

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

A novel active learning technique for multi-label remote sensing image scene classification

Bayable Teshome Zegeye, Begüm Demir

Bayable Teshome Zegeye, Begüm Demir, "A novel active learning technique for multi-label remote sensing image scene classification," Proc. SPIE 10789, Image and Signal Processing for Remote Sensing XXIV, 107890B (9 October 2018); doi: 10.1117/12.2500191

SPIE.

Event: SPIE Remote Sensing, 2018, Berlin, Germany

A Novel Active Learning Technique for Multi-Label Remote Sensing Image Scene Classification

Bayable Teshome Zegeye¹, Begüm Demir²

¹Dept. of Information Engineering and Computer Science, University of Trento, Trento, Italy
²Faculty of Electrical Engineering and Computer Science, Technical University of Berlin, Berlin, Germany

ABSTRACT

This paper presents a novel multi-label active learning (MLAL) technique in the framework of multi-label remote sensing (RS) image scene classification problems. The proposed MLAL technique is developed in the framework of the multi-label SVM classifier (ML-SVM). Unlike the standard AL methods, the proposed MLAL technique redefines active learning by evaluating the informativeness of each image based on its multiple land-cover classes. Accordingly, the proposed MLAL technique is based on the joint evaluation of two criteria for the selection of the most informative images: i) multi-label uncertainty and ii) multi-label diversity. The multi-label uncertainty criterion is associated to the confidence of the multi-label classification algorithm in correctly assigning multi-labels to each image, whereas multi-label diversity criterion aims at selecting a set of un-annotated images that are as more diverse as possible to reduce the redundancy among them. In order to evaluate the multi-label uncertainty of each image, we propose a novel multi-label margin sampling strategy that: 1) considers the functional distances of each image to all ML-SVM hyperplanes; and then 2) estimates the occurrence on how many times each image falls inside the margins of ML-SVMs. If the occurrence is small, the classifiers are confident to correctly classify the considered image, and vice versa. In order to evaluate the multi-label diversity of each image, we propose a novel clustering-based strategy that clusters all the images inside the margins of the ML-SVMs and avoids selecting the uncertain images from the same clusters. The joint use of the two criteria allows one to enrich the training set of images with multi-labels. Experimental results obtained on a benchmark archive with 2100 images with their multi-labels show the effectiveness of the proposed MLAL method compared to the standard AL methods that neglect the evaluation of the uncertainty and diversity on multi-labels.

Keywords: active learning, multi-label uncertainty, multi-label diversity, multi-label classification, remote sensing.

1. INTRODUCTION

With the continuous advances in satellite technology, the past decade has observed a tremendous growth in remote sensing (RS) image archives [1]. Accordingly, classification of RS image scenes, which are usually achieved by direct supervised classification of each image in the archive, has received an extensive attention in RS [2]. Most of the supervised classifiers developed for RS scene classification problems are trained on images annotated with single high-level category labels, which are associated to the most significant content of the image. However, each RS image typically consists of multiple classes and thus can simultaneously be associated with different low-level land-cover class labels (i.e., multi-labels). Thus, supervised classification methods that exploit training images annotated by multi-labels have been recently introduced in RS. Such an approach requires a sufficient number of high quality annotated training images with multi-labels. However, collecting such multi-labels is costly in terms of human time and effort.

An effective approach to reduce the annotation effort is active learning (AL) that aims at finding the most informative images in the archive that, when annotated and included in the set of training images, can significantly improve the classification performance [3], [4]. Human experts are requested to assign labels to a small optimal number of images identified by the classification algorithm. By this way, the collection of images for non-useful (i.e., redundant) images is avoided, leading: i) to reduce the computational complexity for the learning phase (due to the optimized training set with a small number of labeled images), and ii) to increase the classification accuracy (due to the improved class models estimated on a high quality training set defined on the basis of the classification rule of the considered classifier). Most of the AL studies presented in the RS literature select a batch of informative images at each iteration on the basis of two main criteria: 1) uncertainty, and 2) diversity of images. The uncertainty criterion evaluates the confidence of the

