

A Massively Parallel Deep Rule-Based Ensemble Classifier for Remote Sensing Scenes

Xiaowei Gu, Plamen P. Angelov, *Fellow, IEEE*, Ce Zhang, and Peter M. Atkinson

Abstract— In this Letter, we propose a new approach for remote sensing scene classification by creating an ensemble of the recently introduced massively parallel deep (fuzzy) rule-based (DRB) classifiers trained with different levels of spatial information separately. Each DRB classifier consists of a massively parallel set of human-interpretable, transparent 0-order fuzzy IF...THEN... rules with a prototype-based nature. The DRB classifier can self-organize “from scratch” and self-evolve its structure. By employing the pre-trained deep convolution neural network as the feature descriptor, the proposed DRB ensemble is able to exhibit human-level performance through a transparent and parallelizable training process. Numerical examples using benchmark dataset demonstrate the superior accuracy of the proposed approach together with human-interpretable fuzzy rules autonomously generated by the DRB classifier.

Index Terms—deep learning, rule-based classifier, scene classification, fuzzy rules.

I. INTRODUCTION

REMOTE sensing scene classification aims to allocate the sub-regions of fine spatial resolution images to distinct land use categories, a goal which is of paramount importance for many applications, such as urban planning, land resource management, and environmental conservation [1]. At the same time, land use classification is recognized widely as a challenging task because the land use sub-regions are recognised implicitly through their high-level semantic function, where multiple low-level features or land cover classes can appear in one land use category, and identical land cover classes can be shared among different land use categories. These high-level semantics need to be exploited sufficiently using robust and accurate approaches for feature representation.

Currently, deep learning (DL) neural networks (NN) have gained huge popularity amongst researchers as well as amongst the general public, quickly becoming the state-of-art approach in the remote sensing domain, in particular [2]. Several publications have reported very promising results using DL for

spatial and spectral feature learning [2], [3]. Indeed, compared with the low- and mid-level feature-based methods (e.g., GIST [4], histogram of oriented gradient (HOG) [5], bags of visual words (BoVW) [6] and scale invariant feature transform (SIFT) [7], etc.), DL-based methods can learn more abstract and discriminative high-level semantic features and achieve greater accuracy.

Nonetheless, the DL-based approaches [2], [3] and some other state-of-the-art approaches [8]–[10] suffer from several deficiencies and shortcomings including:

1. the training process is opaque, and the classifier has low or no human interpretability (black box type);
2. the training process is limited to offline and requires re-training for samples with feature properties different than the observed samples, as well as for samples from unseen classes;
3. the training process is computationally expensive and requires a lot of training samples.

These deficiencies hinder the performance of these new approaches in real applications.

In this Letter, a new approach based on the ensemble of the recently introduced deep (fuzzy) rule-based (DRB) classifiers [11], [12] is proposed for remote sensing scene classification. The DRB classifier employs a massively parallel set of 0-order fuzzy rules as a learning engine [13]. By self-organizing a fully human-interpretable and transparent IF...THEN... model structure via a fast and nonparametric training process, our previous study shows that the DRB classifier can achieve human-level results in various applications, including handwritten digit recognition [11], [12], face recognition [12], etc. This is despite the fact that only low-level feature descriptors [14], namely GIST [4] and HOG [5], are employed. Further, due to its prototype-based nature, the DRB classifier can start learning and self-organizing “from scratch” and continue to learn from new images. Its training can be highly parallelized thanks to its specific structure [11], [12].

By introducing a pre-trained deep convolutional neural network (DCNN) as the feature descriptor [14] into the DRB classifier, it is able to grasp more abstract and discriminative semantic features within the remote sensing images. By further creating an ensemble of DRB classifiers trained with segments of remote sensing images partitioned with different granularities, the proposed approach is able to utilize spatial information at multiple scales and exhibit highly accurate classification performance.

Preliminary numerical examples demonstrate that the DRB ensemble outperforms the state-of-the-art approaches [2], [3],

Corresponding Author: *Plamen P. Angelov*.

X. Gu and P. P. Angelov are with School of Computing and Communications, Lancaster University, Lancaster, LA1 4WA. P. P. Angelov also holds a Honorary Professorship with Technical University, Sofia, Bulgaria (e-mail: {x.gu3, p.angelov}@lancaster.ac.uk).

C. Zhang, and P. M. Atkinson are with the Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ (e-mail: {c.zhang9, pma}@lancaster.ac.uk).

