A novel fuzzy classifier based on product aggregation operator

Ashish Ghosh *, Saroj K. Meher, B. Uma Shankar

Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India

Abstract

The present article proposes a fuzzy set-based classifier with a better learning and generalization capability. The proposed classifier exploits the feature-wise degree of belonging of a pattern to all classes, generalization in the fuzzification process and the combined class-wise contribution of features effectively. The classifier uses a π-type membership function and *product* aggregation reasoning rule (operator). Its effectiveness is verified with two conventional (completely labeled) data sets and two remote sensing images (partially labeled data sets). The proposed classifier is compared with similar fuzzy methods. Different performance measures are used for quantitative evaluation of the proposed classifier.

Keywords: Pattern recognition; Fuzzy classifier; Aggregation operators; Remote sensing images

1. Introduction

Classification of patterns [1,2] is an important area in a variety of fields including artificial intelligence [3] computer vision [4] and image analysis [5]. In such problems, if *a priori* probabilities and the conditional probability density of all classes are known, then Bayes decision theory produces optimal results [1,2], i.e., it provides minimum expected error. However, in many applications, such knowledge is not available. For these cases algorithms like maximum likelihood (ML) [1], k-nearest neighbor [1,2] and the soft computing tools like neural networks (NNs) [6,7], fuzzy sets [8–10], genetic algorithms [11] are used.

A conventional hard or non-fuzzy classifier assumes that the pattern $\mathbf{x}$ belongs to a particular class only according to the given criteria. The hard classifiers are thus easy to implement and can be used to classify the classes that are well separable, well defined and have distinct boundaries. However, these algorithms may not be useful to classify ill-defined with overlapping classes. For such problems fuzzy classifiers [8,12] are more useful as it allows imprecise class definition and recognize patterns belonging to more than one class [13,14] with varying degree of membership values. Thus the partitions in fuzzy classes are soft and gradual rather than hard and crisp. With the coming of fuzzy sets [8], many research works have been carried out for applications to pattern classification and decision making problems. The most important work done in this area includes fuzzy k-nearest neighbor algorithm by Keller et al. [15], fuzzy rule-based algorithms by Ishibuchi et al. [16], Abe and Lan [17] and fuzzy ML classifier by Wang [18]. In this regard, Pedrycz [19] provided a survey on fuzzy classification methods. Fuzzy techniques are applied successfully to various areas including land cover classification of remote sensing images [20,21]. A summary of different fuzzy classifiers and their applications are described by Kuncheva [12].

Fuzzy rule-based classification systems have become an important research area in the recent past. Many approaches have been proposed for generating and learning fuzzy *if-then* rules from numerical data for classification problems [9,10,16,17,22–24]. A comparative analysis has been made by Ishibuchi and Yamamoto [22] on heuristic criteria that are used for extracting a pre-specified number of fuzzy classification rules from numerical data. In a similar study, Abe and Lan [17] described a method of extracting fuzzy rules directly from numerical input–output data for pattern classification. These rules are extracted from numerical data by recursively resolving overlaps between two classes. Then, optimal number of input variables for the rules are determined using the number of extracted rules as a criterion. In another approach Ishibuchi and Nakashima [25] proposed to use the effects of rule weights in fuzzy rule-based classification systems. Further, Bardossy and Samaniego [23] proposed a fuzzy rule-based classification of remotely sensed images. Here they have used a simulated annealing-based optimization technique to derive the fuzzy rules from training data. A support vector learning for fuzzy rule-based classification systems is proposed by Chen and Wang [24], where they discussed the connection between fuzzy classifiers and kernel machines that establishes a link between fuzzy rules and kernels, and proposes a learning algorithm for fuzzy classifiers. Apart from these, in some approaches aggregation operators are used on the fuzzified value to get an aggregated decision on the available information [9,10,26]. Peneva and Popchev [27] described an application of different fuzzy logic operators in decision making and discussed how to enhance the ability to solve the problem of ranking or choice. However, the selection of an aggregation operator is an important issue in any decision making process. In this regard, Beliakov and Warren [28] suggested a few ways of selecting aggregation operators in fuzzy decision support systems. A large variety of aggregation operators has been proposed by Bloch [29]. She made a classification of these operators used in different data fusion theories with respect to their behavior and the classification provides a guide for choosing an operator for a given problem. In another study, a discussion is provided by Dubois et al. [26] to suggest directions for using the results of mathematical investigations in the structure of aggregation operations. The problem of information generalization in multicriteria decision making, where the information is unified by fuzzy relations, is realized with the help of aggregation operators [27]. However, the proposed classifier is different from the above mentioned approaches.

Two important aspects, namely learning and generalization capabilities, play an important role in any pattern classification problem. Intuitively, these can be achieved through featurewise information extraction, generalization in the fuzzification process and combined contribution of these information to all classes of a pattern because there is a high possibility that various valuable information for different classes may reside in features of a pattern and they supplement each other. The problem becomes more complex if the classes are overlapping and ill-defined. Keeping in view of these aspects, we have designed a classifier and highlighted the method of feature-wise extraction of information and combining/aggregating the features' information to get an improved classification. The actual parameters that are participating in the design and classification process are the membership value of the features to different classes, which in turn shows how much a feature is compatible to a class. The objective of the product aggregation is to assign a pattern to a class where all the features are useful to represent that class properly, rather than the class where only some features are representing it. These characteristics are very useful in remotely sensed image analysis. In some other real life problems this may not be true. In such cases we may use some feature selection methods to choose a set of features, all of which contribute in designing the classifier. Alternatively, we can use different aggregation rules, which may be suitable for the problem at hand. The product aggregation rule is applicable for problems where the fuzzy sets (with respect to each feature) represent properties, all of which contribute to a large extent to the desired class [12,26,29–31]. This is applicable in case of fuzzy sets only and not in case of crisp sets. Due to overlapping nature of the classes that we normally consider in real life applications, this type of decision is very much suitable for fusion of the feature-wise information.

In addition to this, we have also taken care of the generalization capability of classifiers to further enhance the classification performance. In this regard, we have proposed a fuzzy product aggregation reasoning rule-based classifier and applied on two conventional (completely labeled) data sets and two remote sensing images (partially labeled) to justify its potentiality. Various performance measurement parameters such as number of overall misclassification, percentage of overall accuracy, producer's accuracy [32], user's accuracy [32] are considered for completely labeled data sets. For remote sensing images, β index [33] of homogeneity and Xie–Beni (XB) [34] index of compactness are evaluated to validate the superiority of the proposed classifier over others. In addition to these performance measures, Kappa coefficient (KC) [35] is also estimated for completely labeled data sets to corroborate the advantage of the proposed fuzzy classifier. Experimental results showed promising and improved classification performance on the above mentioned data sets.

The objective of the present article is to demonstrate the usefulness of the fuzzy product aggregation reasoning rule in classification of remote sensing images. Therefore, we have used the spectral (band) values as feature values. Each of these feature values is used to generate class-wise membership values which are used as final features. Also we have used two conventional (completely labeled) data sets (i.e., WAVEFORM and BUPA [36]) to demonstrate the effectiveness of the proposed method.

