

NEIGHBOUR-BASED DOMAIN ADAPTATION FOR INVESTIGATION OF TRANSFERABLE ABILITY OF PREVIOUSLY LABELED DATA FOR LAND-COVER CLASSIFICATION OF AERIAL IMAGES.

Ruixi Zhu, Li Yan *

School of Geodesy and Geomatics, Wuhan Univeristy, China--ruixzhu@whu.edu.cn, liyan@sgg.whu.edu.cn

Commission II, ICWG II/III

KEY WORDS: domain adaptation, reliable target pseudo-labelled data, select source labelled data, augment the training dataset.

ABSTRACT:

Existing land-cover classification methods are usually based on adequate labelled data. But annotating enough training samples is hard and time-consuming. Therefore, we need to investigate how existing labelled data can help to increase land-cover classification. Source labelled data are proposed to be selected by calculating the target center of reliable target pseudo-labelled data for each class in this paper. Then we augment the training dataset with reliable target pseudo-labeled data and selected source labelled data to improve the quality and quantity of training dataset. We also investigate the amount of source labelled data that should be selected and the number of limited target labelled data that can produce good transfer learning performance. The UC Merced dataset is employed as the target dataset to evaluate the proposed approach while the NWPU-RESISC45 dataset is considered as the source labelled data. The experimental results show that selected source labelled data and reliable target pseudo-labeled data may improve the land-cover classification performance if selected source labelled data and reliable target pseudo-labeled data are augmented with the limited target labelled data respectively.

1. INTRODUCTION

Aerial images have become one of the most important data sources in land-cover classification since diverse aerial sensors were developed (Mora et al., 2015). The supervised methods that are successful in land-cover classification of aerial images usually depend on the availability of abundant annotated data (Liu et al., 2014). However, annotating adequate samples to obtain satisfactory land-cover classification results is difficult, time consuming, and sometimes even impossible. To address this problem, the abundant previously labelled data can help to classify new aerial images by transferring the knowledge learnt from the previously labelled data to the aerial images to be classified. However, the spectral distributions of the new images may be highly different from those of previously labelled data due to diverse imaging conditions including different natural environments, locations, seasons, and imaging modalities (Zhu et al., 2019).

Domain adaptation can address the spectral distribution bias since it can transfer knowledge learnt from previously labelled data to new aerial images. In domain adaptation, the source domain usually represents the previously labelled data while target domain denotes the new images to be classified that are with limited or no labelled data. In accordance with (Tuia et al., 2016), three types of domain adaptation methods can be used to transfer previously labelled data for the land-cover classification of aerial images.

1. Invariant feature selection methods. In this family of methods, features extracted from previously labelled data and aerial images to be classified are mapped to a common space that are not easily influenced by the bias across domains. However, this family of methods may deliver poor performance when classifying similar land-cover scenes.

2. Classifier adaptation methods. In this type of approach, target labelled or unlabelled data are adapted to the classifier trained

by the previously labelled data so as to make the source classifier better classify target data. These methods might fail when a strong probability distribution bias between previously labelled data and aerial images exists.

3. Data distribution adaptation methods. These methods aim to maintain similar data distributions among the data from different domains so that the classifier trained from previously labelled data can classify the target images with a different feature distribution.

As shown in Fig.1a-c, there may exist strong spectral shift between previously labelled data and aerial images to be classified in some specific classes such as river and airplane since highly different spectral characteristics may exist in these land-cover categories. Moreover, high inter-class similarity between diverse land-cover types may exist in the images to be classified, which may lead to misclassification in similar land-cover categories, as shown in Fig 2. Therefore, the data distribution adaptation methods rather than other two types of methods is investigated in this study. The objective of data distribution adaptation methods is to manage the main cause of failure in the domain adaptation.

The data distribution adaptation methods achieve great success in classifying land-cover types existing in aerial images. This type of method consists of feature-based methods, classifier-based methods and instance-based methods (Ma et al., 2018).

Feature-based methods are aimed to find a common feature space in which the bias between domains can be decreased. Then the source features may distinguish target features well if they are mapped to the shared feature space. This family of approaches include computing the manifold (Ham et al., 2005), correlation index (Yeh et al., 2014), covariance matrix (Sun et al., 2017), subspace eigenvectors (Fernando et al., 2013) between domains or forcing the mean from domains to be close (Pan et al., 2010).