supervised algorithm in correctly classifying the considered image, whereas the diversity criterion aims at selecting a set of unlabeled images that are as more diverse as possible to each other. In [3] an AL method is presented in the context of Support Vector Machine (SVM) classifier, for which the uncertainty and diversity criteria have been applied in two consecutive steps. In the first step the most uncertain images are selected from the archive. To this end, the unlabeled images closest to the current separating hyperplane (those that are the most uncertain) are initially selected by margin sampling (MS). In the second step, the images that are diverse to each other among the uncertain ones are chosen on the basis of the distances estimated between them [3]. In [4], in the first step of the AL method the most uncertain images are selected from the archive on the basis of margin sampling strategy similar to [3]. In the second step the images that are both diverse (i.e., distant) to each other and associated to high density regions of the image feature space in the archive are chosen from the most uncertain images. This step is achieved by a clustering-based strategy.

All of the above-mentioned AL methods presented in RS assume that each image in the archive is categorized under a single image category label and thus not suitable for multi-label RS image classification problems [2]. To address this problem, in this paper we present a novel multi-label active learning (MLAL) technique in the context of multi-label RS scene classification problems. The proposed MLAL technique is developed in the framework of the multi-label SVM classifier (ML-SVM) [5] that consists of a parallel architecture made up of P one versus all SVMs to assign multi-labels to each image in the archive, where P is the total number of land-cover classes present in the archive. Each classifier is trained to solve a binary classification problem defined by one land-cover class against all the others. Unlike the standard AL methods, the proposed MLAL technique redefines AL by evaluating the informativeness of each image based on its multiple land-cover classes. Consequently, the proposed MLAL technique is based on the joint evaluation of two criteria for the selection of the most informative images: i) multi-label uncertainty and ii) multi-label diversity. The multi-label uncertainty criterion is associated to the confidence of the multi-label classification algorithm in correctly assigning multi-labels to each image, whereas multi-label diversity criterion aims at selecting a set of un-annotated images that are as more diverse as possible to reduce the redundancy among them. In order to evaluate the multi-label uncertainty of each image, we propose a novel multi-label margin sampling strategy that: 1) considers the P functional distances of each image to P SVM hyperplanes; and then 2) estimates the occurrence on how many times this image falls inside the margins of P SVMs. If the occurrence is small, the classifiers are confident to correctly classify the considered image, and vice versa. In order to evaluate the multi-label diversity of each image, we propose a novel clustering-based strategy that clusters all the images inside the margins of the P SVMs and avoid selecting the uncertain images from the same clusters. It is worth noting that the number of clusters is determined by applying rule of thumb (which is defined as the square root of the half of the total number of samples within the margin) and thus can be different associated to each SVM. The joint use of the proposed two criteria allows one to enrich the training set of images with multi-labels. Experimental results obtained on a benchmark archive with 2100 images with their multi-labels show the effectiveness of the proposed MLAL method compared to the standard AL methods that neglect the evaluation of the uncertainty and diversity of multi-labels.

The paper is organized into four sections. Section 2 introduces the proposed MLAL method, while Section 3 describes the considered archive and illustrates the experimental results. Finally, Section 4 draws the conclusion of this work.

