Change Detection for Hyperspectral Imagery Based on Multi-layer Cascade Screening Strategy

Lian Liu[1], Danfeng Hong[2] and Lianru Gao[1]

[1]Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, 100094, Beijing, China
[2]Remote Sensing Technology Institute, German Aerospace Center, 82234, Wessling, Germany

Abstract

Change detection (CD) is an important application of remote sensing, which provides information about land cover changes on the earth's surface. Hyperspectral image (HSI) can show more spectral information, which greatly improves the ability of remote sensing to identify change features. The challenge is how to overcome the scarcity of labeled samples and extract the change information of high-dimensional spectra in HSI. To solve the previous problem, a semi-supervised CD with multi-layer cascade screening strategy (MCS^4CD) that uses both the spatial information and active learning is proposed to select highly reliable unlabeled samples to increase the training sets. The MCS^4CD method can effectively use unlabeled samples to improve accuracy. Additionally, a subspace CD method based on iterative slow feature analysis (ISFA) and principal component analysis (PCA) is designed to extract the most temporally invariant component from the high dimensional space. Experimental results on a hyperspectral dataset show that with a small umber of labeled samples, the proposed method achieves a much better performance than existing CD methods.

Keywords

Change detection, hyperspectral image (HSI), iterative slow feature analysis (ISFA), semi-supervised learning, active learning

1. Introduction

The surface ecosystem and human social activities are dynamically developing and evolving[1, 2]. Accurate acquisition of land surface change information is of great significance to better protect the ecological environment, manage natural resources, study social development, and understand the relationship and interaction between human activities and the natural environment[3, 4, 5]. Change detection (CD) is the process of determining the change of the land cover state based on multiple observations at different times. As an advanced and mature technical means, remote sensing earth observation can quickly, macroscopically and dynamically obtain surface images, which provides important data support for solving the CD of land cover. Therefore, using multi-temporal or bi-temporal remote sensing data to obtain the CD of surface features has become one of the most widely used research fields of remote sensing technology[6, 7]. The purpose of CD research is to find interesting change information and filter out irrelevant change information as an interference [8].

Over the past decades, land-use and land-cover CD tasks of optical remote sensing imagery has received increasing attention in the supervised, unsupervised, and semi-supervised algorithms. Supervised learning (SL) relies on prior knowledge to supply the learning algorithm with labeled class data, unsupervised learning (UL) classifies data samples based on the features inherent within the data, and semi-supervised learning (SSL) uses a combination of both.However, obtaining such large labeled training samples requires a considerable amount of resources, time, and effort[9, 10]. In many practical situations, the scarcity of labeled training samples is caused by bad weather areas, which limits the adoption and application of many SL learning methods. In a search for more efficient unlabeled samples to overcome the need for large labeled datasets, there is a rising research interest in SSL and its applications to CD to reduce the amount of labeled data required, by either developing novel classification algorithms or adopting existing SSL frameworks.

At the same time, there are mainly three CD feature extraction algorithms of spectral characterization for classification: 1) Algebraic operation: including image differencing, image ratioing, image regression, and change vector analysis (CVA)[11]; 2) Image transformation: including multivariate alteration detection (MAD)[12], iteratively reweighted MAD (IR-MAD)[13], differential principal component analysis (DPCA)[14], and iterative slow feature analysis (ISFA)[15]; 3) Post classification comparison. Among all these CD algorithms, image transformation methods have been extensively studied and applied. The ISFA algorithm has obtained good experimental results on two groups of real multi-spectral datasets. Although ISFA methods can make use of spectral information, they are not suitable for the continuous high-dimensional spectral features derived from hyper-

CDCEO 2021: 1st Workshop on Complex Data Challenges in Earth Observation, November 1, 2021, Virtual Event, QLD, Australia.
liulian0603@126.com (L. Liu); danfeng.hong@dlr.de (D. Hong); gaolr@aircas.ac.cn (L. Gao)
https://sites.google.com/view/danfeng-hong (D. Hong)

spectral image (HSI).

Based on the above analysis of the current hyperspectral CD problems, it is obvious that we need to explore CD algorithms by focusing on two main points. Firstly, it is often difficult for ISFA algorithm to separate the changed and unchanged pixels in HSI classification especially with limited small training samples. In this paper, a novel semi-supervised classification algorithm based on multi-layer cascade screening strategy (MCS^{4}CD) is put forward. In the semi-supervised process, the spatial neighborhood of labeled training samples is combined with active learning (AL) to select the most helpful unlabeled samples[16], which is used as the pseudo labeled set to retrain the support vector machine (SVM) classifier. Secondly, the performance of ISFA algorithm is degraded due to the band redundancy of HSI. To solve this problem, we designed a new CD algorithm using PCA and ISFA.