* Corresponding author

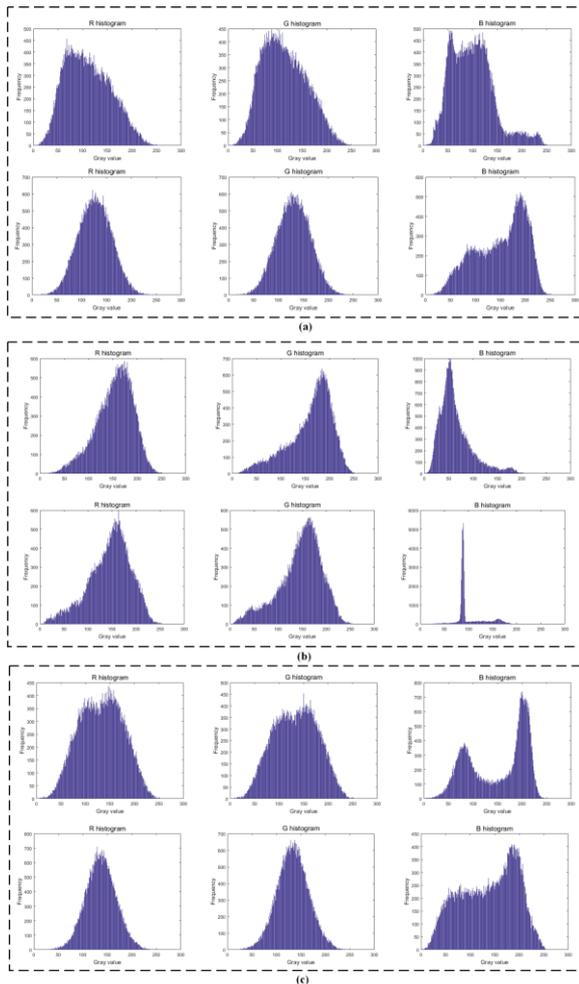


Figure 1. The spectral distribution difference between NWPU-RESISC45 and UC Merced dataset in some specific classes. (a) Airplane. (b) River and (c) Chaparral.

Classifier-based methods are aimed at learning knowledge from the classifier trained by source labelled data so as to obtain a more discriminative target classifier. The support vector machine (SVM) classifier is commonly used in this type of approaches. We can learn a more discriminative target classifier if we augment the training dataset with some high-quality target samples (Bahirat et al., 2011), design an object function considering the discriminative ability of target features (Yang et al., 2007) or combine source vectors with the target samples (Jiang et al., 2008).

Instance-based approaches reweight or choose the previously labelled data so as to decrease the spectral distribution bias between domains. The instance-based approaches can be realized by decreasing the distance between features from diverse domains (Jiang et al., 2008), clustering cross-domain data (Zhong et al., 2009) and calculating the transformation matrix restricted by a row-sparsity constraint (Yang et al., 2007).

Although numerous researches have been done on instance-based methods, seldom attention has been paid to how target pseudo labelled data and source samples contribute to classifying target images to be classified. As a result, we propose to select source labelled data by calculating the center of reliable target pseudo-labelled data for each class. We also investigate the amount of source labelled data that should be selected and the number of limited target labelled data that can produce good transfer learning performance.

In this paper, the major contributions can be described as follows:

1. The source labelled data is proposed to be selected based on the distance to centers of those reliable target pseudo-labelled data iteratively in order to decrease the spectral shift between domains.
2. We investigate the amount of limited target samples and source samples that can produce satisfactory domain adaptation performance.
3. We investigate the proposed selection method on different CNN architectures such as VGG-16, Resnet-18 and Resnet-101.