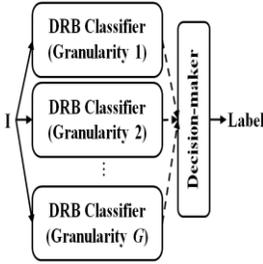


Fig. 1. General architecture

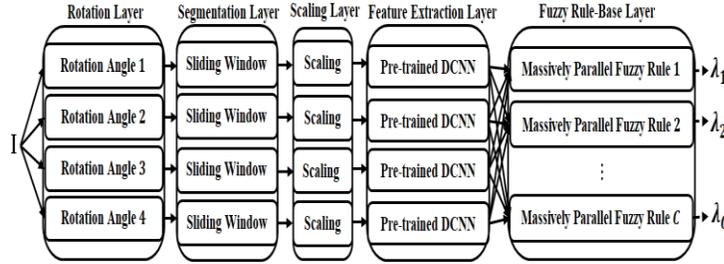


Fig. 2. The structure of a DRB classifier in a modular/layered form

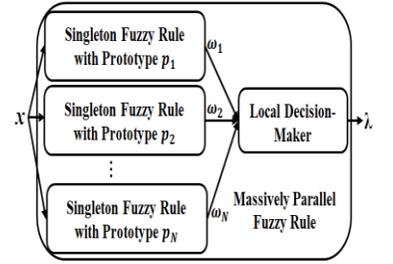


Fig. 3. Zoomed-in structure of a fuzzy rule

[8]–[10] in terms of classification accuracy and can serve as a competitive alternative to the traditional DCNNs [2], [3] with its highly transparent, human-interpretable, parallelizable system structure.

II. ARCHITECTURE OF THE PROPOSED APPROACH

The general architecture of the proposed DRB ensemble is depicted in Fig. 1, which consists of multiple deep (fuzzy) rule-based (DRB) classifiers trained with the segments of remote sensing images at different levels of granularity (G in total). The architecture of a multi-layer DRB classifier is presented in a modular/layered form in Fig. 2. One can see that each classifier consists of the following layers [11], [12]: 1) Rotation layer; 2) Segmentation layer; 3) Scaling layer; 4) Feature extraction layer and 5) Massively parallel IF...THEN... rule-based system.

The rotation layer of the DRB rotates the images at 1) 0° ; 2) 90° ; 3) 180° and 4) 270° around the centre point. The rotation operation significantly increases the generalizability and reduces the overfitting.

The segmentation layer uses a sliding window to partition the remote sensing images into smaller pieces for local information extraction. By changing the size of the sliding window, the level of granularity of the segmentation result can be changed accordingly, which results in different scales of spatial information. A larger sliding window size allows the DRB to capture coarse scale spatial information at the cost of losing fine scale spatial information and, similarly, a smaller sliding window size results in more fine scale detail. In this Letter, we use sliding windows of different sizes and partition the images into segments of different levels of granularity, train a number of DRB classifiers (one per level of granularity) in parallel, and create an ensemble for decision-making at the validation stage.

The DRB ensemble proposed in this Letter employs one of the currently best performing pre-trained deep convolutional neural network models, namely, VGG-VD-16 [15] as the feature descriptor due to its simple structure and high performance [14]. However, we stress that there is no further tuning involved, and alternative feature descriptors can also be used, i.e. GIST [4], HOG [5], or GoogLeNet [16].

As the segmentation layer produces image pieces of different sizes, a scaling layer is involved in the DRB classifier to rescale the segments into the uniform size of 227×227 pixels required by the VGG-VD-16 model [15]. Then, following common practice, the 1×4096 dimensional activations from the first fully connected layer are extracted as the feature vectors of the image

segments (one 1×4096 dimensional vector per segment).

The proposed DRB ensemble employs the self-organizing evolving fuzzy rule-based (FRB) systems of AnYa type with singletons in the consequent part [13] as its “learning engine”, and it involves a two-level decision-making process.