2. PROPOSED MULTI-LABEL ACTIVE LEARNING TECHNIQUE

Let $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$ be an archive of N RS images, where \mathbf{X}_i is the i -th image. We assume that an initial training set T is available with a small number of annotated images with multi-labels. We propose a novel multi-label AL (MLAL) method to expand the initial training set in RS scene classification applications. The aims of the proposed AL method are as follows: i) to achieve a training set of annotated images with multi-labels as small as possible; and ii) to classify the images with high accuracy. The proposed MLAL method is defined in the context of ML-SVMs and selects a batch $S = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_h\}$ of h images at each AL iteration based on the two criteria: i) multi-label uncertainty; and ii) multi-label diversity. The multi-label uncertainty of images is assessed according to the novel multi-label margin sampling (MS) strategy, whereas diversity of image is evaluated by a novel clustering-based strategy. At each iteration, the proposed AL method jointly evaluates the above-mentioned two criteria based on two consecutive steps to select the batch S of images. In the first step, multi-label uncertainty scores are assigned to the images according to the proposed multi-label MS technique. In the second step the most diverse h images among the most uncertain (i.e., ambiguous) images are chosen. Fig. 1 shows the general block scheme of the proposed MLAL method.

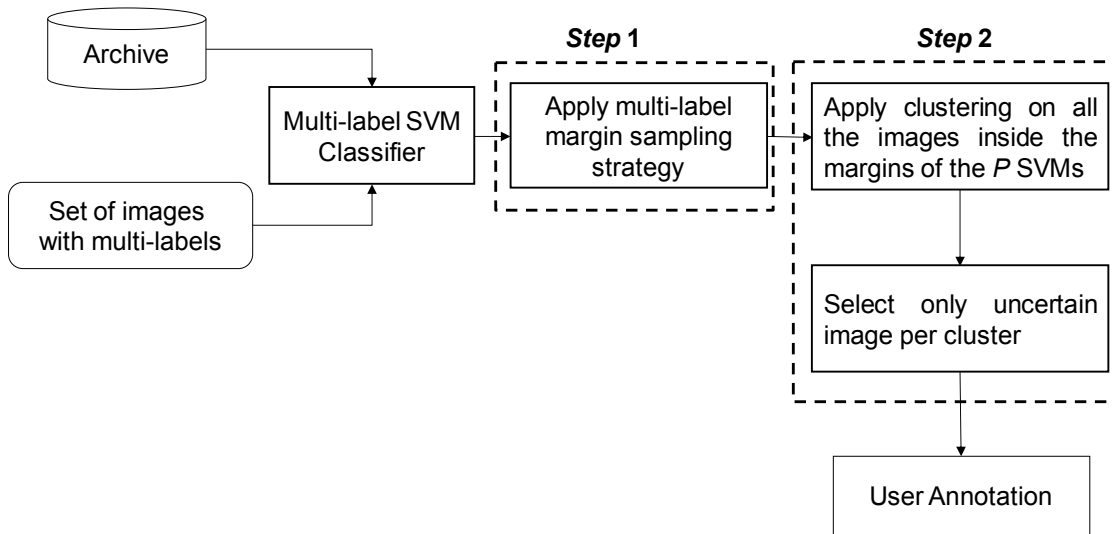


Figure 1. Block scheme of the proposed MLAL method.

ML-SVM consists of a parallel architecture made up of P one versus all SVMs to assign multi-labels to each image in the archive, where P is the total number of land-cover classes present in the archive. Each classifier is trained to solve a binary classification problem defined by one land-cover class against all the others. Fig. 2 shows a qualitative example of ML-SVM when $P=3$, aiming to assign existence/inexistence information of three classes to each image (sample). From this qualitative example, one can see that for the circle in pink all three classes are present, while only the first class is present for the circle in blue.