The organization of the article is as follows. Section 2 describes the proposed classifier and discusses the advantages. Different performance measurement parameters are discussed in Section 3. A brief description on the data sets used is given in Section 4. In Section 5, a complete discussion on the implementation and results are given. Finally, the concluding remarks are given in Section 6.

2. Proposed fuzzy classifier

The proposed fuzzy classifier has three steps of operation as illustrated in Fig. 1. The first step fuzzifies the input feature vector using a π-type membership function (MF) [8] that explores the information of different features for each pattern and collects the hidden or interrelated information to provide a better classification accuracy. The advantage of using π-type MF is that it has a parameter, called fuzzifier (N), which can be tuned easily according to the requirement of the problem. This provides more flexibility for classification and hence the generalizational capability can be enhanced by selecting a proper value of N. The fuzzified feature values are then aggregated using product aggregation reasoning rule ($PARR$) in the second

Fig. 1. Block diagram of the proposed fuzzy classifier.

step to get a combined contribution of the membership value of features to a particular class. The reason behind the selection of the *PARR* is that *MIN* or *MAX RR* is not good enough for the problems with overlapping classes, where different features reserve valuable information regarding the class belongingness of a pattern. Each of these features contributes significantly to the desired class and thus the combined effect is high to represent the desired class properly [12,29,30]. It is widely known that the *PARR* provides improved aggregation results compared to MIN [9], tested on different cases, and these advantages are being exploited in the proposed classifier. In this regard, we have tried with different aggregation methods [9,10] to have a combined effect of various features and found that the product (*PROD*) and geometric mean (*GM*) perform better than other aggregation *RRs*. However, the *GM* provides the same results with more computational complexities than *PARR*. Note that the geometric mean is equivalent to the product aggregation risen to the power $1/D$ (D number of features), is a monotone transformation that does not depend on the class label c (where $c = 1, \ldots, C$) and therefore will not change the order of membership values $F_c(x)$; the winning label obtained from the product aggregation *RR* will be the same as the winning label from the geometric mean [37]. Further, the computational time difference between GM and *PARR*, for one pattern, may be small but for a full image of 512×512 with six or more classes the time difference will be very high. Thus the combination of π-type *MF* along with *PARR* can be a better classifier, which is incorporated in the proposed classifier. In the final step the output values are defuzzified using a *MAX* operation.

2.1. Fuzzification

At first the π-type membership function (*MF*) has been used to get the degree of belonging of a pattern into different classes based on different features. The membership value $f_{d,c}(x_d) = \pi_{d,c}(x_d)$, thus generated, expresses the degree of belonging of dth feature to cth class of a pattern $\mathbf{x}$, where $\mathbf{x} = [x_1, x_2, \ldots, x_d, \ldots, x_D]^{\mathrm{T}}, d = 1, 2, \ldots, D$ and $c = 1, 2, \ldots, C$. The π-type *MF* is given by

$$\begin{aligned}\pi(x; a, r, b) &= 0, \quad x \leqslant a,\\ &= 2^{N-1}[(x-a)/(r-a)]^N, \quad a < x \leqslant p,\\ &= 1 - 2^{N-1}[(r-x)/(r-a)]^N, \quad p < x \leqslant r,\\ &= 2^{N-1}[(x-r)/(b-r)]^N, \quad r < x \leqslant q,\\ &= 1 - 2^{N-1}[(b-x)/(b-r)]^N, \quad q < x < b,\\ &= 0, \quad x \geqslant b,\end{aligned} \tag{1}$$

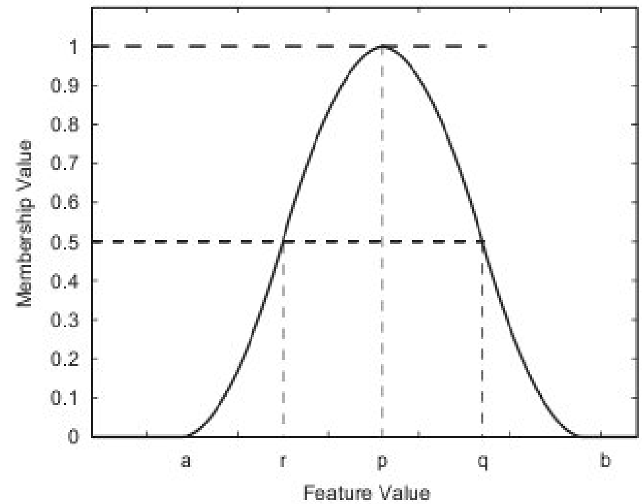

Fig. 2. π-Type membership function.

where N is the fuzzifier [8,38] of the *MF* (Fig. 2). The *MF* can be estimated with center at $r = (p+q)/2$, where p and q are the two crossover points. The membership values at the crossover points are 0.5 and at the center r its value is maximum (i.e., 1). Assignment of membership value is made in such a way that a training data gets a value closer to 1 when it is nearer to the center of *MF* and a value closer to 0.5 when it is away from the center. In the present study for one experiment (Table 3) we used different value of N to demonstrate the role of it. It is quite difficult to define the optimal value for N, but a value in the interval 2.0 to 3.0 is reasonable. The exact value of N will depend on the problem and the training set. In the present article in all our experiments we have used a value of *2* (a popular choice) for the fuzzifier N.

For the determination of the *MF*, we have considered *mean* as the center "r". Here the $mean = mean(y)$ (i.e., *mean* is the average of y), and the two crossover points p and q are estimated as $p = mean(y) - [max(y) - min(y)]/2$, and $q = mean(y) + [max(y) - min(y)]/2$, where *min* and *max* are the minimum and maximum value of the data set for a feature y. Thus for a multi-featured pattern $\mathbf{x}$, the membership matrix after the fuzzification process (evaluated by $f_{d,c}(x_d)$ for dth feature/attribute to cth class) can be expressed as

$$F(\mathbf{x}) = \begin{bmatrix} f_{1,1}(x_1) & f_{1,2}(x_1) & \ldots & f_{1,C}(x_1) \\ f_{2,1}(x_2) & f_{2,2}(x_2) & \ldots & f_{2,C}(x_2) \\ \ldots & \ldots & \ldots & \ldots \\ f_{D,1}(x_D) & f_{D,2}(x_D) & \ldots & f_{D,C}(x_D) \end{bmatrix}. \tag{2}$$

The above membership matrix provides the membership grade of the features of a pattern $\mathbf{x}$ to different classes. For performance comparison of π-type MF, we have taken Gaussian MF with its parameters (μ and σ) as it has some resemblance to π-type MF. Using Gaussian MF also the membership value for each feature to different classes (i.e., membership matrix) is computed.