As shown in Fig. 2, each DRB classifier is trained with the segments of the remote sensing images of one different granularity. Each DRB classifier consists of C 0-order fuzzy rule-based subsystems of AnYa type trained in parallel corresponding to the C classes from the image set (one per class). These FRB subsystems are entirely independent from each other, and each subsystem can be updated or removed without influencing others. Each FRB subsystem contains one massively parallel fuzzy rule set formulated around the prototypes generalized or learned from the segments of the corresponding class. These rules have the following form [12]:

$$IF(\mathbf{sg} \sim \mathbf{P}_{c,1}) OR \dots OR (\mathbf{sg} \sim \mathbf{P}_{c,N_c}) THEN (class\ c), \quad (1)$$

where \sim denotes similarity, which can also be seen as a fuzzy degree of satisfaction or membership; \mathbf{sg} is a particular segment of an image, and its corresponding feature vector is denoted as \mathbf{x} ; $\mathbf{P}_{c,j}$ ($j=1,2,\dots,N_c$, $c=1,2,\dots,C$) denotes the j^{th} prototype of the c^{th} class with the corresponding feature vector $\mathbf{p}_{c,j}$; $\mathbf{P}_{c,j}$ is of the same size as \mathbf{sg} ; \mathbf{x} and $\mathbf{p}_{c,j}$ have the same dimensionality; N_c corresponds to the number of prototypes of the c^{th} class. Each fuzzy IF...THEN... rule contains a number of prototypes identified from the segments of the images from the same class, which are connected by a local decision-maker using the “winner-takes-all” principle. Therefore, each massively parallel AnYa type fuzzy rule [13] can be represented as a series of simpler fuzzy rules with a single prototype connected by a logical “OR” operator. The zoomed-in structure of a particular massively parallel fuzzy rule is depicted in Fig. 3.

The 0-order massively parallel fuzzy rules of AnYa type are nonparametric and transparent. Their identification process [11], [12] (as briefly described in the form of pseudo-code in section III.A to make this Letter self-contained) is autonomous (does not require any user input) and can also be self-evolving (new rules and prototypes can be added, merged or removed). The operating mechanism of each IF...THEN... rule during the validation process will be described in section III.B.

The final layer of the proposed DRB ensemble is the overall decision-maker that decides the winning label of the validation images based on the suggestions of the individual (per class) IF... THEN... rules of the DRB classifiers within the ensemble. The operating mechanism of the decision-maker will be described in section IV.

III. DEEP RULE-BASE

This section describes briefly the training process of the fuzzy rules in the form of pseudo-code as well as the validation process within each fuzzy rule.

A. Training Process

The DRB classifier identifies prototypes from the segments of the observed images of each class autonomously, in a nonparametric manner, and forms data clouds around the prototypes from similar segments of the same class.

As described in section II, the remote sensing images are partitioned into G segment sets of different levels of granularity, which correspond to three scales of spatial information. In this way, $C \cdot G$ 0-order massively parallel fuzzy rules of AnYa type in total are formed (learned) through the training processes independently based on the identified prototypes (one rule per class per level of granularity). The detailed training process of the FRB subsystems is described in [11], [12], and the main procedure of the training process is summarized in the form of pseudo-code in Table I.

Once the training process is finished, the classification of new images can be performed using the identified FRBs.

B. Validation Process

During the validation process, for a particular segment of a testing image with its feature vector denoted as \mathbf{x}_{ij} (the j^{th} segment under the i^{th} segmentation granularity), one can obtain C scores of confidence using the corresponding C 0-order massively parallel AnYa type fuzzy rules identified from the segments of the same size through the training process. The score of confidence produced by the local decision-maker of the c^{th} fuzzy rule, denoted by $\lambda_{i,c}(\mathbf{x}_{ij})$, is expressed as:

$$\lambda_{i,c}(\mathbf{x}_{ij}) = \arg \max_{l=1,2,\dots,N_{l,c}} \left(\omega_{i,c,l}(\mathbf{x}_{ij}) \right), \quad (2)$$

where $\omega_{i,c,l}(\mathbf{x}_{ij}) = \exp\left(-\frac{1}{2}\left\|\mathbf{x}_{ij}/\|\mathbf{x}_{ij}\| - \mathbf{p}_{i,c,l}\right\|^2\right)$; $c=1,2,\dots,C$; $i=1,2,\dots,G$; $j=1,2,\dots,M_i$ and M_i denotes the number of segments of the i^{th} granularity. Therefore, for a particular remotely sensed image, one can obtain $M_1 + M_2 + \dots + M_G$ segments, which result in $C \times (M_1 + M_2 + \dots + M_G)$ confidence score vectors in total, denoted by $\boldsymbol{\lambda}_i(\mathbf{x}_{ij}) = [\lambda_{i,1}(\mathbf{x}_{ij}), \dots, \lambda_{i,C}(\mathbf{x}_{ij})]$ ($i=1,2,\dots,G$; $j=1,2,\dots,M_i$). They serve as the inputs of the decision-maker.