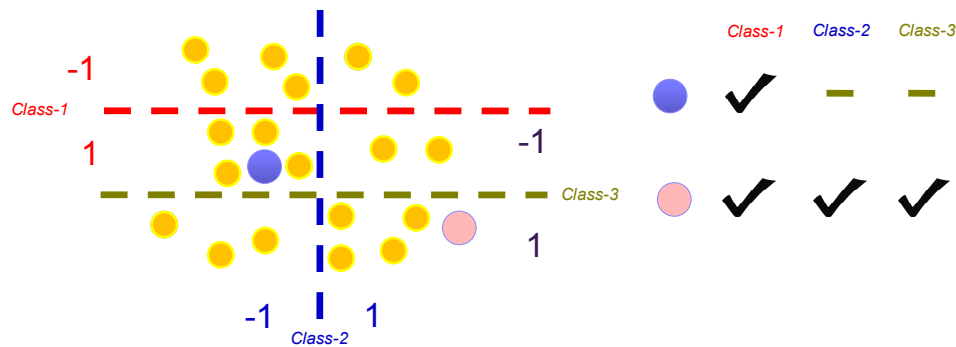


Figure 2. Qualitative example of ML-SVM in a 2-D feature space for three classes. All circles denote the unannotated images, while class existence information is given on the right for circles in blue and pink.

The first step of the proposed MLAL is devoted to evaluate the multi-label uncertainty and assign uncertainty scores to all unannotated images on their correct multiple classes according to the ML-SVM classification properties. The basic idea behind this concept is that images, which have the lowest probability to be accurately classified by the considered multi-label classifier, are the most beneficial to be included in the training set. It is worth noting that for SVM classification the images closest to the separating hyperplane have low confidence to be correctly classified. When standard SVM is used, one of the most popular AL method is MS, which selects the unlabeled images closest to the separating hyperplane, as they are the samples considered with the lowest confidence (i.e., those that have the maximal

uncertainty on the true information class) [4]. This approach is simple, effective and has fast implementation, whereas it evaluates and selects the most uncertain images by considering that each image is associated to a single category label. Thus, it is not suitable for multi-label classification problems. Inspired from this strategy, in this paper we introduce a method that selects the most uncertain images according to a uncertainty value $c(\mathbf{X}_j)$ which is defined based on a novel multi-label MS strategy. This strategy considers the functional distance $f_i(\mathbf{X}_j)$, $i=1, \dots, P$ of each image to the P decision boundaries of the binary SVM classifiers included in the ML-SVM architecture. To this end, the distance of each image to each hyperplane is initially calculated and a set of P distance values $\{f_1(\mathbf{X}_j), f_2(\mathbf{X}_j), \dots, f_P(\mathbf{X}_j)\}$ is obtained. Then, the uncertainty value $c(\mathbf{X}_j)$ is calculated as the occurrence on how many times this image falls inside the margins of P SVMs. If this value is high, the image \mathbf{X} is assigned to its classes with low confidence. Thus, this image is considered as uncertain and selected for better modeling the decision functions in the corresponding positions of the feature space. On the contrary, if $c(\mathbf{X}_j)$ is small, the decisions are reliable (i.e., the image is very certain). If there are images with equal number of uncertainty scores, we consider the average of the functional distances for these images. Fig. 3 shows qualitative example on evaluation of the multi-label uncertainty of the images. Circle in light blue is the most uncertain, since it has the highest uncertainty score that is 3. Circles in dark blue and in red are the subsequent most uncertain ones, having the equal uncertainty scores that are 2. Thus, it is required to check their average functional distances to the hyperplane to sort them among themselves. The circle associated with the smallest average distance is the most uncertain. The uncertainty scores related to circles in green and in orange are also equal (which is 1), and thus also for these samples their average distances should be evaluated.