The main contributions of this paper are summarized as follows:

1) The reliability of selected unlabeled samples is increased with the proposed MCS^{4}CD strategy that utilizes the spatial information and AL algorithm.

2) The MCS^{4}CD strategy takes into account the positive effect of neighborhood spatial information in the semi-supervised classification process.

3) The PCA+ISFA method is designed for extracting the unchanged features of bi-temporal HSI data.

2. Semi-supervised Classification based on Multi-layer Cascade Screening Strategy

The proposed MCS^{4}CD strategy combines spatial neighborhood information (SNI) extraction strategy with AL algorithm to select the most informative unlabeled samples as the pseudo labeled set to further improve classification performance. The specific procedures of MCS^{4}CD are shown in Figure 1.

2.1. Semi-supervised Classification based on Spatial Neighborhood Information

If the label categories of unlabeled samples are determined only by the primary SVM classification map, it is difficult to ensure satisfactory accuracy. This is because the primary classification accuracy is not high, the labels of the candidate set are misclassified. Therefore, the subsequent SSL process will be affected by the error labels, resulting in error accumulation. In this paper, MCS^{4}CD strategy is constructed to help the SVM classifier label the selected unlabeled samples.

Step 1): Circular neighborhood (CN): We empirically find out that 4 or 8 neighborhood are usually used to obtain the SNI. It only covers a small area within a fixed radius, which obviously cannot meet the needs of different sizes. 4 or 8 neighborhood windows are too small for searching useful unlabeled samples. In this paper, we adopt a CN window, which can adjust the search radius d to find the optimal size (including 4- and 8-neighborhood).

Step 2): SNI extraction strategy: "Tobler's First Law of Geography" gives us an important assumption that the label category of unlabeled samples should be consistent with the existing training sample categories in the spatial neighborhood area. However, in the process of determining unlabeled samples, the positive effect of SNI on SSL method is often ignored. Based on the initial classification results and SNI, the labels of unlabeled samples are screened for the second time. The second screening strategy eliminates the negative effects of low quality and wrong labels.

2.2. Selecting the Most Informative Unlabeled Samples based on Active Learning

In the SSL process, a large number of unlabeled data points are selected as the candidate sets. To further simplify the samples, AL strategy is used to select the most informative unlabeled samples from the candidate sets. AL aims to carefully choose the samples to be labeled to achieve a higher accuracy while using as few requests as possible, thereby minimizing the cost of obtaining labeled data. After the third screening step, the unlabeled samples with high confidence or most informative are selected as the final pseudo samples through AL.

In this paper, breaking ties (BT) query strategy of AL algorithm is used to collect the most informative unlabeled samples. The decision criterion of BT is:

$$\hat{\boldsymbol{x}}_i^{BT} = \arg \min_{x_{i,i \in X_U}} \left\{ \max_{k \in \mathscr{K}} p(y_i = k|\boldsymbol{x}_i) - \max_{k \in K \setminus \{k^+\}} p(y_i = k|\boldsymbol{x}_i) \right\} \tag{1}$$

$$p(y = k|\boldsymbol{x}_i) = \frac{1}{1 + \exp(Af(\boldsymbol{x}) + B)} \tag{2}$$

Where $k^+ = \arg\max_{k \in K} p(y_i = k|\boldsymbol{x}_i)$ represents the class label corresponding to the largest posterior probability for sample $\boldsymbol{x}_i$, and $k \in K \setminus \{k^+\}$ represents the interested class labels excluding k^+. p is provided the probabilistic outputs by the probability model-based SVM.