2.2. Reasoning rule

Aggregation operation on fuzzy sets [9,10] are operations by which several fuzzy sets are combined in a desirable way to produce a single fuzzy set. Intersection and union (standard aggregation operators) are generally not good enough for problems where all features are contributing reasonably to the desired class and collaborate with each other in decision making process. In the present work we are representing different features/properties of data sets by fuzzy sets. Thus it will be meaningful to use an aggregation operation which takes into account the effects of various features for deciding about the class belongingness. There are various types of aggregation operations [9,10]. For the present work we have used some of them like minimum, maximum, product, average (arithmetic mean, geometric mean and harmonic mean).

In this article we propose to use a product aggregation RR (*PARR*). It is applied on the fuzzy membership matrix to get the combined membership grade and compute overall degree of belonging of features of a pattern to various classes. After applying the *PARR*, we obtain the output as a vector given by

$$F'(\mathbf{x}) = [F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_c(\mathbf{x}), \ldots, F_C(\mathbf{x})]^{\mathrm{T}}, \quad (3)$$

where $\mathbf{x}$ is a multi-featured input pattern and

$$F_c(\mathbf{x}) = \prod_{d=1}^{D} f_{d,c}(x_d), \quad (4)$$

with $d = 1, 2, \ldots, D$, $c = 1, 2, \ldots, C$.

2.3. Defuzzification operation

The last step of the proposed classifier is a hard classification and is obtained through a *MAX* operation to defuzzify the output of the PARR. Here the pattern is classified to class c with the highest class membership value. Mathematically the expression for this operation is given as

$$\forall_{j\in 1,2,\ldots,C} \quad \text{and } j \neq c, \quad F_c(\mathbf{x}) \geqslant F_j(\mathbf{x}), \quad (5)$$

where $F_j(\mathbf{x})$ is the membership value for the jth class. *MAX* defuzzification operation is normally used for the problem of classification in order to get a hard class label. The other defuzzification methods like centroid of area, mean of maximum, etc, are used in other problems (e.g., in the problem of control system [39]). However, one can use the fuzzy class label also for higher level of analysis, but in that case a normalization of the output may be necessary.

3. Performance measurement indices

To examine the practical applicability of the proposed classifier, various performance measures are used for conventional data sets. These are number of misclassification (*MC*), percentage of overall accuracy (*PA*), producer's and user's accuracy and Kappa coefficient (*KC*). The *MC* value in the classification process is the number of overall samples/patterns that are wrongly classified. The *PA* value is the percentage of samples that are correctly classified. The *MC* and *PA* parameters are calculated with respect to the total number of patterns having true class labels. These two parameters can be expressed in a tabular form, referred to as confusion (or error) matrix (CM) [32,40]. A CM is a square assortment of numbers defined in rows and columns that represent the number of samples (i.e., patterns) assigned to a particular category relative to the true category. However, we have discussed the significance of CM in a different point of view. In this regard, the result of accuracy with respect to individual class is analyzed. Sometimes a distinction is made between errors of omission and errors of commission, particularly when only a small number of class type is of interest [32,41]. Errors of omission correspond to those pattern belonging to the class of interest that the classifier has failed to recognize whereas errors of commission are those that correspond to pattern from other classes that the classifier has labeled as belonging to the class of interest. The former refers to columns of the CM, whereas the latter refers to rows (if the CM is developed with rows representing the assigned/estimated and column as true/reference class labels). Thus interpreting a CM from a particular class point of view, it is important to notice that different indications of class accuracies will result differently according to whether the number of correct pattern for a class is divided by the total number of true (reference) patterns for the class or the total number of patterns the classifier attributes to the class. The former is normally known as producer's accuracy and the latter as user's accuracy [32].

However, the overall classification accuracy does not provide the class-wise agreement between the true and estimated class labels and the producer's and user's accuracy gives the results for individual class. To get an overall class agreement based on the individual class accuracy, we have used Kappa coefficient (*KC*) estimation to validate the superiority of the classifiers effectively. The *KC* measure was introduced by the psychologist Cohen [35] and adapted for accuracy assessment in the remote sensing field by Congalton and Mead [42]. The *KC* and classification accuracy are not proportional, which means a good percentage of accuracy may lead to a poor *KC* value because it provides the measurement of class-wise agreement between the true and estimated class labels. The More the coefficient value, the better the agreement. Thus a good *KC* value signifies better agreement of the estimated data with the true one. The *KC* value is normally estimated from a CM [32,40]. Many statistical measures of class accuracies like *KC* and overall classification accuracy can be computed from CM. The overall classification accuracy is computed from CM by taking the sum of the diagonal elements divided by the total number of samples.

The *KC* can be computed as follows:

$$KC = \frac{M\sum_{i=1}^{r} Y_{ii} - \sum_{i=1}^{r}(Y_{i+} \times Y_{+i})}{M^2 - \sum_{i=1}^{r}(Y_{i+} \times Y_{+i})}, \tag{6}$$

where

r = the number of rows in the error matrix,
Y_{ii} = the number of observations in row i and column i,
Y_{i+} = the total observation in row i,
Y_{+i} = the total observation in column i and
M = the total number of observations included in the matrix. (7)

A *KC* value (> 0) indicates the amount of agreement between the two observers (true and estimated). A value of 1 indicates perfect agreement (when all the values are falling on the diagonal) [35].

For the remote sensing image data, a very small set of training patterns is picked up from the known regions according to the available ground truth for classifying rest of the image (partially labeled data sets). Therefore, it is not possible to assess the results with the indices described above. Here we have used two indices, one is β index [33] of *homogeneity* and the other Xie–Beni (*XB*) index [34] of *compactness and separability* as discussed below.

3.1. β index of homogeneity

The β index has been successfully used in the assessment of image segmentation quality in [33]. β is defined as (Eq. (8)) the ratio of the total variation and within-class variation. Since the numerator is constant for a given image, β value is dependent only on the denominator. *The denominator decreases with increase in homogeneity within the class for a fixed number of classes* (C). Thus for a given image and given number of classes, *the higher the homogeneity within the classes, the higher would be the β value*. Mathematically β is represented as

$$\beta = \frac{\sum_{c=1}^{C}\sum_{j=1}^{M_c}(\mathbf{x}_{cj} - \overline{\mathbf{x}})^2}{\sum_{c=1}^{C}\sum_{j=1}^{M_c}(\mathbf{x}_{cj} - \overline{\mathbf{x}}_c)^2}, \tag{8}$$

where $\overline{\mathbf{x}}$ is the mean grey value of all the pixels in an image (pattern vectors), M_c is the number of pixels in cth ($c = 1, 2, \ldots, C$) class, $\mathbf{x}_{cj}$ is the grey value of the jth pixel ($j = 1, 2, \ldots, M_c$) in class c, and $\overline{\mathbf{x}}_c$ is the mean of M_c pixels of the cth class.