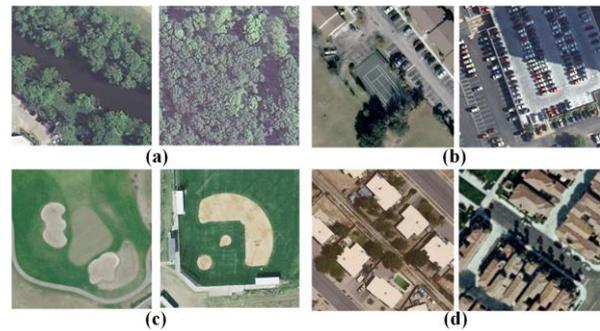


Figure 2. The high inter-class similarity in UC Merced dataset. (a) River and forest. (b) Tennis court and parking lot. (c) Baseball diamond and golf course. (d) Medium residential and mobile home park.

2. METHODOLOGY

2.1 The overall architecture

In order to select source labelled data share similar feature distributions with target data when limited target labelled data are available, we propose a predicted center based source sample selection method as shown in Figure 3. Algorithm 1 summarizes the proposed method.

Algorithm 1 The process of the proposed source sample selection method.

Input: Source labelled data X_s and their labels Y_s . Target limited labelled data X_{tl} and their true labels Y_{tl} , target unlabelled data X_{tu} .

Output: The selected source labelled data X_s^l .

1: Extract features for data from both domains based on the pre-trained CNN models such as VGG-16 or Resnet-101

2: Train a SVM classifier with limited target labelled data X_{tl} and get initial pseudo labels Y_{tu}^0 for unlabelled target data X_{tu} .

3: Select reliable target pseudo-labelled data X_{tur}^0 according to the pseudo labels of target unlabelled data and their nearest neighbours in the feature space.

4: Calculate the target center C_i^0 of i -th class by averaging all reliable target pseudo-labelled data belonging to i -th class.

5: Select source labelled data X_s^0 for i -th class that are near to the center of i -th class.

6: Augment the limited target labelled data with target reliable target pseudo-labelled data.

7: Train a new SVM classifier with augmented training data and get new pseudo labels Y_{tu}^l .

8: Repeat step 3 to step 7 until $\|Y_{tu}^{l+1} - Y_{tu}^l\| \leq 10^{-3}$

9: Return X_s^l

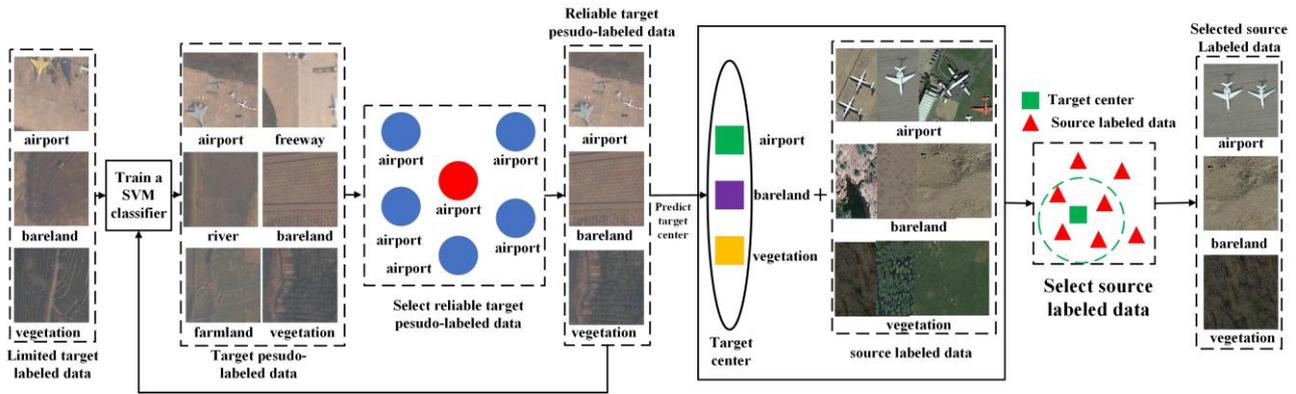


Figure 3. The overall architecture of the proposed method

2.2 Predict target center based on selected reliable target pseudo-labelled data

The limited target labelled data may not reflect the feature distributions of the whole target data due to the small number of training data. The pseudo labels of target unlabelled data can provide valuable hidden information for predicting the feature distributions of the whole target data. As a result, the reliable target pseudo labels that are more likely to be classified need to be found out in order to predict the feature distributions of target data. It is evidential that the nearest neighbours of a correctly labelled target data are likely to be classified. In this paper, we assume that j -th reliable target data x_{tr}^j are those share the same pseudo labels with most of their corresponding nearest neighbours M in the feature space, as shown in Eq. (1) and Figure 4.

