numbers $\mu_{A_{i,1}}(X_i), \ldots\ldots\ldots, \mu_{A_{i,ki}}(X_i)$ in [0,1], showing how well the value Xi matches each of the Ki linguistic terms.

B. Fuzzy Inference

The inference engine uses a set of M fuzzy if-then rules (a fuzzy rule base), built up in advance. The rule base can be specified by a domain expert or extracted from input-output data pairs. Each rule has an antecedent part (if-part, or premise), and a consequent part (then part, or consequent). The antecedent part is a Boolean expression of simple clauses on the individual features X_1, ... , X_N. A simple clause is, for example, "X_i is small'. There are two basic types of fuzzy if-then systems:

– The Mamdani-Assilian (MA) model (the logical model). In MA systems, both the input and the output are represented by linguistic terms. The antecedent and the consequent of an if-then rule are typically Boolean expressions of simple clauses.
– The second type of fuzzy systems is the Takagi-Sugeno-Kang (TSK) model (the functional model). The antecedent part of each rule is again a Boolean expression of simple clauses but the consequent is a function of the input x (most often a polynomial).

In Mamdani-Assilian (MA) fuzzy systems, the antecedent part in this model is a conjunction of simple clauses. For the time being we shall assume that all features are used in each rule (this means that the antecedent part is a conjunction of n clauses). The k-th rule in the database of an MIMO fuzzy if-then system has the following general form:

$$\begin{aligned}&R_k : \text{IF } x_1 \text{ is } A_{1,i(1,k)} \text{ AND } x_2 \text{ is } A_{2,i(2,k)} \text{ AND } \ldots \text{ AND } x_n \text{ is } A_{n,i(n,k)}\\&\text{THEN } y_1 \text{ is } B_{o(1,k)} \text{ AND } \ldots \text{ AND } y_c \text{ is } B_{o(c,k)},\\&\text{for } k = 1, \ldots, M.\end{aligned}$$

In this expression, the subscript i(j, k) specifies which linguistic term on the j-th feature is used in the k-th rule. The notations i(.,.) and 0(.,.) are used as index functions for the input and output part of rules.

The firing strength $\tau_k(X)$ of rule R_K, k = 1 ... , M, for an input x is calculated by a conjunction type of aggregation operation on input features or can be computed by multiplying their membership function values. Thus, the firing strength reflects to what extent x satisfies all antecedent clauses. The minimum and the product are the two most widely used operations.

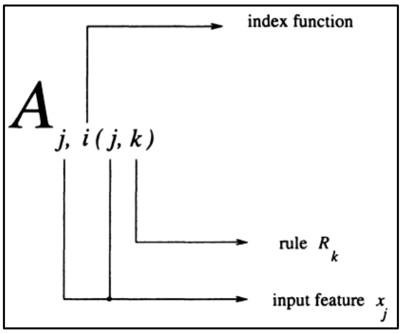

Fig. 2: Notation for the jth antecedent clause of the kth rule in a fuzzy rule-based system

Takagi-Sugeno-Kang (TSK) fuzzy systems. A MIMO Takagi-Sugeno Kang fuzzy system has the following type of rules:

R_k : IF x_1 is $A_{1,i(1,k)}$ AND x_2 is $A_{2,i(2,k)}$ AND … AND x_n is $A_{n,i(n,k)}$
THEN $\mathbf{y} = \mathbf{f}_k(\mathbf{x})$,
$k = 1, \ldots, M$.

The firing strength is calculated as in the MA model. The overall output of a TSK fuzzy system is

$$\mathbf{y}(\mathbf{x}) = \frac{\sum_{k=1}^{M} \tau_k(\mathbf{x}) \mathbf{f}_k(\mathbf{x})}{\sum_{k=1}^{M} \tau_k(\mathbf{x})}.$$