3.2. Xie–Beni (XB) index of compactness and separability

This index, popularly known as the *Xie–Beni* (*XB*), was first proposed in Ref. [34]. The *XB* index provides a validity criterion based on a validity function that identifies overall compact and separate fuzzy partitions without any assumptions to the number of substructures inherent in the data. The index depends on the data set, geometric distance measure, distance between class centroid, and more importantly on the fuzzy partition generated by the fuzzy algorithms used. It is mathematically expressed as the ratio of compactness (θ) to the separation (ξ), i.e.,

$$XB = \frac{\theta}{\xi}, \tag{9}$$

where θ is defined as

$$\theta = \frac{v}{M}, \tag{10}$$

with v as the total variation in the data set with respect to the fuzzy partition and is defined as

$$v = \sum_{c=1}^{C}\sum_{j=1}^{M}\mu_{cj}^2\|\mathbf{x}_j - \overline{\mathbf{x}}_c\|^2, \tag{11}$$

where $\mathbf{x}_j$ is the jth pattern, $\overline{\mathbf{x}}_c$ is the centroid of the cth class and M is the total number of data points. μ_{cj} represents the degree of membership of jth pattern to cth class. Similarly, ξ (Eq. (9)) is defined as the separation of the fuzzy partitions and calculated as the minimum distance between class centroid, i.e.,

$$\xi = \min_{c \neq k}\|\overline{\mathbf{x}}_c - \overline{\mathbf{x}}_k\|^2. \tag{12}$$

Thus the expression for (*XB*) becomes

$$XB = \frac{1}{M}\frac{\sum_{c=1}^{C}\sum_{j=1}^{M}\mu_{cj}^2\|\mathbf{x}_j - \overline{\mathbf{x}}_c\|^2}{\min_{c \neq k}\|\overline{\mathbf{x}}_c - \overline{\mathbf{x}}_k\|^2}. \tag{13}$$

The *XB* index measures the compactness among the data points within a class and simultaneously provides the information regarding the separation between the classes. The *smaller* the *XB* value, the *better* the partitioning [34].

4. Description of the data set used

A short description of two conventional data sets, namely, WAVEFORM [36] and BUPA (liver disorder) [36], are used in the present study and are provided in Table 1. We have used data sets bearing different number of features and classes. Also the data sets are selected from the group of both small and large number of labeled patterns. In addition to these, two multispectral remote sensing images (size 512×512) obtained from two different satellites are used for the present simulation study: one from IRS-1A [43] and one from SPOT [43] satellite. The actual classes (land covers) present in the input images are not visible clearly and hence presented the enhanced images in Figs. 3a and b. However, the algorithms are implemented on actual (original) images.

The IRS-1A image is taken from the linear imaging self scanner (LISS-II) which has a spatial resolution of

Table 1
A brief description of the data sets used

Sl. no.	Name of the data set	Number of classes	Number of features available	Number of patterns
1	WAVEFORM	3	21	5000
2	BUPA	2	6	345

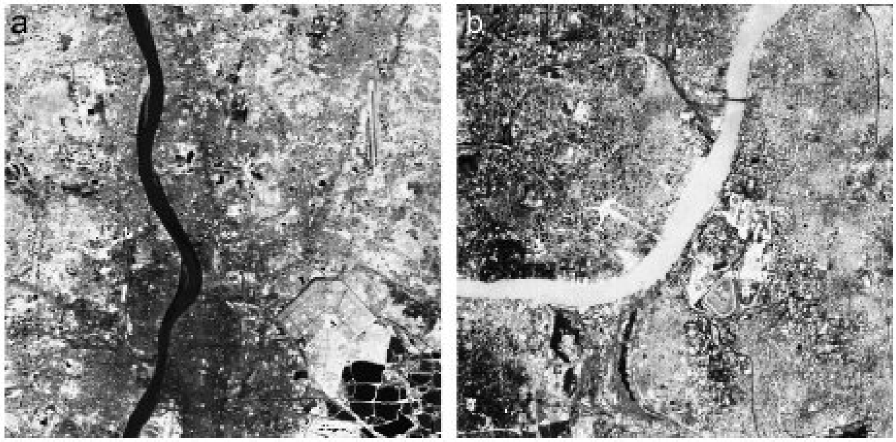

Fig. 3. Original (a) IRS-1A (band-4) and (b) SPOT (band-3) images.

$36.25\,\text{m} \times 36.25\,\text{m}$ and works in the wavelength range of 0.45–0.86 µm. The whole spectrum range is decomposed into four spectral bands, namely blue band (band-1), green band (band-2), red band (band-3) and near-infrared band (band-4) with wavelengths 0.45–0.52, 0.52–0.59, 0.62–0.68 and 0.77-0.86 µm, respectively. The image in Fig. 3a covers an area around the city of Calcutta in the near infrared band having six major land cover classes. These are *pond or fishery water* (PW), *turbid water* (TW), *concrete area* (CA), *habitation* (HAB), *vegetation* (VEG) and *open spaces* (OS). In India, these happen to be the major land cover types. PW class contains pond water, fisheries, etc. River water, etc., where the soil content is more belongs to TW class. CA class consists of buildings, roads, airport runways, bridges, etc. Suburban and rural habitation, i.e., concrete structures are comparatively less in density than the previous class (CA) and come under HAB class. VEG class essentially represents crop and forest areas. OS class contains the barren land.

The SPOT image shown in Fig. 3b is obtained from SPOT (Systeme Pour d'Observation de la Terre) satellite [43] and has been acquired from the *High Resolution Visible* (HRV) that uses the wavelength range of 0.50–0.89 µm. The whole spectrum range is decomposed into three spectral bands, namely green band (band-1), red band (band-2) and near-infrared band (band-3) of wavelengths 0.50–0.59, 0.61–0.68 and 0.79–0.89 µm, respectively, and a higher spatial resolution of $20\,\text{m} \times 20\,\text{m}$. The same six classes of land cover as in IRS-1A image is taken for this image also.

5. Results and discussion

5.1. Strategy of selecting the training data set

The conventional data sets are divided into two parts. First part is taken for estimation of the parameters of the classifier (training data). The second part is taken for testing the performance (test data). We have taken three different sizes of data as training: these are 10%, 20% and 50% and the rest 90%, 80% and 50% considered as test data. Selection of the training data is random in nature and an equal percent of data is collected from each class.

For remote sensing images training samples are selected according to a prior information of the land cover regions obtained from the available ground truth. These training samples are used to estimate the parameters of the classifiers. After learning the classifier, it is used to classify the land covers of the whole image.

5.2. Implementation and results with discussion

In the present investigation the proposed method exploits the advantage of combining π-type *MF* and *PARR*. The performance of this classifier is compared with other similar methods that incorporate π-type *MF* with *MIN RR* and Gaussian *MF* with *PROD* and *MIN RRs*. These two aggregation operators (*MIN* and *PROD*) are selected as they provided improved classification accuracy compared to others experimented on various data sets. Also the Gaussian *MF* is used here for its resemblance in shape with π-type *MF* and thus selected for comparison of performances. As mentioned previously (Section 1), for some data sets, particularly for remote sensing image data, all the features (spectral values) contribute reasonably to the desired class and collaborate with each other in the decision making process. Therefore fuzzy product aggregation is a perfectly suitable operator for this problem; and it is demonstrated here with experimental results. In the experiments with π-type *MF* for all the data sets, we have taken the N (fuzzifier) value as 2 (Section 2.1). Various performance comparison measures are estimated for all data sets as mentioned above.

