ACCESSIBILITY-FREE ACTIVE LEARNING FOR HYPERSPECTRAL IMAGE CLASSIFICATION

Chenying Liu[1]*, Jun Li*[1]*, Mercedes E. Paoletti*[2]*, Juan M. Haut*[2]*, Antonio Plaza*[2]*, and Qian Shi*[1]

[1]Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation,
School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China
[2]Hyperspectral Computing Laboratory, University of Extremadura, E-10003 Caceres, Spain

ABSTRACT

This work proposes a new collaborative active and semi-supervised learning approach, named accessibility-free active learning (AFAL), for hyperspectral imaging classification. The proposed approach aims to tackle an existing problem in traditional active learning methods, that is, the fact that some selected samples are not accessible by oracles for assigning them pseudo labels, *i.e.*, confident predictions for the classifier. The proposal specifically addresses this problem using superpixels in a self-training context. Specifically, AFAL first generates a set of candidates locally around the labeled pixels and then expands them to other subregions via a density peak-based augmentation strategy, in order to guarantee the confidence of pseudo labels. Our experimental results, obtained on two real and well-used hyperspectral images, reveal that the proposed scheme can lead to state-of-the-art performance.

Index Terms— Hyperspectral image classification, active learning (AL), semi-supervised learning (SSL), superpixels.

1. INTRODUCTION

In standard supervised hyperspectral imaging (HSI) classification, predictive models are learned from some *a priori* labeled samples. Their quantity and quality are crucial to obtain satisfactory performance. However, in many tasks it is difficult to get such information due to the high cost and time needed in the data labeling process. A small amount of labeled samples is generally insufficient to train a good classifier. To address this problem, active learning (AL) is an effective approach that iteratively selects the most informative unlabeled samples (*i.e.* queries) and then asks an oracle to annotate them [1, 2, 3, 4]. As a result, the classification performance can be greatly improved with a minimum number of labeled samples.

In traditional AL methods, it is assumed that the labeling cost only depends on the number of queries [5], that is, all the samples can be annotated with the same effort, which is not always the case in practice. There are generally two annotation strategies in remote sensing: i) *in situ* field surveys and ii) visual photo-interpretation [6]. The former one is mandatory when extremely detailed land cover maps are required, while its cost is much higher than that of the latter one. Recent researches on cost-sensitive AL (CSAL) have taken into consideration the physical traveling costs involved [7, 8] On the contrary, photo-interpretation is easier and cheaper, yet subjective and dependent on the knowledge of the analyst [6]. In practice, a hybrid strategy is preferred, where photo-interpretation is complemented by limited field surveys in order to lower the cost and guarantee the confidence of labels. However, some labels are inevitably unaccessible in the aforementioned cases. On the one hand, the large spatial extent and the complex terrain can make many sites unreachable in field surveys. On the other hand, highly mixed samples (which often occur around class boundaries) are difficult to interpret, especially in hyperspectral images with a low spatial resolution. Existing AL methods (including CSAL) generally skip or avoid these samples, leading to a loss of highly informative training samples.

In contrast, semi-supervised learning (SSL) trains the classifier using both labeled and unlabeled data [9]. Self-training avoids the sample accessibility by supplementing the training set with pseudo-labeled samples, automatically annotated by the classifier rather than by human experts [10, 11].

Based on the above observations, in this work we develop an accessibility-free active learning (AFAL) approach for HSI data classification, where the queries out of accessibility are labeled by the classifier under a superpixel assumption in a self-training context, thus making the labeling more flexible. Different from traditional AL methods, that select samples globally, AFAL first selects samples locally around the labeled ones and then gradually expands the candidate set using a density peak (DP)-based augmentation strategy introduced by [11] in order to guarantee the confidence of pseudo labels. Furthermore, we utilize the multinomial logistic regression model with a Markov random field regularizer (MLR-MRF) [2, 12] as the classifier and the *breaking ties* (BT) [1] function as the sampling criterion. Notice that the proposed AFAL is virtually a collaborative active and semi-supervised learning method.

The paper is organized as follows. The proposed AFAL method is introduced in section 2. The experimental results and their discussion are presented in section 3. Finally, section 4 concludes the paper with a few remarks and future perspectives.