C. Defuzzification

"Defuzzification" is a procedure to calculate the single representative value of a fuzzy set. Ideally, this element should best characterize the fuzzy set. This operation is needed to find the system's output Y from the resultant fuzzy set C. The two most popular defuzzification methods are the Center of Gravity (COG), and the Mean of Maxima (MOM). If the output fuzzy sets are defined over a finite universal set $Y = \{y_1, \ldots \ldots \ldots y_S\}$, the resultant fuzzy set C is also defined on Y, and the COG defuzzified value is,

$$y = \frac{\sum_{i=1}^{S} \mu_C(y_i) y_i}{\sum_{i=1}^{S} \mu_C(y_i)}.$$

IV. General Design of Fuzzy Rule-Based Classification Systems

Consider a classification problem with a data set of m patterns, DS = {(Xp;Yp), p = 1..m}. For pth pattern, if input feature vector of variables, Xp, is n-dimensional. That is, Xp = [$xp_1, \ldots, xp_n$] with feature labels taken as {fi, i = 1, …, n}. The output variable, Yp, is a class label in M classes such that Yp ∈{$c_1, \ldots, c_M$ }. We assume that each input variable, xp_i, is rescaled to unit interval [0,1] using a linear transformation that preserves the uniformity in distribution of the data set. In [7], the classical single model architecture of fuzzy classifiers is utilized to handle the multiclass classification problems. The benefits of this model are simplicity, transparency and more interpretability of the designed classifiers[21]. The general form of fuzzy if–then rules is given as:

Rule Rj: if x_1 is A_{j_1} and. .. . . .and x_n is A_{j_n} then class C_j, for j = 1, . . . , n (1)

Where X = [$x_1, \ldots, x_n$] is an input vector and A_{j_i} for (i = 1, …, n) shows the fuzzy set on variable xi in the antecedent part of some rule Rj, Cj is the consequent class (that is, Cj∈{$c_1, \ldots, c_M$ }. and N is the number of fuzzy rules. Herein, the fuzzy rule Rj is abbreviated as Aj ⇒ class Cj where Aj = $A_{j_1} \times \cdots \times A_{j_n}$. Generally, designing a fuzzy classifier can be described as generating a set of N fuzzy rules in the form of (1).

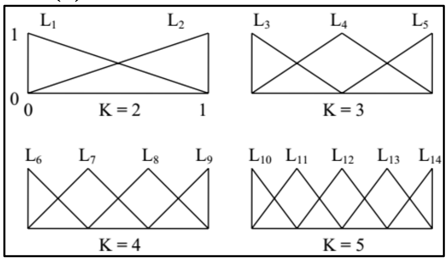

Fig. 3: Fourteen fuzzy sets of each input variable

The first step in generating fuzzy rules is partitioning the pattern space into fuzzy subspaces. If a subspace contains some patterns, a fuzzy rule will refer to it. Partitioning is usually done using K suitable membership functions. The most common type of membership functions is triangular because they are simpler and easily understandable by humans.However under some conditions , the fuzzy partitions built out of the triangular membership functions lead to entropy equalization [22]. Figure 3 shows these membership functions for four different values of K. Number of membership functions should be upto nine, used as upper bound on the fuzzy sets in fuzzy modeling techniques [23]. Various approaches for generating fuzzy classification rules have been suggested in[7] and [24]. The approach in[24] used the fuzzy set don't care (with membership function μdon′t care(x_i) = 1, ∀x_i ∈ [0, 1]) beside the 14 triangular fuzzy sets in Fig. 3. This don't care fuzzy set for a variable in the antecedent part of a rule will lead to reduction its length as that variable will be removed. The consequent class C_j of fuzzy rule R_j in (1) is determined using the patterns in the corresponding fuzzy subspace.

The compatibility grade or firing strength of training pattern Xp = [$xp_1, \ldots, xp_n$] is defined with the antecedent part Aj = $A_{j_1} \times \cdots \times A_{j_n}$ of rule R_j as:

$$\mu j(Xp) = \prod_{i=1}^{n} \mu j_i(xp_i) \quad (2)$$

Where $\mu j_i(x_i)$ is the membership function of the antecedent fuzzy set A_{j_i} on variable x_i. One of the methods for selecting the consequent class of a rule is based on confidence [24].