5.3. Classification of conventional (completely labeled) data sets

5.3.1. WAVEFORM Data

WAVEFORM data set consists of three classes of waves with 21 number of attributes. Each class of the data set is generated from a combination of base waves. There are no missing attribute/feature values in the data set and all the features are

Table 2
Performance comparison of the proposed classifier for WAVEFORM data set

% Train. data		10		20		50	
Mem. fun.		Gaussian	π-Type	Gaussian	π-Type	Gaussian	π-Type
MIN	*MC*	1169	1166	982	970	577	577
	PA	74.01	74.08	75.43	75.73	76.91	76.91
PROD	*MC*	764	**758**	673	**645**	410	**380**
	PA	83.01	**83.15**	83.16	**83.86**	83.59	**84.79**

The bold values highlight proposed method.

corrupted with noise (mean 0, variance 1). Class distribution of the patterns present in the data set is made with 33% for each class.

The classification results for this data set is depicted in Table 2 with center of the MF at *mean*. It is observed that for all percentages of training data the number of *MC* and *PA* values are similar with a small variation of three to five numbers in *MC* value for the π-type and Gaussian *MF* with *MIN RR*. However, there is a large difference of performance using *PROD RR* for both the functions. For example, with 10% of training data the method using Gaussian *MF* with *MIN RR* provided *MC* value of 1169 with *PA* value of 74.01, whereas with *PROD RR* the *MC* value reduced to 764 and *PA* value increased to 83.01.

In case of π-type *MF* with *MIN* and *PROD RR*, it gave *MC* value as 1166 with *PA* value 74.08 and *MC* value as 758 with *PA* value 83.15, respectively. Thus among the four combinations of classification methods, the method that incorporates π-type *MF* and *PARR* provided better accuracy. It is observed that the performance difference between the Gaussian and π-type *MF*-based methods with the same *RR* is varying little. The performance is almost similar for 20% and 50% of the training data as in the case of 10%. The classification accuracy of the π-type *MF* with *maximum*, *arithmetic mean* and *harmonic mean* instead of *PROD* is also tested. We have not shown the results with other aggregation operators in the tables, as their performances compared to *MIN* and *PROD RRs* are poor. This clearly demonstrated the better classification capability of the proposed classifier. However, both the *MFs* (i.e., π-type and Gaussian) along with the *PARR* performed similarly, with a little variation in performance. This shows the group contribution efficiency of the *PARR* over *MIN RR*.

In another study the advantages of using the π-type *MF* over the Gaussian are realized and shown in Table 3. It is obvious from the table that with the increase in the fuzzifier value (N) of the π-type *MF*, the classification accuracy further increased. We have seen that for $N = 1$ and with 10% training data, the *MC* values are 1174 and 766 with *MIN* and *PROD RR*, respectively. These values decreased to 1163 and 748 for *MIN* and *PROD RR*, respectively, for $N = 3.00$. Similarly, the *PA* increased from 73.90 to 74.14 and 82.97 to 83.37 in case of *MIN* and *PROD RR*, respectively. It is also observed that the increase in accuracy with increase in fuzzifier is more in case of *PROD RR* compared to *MIN RR*. The same explanation can be put for 50% of training data as in the case of 10%. The increase

Table 3
Classification result of π-type *MF* for different *N* value for WAVEFORM data

% Train. data	10				50			
RR	*MIN*		*PROD*		*MIN*		*PROD*	
Value of *N*	*MC*	*PA*	*MC*	*PA*	*MC*	*PA*	*MC*	*PA*
1.00	1174	73.90	766	82.97	582	76.71	388	84.47
1.25	1174	73.90	763	83.04	582	76.71	388	84.47
1.50	1171	73.97	760	83.10	580	76.79	384	84.63
1.75	1169	74.01	758	83.15	577	76.91	382	84.71
2.00	1169	74.01	758	83.15	577	76.91	380	84.79
2.25	1165	74.10	755	83.21	574	77.03	380	84.79
2.50	1163	74.14	752	83.28	574	77.03	375	84.99
2.75	1163	74.14	748	83.37	572	77.11	374	85.03
3.00	1163	74.14	748	83.37	572	77.11	371	85.15

in classification accuracy that is obtained with the tuning of fuzzifier value is studied for all other types of data sets. As the performances with all the data sets are similar to WAVEFORM data, the results for this WAVEFORM data is only reported in the article. As mentioned is Section 2.1, we have used $N = 2$ for all other experiments for uniformity.

In addition to the above comparison, the performance measures like producer's and user's accuracy for the WAVEFORM data set have been calculated for all the classifiers based on different combination of *MFs* and *RRs*. Table 4 depicts the results of these accuracies for the classification methods using Gaussian *MF* with *MIN RR* and π-type *MF* with *PROD RR*. These two methods are selected for comparison because they provided the best and worst results among the selected combinations of *MF* and *RR*. The measurement is also made for the case of 50% training data. It is observed from Table 4 that for all classes and with both the accuracy measurements the proposed classifier performed better than others in class-wise and overall comparison. It is valid for both the percentages of training data set. As this trend of improvement of accuracy is observed in both types of data sets considered in the present study, we have only shown the results with WAVEFORM data. The superiority of the proposed classifier is also validated with the *KC* as shown in Table 5. The performance comparison of the proposed classifier with other three combinations of *MFs* and *RRs* is also made. However, the results of Gaussian *MF* with *MIN RR* and π-type *MF* with PROD RR are provided in the table (Table 5). The *KC* value obtained for WAVEFORM data set with the proposed classifier for 10% training data is

Table 4
Producer's versus user's class accuracy of the proposed (π-type MF with *PARR*) method and Gaussian *MF* with *MIN RR* for WAVEFORM data set

% Train. data	10				50			
	Gaussian MF with *MIN RR*		π-Type MF with *PARR*		Gaussian MF with *MIN RR*		π-Type MF with *PARR*	
Class number	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
1	75.39	67.18	76.30	92.15	72.58	71.80	70.71	95.68
2	77.73	75.59	91.43	83.08	74.85	79.08	93.56	81.31
3	69.27	81.18	92.79	81.66	80.07	76.90	93.75	79.79

Table 5
Performance comparison in terms of *KC* for all data sets

% Train. data	10		20		50	
Classification methods	Gaussian *MF MIN RR*	π-Type *MF PROD RR*	Gaussian *MF MIN RR*	π-Type *MF PROD RR*	Gaussian *MF MIN RR*	π-Type *MF PROD RR*
WAVEFORM	0.6001	**0.7214**	0.6113	**0.7461**	0.6207	**0.7755**
BUPA	0.3421	**0.3802**	0.3551	**0.3906**	0.3662	**0.3978**

The bold values highlight proposed method.