2. METHODOLOGIES

Let $\mathbf{X} \equiv \{\mathbf{x}_1, \cdots, \mathbf{x}_n\} \in \mathbb{R}^{d\times n}$ be a hyperspectral image with n samples and d bands; also, let $\mathbf{y} \equiv \{y_1, \cdots, y_n\} \in \mathbb{R}^n$ be the corresponding image of labels where each y_i falls into one of the possible K categories, $y_i \in \mathcal{K} \equiv \{1, \ldots, K\}$. The class densities can be modeled as

$$p(y_i = k|\mathbf{x}_i, \boldsymbol{\omega}) = \frac{\exp(\boldsymbol{\omega}^{(k)^T}\mathbf{h}(\mathbf{x}_i))}{\sum_{c=1}^{K}\exp(\boldsymbol{\omega}^{(c)^T}\mathbf{h}(\mathbf{x}_i))}, \quad (1)$$

where $\boldsymbol{\omega} \equiv [\boldsymbol{\omega}^{(1)}, \cdots, \boldsymbol{\omega}^{(K-1)}]$ is the logistic regressor set with $\boldsymbol{\omega}^{(K)} = \mathbf{0}$, since the density in Eq. (1) does not depend on translations on the regressors $\boldsymbol{\omega}^{(k)}$; $\mathbf{h}(\mathbf{x}_i) \equiv [h_1(\mathbf{x}_i), \cdots, h_l(\mathbf{x}_i)]^T$

This work was supported by National Natural Science Foundation of China under Grant 61771496, Guangdong Provincial Natural Science Foundation under Grant 2016A030313254, National Key Research and Development Program of China under Grant 2017YFB0502900.

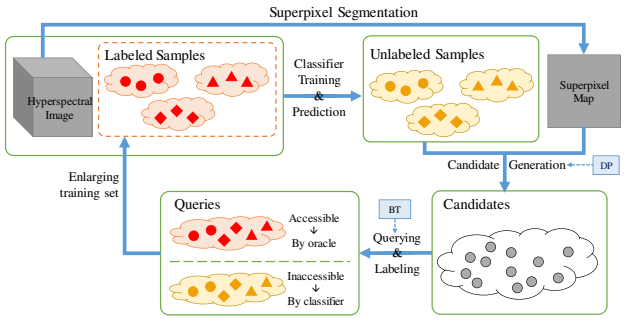


Fig. 1. Graphical illustration for the proposed AFAL approach.

is a feature vector composed of l fixed functions of the input, which is constructed by the nonlinear Gaussian radial basis function (RBF) in the form of $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2)$ using σ as the scale parameter [13]. In this context, suppose $\mathcal{D}_T \equiv \{(y_1, \mathbf{x}_1), \ldots, (y_t, \mathbf{x}_t)\}$ is a training set with t samples and $p(\boldsymbol{\omega})$ is *a priori* distribution of $\boldsymbol{\omega}$, then the learning of $\boldsymbol{\omega}$ is formulated as

$$\widehat{\boldsymbol{\omega}} = \arg\max_{\boldsymbol{\omega}} \ \ell(\boldsymbol{\omega}|\mathcal{D}_T)p(\boldsymbol{\omega}). \tag{2}$$

In this work, a new AFAL method has been proposed to tackle the difficulty of limited training samples and unaccessible manual labels. As illustrated in Fig. 1, the proposed AFAL consists of three main steps which are described below in details.

2.1. Candidate Generation

At each iteration in AFAL, the candidate set is generated under the superpixel assumption with DP augmentation [11] in order to guarantee the confidence of pseudo labels once the selected samples are not accessible.

Then, the neighborhood of a sample can be defined as the superpixel it belongs to, based on the fact that superpixels are homogeneous regions in the image $\mathbf{X}$. Thus, it is reasonable to assume that the pixels within such neighborhood share the same label. In this work, the regional clustering based spatial preprocessing (RC-SPP) method [14] has been employed for HSI data segmentation due to its capacity to generate high-quality, compact, and nearly uniform superpixels. Furthermore, from a spectral point of view, similar features generally belong to the same class. Based on the two aforementioned observations, the candidate set is confined within those pixels that are both located in the same superpixel and predicted with the same labels. As a result, though out of accessibility, they can be assigned to a relatively confident pseudo label by the classifier. However, the unlabeled samples will be constrained into very few superpixels. In order to expand the candidate set, DP augmentation [11] has been considered, which is defined by a clustering technique and a distance criterion to select modes from unlabeled superpixels, *i.e.*, those without initial labeled ones, in which case the modes are the most representative ones within the superpixels. After that, the same superpixel-based candidate generation strategy is implemented in terms of the selected modes.