IV. DECISION-MAKING MECHANISM

During the validation stage, for each testing image, the overall decision-maker of the proposed DRB ensemble firstly integrates the confidence scores obtained by the local decision-makers of the low-level decision-making committees:

TABLE I
SUMMARY OF THE TRAINING PROCESS OF THE FRB SUBSYSTEM

While new feature vector \mathbf{x}_k of the c^{th} class is available	
System Initialization	i. $\mathbf{x}_k \leftarrow \mathbf{x}_k / \ \mathbf{x}_k\ $ ($\ \mathbf{x}_k\ $ is the norm of \mathbf{x}_k); ii. If ($k=1$) Then 1. $\boldsymbol{\mu}_c \leftarrow \mathbf{x}_k$ ($\boldsymbol{\mu}_c$ is the global mean); 2. $N_c \leftarrow 1$ (N_c is the number of prototypes); 3. $\mathbf{p}_{c,1} \leftarrow \mathbf{x}_k$ ($\mathbf{p}_{c,1}$ is the first prototype); 4. $S_{c,1} \leftarrow 1$ ($S_{c,1}$ is the corresponding support); 5. $r_{c,1} \leftarrow r_o$ ($r_{c,1}$ is the radius, $r_o = \sqrt{2 - 2\cos\left(\frac{\pi}{6}\right)}$); iii. Else 1. Update global mean: $\boldsymbol{\mu}_c \leftarrow \frac{k-1}{k}\boldsymbol{\mu}_c + \frac{1}{k}\mathbf{x}_k$; 2. Calculate the density of \mathbf{x}_k : $D(\mathbf{x}_k) = \frac{1}{1 + \frac{\ \mathbf{x}_k - \boldsymbol{\mu}_c\ ^2}{1 - \ \boldsymbol{\mu}_c\ ^2}}$; 3. Update the densities of prototypes: $D(\mathbf{p}_{c,j}) = \frac{1}{1 + \frac{\ \mathbf{p}_{c,j} - \boldsymbol{\mu}_c\ ^2}{1 - \ \boldsymbol{\mu}_c\ ^2}}$; $j=1,2,\dots,N_c$
System Update	4. If ($D(\mathbf{x}_k) > \max_{j=1,2,\dots,N_c} (D(\mathbf{p}_{c,j}))$) Or ($D(\mathbf{x}_k) < \min_{j=1,2,\dots,N_c} (D(\mathbf{p}_{c,j}))$) Then $N_c \leftarrow N_c + 1$; $\mathbf{p}_{c,N_c} \leftarrow \mathbf{x}_k$; $S_{c,N_c} \leftarrow 1$; $r_{c,N_c} \leftarrow r_o$; 5. Else - Find the nearest prototype: $\mathbf{p}_{c,n} = \arg \min_{j=1,2,\dots,N_c} (\ \mathbf{x}_k - \mathbf{p}_{c,j}\)$; - If ($\ \mathbf{x}_k - \mathbf{p}_{c,n}\ > r_{c,n}$) Then $N_c \leftarrow N_c + 1$; $\mathbf{p}_{c,N_c} \leftarrow \mathbf{x}_k$; $S_{c,N_c} \leftarrow 1$; $r_{c,N_c} \leftarrow r_o$; - Else $\mathbf{p}_{c,n} \leftarrow \frac{S_{c,n}}{S_{c,n} + 1} \mathbf{p}_{c,n} + \frac{1}{S_{c,n} + 1} \mathbf{x}_k$; $S_{c,n} \leftarrow S_{c,n} + 1$; $r_{c,n} \leftarrow \sqrt{\frac{1}{2}(r_{c,n})^2 + \frac{1}{2}(1 - \ \mathbf{p}_{c,n} \ ^2)}$; - End If 6. End If iv. End If v. Generate/update the AnYa type fuzzy rule; End While
Preparation	

$$\Lambda_{i,c} = \frac{1}{M_i} \sum_{j=1}^{M_i} \lambda_{i,c}(\mathbf{x}_{i,j}); \quad i=1,2,\dots,G; \quad c=1,2,\dots,C. \quad (3)$$