Table 6
Performance comparison of the proposed classifier for BUPA (liver disorder) data set

% Train. data		10		20		50	
Mem. fun.		Gaussian	π-Type	Gaussian	π-Type	Gaussian	π-Type
MIN	*MC*	121	120	105	105	64	64
	PA	60.96	61.29	61.95	61.95	62.79	62.79
PROD	*MC*	113	**111**	97	**96**	59	**59**
	PA	63.54	**64.19**	64.85	**65.21**	65.69	**65.69**

The bold values highlight proposed method.

0.7214, which is superior to the value 0.6001 obtained using Gaussian *MF* and *MIN RR*. The similar improvement of *KC* value is obtained with 20% and 50% of training data. It is also seen that the Gaussian *MF* with *MIN* and *PROD* aggregation *RRs* performs similarly, with a little variation of *KC*. Thus the performance of the proposed classifier has clearly revealed its superiority as the *KC* is higher over others. It is interesting to observe that an explicit fuzzy classifier discussed by Melgani et al. [44] is similar to the method that incorporates Gaussian *MF* and *MIN RR*.

5.3.2. BUPA Data

The data are collected from the patient with two types of liver disorder. Each record/sample in the data set constitutes the record of a single male individual. Six features of each sample are results from blood tests which are thought to be sensitive to liver disorders that may arise from excessive alcohol consumption, as well as the number of drinks per day. This database contains 345 number of instances with six different features separated into two classes of liver disorder.

This data set is being used for the comparative analysis of the four classification methods. The performance comparison results are shown in Table 6. From the table it is seen that for 10% of training data and with *MIN RR*, π-type *MF* provided an equivalent performance compared to the Gaussian *MF*. However, both the *MFs* performed better with *PARR*, e.g., with 10% training data, *MC* values are 113 and 111 with Gaussian and π-type *MF*, respectively. These values are 121 and 120 with *MIN RR* for Gaussian and π-type *MF*, respectively. Accordingly the *PA* values using Gaussian and π-type *MF* with *MIN RR* are 60.96 and 61.29 for both the cases that increased to 63.54 and 64.19 with *PROD RR*. However, the proposed classifier performs better than the method that uses Gaussian *MF* and *PARR*. The similar improvement of the π-type *MF* with *PROD RR* over other combinations for the 20% and 50% of training data can be observed from Table 6. The improvement of the proposed classifier over the other three methods in terms of both the accuracies (producer's and user's) is also justified (Table 4). The *KC* value revealed the same conclusion about the proposed classifier as above (Table 5). From Table 5, it is obvious that for all the set of training data the proposed classifier performed superior compared to others. For example, with 10% of training data, the *KC* values are 0.3802 and 0.3421 for the proposed and Gaussian *MF* with *MIN RR*-based method, respectively. The improvements of *KC* values are similar for 20% and 50% of training data as in the case of 10%.

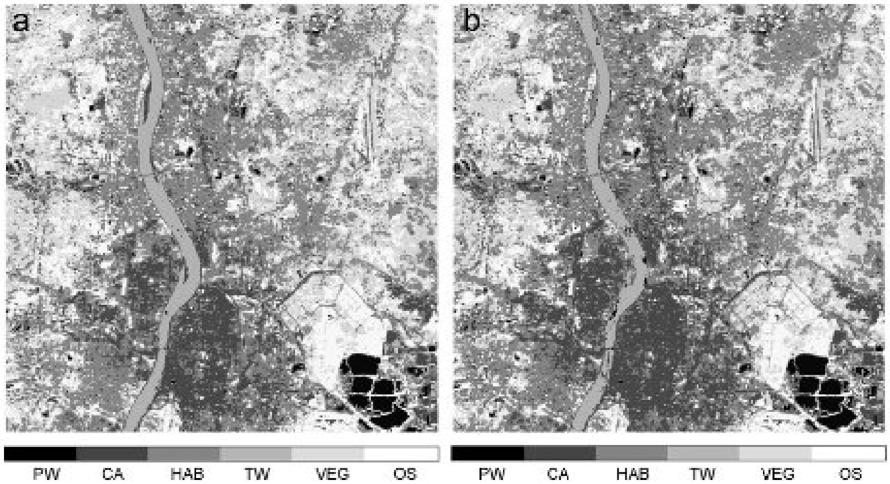

Fig. 4. Classified IRS-1A image with (a) Gaussian *MF* with *MIN RR* and (b) π-type *MF* with *PROD RR*.

Table 7
β values for different classification methods

Classification method	IRS Cal	SPOT Cal
Training patterns	9.4212	9.3343
π-Type *MF* + *PROD RR*	**8.1717**	**8.1078**
Gaussian *MF* + *PROD RR*	8.1001	8.0021
π-Type + *MIN RR*	7.1973	7.1267
Gaussian *MF* + *MIN RR*	7.1312	7.0137

The bold values highlight proposed method.

Table 8
XB values for different classification methods

Classification method	IRS Cal	SPOT Cal
π-Type *MF* + *PROD RR*	**0.8310**	**2.1021**
Gaussian *MF* + *PROD RR*	0.8498	2.1236
π-Type + *MIN RR*	0.8901	2.2695
Gaussian *MF* + *MIN RR*	0.9012	2.3031

The bold values highlight proposed method.

5.4. Classification of remote sensing image (partially labeled) data sets

5.4.1. IRS-1A Calcutta image

The six land covers of IRS-1A Calcutta image is classified using the four classifiers. From the visualization point of view, it is observed that the proposed classifier performs better in classifying the land covers compared to the remaining three. Hence we have shown only the results obtained by the proposed classifier in Fig. 4b. However, for a comparison point of view we have provided the classified results (Fig. 4a) for the method based on Gaussian *MF* and *MIN RR*. Various objects like *Airport runways, Stadium, Bridges* and different land cover classes are clearly visible with the proposed classifier. Also it is observed that regions are more clear and distinct using the proposed classifier. With the use of π-type *MF* and *PROD RR*, the classes became more separated and well identified. These objects are more or less visible in case of other classifiers based on Gaussian *MF* with *MIN* and *PROD RR* and π-type *MF* with *MIN RR*.

Further, a concrete distinction between various classes obtained by different classifiers is justified with the estimation of quantitative index rather than only visualizing the regions. Two quantitative indices, namely β and *XB*, as discussed in Section 3, have been used to justify these findings. Table 7 depicts the results of β. As expected, the β value is the highest for the training data (9.4212) in the classification of IRS-1A Calcutta image. Its value is 8.1717 for the proposed classifier, which is the highest among others. This value of β is closer to 8.1001 obtained with Gaussian *MF* and *PROD RR*, which justified the superiority of *PARR* compared to *MIN RR*. It can be seen that the objects are segregated better in Fig. 4b.

Similar to β measure, *XB* measure also supported the superiority of the proposed classifier. Values for *XB* measure obtained with this classified image are depicted in Table 8. It is seen that a better compaction and separation of different regions of the images are obtained with the proposed classifier compared to others. The *XB* value obtained using the proposed classifier is found to be 0.8310. The *XB* values using other classifiers are 0.8498, 0.8901 and 0.9012, respectively, as shown in Table 8.