2.2. Querying and Labeling

After the candidate generation, the most informative samples are selected from the candidate set and then labeled by the oracle or the classifier. In the querying step, the queries are first identified in accordance with some sampling criterion, *i.e.*, the BT function formulated as follows:

$$\begin{aligned}\mathbf{x}_i' = \arg\min_{\mathbf{x}_i \in \mathcal{D}_C} \big[&\max_{k \in \mathcal{K}} p(y_i = k|\mathbf{x}_i, \widehat{\boldsymbol{\omega}}) \\ &- \max_{k \in \mathcal{K}\setminus\{k^+\}} p(y_i = k|\mathbf{x}_i, \widehat{\boldsymbol{\omega}}) \big],\end{aligned} \tag{3}$$

where $\mathcal{D}_C$ is the candidate set. Obviously, the queries minimize the distance between the first two most probable classes indicating that they are close to boundaries.

Then, in the labeling step, the queries are labeled by the oracle or by the classifier if out of accessibility, that is,

$$y_i' = \begin{cases} y_i & y_i \text{ is accessible} \\ \widehat{y}_i & y_i \text{ is inaccessible} \end{cases}, \tag{4}$$

where y_i is the manual label of $\mathbf{x}_i'$, while $\widehat{y}_i = \max_k p(y_i = k|\mathbf{x}_i, \widehat{\boldsymbol{\omega}})$ is the prediction. As a consequence, the labeling process is free of the sample accessibility problem.

2.3. MLR-MRF Classification

Finally, and following [2], we use the MLR via variable splitting and augmented Lagrangian (LORSAL) algorithm [15] followed by an MRF regularizer [11], called MLR-MRF, to solve Eq. (2). The LORSAL algorithm has been shown to be effective for dealing with large quantities of training samples, while the MRF scheme incorporates piecewise smoothness into the final classification results, taking into account that spatially adjacent pixels generally exhibit the same class label in a real scene.

3. EXPERIMENTAL RESULTS

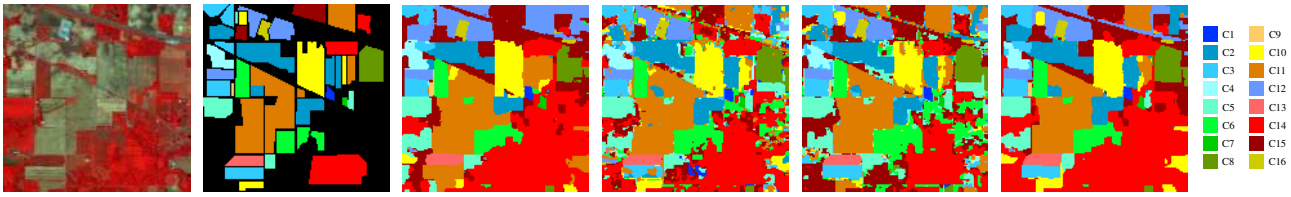
In this section, the proposed AFAL approach is evaluated on two real and well-used HSI scenes. For comparison, another three cases have been considered in experiments. The first one is the superpixel based semi-supervised active learning method with DP augmentation (SSAL-SDP) in [11], which can be regarded as the semi-supervised version of AFAL. In the second one, the classifier has been retrained only using the manually labeled samples in AFAL (hereinafter referred as AFAL-True). The last one is random sampling (RS) where manually labeled training samples have been randomly selected with the same size. In the following, the overall accuracy (OA), average accuracy (AA), class individual accuracy (CA), and kappa statistic (κ) obtained from five Monte Carlo runs have been considered for quantitative evaluation. To generate the initial training set $\mathcal{D}_T$, 5 samples per class have been randomly selected from the ground truth for each scene.

3.1. AVIRIS Indian Pines scene

The first HSI scene is the public **AVIRIS Indian Pines scene**[1] composed of 145×145 samples and 200 bands after removing the noise and water absorption ones. The available ground truth contains 10366 labels categorized in 16 classes (the class sizes are listed in Table 1). For SSAL-SDP and AFAL, we added 25 samples into the training set and 5 modes for candidate generation at each iteration.

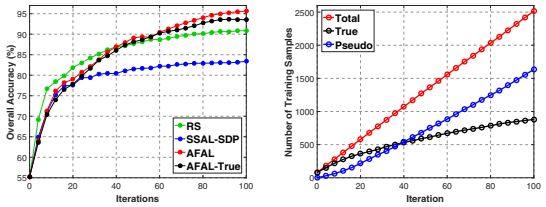
Fig. 2 and Table 1 list the OAs and the numbers of training samples obtained by the considered methods. It can be seen that

[1] http://dynamo.ecn.purdue.edu/biehl/MultiSpec



(a) False color map (b) Ground truth (c) SDP (80.39%) (d) RS (91.96%) (e) AFAL-True (94.38%) (f) AFAL (97.14%)

Fig. 3. Classification maps (along with the corresponding OAs) obtained after 100 iterations for the AVIRIS Indian Pines scene, where a false color map (a) and the ground truth (b) are also provided for reference.