The confidence of the fuzzy rule Aj ⇒ class c is defined as:

$$\mathrm{Conf}(Aj \Rightarrow \text{class } c) = \frac{\sum_{Xp \in \text{class } c} \mu j(Xp)}{\sum_{p=1}^{m} \mu j(Xp)} \quad (3)$$

The consequent class Cj of fuzzy rule Rj can be obtained by calculating the class with the maximum confidence as:

$$Cj = \arg_c \max\{\mathrm{Conf}(Aj \Rightarrow \text{class } c) | c \in \{c_1, \ldots c_M\}\} \quad (4)$$

In, some heuristic measures have been used for evaluating the candidate rules.The basic criterion is difference between positive and negative samples. Its fuzzy version is specifed as:

$$\mathrm{Eval}(Aj \Rightarrow \text{class } Cj) = \sum_{Xp \in c} \mu j(Xp) - \sum_{Xp \notin c} \mu j(Xp) \quad (5)$$

A Single winner (that is, winner-takes-all approach) is the most popular reasoning method in fuzzy rule-based classifers [25] because of its simplicity and intuition for human users. Using this method, a new pattern x_t = {$x_{t_1}, \ldots, x_{t_n}$} is classified according to the consequent class of the winner rule R_w. Indeed, the winner rule has the maximum compatibility grade with x_t among the fired rules. This can be stated as:

$$\mu w(x_t) = \max\{\mu j(x_t), j = 1, \ldots, n\} \quad (6)$$

Where $\mu j(x_t)$ is the compatibility grade of rule Rj with pattern in x_t (2).

V. Current Hot Issues and Future Directions

1) Large Data Sets. (e.g., 1,000,000 patterns)
2) Imbalanced Data Sets (e.g., Class 1: 100,000 patterns, Class 2: 100 patterns)
3) Semi-Supervised Learning (e.g., Data Set = 100 labeled patterns + 9,900 unlabeled)
4) On-Line Learning (E.g. New training patterns come every minute)
5) Some other issues included are but are not limited to attribute/feature selection, efficient fuzzy set deriviation, adoption of a specific approach/algorithm, evaluation of the classifier performance, etc.

VI. Conclusion

Fuzzy rule based classifiers can over-perform other classification techniques when it came to human interpretation and understand ability. According to reviewed literature we can say that it provide quite accurate and fast retrieval also. But still the question of information loss is there as while designing fuzzy rule base incomplete information [9] can be a problem. Recent approaches of learning classification rules via genetic algorithms [2], efficient fuzzy feature selection [7] and boosting[5] in case of high dimensional data and context sensitive inferencing [11] can be helpful in enhancing classification performance.