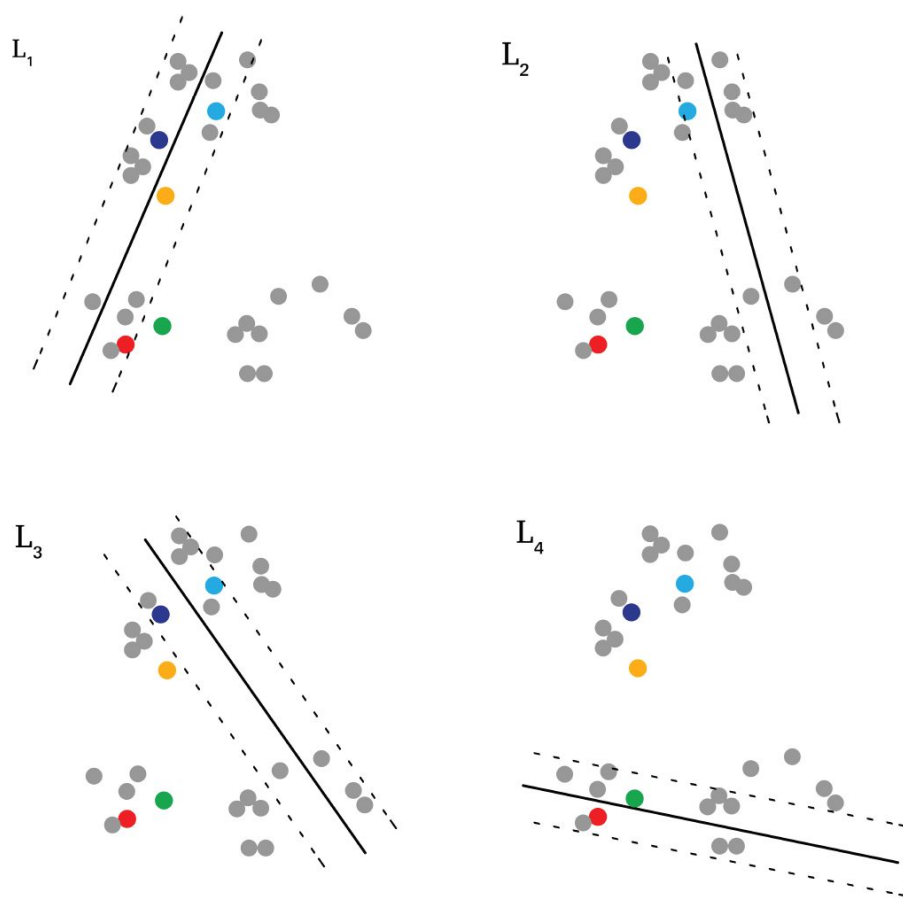


Figure 3. Qualitative example on multi-label uncertainty evaluation of the images.

After assigning the confidence scores, the second step is devoted to select a batch $S = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_h\}$ of h images that are diverse among the uncertain ones. In order to evaluate the multi-label diversity of each image, we propose a novel clustering-based strategy that: i) clusters all the images inside the margins of the P SVMs; ii) selects the sample with the highest uncertainty score first and initializes a batch; and iii) avoids selecting the uncertain images from the same clusters. It is worth noting that we have considered clustering, since it is an effective way to assess the diversity of samples as the patterns within the same cluster are correlated and vice versa [4]. The number of clusters is determined by applying rule of thumb and thus can be different associated to each SVM. Fig. 4 shows qualitative example on multi-label diversity evaluation of the images. According to the proposed strategy, in this qualitative example, the most uncertain image that is shown by a circle in light blue is selected first. Let us assume that the circle in dark blue is associated with the smallest average distance and thus the most uncertain second sample. Since this sample is in the same cluster with the already selected sample (i.e., circle in light blue), this sample will not be selected. The algorithm continues and select circle in red, because this sample is not located within the same cluster of the already selected sample (i.e., the circle in light blue). These processes continue until h samples are selected. Finally, the selected samples are annotated by the supervisor and added to the existing labeled training set.