Then, the label of the testing image is allocated following the “winner-takes-all” principle based on the overall outputs of the three DRB classifiers:

$$\text{Label} = \arg \max_{c=1,2,\dots,C} \left(\sum_{i=1}^G \Lambda_{i,c} \right). \quad (4)$$

V. NUMERICAL EXAMPLES

To evaluate the performance of the proposed DRB ensemble, in this Letter, we use the well-known UC-Merced (UCM) dataset [6] as the standard use case. All the experiments in this Letter are conducted with Matlab2017a on a PC with dual core i7 processor each with a clock frequency of 3.4GHz and 16GB RAM. The RGB images are used directly by the DRB classifiers. The classification experiments are repeated five times under the same ratio of training-to-testing sample sets and the average result is reported as in [8].

The UCM dataset consists of fine spatial resolution remote sensing images of 21 challenging scene categories (including airplane, beach, building, etc.). Each category contains 100 images of the same image size (256×256 pixels). Following the common experiment protocol [6], 80% of the images in each category were selected for training and the rest used for testing.

Considering the input image size of the VGG-VD-16 model [15], in this Letter, for simplification, each image is divided using an 8×8 grid net with the size of each grid equal to 32×32 pixels (an illustration is given in Fig. 4). Sliding windows of four different window sizes (4×4 grids, 5×5 grids, 6×6 grids and 7×7 grids) and two different step sizes (1 grid and 2 grids in both horizontal and vertical directions) are used. As a result, in total, eight DRB classifiers are trained, which correspond to

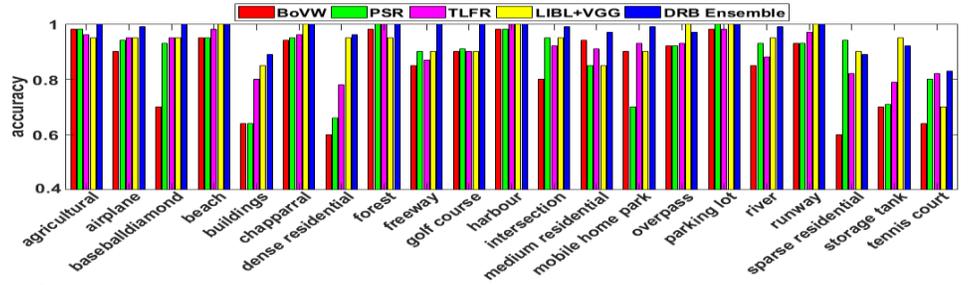
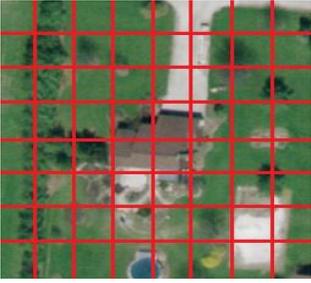


Fig.4. Segmentation using an 8x8 grid net Fig.5. Category wise performance comparison

eight different sliding windows. The corresponding overall classification accuracy and the training time per massively parallel rule are reported in Table II.

From Table II one can see that the DRB classifier can achieve 96%+ accuracy with a parallel training process of less than 6 seconds. It is noticeable that both the window size and the step size of the sliding window can influence the performance and the efficiency of the DRB classifier. The training process varies from 2 seconds to 15 minutes per rule/class and the classification accuracy varies from 94.62% to 96.19%. In general, with the same segmentation granularity, the smaller step size can allow the DRB classifier to gain more details of the image and, thus, leads to higher classification accuracy; meanwhile, the training time required by the DRB classifier will increase. Especially, for a DRB classifier with a smaller granularity, a smaller step size can increase the classification accuracy further. In contrast, the larger granularity results in higher computational efficiency, but it may decrease performance as the feature descriptor focuses only on the large scale spatial information. A DRB classifier with a smaller granularity relies mainly on small scale spatial information for the classification, it requires more training time, and its performance may also deteriorate due to the loss of large scale spatial information.

By creating an ensemble of DRB classifiers trained with segments of different granularities, more accurate classification can be expected. In this Letter, the DRB ensembles consisting of the best two ($G=2$), the best three ($G=3$) and all four DRB classifiers ($G=4$) using the sliding windows of the same step size are considered, and the corresponding classification accuracies are given in Table III.