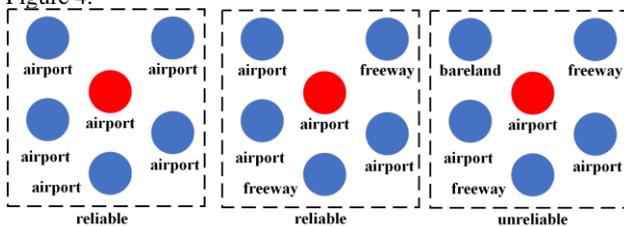


Figure 4. The definition of reliable target pseudo-labelled data.

$$\exists \{x_{tr}^j, x_{tr}^1, x_{tr}^2, x_{tr}^3, x_{tr}^4, x_{tr}^5\}, \{x_{tr}^j, x_{tr}^1, x_{tr}^2, x_{tr}^3\} \in c_i \quad (1)$$

Where c_i represents the i -th class. Nearest neighbours M of target pseudo-labelled data x_{tr}^j are calculated according to the Euclidean distances between target features x_{tr}^j and x_{tr}^i as we can see in Eq. (2). Five target data $M = \{x_{tr}^1, \dots, x_{tr}^5\}$ that are close to the x_{tr}^j are considered as the nearest neighbours of target data.

$$\arg \min_{top 5} \|x_{tr}^j - x_{tr}^i\|_2 \quad (2)$$

Then the predicted target center C_i of i -th class c_i can be calculated based on all reliable target pseudo-labelled data.

$$C_i = \frac{\sum_{y_{tr}^j \in c_i} x_{tr}^j}{N_{tr}^{c_i}} \quad (3)$$

Where $N_{tr}^{c_i}$ represents the number of reliable target pseudo-labelled data that belong to land-cover class c_i . x_{tr}^j represents j -th reliable target data.

2.3 Source sample selection based on predicted target center

When performing domain adaptation, source labelled data whose feature distributions are different from target data may have a negative influence on classifying target data. Therefore, effective source labelled data that share similar feature distributions with target data should be selected in order to enhance the domain adaptation performance. Therefore, we select source samples that are close to the predicted target center in each class according to Eq. (4).

$$\arg \min_{top n} \|X_s^i - C_i\|_2 \quad (4)$$

Where X_s^i represents the source samples that belong to class C_i and n represents the amount of selected source labelled data for each land-cover type. When the limited target labelled data are augmented with the reliable target pseudo-labelled data, the classification accuracy in the target domain will be increased. Then the predicted target center will be more accurate if they are calculated from the new prediction results. The selection continues until the prediction results are stable.

3. EXPERIMENTAL RESULTS AND ANALYZES

Dataset description, implementation details and experimental results are provided in this section.

3.1 Description of datasets

Source domain: Scene images corresponding to 21 land-cover classes in the UC Merced dataset (Yang et al., 2010) are chosen from 45 land-cover types in NWPU-RESISC45 dataset (Cheng et al., 2017). In order to maintain the class balance, there exist 700 samples in each class. The examples of each class are shown in Figure 5. The spatial resolution of this dataset varies from 0.2m to 30m.



Figure 5. Examples of each class in the NWPU-RESISC45 dataset for single-source domain adaptation.



Figure 6. Examples of each class in the target dataset

Target domain: UC Merced dataset is regarded as the target dataset. It is a publicly available aerial land-cover classification dataset. The dataset consists of 100 images for each class in 21 classes. The spatial resolution of this dataset is 0.3m. Figure 6 shows examples of each class in the UC Merced dataset.