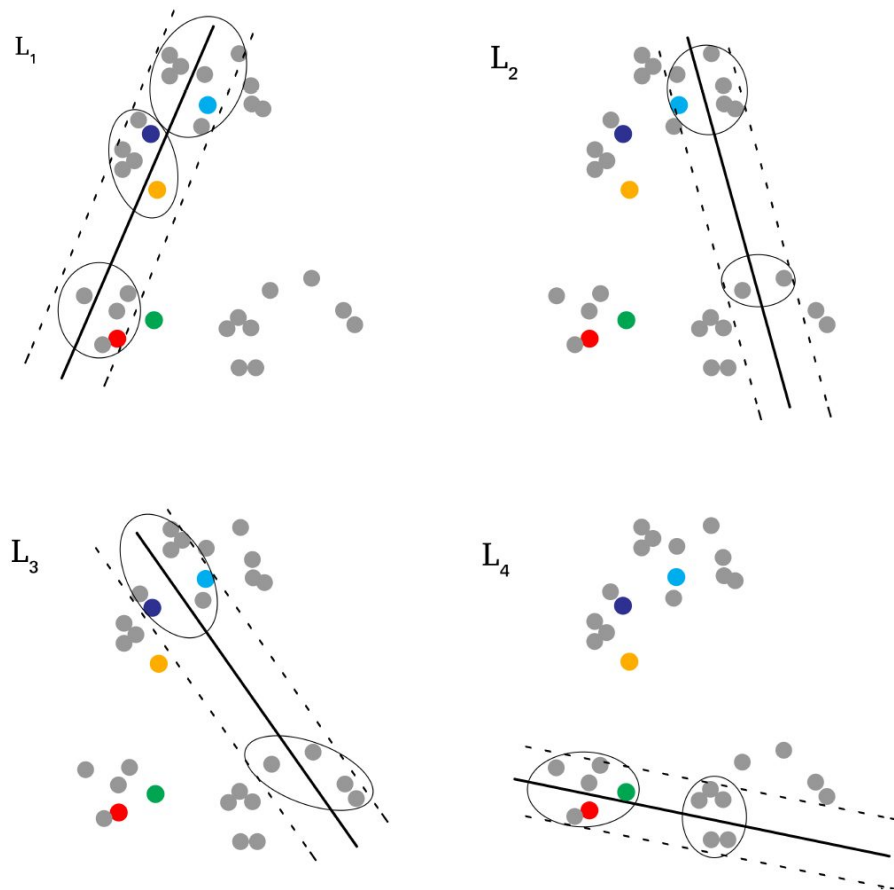


Figure 4. Qualitative example on multi-label diversity evaluation of the images.

3. EXPERIMENTAL RESULTS

In our experiments the UCMERGED benchmark archive, which consists of 2100 images selected from aerial orthoimagery with a spatial resolution of 30 cm [6] was utilized. To assess the performance of the proposed method, we have used the annotations of multiple classes of images that are described in [7] and available at

'<http://bigearth.eu/datasets>'. The total number of considered class labels is 17 (which are: airplane; bare-soil; buildings; cars; chaparral; court; dock; field; grass; mobile-home; pavement; sand; sea; ship; tanks; trees; water), while the number of labels associated with each image varies between 1 and 7.

The available images in the archive (2100 images) were randomly divided to derive a test set with 1050 images and a pool set with 1050 images. 40 images in the pool are randomly selected as initial training images, and the rest are considered as unlabeled images. The images in the archive are modeled to bag-of-visual-words (BOVW) representation of histogram-based features extracted by scale invariant feature transform (SIFT). In our experiments, we used an SVM classifier with chi-square kernel [14]. The values for the regularization parameter was chosen performing a grid-search model selection only at the first iteration of the AL process as suggested in [4]. In the experiments, clustering is implemented using the kernel k-means clustering method, because of its efficiency for non-linearly separable data in the kernel space where SVM operates. We compared the proposed method that consists of multi-label MS and clustering based multi-label diversity strategies (denoted as proposed MMS-CBMD in the experiments) with: i) the random sampling (denoted as random sampling) that selects the samples randomly at each iteration, and ii) an AL method that includes margin sampling and clustering based diversity (denoted as MS-CBD) [4]. We carried out the experiments by adding 10 images at each iteration of AL. Results of each method are provided as learning rate graphs, which show the average recall on 10 trials (which are associated to the random selection of 10 different initial training sets) versus the number of AL iteration. Average recall is defined as the average ratio of the total number of intersecting labels between the estimated and true class labels for each test image.