(a) OAs (b) Numbers of training samples

Fig. 2. OAs and the number of training samples expressed as a function of the number of iterations for the AVIRIS Indian Pines scene. Here, true and pseudo represents the manually labeled and pseudo-labeled samples, respectively.

the proposed AFAL obtained very competitive and more robust results when compared to the others. On the one hand, AFAL-True obtained higher OAs than RS meaning that AFAL has the ability to select more informative samples. On the other hand, AFAL outperformed AFAL-True indicating that the pseudo labels benefit the training. The improvement is more significant with increased iterations. It is expected since the sampling covered more zones and the predictions became more accurate. However, focusing on Fig. 2, AFAL could not perform better than RS at the beginning due to the fact that AFAL first generated the candidate set locally. This limitation was relaxed as the number of candidates expanded.

For illustrative purposes, Fig. 3 shows the classification maps along with their OAs obtained by the considered methods. It is worth noting that AFAL generated the most accurate and smooth classification maps in comparison to the other tested methods. The map produced by SSAL-SDP, though smooth enough, contains many misclassified subregions corresponding to some low CAs listed in Table 1, while those by RS and AFAL-True are affected by some pepper-and-salt noises. Thus, we can conclude that the proposed AFAL approach is practically valuable to generate high-quality land cover maps.

3.2. OMIS Zaoyuan scene

The second image used in experiments is the **OMIS Zaoyuan scene**, with a size of 137×202 samples and 80 spectral bands. The ground truth contains 23821 labels in 8 classes covering more than 85% of this scene. On this scene, 10 training samples and 5 modes have been added per iteration for SSAL-SDP and AFAL. First, accuracy results and numbers of training samples of all the considered cases are provided in Fig. 4 and Table 2. It can be observed that the proposed AFAL also obtained competitive results on this scene, particularly when the number of pseudo labels increased. It is noticeable that the pseudo labels still contributed to improving the performance of

Table 1. CAs[%] (total number of available labeled samples in brackets), OAs[%], AAs[%] and κs[%] along with standard deviations[%] after 100 iterations for the AVIRIS Indian Pines scene, where we also list the number of training samples including the total ones (Total), those of true labels (True) and those of pseudo labels (Pseudo).

		SSAL-SDP	RS	AFAL-True	**AFAL**
CA	C1 (54)	**90.94±6.38**	62.24±28.13	84.63±14.27	85.12±13.85
	C2 (1434)	75.05±9.46	88.68±4.23	89.96±7.53	**95.99±3.80**
	C3 (834)	70.73±18.35	79.61±8.61	**97.19±1.55**	96.70±2.25
	C4 (234)	**85.34±15.39**	85.01±12.74	71.80±21.89	79.53±22.62
	C5 (497)	82.84±6.46	90.89±1.69	**94.75±2.17**	89.29±7.93
	C6 (747)	95.93±2.09	98.33±1.22	**99.24±0.60**	98.60±1.27
	C7 (26)	**95.60±2.48**	81.66±14.26	90.27±6.84	90.27±6.84
	C8 (489)	92.44±9.38	99.07±0.78	96.19±7.58	**99.70±0.32**
	C9 (20)	**100.0±0.00**	89.79±7.49	93.85±8.43	**100.0±0.00**
	C10 (968)	77.41±5.79	88.31±4.09	86.52±7.03	**89.76±5.60**
	C11 (2468)	83.76±6.88	94.09±1.26	96.98±2.16	**98.73±1.71**
	C12 (614)	81.00±21.12	90.43±4.80	98.41±1.03	**98.85±0.87**
	C13 (212)	98.23±1.28	99.17±0.77	**99.41±0.41**	**99.41±0.41**
	C14 (1294)	85.62±10.18	**98.55±0.73**	90.86±4.66	92.65±3.48
	C15 (380)	**99.93±0.16**	66.49±7.13	89.50±13.18	96.11±7.28
	C16 (95)	**99.19±1.22**	74.58±17.49	97.93±1.98	98.41±1.51
OA		83.40±3.13	90.86±1.77	93.51±1.36	**95.62±1.28**
AA		88.37±2.08	86.68±2.96	92.34±1.84	**94.32±2.44**
κ		81.08±3.53	89.54±2.04	92.57±1.58	**94.99±1.47**
Num	Total	2518	880	2515	2515
	True	0	880	880	880
	Pseudo	2518	0	0	1635

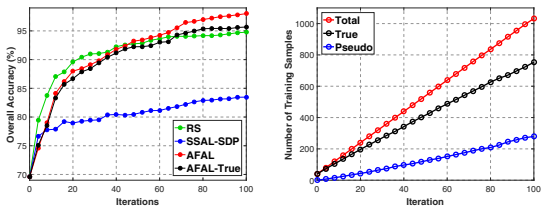
AFAL-True with a 2.3% improvement and a 0.76% lower standard deviation of OAs, while they were only around one fifth of all the training samples. The same conclusions can also be drawn from Fig. 5, where the classification map obtained by AFAL is more accurate and smooth than the rest.