2.3. Procedure of the Proposed MCS^{4}CD

To increase the reliability of selected unlabeled samples, a SSL method that is based on the MCS^{4}CD strategy is adopted. The detailed strategy is described as follows:

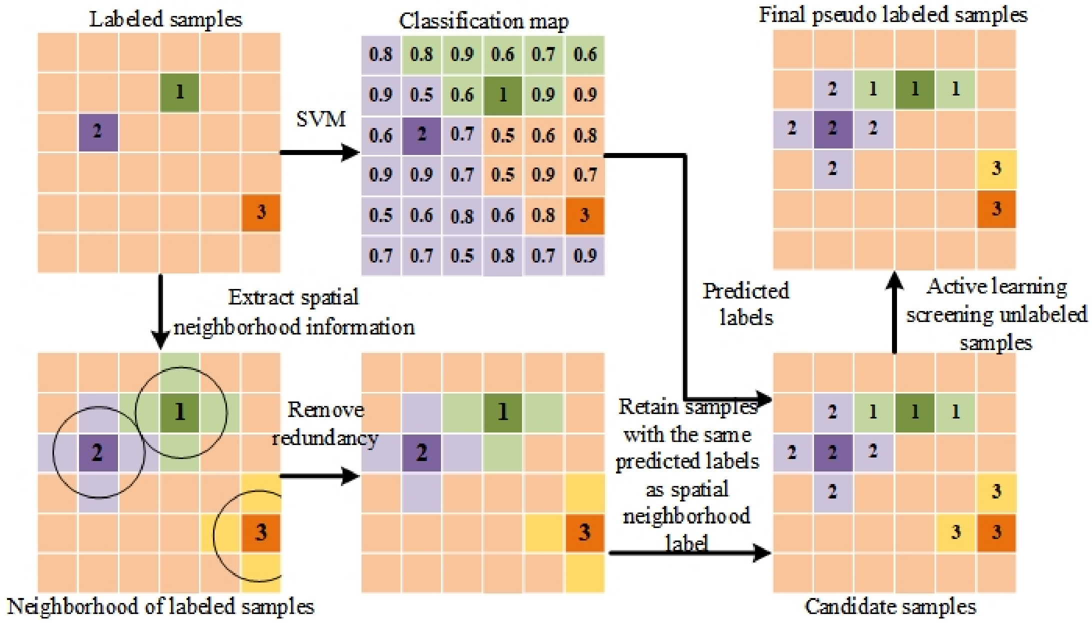


Figure 1: Flowchart of the proposed MCS^{4}CD method

Step 1) Initialize training samples $D_L = \{(x_1, y_1), ...(x_1, y_l)\}$, and set parameters: spectral dimension of subspace feature ω, radius d of CN, the number of training samples for each class p;

Step 2) Extract subspace spectral feature $\boldsymbol{x}^{\omega} \in \mathbb{R}^s$ using the PCA method;

Step 3) Extract changed and unchanged features $\boldsymbol{x}^{\varpi} \in \mathbb{R}^{\varpi}$ using the ISFA method;

Step 4) Train SVM probability model to predict the label $\hat{\boldsymbol{y}} = \{\hat{y}_i, i = 1, ..., n_{pre}\}$ of unlabeled samples;

Step 5) Select a circle neighborhood which takes the selected $\boldsymbol{D}_L$ as the center and remove redundancy samples (including background information and repeated selection of training samples). Retain samples with the same labels as spatial neighborhood label. The extracted candidate samples are denoted as $\boldsymbol{D}_U = \{(x_1, y_1), ...(x_1, y_{n_d})\}, i = 1, ..., n_d$;

Step 6) Simplify $\boldsymbol{D}_U$ to $\boldsymbol{D}_C$ by using the AL algorithm. The $\boldsymbol{D}_L$ is expanded with each iteration of BT. Then, update the labeled sample sets $\boldsymbol{D}_{C+L} = \{\boldsymbol{D}_C, \boldsymbol{D}_L\}$;

Step 7) Test the performance of the final pseudo labeled samples using the SVM classifier.

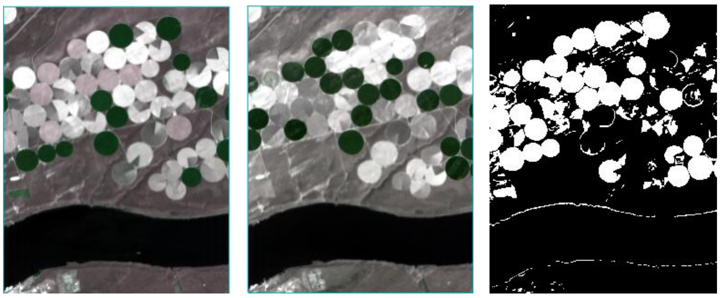
3. Experimental results

3.1. DataSet

The USA Dataset illustrates an irrigated agricultural field of Hermiston city in Umatilla County, Oregon, OR, the USA, which was collected on May 1, 2004, and May 8, 2007, respectively. The size of this dataset is 307 lines by 241 samples, with 154 spectral bands. The land cover types include soil, irrigated fields, river, cultivated land and grassland. For this dataset, all changes related to the type of land cover and river. The true color composite image of the USA dataset and its corresponding land-cover CD map are shown in Figure. 2.