3.2 Implementation details

Three different CNN architectures pre-trained on ImageNet including VGG-16 (Simonyan et al., 2014), Resnet-18(He et al., 2016) and Resnet-101(He et al., 2016) are used for extracting features for each source and target image. All methods have been implemented in Caffe with MATLAB 2016a. The training ratios of target images are set to 5%, 10% and 20% since it is assumed that target images are with limited training data. Each training ratio is repeated five times to prove the stability of the proposed algorithm. The number of selected source samples is set according to cross-validation while the number of nearest neighbours is empirically set to 3. The SVM classifier is used for classification in this paper and implemented according to the vlfeat toolbox. The SVM classifier trained from limited target labelled data and reliable pseudo-labelled data are used to provide pseudo labels. The maximum iteration is set to 15 according to experience.

3.3 Experimental results and analyses

3.3.1 Analysis of chosen source sample amount

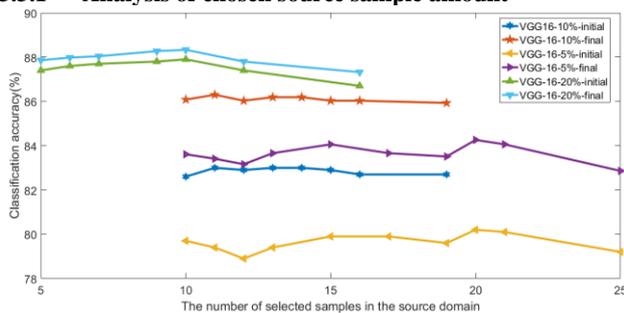


Figure 7. The experimental results of VGG-16 architecture with diverse chosen source sample number.

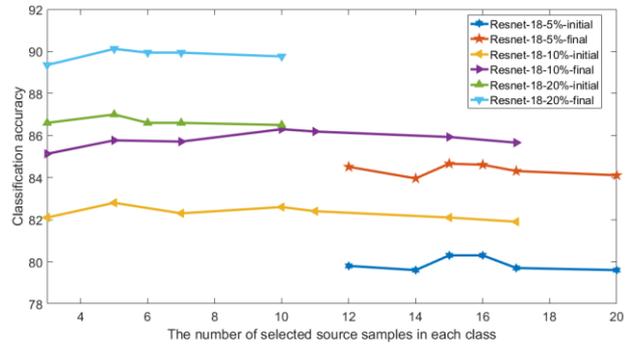


Figure 8. The experimental results of ResNet-18 architecture with diverse chosen source sample number.

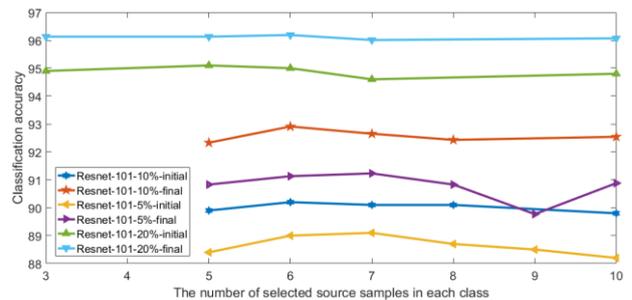


Figure 9. The experimental results of Resnet-101 architecture with diverse chosen source sample number.

Figure 7 to 9 show the influence of the selected source sample number on the results of the proposed approaches with three existing CNN architectures. As can be seen in Figure 7 to 9, VGG-16 require the largest number of selected source samples while ResNet-101 needs the least number since ResNet-101 can extract the most discriminative features due to deep CNN architectures. The influence of selected source sample number is highest on 5% training ratio and lowest on 20% training ratio since 5% training samples cannot produce accurate target pseudo labels.

3.3.2 Analysis of selected source samples and reliable target pseudo-labeled data

In order to investigate how target pseudo-labeled data and source samples can contribute to the proposed algorithm, diverse experiments have been performed as shown in Table 1 to 3.

Methods	5%	10%	20%
Only limited target labelled data	71.2%	75.3%	82.7%
selected source samples+ limited target samples + reliable target pseudo labels	80.8%	83.4%	87.4%
Selected source samples	65.2%	66.0%	66.5%
Selected source samples +limited target labelled data	74.3%	77.6%	84.2%
Limited target labelled data+ reliable target pseudo-labelled data	80.2%	83.5%	88.0%
All source labelled data	62.4%	62.1%	61.3%

Table 1. The classification accuracy under different training ratios when using VGG-16 architecture.