4. CONCLUSIONS

In this work, a new accessibility-free active learning (AFAL) approach has been developed for HSI data classification. It aims at tackling the problem that many labels are out of accessibility for oracles, and thus skipped in traditional active learning methods. This newly proposed AFAL, which is actually a collaborative active and semi-supervised learning approach, labels these samples with the predictions of the classifier in a self-training context, where a su-

(a) False color map (b) Ground truth (c) SDP (89.00%) (d) RS (94.27%) (e) AFAL-True (94.99%) (f) AFAL (98.48%)

Fig. 5. Classification maps (along with the corresponding OAs) obtained after 100 iterations for the OMIS Zaoyuan scene, where a false color map (a) and the ground truth (b) are also provided for reference.



(a) OAs (b) Numbers of training samples

Fig. 4. OAs and the number of training samples as a function of the number of iterations for the OMIS Zaoyuan scene.

Table 2. Statistical results obtained after 100 iterations for the OMIS Zaoyuan scene.

		SSAL-SDP	RS	AFAL-True	**AFAL**
Num	Total	1029	753	753	1032
	True	0	753	753	753
	Pseudo	1029	0	0	279
OA		83.47±6.39	94.77±0.61	95.68±1.17	**98.01±0.41**
AA		90.37±3.41	92.91±0.52	93.99±1.30	**97.10±0.77**
κ		79.47±7.50	93.12±0.80	94.26±1.56	**97.37±0.54**

perpixel assumption has been employed coupled with a DP augmentation strategy to generate the candidates, in order to guarantee the confidence of pseudo labels. The experimental results, conducted using two real HSI scenes, indicate that the proposed approach can lead to state-of-the-art performance (especially from a visual standpoint), making it useful for high-quality land cover map generation. In the future, we will explore how to make a better use of pseudo labels and refine the manually labeled set for AFAL, applying it to transfer learning problems.

5. REFERENCES

[1] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new bayesian approach with active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3947–3960, Oct 2011.

[2] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 844–856, Feb 2013.

[3] C. Liu, L. He, Z. Li, and J. Li, "Feature-driven active learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 341–354, Jan 2018.

[4] J. M. Haut, M. E. Paoletti, J. Plaza, J. Li, and A. Plaza, "Active learning with convolutional neural networks for hyperspectral image classification using a new bayesian approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6440–6461, Nov 2018.

[5] Z. Zhou, "A brief introduction to weakly supervised learning," *NATIONAL SCIENCE REVIEW*, vol. 5, no. 1, pp. 44–53, Jan. 2018.

[6] M. M. Crawford, D. Tuia, and H. L. Yang, "Active learning: Any value for classification of remotely sensed data?," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 593–608, March 2013.

[7] G. Jun A. Liu and J. Ghosh, "Active learning with spatially sensitive labeling costs," in *Proc. NIPS Workshop Cost-Sensitive Learn.*, 2008.

[8] B. Demir, L. Minello, and L. Bruzzone, "Definition of effective training sets for supervised classification of remote sensing images by a novel cost-sensitive active learning method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 1272–1284, Feb 2014.

[9] X. Zhu, "Semi-supervised learning literature survey," Tech. Rep. 1530, Computer Sciences, University of Wisconsin-Madison, 2005.

[10] I. Dópido, J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas Dias, and J. A. Benediktsson, "Semisupervised self-learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 4032–4044, July 2013.

[11] C. Liu, J. Li, and L. He, "Superpixel-based semisupervised active learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–14, *Accepted*, 2018.

[12] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectralspatial hyperspectral image classification: An overview and new guidelines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1579–1597, March 2018.

[13] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, June 2005.

[14] X. Xu, J. Li, C. Wu, and A. Plaza, "Regional clustering-based spatial preprocessing for hyperspectral unmixing," *Remote Sensing of Environment*, vol. 204, pp. 333–346, Jan. 2018.

[15] J. M. Bioucas-Dias and M Figueiredo, "Logistic regression via variable splitting and augmented lagrangian tools," Tech. Rep., Instituto Superior Técnico,, TULisbon, 2009.