3.2. Comparison with other change detection methods

To demonstrate the effectiveness of the proposed method, we will display the numerical results on the USA dataset, as shown in Figure 3. As can be observed in table 1, MCS^{4}CD is capable of building better classification performance that compensates for the lack of labeled training data. MCS^{4}CD projects the original data into a new transformed space to better separate the changed and unchanged pixels.



(a) May 1, 2004 (b) May 8, 200 (c) ground truth

Figure 2: The USA Dataset and land-cover ground truth map

Table 1
Classification accuracy in relation to different change detection methods for the USA dataset

Method	CVA	MAD	IRMAD	ISFA	DPCA	PCA+ISFA	MCS^{4}CD
OA	92.11	83.13	74.83	81.05	89.42	91.58	**94.46**
Kappa	0.7911	0.5226	0.3495	0.5489	0.7176	0.7743	**0.8458**
MA	0.0845	0.1139	0.2068	0.2057	0.0983	0.0822	**0.0476**
FA	**0.0186**	0.1050	0.1296	0.0467	0.0407	0.0280	0.0245

3.3. Effect of the suitable search radius d

For the MCS^{4}CD, one of the key questions is how to confirm the suitable radius d, which influences the accuracy and the numbers of selected unlabeled samples. Figure. 4 shows the result of USA data. When the number of pseudo labeled samples is 0, 20, 40, 60, 80,100, 120 and the search radius d ranges from 1 to 3, the results corresponding d =2 is the best.

4. Conclusion

This study proposed a semi-supervised CD method with slow feature analysis, including multi-layer cascade screening strategy and data transformation strategy in spectral subspace domain, for HSI classification with a small number of labeled samples. In the semi-supervised process, the slow feature extraction in high-dimensional space, the use of BT algorithm, and the decision strategy for the label of unlabeled samples are all key points. On the one hand, we use PCA method to reduce too many spectral bands, which has a negative impact on the expected CD performance. On the other hand, BT, circular neighbor and SVM are combined together to improve the judgment accuracy of unlabeled samples. Experimental results with HSI indicate that the proposed MCS4CD approach can obtain well performance.

References

[1] Z. Zhu, Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications, ISPRS Journal of Photogrammetry and Remote Sensing 130 (2017) 370–384.

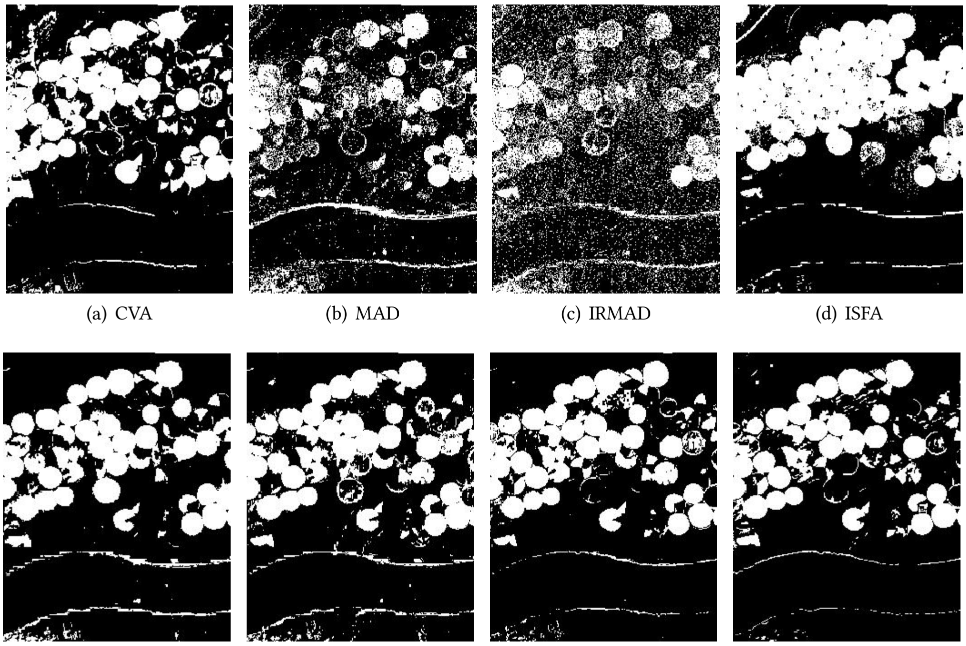
[2] S. Liu, Q. Du, X. Tong, A. Samat, L. Bruzzone, Unsupervised change detection in multispectral remote sensing images via spectral-spatial band expansion, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12 (2019) 3578–3587.