Methods	5%	10%	20%
Only limited target labelled data	70.4%	79.5%	82.1%
selected source samples+ limited target samples + reliable target pseudo labels	81.0%	86.1%	87.4%
Selected source samples	68.8%	69%	70.7%
Selected source samples +limited target labelled data	76.9%	80.1%	82.9%
Limited target labelled data+ reliable target pseudo-labelled data	80.5%	86.1%	86.5%
All source labelled data	62.8%	61.2%	62.9%

Table 2. The classification accuracy under different training ratios when using Resnet-18 architecture.

	5%	10%	20%
Only limited target labelled data	75.9%	84.5%	88.9%
selected source samples+ limited target samples + reliable target pseudo labels	86.5%	89.6%	92.3%
Selected source samples	67.6%	69.3%	70.8%
Selected source samples +limited target labelled data	80.4%	85.1%	89.3%
Limited target labelled data+ reliable target pseudo-labelled data	85.7%	89.4%	92.7%
All source labelled data	66.9%	68.3%	69.6%

Table 3. The classification accuracy under different training ratios when using Resnet-101 architecture.

As we can see in Table 1 to 3, both the selected source labelled data and reliable target pseudo-labelled data may enhance the land-cover classification performance of unlabeled target images. The reliable target pseudo-labelled data can provide a higher increase in classification accuracy compared with selected source samples since reliable target pseudo-labelled data share similar feature distributions with limited target labelled data while selected source samples do not. The augmented training dataset including limited target labelled data, selected source samples and reliable target pseudo-labelled data can provide the highest classification accuracy because they increase both the quality and quantity of training dataset.

It can also be seen in Table 1 to 3 that the source labelled data can increase the highest classification accuracy with when 5% target labelled data are available while selected source labelled data can contribute to 20% target limited labelled data little since 20% limited target labelled data can produce satisfactory land-cover classification performance. It can also be seen in Table 1 to 3 that most of the source labelled data are not beneficial to the classification of target images since they are different feature distributions. The source labelled data that are close to the predicted target centers are effective in improving land-cover classification performance when employing three CNN architectures.

3.3.3 Analysis of iteration

Figure 10 to 18 show the accuracy in each iteration under diverse CNN architectures and training ratios to demonstrate the role of iteration in the proposed approach. Selected source

samples and reliable tend to increase the classification accuracy of target images in the first several iterations before the classification accuracy decreases as shown from Figure 10 to 18. That is because the initial target pseudo contain some mistakes, which may lead to the decreased quality of training dataset. The iteration can increase the accuracy of target pseudo labels created by the SVM classifier trained from source samples. But after some iterations, the classification accuracy reaches its limits. The highest classification accuracy on each curve corresponds to the accuracy shown in Table 1 to 3.

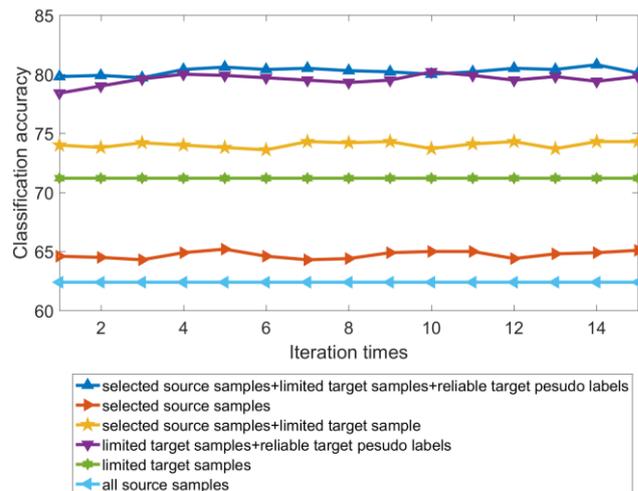


Figure 10. The experimental results of VGG-16 architecture under 5% training ratio.