5.4.2. SPOT Calcutta image

For SPOT Calcutta image, the classified regions of the images are shown in Figs. 5b and 5a for the proposed and Gaussian *MF* with *MIN RR*-based methods. From the figures it is observed that all the classes are separated properly and distinctly with the proposed classifier compared to Gaussian *MF* with *MIN RR*-based method. It is evident that the results of proposed classifier (Fig. 5b) produced a well-structured and proper shaped region compared to others. From the figure it is observed that there is a clear separation of different classes and some man made objects like *Race Course*, *Bridges*, *Canals*, *Dockyard*, *Ponds* by the proposed classifier. However, a better performance comparison with the help of β value can be seen from Table 7. The β value for the training data set is 9.3343. Its values are 8.1078, 8.0021, 7.1267 and 7.0137 for the classified

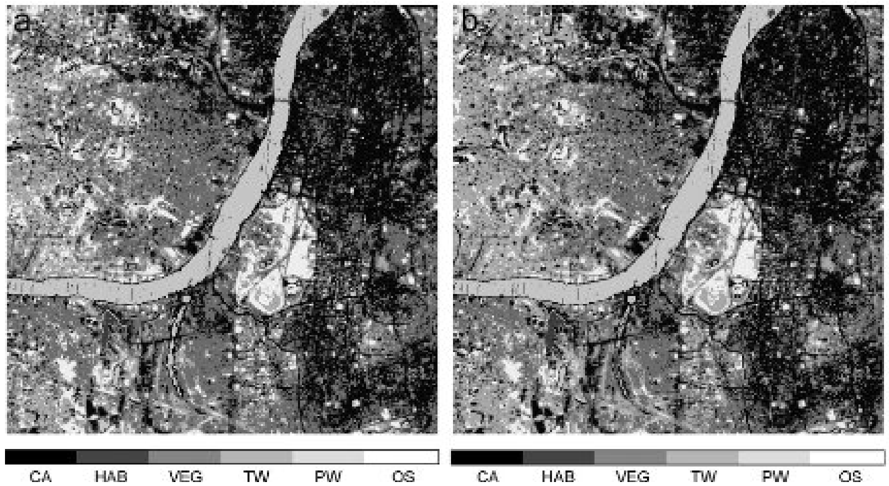

Fig. 5. Classified SPOT image with (a) Gaussian *MF* with *MIN RR* and (b) π-type *MF* with *PROD RR*.

images using the proposed and other three methods, respectively. It is seen that β is the highest for the proposed method and reveals its supremacy. Like the β index, the *XB* values as shown in Table 8 also corroborate the above findings. The corresponding *XB* value with the proposed classifier is the minimum among all and justified its superiority. It is observed that a better compactness and separability of land covers are obtained with the proposed classifier compared to others.

6. Conclusion

In the present article we have proposed a fuzzy classifier (based on fuzzy sets) that explored three important aspects. These are (i) extracted feature-wise information for different classes, (ii) generalization capability and (iii) combined contribution of individual features to a particular class. It is observed that individually the *PROD* aggregation reasoning rule (*RR*) has a better classification capability compared to other RRs. This is because of the fact that the fuzzy product aggregation operator works better with features which collaborate with each other in decision making process. Also with the use of π-type *MF* instead of Gaussian *MF* (used for the comparison) the classification accuracy can be enhanced and this expresses its better generalization capability.

Experimental study performed on two conventional (completely labeled) data sets and two remote sensing images (partially labeled) verified the potentiality of proposed classifier quantitatively using various performance measures.

Acknowledgments

The authors would like to thank the reviewers for their valuable and constructive suggestions. Thanks are also due to the Department of Science and Technology, Government of India, under which a project titled “Advanced Techniques for Remote Sensing Image Processing” is being carried out at the Machine Intelligence Unit, Indian Statistical Institute, Kolkata.

References

[1] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second ed., Wiley Interscience Publications, 2001.

[2] J.T. Tou, R.C. Gonzalez, Pattern Recognition Principles, Addition-Wesley, Reading, MA, 1974.

[3] D. Tveter, The Pattern Recognition Basis of Artificial Intelligence, Wiley-IEEE Computer Society Press, 1998.

[4] C.H. Chen, L.F. Pau, P.S.P. Wang, Handbook of Pattern Recognition and Computer Vision, World Scientific, Singapore, 1993.

[5] A.B. Watson, Digital Images and Human Vision, The MIT Press, Cambridge, 1993.

[6] S. Haykin, Neural Networks: A Comprehensive Foundation, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1998.

[7] B.D. Ripley, Pattern Recognition and Neural Networks, Cambridge University Press, Cambridge, 1996.

[8] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.

[9] H.J. Zimmermann, Fuzzy Set Theory and its Applications, fourth ed., Kluwer Academic Publishers, Dordrecht, 2001.

[10] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, India, 1997.

[11] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989.

[12] L.I. Kuncheva, Fuzzy Classifier Design, Springer, Berlin, 2000.

[13] F. Maselli, A. Rodolfi, C. Copnese, Fuzzy classification of spatially degraded thematic mapper data for the estimation of sub-pixel components, Int. J. Remote Sensing 17 (3) (1996) 537–551.

[14] D.P. Mandal, C.A. Murthy, S.K. Pal, Formulation of a multivalued recognition system, IEEE Trans. Syst. Man Cybern. 22 (4) (1992) 607–620.

[15] J.M. Keller, M. Gray, J. Givens, A fuzzy k-nearest neighbor algorithm, IEEE Trans. Syst. Man Cybern. 15 (4) (1985) 580–585.

[16] H. Ishibuchi, K. Nozaki, H. Tanaka, Distributed representation of fuzzy rules and its application to pattern classification, Fuzzy Sets Syst. 52 (1) (1992) 21–32.

[17] S. Abe, M.S. Lan, A method for fuzzy rules extraction directly from numerical data and its application to pattern classification, IEEE Trans. Fuzzy Syst. 3 (1) (1995) 18–28.

[18] F. Wang, Fuzzy supervised classification of remote sensing images, IEEE Trans. Geosci. Remote Sensing 28 (2) (1990) 194–201.

[19] W. Pedrycz, Fuzzy sets in pattern recognition: methodology and methods, Pattern Recognition 23 (1/2) (1990) 121–146.

[20] B. Tso, P.M. Mather, Classification Methods for Remotely Sensed Data, Taylor and Francis, London, 2001.

[21] S.K. Meher, A. Ghosh, B. Uma Shankar, Wavelet feature based classifier for multispectral remote sensing images, IEEE Trans. Geosci. Remote Sensing 45 (6) (2007) 1881–1886.

[22] H. Ishibuchi, T. Yamamoto, Comparison of heuristic criteria for fuzzy rule selection in classification problems, Fuzzy Optim. Decision Making 3 (2) (2004) 119–139.

[23] A. Bardossy, L. Samaniego, Fuzzy rule-based classification of remotely sensed imagery, IEEE Trans. Geosci. Remote Sensing 40 (2) (2002) 362–374.

[24] Y. Chen, J.Z. Wang, Support vector learning for fuzzy rule-based classification systems, IEEE Trans. Fuzzy Syst. 11 (6) (2003) 716–728.