Fig. 5 shows the behavior of the average (on 10 trials) recall versus the number of annotated images during AL iterations. In the figure, we compare the effectiveness of the proposed MMS-CBMD with those of the random sampling and of the MS-CBD method presented in [4]. From the figure, one can see that the proposed method leads to the highest recall at most of the AL iterations and significantly outperforms both the MS-CBD and the random sampling. Only at early iterations the MMS-CBD provides similar precision to the MS-CBD and random sampling, whereas it considerably increases the retrieval performance at the later iterations. As an example, the proposed MMS-CBD achieves an average recall improvement of 0.7% over the MS-CBD and random sampling when the number of total annotated images is 300 (see Fig. 5). In details, both the MMS-CBD and the MS-CBD methods achieve higher accuracies than the random sampling at most of the iterations. Moreover, the MMS-CBD method provides the same recall achieved by the MS-CBD with a smaller number of annotated images. For example, the MMS-CBD method obtains a recall of 82% with 400 annotated images, whereas the MS-CBD reaches similar recall with around 600 annotated images (see Fig. 5). These results show that selecting both uncertain and diverse unannotated images based on their multi-labels is very important. All these results point out the robustness and effectiveness of the proposed MMS-CBD method. Fig. 6 shows examples of images with their correct class labels and also with the predicted labels by the MMS-CBMD, the MS-CBD and random sampling methods. By analyzing the figure, one can see that the proposed method assigned much more accurate labels than the other methods.

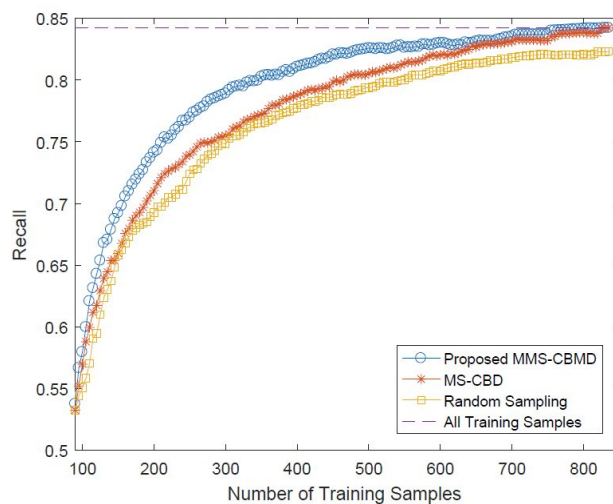


Figure 5. Average (on ten trials) recall (in %) obtained by the MMS-CBMD, the MS-CBD and random sampling methods. The dashed line “All training samples” shows the accuracy obtained including all unlabeled images in the training set after manual annotation.

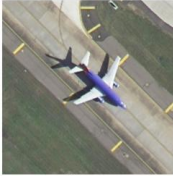


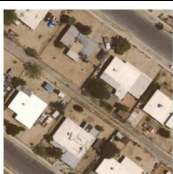




Image	True Class Labels	Predicted Labels		
		Proposed MMS-CBMD	MS-CBD	Random Sampling
 airport	airplane, grass, pavement	airplane, grass, pavement	grass, pavement	bare soil, grass, pavement
 golf course	grass, pavement, trees	grass, pavement, sand, trees	pavement, trees	grass, sand, trees
 intersection	bare soil, buildings, cars, grass, pavement, trees	bare soil, buildings, cars, grass, pavement, trees	bare soil, cars, grass, pavement, trees	bare soil, cars, pavement, trees
 medium residential	bare soil, buildings, cars, pavement, trees	bare soil, buildings, cars, grass, pavement, trees	buildings, cars, pavement, trees	buildings, cars, grass, pavement, trees
 overpass	cars, grass, pavement	cars, grass, pavement, trees	bare soil, cars, pavement, trees	bare soil, cars, pavement, trees
 river	grass, sand, trees, water	grass, sand, trees, water	trees, water	grass, trees
 storage tanks	bare soil, buildings, cars, grass, pavement, tanks, trees	bare soil, buildings, cars, grass, pavement, tanks	bare soil, buildings, grass, pavement, tanks	bare soil, buildings, grass, pavement, tanks
 tennis area	bare soil, court, grass, trees	bare soil, court, grass, trees	court, grass, trees	court, grass, pavement, trees

Figure 6. Examples of images with class labels obtained by the MMS-CBMD, the MS-CBD and random sampling methods.