One can see from Table III that, the more DRB classifiers involved in the ensemble, the better the performance of the ensemble. This is because the DRB ensemble can effectively

integrate the multiple scales of spatial information into the decision-making process. Meanwhile, creating an ensemble of more DRB classifiers with different levels of granularity can further increase the classification accuracy, but also costs more computation- and memory-resources. In real applications, it is necessary to take both performance and cost into consideration. Due to the limited space of this Letter, we considered only the ensemble of maximum four DRB classifiers. Nonetheless, different experimental settings of the DRB ensemble can be considered as well.

As the ensemble of the four DRB classifiers using a smaller step size achieved the highest accuracy (97.10%), in the remainder of this section this particular DRB ensemble is considered by default. The category-wise performance of the proposed DRB ensemble is depicted in Fig.5. As we can see, the DRB ensemble achieves 100% accuracy in classifying the following 10 categories: “agricultural”, “baseball diamond”, “beach”, “chaparral”, “forest”, “freeway”, “golf course”, “harbour”, “parking lot” and “runway”. There are only four categories with classification accuracy below 95%, namely; 89% for “buildings”, 89% for “sparse residential”, 92% for “storage tanks” and 83% for “tennis court”, all of which are categories demonstrating distinctive contents and a wide variety of textures. In particular, we observe that the misclassified classes in these four categories (“buildings”, “sparse residential”, “storage tanks” and “tennis court”) are categorized mainly as “dense residential”, “medium residential”, “mobile home park” and “intersection”. It is obvious that these misclassifications are caused by the visual similarity in the local textures shared by the images of the eight classes and the wide variety of details within the images themselves.

The accuracy of the proposed DRB ensemble is contrasted with the following state-of-the-art approaches for benchmark comparison. These algorithms are spatial pyramid co-variance (SPCV) and bags of visual words (BoVW), both of which were implemented in [17], multipath unsupervised feature learning (MUFL) in [9], pyramid of spatial relations (PSR) in [18], two-level feature representation (TLFR) as implemented in [8], linear SVM with pre-trained CaffeNet (SVM+Caffe) in [19]. We also applied the classical k -nearest neighbours classifier ($k=1$) trained with GIST and HOG features extracted from the grey-level images (KNN+GIST and KNN+HOG) as well as the LIBLINEAR classifier [14] and the linear SVM trained with the high-level features extracted from the original RGB images by VGG-VD-16 model (LIBL+VGG and SVM+VGG) for comparison. To be more specific, the codebook size in SPCV

TABLE II
PERFORMANCE OF DIFFERENT DRB CLASSIFIERS

Individual DRB		Window Size			
		4x4	5x5	6x6	7x7
Step 1	Accuracy	0.9543	0.9610	0.9619	0.9614
	Time (s)	807.38	272.57	68.05	11.84
Step 2	Accuracy	0.9462	0.9538	0.9614	0.9519
	Time (s)	164.27	26.06	5.75	1.70

TABLE III
PERFORMANCE OF DIFFERENT DRB ENSEMBLES

Accuracy		DRB Ensemble		
		G=2	G=3	G=4
Step 1		0.9638	0.9676	0.9710
Step 2		0.9629	0.9633	0.9667

and BoVW was 100 and 1000, respectively. PSR used three hierarchical levels with the codebook size of 5000 and the dictionary size of 300. The codebook size, dictionary size, fixed number and the arrangement for the grid search strategy of TLFR were set as 300, 128, 10 and 10, respectively. The grey-level images of the original size are used in SPCV, BoVW, PSR and TLFR. MUFL adopted the same parameter setting as used in [20] and a linear SVM was used after the features were learnt from the RGB images resized to 300×300 pixels. The accuracies of the DRB ensemble and the comparative algorithms are tabulated in Table IV. The category-wise comparison of the selected algorithms is also depicted in Fig. 5.

From Table IV and Fig. 5 one can see that, the proposed DRB ensemble achieves the highest overall classification accuracy among the state-of-the-art approaches, including the DL-based approaches [14], [19]. It can produce the highest category-wise classification accuracy for 18 out of 21 classes except for the categories “overpass”, “sparse residential” and “storage tank”. One possible reason is that some segments of the images of these categories confuse the DRB ensemble as they can be highly similar to the segments of other categories.

One appealing aspect of the proposed DRB ensemble is its fully human-interpretable and transparent IF...THEN... model structure generated after the training process. For a better illustration, examples of the 0-order AnYa type fuzzy rules generated autonomously from the data based on the segments of the remote sensing images with three levels of granularity are given in Table V, where the segments are reshaped into the same sizes for visual clarity.