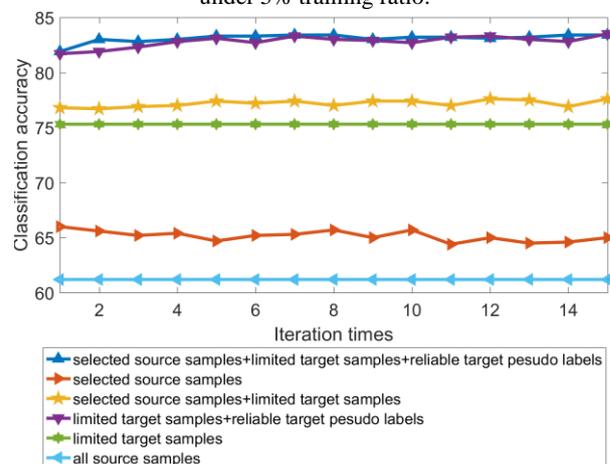


Figure 11. The experimental results of VGG-16 architecture under 10% training ratio.

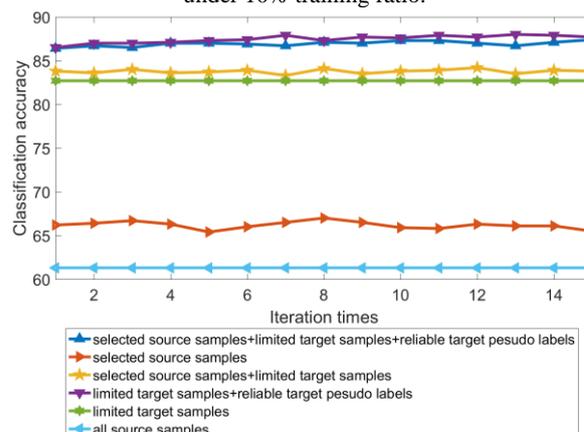


Figure 12. The experimental results of VGG-16 architecture with 20% training ratio.

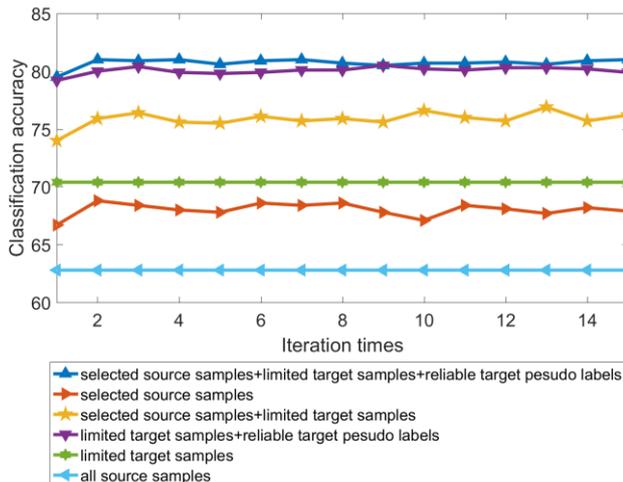


Figure 13. The experimental results of ResNet-18 architecture with 5% training ratio.

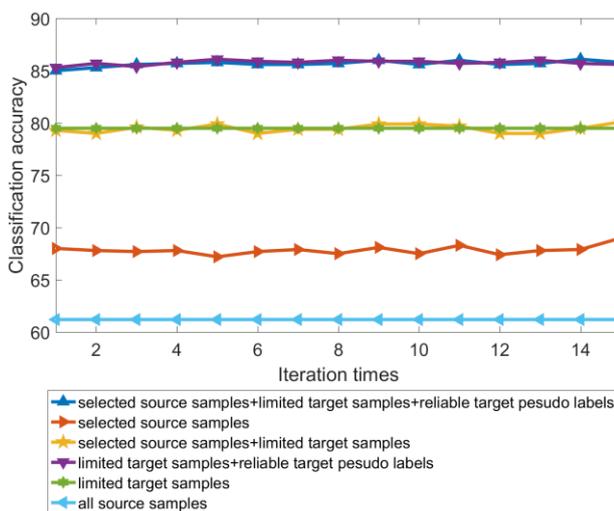


Figure 14. The experimental results of ResNet-18 architecture with 10% training ratio.