[25] H. Ishibuchi, T. Nakashima, Effect of rule weights in fuzzy rule-based classification systems, IEEE Trans. Fuzzy Syst. 9 (4) (2001) 506–515.

[26] D. Dubois, H. Prade, R. Yager, Merging fuzzy information, in: J. Bezdek, D. Dubois, H. Prade (Eds.), Fuzzy Sets in Approximate Reasoning and Information Systems, Kluwer, Boston, 1999, pp. 335–401.

[27] V. Peneva, I. Popchev, Fuzzy logic operators in decision-making, Cybern. Syst. 30 (8) (1999) 725–745.

[28] G. Beliakov, J. Warren, Appropriate choice of aggregation operators in fuzzy decision support systems, IEEE Trans. Fuzzy Syst. 9 (6) (2001) 773–784.

[29] I. Bloch, Information combination operators for data fusion: a comparative review with classification, IEEE Trans. Syst. Man Cybern. A Syst. Hum. 26 (1) (1996) 52–67.

[30] D. Dubois, H. Prade, A review of fuzzy set aggregation connectives, Info. Sci. 36 (1985) 85–121.

[31] D. Dubois, H. Prade, On the use of aggregation operations in information fusion processes, Fuzzy Sets Syst. 142 (1) (2004) 143–161.

[32] R.G. Congalton, A review of assessing the accuracy of classifications of remotely sensed data, Remote Sensing Environ. 37 (1) (1991) 35–46.

[33] S.K. Pal, A. Ghosh, B. Uma Shankar, Segmentation of remotely sensed images with fuzzy thresholding, and quatitative evaluation, Int. J. Remote Sensing 21 (11) (2000) 2269–2300.

[34] X.L. Xie, G. Beni, A validity measure for fuzzy clustering, IEEE Trans. Pattern Anal. Mach. Intell. 13 (8) (1991) 841–847.

[35] J. Cohen, A coefficient of aggrement for nominal scale, Educ. Psychol. Meas. 20 (1) (1960) 37–46.

[36] D.J. Newman, S. Hettich, C.L. Blake, C.J. Merz, UCI repository of machine learning databases, Department of Information and Computer Sciences, University of California, Irvine, ⟨http://www.ics.uci.edu/~mlearn/MLRepository.html⟩, 1998.

[37] L.I. Kuncheva, Combining Pattern Classifier: Methods and Algorithms, Wiley-Interscience, 2004.

[38] S.K. Pal, D.D. Majumder, Fuzzy Mathemetical Approach to Pattern Recognition, Wiley Eastern Limited, 1986.

[39] A.V. Patel, Simplest fuzzy PI controllers under various defuzzification methods, Int. J. Comput. Cognit. 3 (1) (2005) 21–34.

[40] D.H. Card, Using known map category marginal frequencies to improve estimates of thematic map accuracy, Photogramm. Eng. Remote Sensing 48 (3) (1982) 431–439.

[41] J.A. Richards, X. Jia, Remote Sensing Digital Image Analysis: An Introduction, fourth ed., Springer, Berlin, Heidelberg, Germany, 2006.

[42] R.G. Congalton, R. Mead, A quantitative method to test for consistency and correctness in photointerpretation, Photogramm. Eng. Remote Sensing 49 (1) (1983) 69–74.

[43] NRSA, IRS data users hand book, Technical Report, Document No. IRS/NRSA/NDC/HB-02/89, 1989.

[44] F. Melgani, B.A.R. Al Hashemy, S.M.R. Taha, An explicit fuzzy supervised classification method for multispectral remote sensing images, IEEE Trans. Geosci. Remote Sensing 38 (1) (2000) 287–295.

About the Author—ASHISH GHOSH is a Professor of the Machine Intelligence Unit at the Indian Statistical Institute, Kolkata. He received the B.E. degree in Electronics and Telecommunication from the Jadavpur University, Kolkata, in 1987, and the M.Tech. and Ph.D. degrees in Computer Science from the Indian Statistical Institute, Kolkata, in 1989 and 1993, respectively. He received the prestigious and most coveted *Young Scientists* award in Engineering Sciences from the Indian National Science Academy in 1995, and in Computer Science from the Indian Science Congress Association in 1992. He has been selected as an *Associate* of the Indian Academy of Sciences, Bangalore, in 1997. He visited the Osaka Prefecture University, Japan, with a Post-doctoral fellowship during October 1995 to March 1997, and Hannan University, Japan, as a *Visiting Faculty* during September to October, 1997 and September to October, 2004. He has also visited Hannan University, Japan, as *Visiting Professor* with a fellowship from Japan Society for Promotion of Sciences (JSPS) during February to April, 2005. During May 1999, he was at the Institute of Automation, Chinese Academy of Sciences, Beijing, with CIMPA (France) fellowship. He was at the German National Research Center for Information Technology, Germany, with a German Government (DFG) Fellowship during January to April, 2000. During October to December 2003 he was a *Visiting Professor* at the University of California, Los Angeles; and during December 2006 to January 2007 he was at the Computer Science Department of Yonsei University, Korea. His visits to University of Trento and University of Palermo (Italy) during May to June 2004, March to April 2006, and May to June 2007 were in connection with collaborative international projects. He also visited various Universities/Academic Institutes and delivered lectures in different countries including South Korea, Poland and The Netherland.
His research interests include *Pattern Recognition and Machine Learning*, *Data Mining*, *Image Analysis*, *Remotely Sensed Image Analysis*, *Soft Computing*, *Fuzzy Sets and Uncertainty Analysis*, *Neural Networks*, *Evolutionary Computation and Bioinformatics*. He has already published about 100 research papers in internationally reputed journals and refereed conferences, has edited six books and is acting as a member of the editorial board of various international journals. He is a member of the founding team that established a *National Center for Soft Computing Research* at the Indian Statistical Institute, Kolkata, in 2004, with funding from the Department of Science and Technology (DST), Government of India.

About the Author—SAROJ K. MEHER received the Ph.D. degree in electronics engineering from National Institute of Technology, Rourkela, India, in 2003. He is a Post-Doctoral Fellow in the Machine Intelligence Unit, Indian Statistical Institute, Kolkata, India. His research interests include pattern recognition, soft computing methods, fuzzy sets and digital signal processing. He has published many research articles in internationally reputed journals and refereed conferences.

About the Author—B. UMA SHANKAR received the M.Sc. degree in statistics from Bhagalpur University, India, in 1985, and the Ph.D. degree in Science from Jadavpur University, India, in 2006. He is an Associate Scientist in the Machine Intelligence Unit, Indian Statistical Institute, Kolkata, India. He was at Stanford University, USA, from September 1992 to March 1993; at CIMPA, Nice, France, from June to July 1990, and June to July 1993, and at the University of Trento, Italy from January to March 2007. His research interests include pattern recognition and machine learning, analysis of remotely sensed images, fuzzy sets, rough sets and neural networks. He has published many research papers in international journals and refereed conferences and is acting as a reviewer of international journals.