VI. CONCLUSION

In this Letter, a novel approach based on an ensemble of the recently introduced DRB classifiers is proposed for remote sensing scene classification. Numerical examples demonstrate the very high accuracy and efficiency of the proposed DRB ensemble with its transparent, parallelizable, human-interpretable IF...THEN... structure as an appealing alternative to the state-of-the-art approaches.

As future research, we would like to study the performance of the DRB ensemble on different remote sensing benchmark problems and apply it to remote sensing images containing multiple sub-regions of different land use classes.

REFERENCES

- [1] Q. Weng, Z. Mao, J. Lin, and W. Guo, “Land-use classification via extreme learning classifier based on deep convolutional features,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 704–708, 2017.
- [2] L. Zhang, L. Zhang, and V. Kumar, “Deep learning for remote sensing data,” *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, 2016.
- [3] G. J. Scott, M. R. England, W. A. Starns, R. A. Marcum, and C. H. Davis, “Training deep convolutional neural networks for land-cover classification of high-resolution imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 549–553, 2017.
- [4] A. Oliva and A. Torralba, “Modeling the shape of the scene: A holistic representation of the spatial envelope,” *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, 2001.
- [5] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886–893.

TABLE IV
PERFORMANCE COMPARISON

Approach	Accuracy	Approach	Accuracy
KNN+HOG	53.81%	PSR [18]	89.10%
KNN+GIST	67.57%	TLFR [8]	91.12%
SPCV [17]	74.00%	SVM+ Caffe [19]	93.42%
BoVW [6]	76.80%	SVM+VGG	94.48%
SIFTSC [10]	81.67%	LIBL+VGG [14]	95.21%
MUFL [9]	88.08%	DRB Ensemble	97.10%

TABLE V
EXAMPLE OF FUZZY RULES

Fuzzy rule			
			<i>IF (sg-) OR (sg-) OR ... OR (sg-) THEN (airplane)</i>
			<i>IF (sg-) OR (sg-) OR ... OR (sg-) THEN (dense residential)</i>
			<i>IF (sg-) OR (sg-) OR ... OR (sg-) THEN (intersection)</i>
			<i>IF (sg-) OR (sg-) OR ... OR (sg-) THEN (tennis court)</i>

- [6] Y. Yang and S. Newsam, “Bag-of-visual-words and spatial extensions for land-use classification,” in *SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2010, pp. 270–279.
- [7] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [8] J. Gan, Q. Li, Z. Zhang, and J. Wang, “Two-level feature representation for aerial scene classification,” *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 11, pp. 1626–1630, 2016.
- [9] J. Fan, T. Chen, and S. Lu, “Unsupervised feature learning for land-use scene recognition,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 4, pp. 2250–2261, 2017.
- [10] A. M. Cheryadat, “Unsupervised feature learning for aerial scene classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 439–451, 2014.
- [11] P. P. Angelov and X. Gu, “MICE: Multi-layer multi-model images classifier ensemble,” in *IEEE International Conference on Cybernetics*, 2017, pp. 436–443.
- [12] P. P. Angelov and X. Gu, “Deep rule-based classifier with human-level performance and characteristics,” *Inf. Sci. (Ny)*, under review, 2017.
- [13] P. Angelov and R. Yager, “A new type of simplified fuzzy rule-based system,” *Int. J. Gen. Syst.*, vol. 41, no. 2, pp. 163–185, 2011.
- [14] G. S. Xia, J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, and L. Zhang, “AID: A benchmark dataset for performance evaluation of aerial scene classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3965–3981, 2017.
- [15] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *International Conference on Learning Representations*, 2015, pp. 1–14.
- [16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, C. Hill, and A. Arbor, “Going deeper with convolutions,” in *IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [17] Y. Yang and S. Newsam, “Spatial pyramid co-occurrence for image classification,” in *IEEE International Conference on Computer Vision*, 2011, pp. 1465–1472.
- [18] S. Chen and Y. Tian, “Pyramid of spatial relations for scene-level land use classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1947–1957, 2015.
- [19] A. B. Penatti, K. Nogueira, and J. A. Santos, “Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 44–51.
- [20] L. Bo, X. Ren, and D. Fox, “Multipath sparse coding using hierarchical matching pursuit,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2013, pp. 660–667.