References

[1] L. a. Zadeh, “Fuzzy sets,” Inf. Control, vol. 8, no. 3, pp. 338–353, 1965.
[2] H. Ishibuchi, Y. Kaisho, and Y. Nojima, “Designing fuzzy rule-based classifiers that can visually explain their classification results to human users,” 2008 3rd Int. Work. Genet. Evol. Fuzzy Syst. GEFS, no. March, pp. 5–10, 2008.
[3] Y. S. Avrithis and S. D. Kollias, “Fuzzy image classification using multiresolution neural networks with applications to remote sensing,” in Proceedings of 13th International Conference on Digital Signal Processing, 1997, vol. 1, pp. 261–264.
[4] S. Y. Ho, H. M. Chen, S. J. Ho, and T. K. Chen, “Design of Accurate Classifiers With a Compact Fuzzy-Rule Base Using an Evolutionary Scatter Partition of Feature Space,” IEEE Trans. Syst. Man, Cybern. Part B Cybern., vol. 34, no. 2, pp. 1031–1044, 2004.
[5] M. Korytkowski, L. Rutkowski, and R. Scherer, “Fast image classification by boosting fuzzy classifiers,” Inf. Sci. (Ny)., vol. 327, pp. 175–182, Jan. 2016.
[6] T. Nakashima, G. Schaefer, Y. Yokota, and H. Ishibuchi, “A weighted fuzzy classifier and its application to image processing tasks,” Fuzzy Sets Syst., vol. 158, no. 3, pp. 284–294, 2007.
[7] E. G. Mansoori and K. S. Shafiee, “On fuzzy feature selection in designing fuzzy classifiers for high-dimensional data,” Evol. Syst., vol. 7, no. 4, pp. 255–265, Dec. 2016.
[8] A. Ralescu and S. Visa, “Fuzzy Classifiers – Opportunities and Challenges –,” in Scalable Uncertainty Management, Dayton,USA: Springer Berlin Heidelberg, 2011, pp. 75–80.
[9] L. X. Wang and J. M. Mendel, “Generating Fuzzy Rules by Learning from Examples,” IEEE Trans. Syst. Man Cybern., vol. 22, no. 6, pp. 1414–1427, 1992.
[10] M. Sugeno and T. Yasukawa, “A fuzzy-logic-based approach to qualitative modeling,” IEEE Trans. Fuzzy Syst., vol. 1, no. 1, p. 7, Feb. 1993.
[11] N. R. Pal, A. Laha, and J. Das, “Designing fuzzy rule based classifier using self-organizing feature map for analysis of multispectral satellite images,” Int. J. Remote Sens., vol. 26, no. 10, pp. 2219–2240, 2009.
[12] A. Laha and N. R. Pal, “Some novel classifiers designed using prototypes extracted by a new scheme based on self-organizing feature map,” IEEE Trans. Syst. MAN Cybern. PART B- Cybern., vol. 31, no. 6, pp. 881–890, 2001.
[13] A. Laha and N. R. Pal, “Dynamic generation of prototypes with self-organizing feature maps for classifier design,” Pattern Recognit., vol. 34, no. 2, pp. 315–321, 2001.
[14] A. Bardossy and L. Samaniego, “Fuzzy rule-based classification of remotely sensed imagery,” IEEE Trans. Geosci. Remote Sens., vol. 40, no. 2, pp. 362–374, 2002.
[15] P. Angelov, X. Zhou, and F. Klawonn, “Evolving Fuzzy Rule-based Classifiers,” in 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing, 2007, no. Ciisp, pp. 220–225.
[16] T. Jaworska, “Application of fuzzy rule-based classifier to CBIR in comparison with other classifiers,” in 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2014, pp. 119–124.
[17] A. Hamilton-Wright and D. W. Stashuk, “Constructing a Fuzzy Rule Based Classification System Using Pattern Discovery,” in NAFIPS 2005 - 2005 Annual Meeting of the North American Fuzzy Information Processing Society, 2005, pp. 460–465.
[18] J. K. Dash, S. Mukhopadhyay, and R. Das Gupta, “Content-based image retrieval using fuzzy class membership and rules based on classifier confidence,” IET Image Process., vol. 9, no. 9, pp. 836–848, 2015.
[19] H. Ishibuchi, T. Nakashima, and T. Morisawa, “Simple fuzzy rule-based classification systems perform well on commonly used real-world data sets,” 1997 Annu. Meet. North Am. Fuzzy Inf. Process. Soc. - NAFIPS (Cat. No.97TH8297), pp. 251–256, 1997.
[20] H. Ishibuchi, T. Nakashima, and T. Murata, “A fuzzy classifier system that generates fuzzy if-then rules for pattern classification problems,” Proc. 1995 IEEE Int. Conf. Evol. Comput., vol. 2, pp. 759–764, 1995.
[21] E. Lughofer and O. Buchtala, “Reliable all-pairs evolving fuzzy classifiers,” IEEE Trans. Fuzzy Syst., vol. 21, no. 4, pp. 635–641, 2013.
[22] W. Pedrycz, “Why triangular membership functions?,” Fuzzy Sets Syst., vol. 64, no. 1, pp. 21–30, 1994.
[23] M. J. Gacto, R. Alcalá, and F. Herrera, “Interpretability of linguistic fuzzy rule-based systems: An overview of

interpretability measures," Inf. Sci. (Ny)., vol. 181, no. 20, pp. 4340–4360, Oct. 2011.

[24] H. Ishibuchi and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining," vol. 141, pp. 59–88, 2004.

[25] H. Ishibuchi, T. Nakashima, and T. Murata, "Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems," IEEE Trans. Syst. Man Cybern. Part B, vol. 29, no. 5, pp. 601–618, 1999.