4. CONCLUSION

In this paper, we have presented a novel multi-label active learning method to the definition of effective training sets for the classification of remote sensing images with multi-label SVMs (ML-SVMs). The proposed MLAL method evaluates the multi-label uncertainty and multi-label diversity in order to select a batch of most informative images to be annotated and included in the training set for multi-label RS scene classification problems. The multi-label uncertainty of images is assessed by the proposed multi-label margin sampling technique that evaluates the occurrence on how many times each image falls inside the margins of ML-SVMs, whereas the multi-label diversity is measured based on a clustering-based strategy that aims at avoiding selection of the the uncertain images from the same clusters. In order to evaluate the above-mentioned two criteria, a two steps procedure is adopted in the paper. The first step is devoted to assign the multi-label uncertainty scores to all the images in the archive, whereas the second step aims at choosing the most diverse images among the most uncertain ones. In the experimental analysis, we also compared the proposed method with the most promising state-of-the-art AL method in RS (which neglects the evaluation of the uncertainty and diversity of multi-labels). By this comparison, we observed that the proposed method allows one to significantly reduce the cost of the collection of multi-labels to reach the desired classification accuracy compared to the state-of-the-art AL method.

As a final remark, we would like to point out that the use of efficient techniques for the exploitation of AL methods in multi-label RS scene classification problems is becoming an important topic. In this context, the proposed method is very promising as it allows optimizing the definition of a training set with multi-labels, decreasing significantly the cost and effort required for collection of multi-labels.

ACKNOWLEDGMENT

This work was supported by the European Research Council under the ERC Starting Grant BigEarth-759764.

REFERENCES

- [1] Liu, Q., Hang, R., Song, H., and Li, Z., "Learning multiscale deep features for high-resolution satellite image scene classification", *IEEE Transactions on Geoscience and Remote Sensing*, 56 (10), 117-126 (2018).
- [2] Chaudhuri, B., Demir, B., Bruzzone, L., and Chaudhuri, S., "Region-based retrieval of remote sensing images using an unsupervised graph-theoretic approach," *IEEE Geoscience and Remote Sensing Letters*, 13 (7), 987-991(2016).
- [3] Ferecatu, M., and Boujmaa, N., "Interactive remote-sensing image retrieval using active relevance feedback," *IEEE Transactions on Geoscience and Remote Sensing*, 45 (40), 818-826 (2007).
- [4] Demir, B., and Bruzzone, L., "A novel active learning method in relevance feedback for content based remote sensing image retrieval", *IEEE Transactions on Geoscience and Remote Sensing*, 53 (5), 2323-2334 (2015).
- [5] Nasierding G., and Kouzani, A. Z., "Empirical study of multi-label classification methods for image annotation and retrieval," in *Int. Conf. Digital Image Computing: Techniques Applications*, 2010, pp. 617622.
- [6] Yang, Y., and Newsam, S., "Geographic image retrieval using local invariant features," *IEEE Transactions on Geoscience and Remote Sensing*, 51 (2), 818-832 (2013).
- [7] Chaudhuri, B., Demir, B., Bruzzone, L., and Chaudhuri, S., "Multi-label Remote Sensing Image Retrieval using a Semi-Supervised Graph-Theoretic Method," *IEEE Transactions on Geoscience and Remote Sensing*, 56 (2), 1144-1158 (2018).