[3] P. Xiao, X. Zhang, D. Wang, M. Yuan, X. Feng, M. Kelly, Change detection of built-up land: A framework of combining pixel-based detection and object-based recognition, ISPRS Journal of Photogrammetry and Remote Sensing 119 (2016) 402–414.

[4] D. Marinelli, F. Bovolo, L. Bruzzone, A novel change detection method for multitemporal hyperspectral images based on binary hyperspectral change vectors, IEEE transactions on geoscience and remote sensing 57 (2019) 4913–4928.

[5] S. Liu, D. Marinelli, L. Bruzzone, F. Bovolo, A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges, IEEE Geoscience and Remote Sensing Magazine 7 (2019) 140–158.

[6] B. Du, L. Ru, C. Wu, L. Zhang, Unsupervised deep slow feature analysis for change detection in multi-



(a) CVA (b) MAD (c) IRMAD (d) ISFA

(e) DPCA (f) PCA+ISFA (g) MCS^{4}CD (h) Ground truth

Figure 3: Classification results of different change detection methods for the USA dataset.

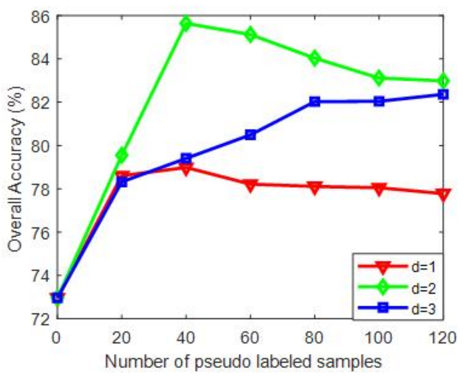


Figure 4: Classification accuracy in relation to different numbers of pseudo labeled samples.

temporal remote sensing images, IEEE Transactions on Geoscience and Remote Sensing 57 (2019) 9976–9992.

[7] C. Wu, L. Zhang, B. Du, Hyperspectral anomaly change detection with slow feature analysis, Neurocomputing 151 (2015) 175–187.

[8] M. Hu, C. Wu, L. Zhang, B. Du, Hyperspectral anomaly change detection based on autoencoder, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021) 3750–3762.

[9] D. Hong, W. He, N. Yokoya, J. Yao, L. Gao, L. Zhang, J. Chanussot, X. Zhu, Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing, IEEE Geoscience and Remote Sensing Magazine 9 (2021) 52–87.

[10] D. Hong, J. Hu, J. Yao, J. Chanussot, X. X. Zhu, Multimodal remote sensing benchmark datasets for land cover classification with a shared and specific feature learning model, ISPRS Journal of Photogrammetry and Remote Sensing 178 (2021) 68–80.

[11] R. D. Johnson, E. Kasischke, Change vector analysis: A technique for the multispectral monitoring of land cover and condition, International journal of remote sensing 19 (1998) 411–426.

[12] A. A. Nielsen, K. Conradsen, J. J. Simpson, Multivariate alteration detection (mad) and maf postprocessing in multispectral, bitemporal image data: New approaches to change detection studies, Remote Sensing of Environment 64 (1998) 1–19.

[13] A. A. Nielsen, The regularized iteratively reweighted mad method for change detection in multi-and hyperspectral data, IEEE Transactions on Image processing 16 (2007) 463–478.

[14] J. Deng, K. Wang, Y. Deng, G. Qi, Pca-based land-

use change detection and analysis using multitemporal and multisensor satellite data, International Journal of Remote Sensing 29 (2008) 4823–4838.
[15] C. Wu, B. Du, L. Zhang, Slow feature analysis for change detection in multispectral imagery, IEEE Transactions on Geoscience and Remote Sensing 52 (2013) 2858–2874.
[16] C. Liu, J. Li, L. He, Superpixel-based semisupervised active learning for hyperspectral image classification, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12 (2018) 357–370.