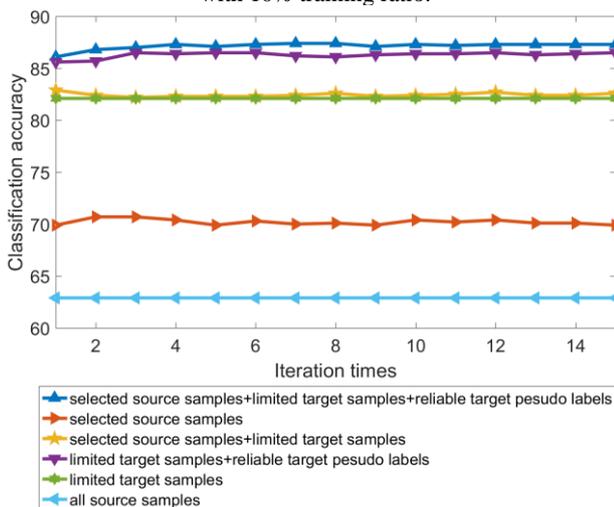


Figure 15. The experimental results of ResNet-18 architecture with 20% training ratio.

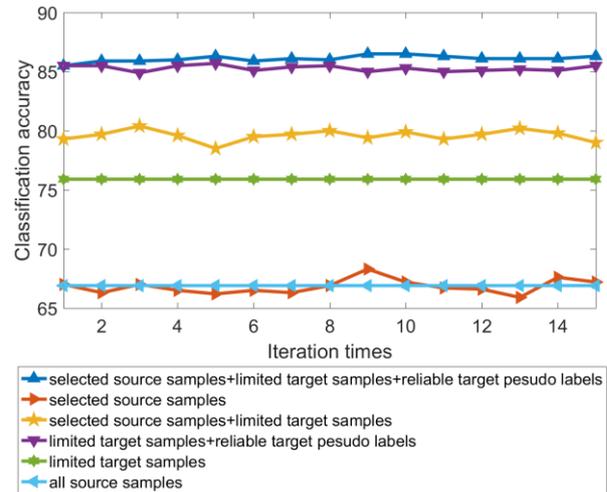


Figure 16. The experimental results of ResNet-101 architecture with 5% training ratio.

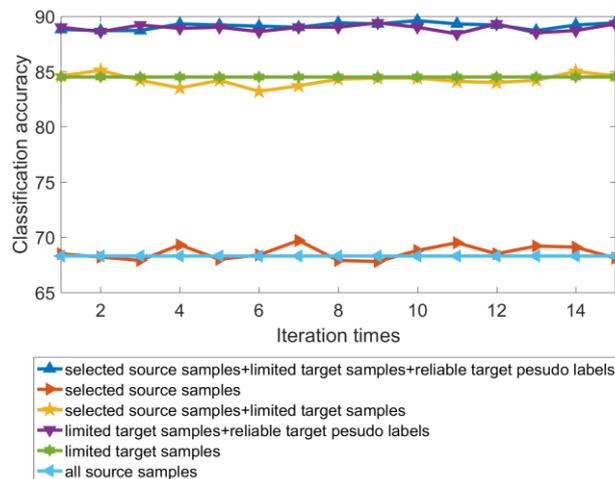


Figure 17. The experimental results of ResNet-101 architecture with 10% training ratio.

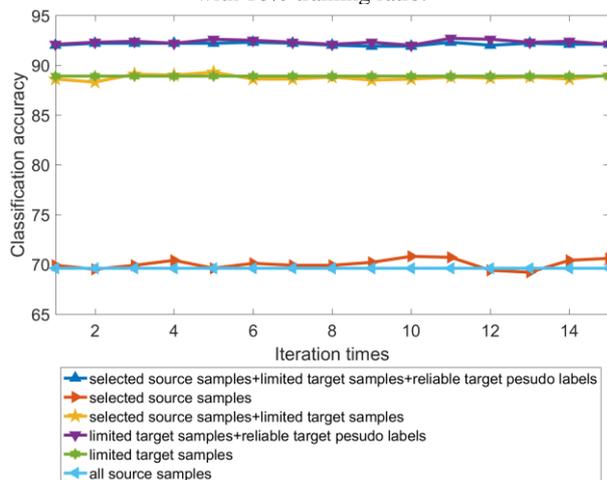


Figure 18. The experimental results of ResNet-101 architecture with 20% training ratio.

4. CONCLUSIONS AND FUTURE DIRECTIONS

The domain adaptation methods can transfer the knowledge learnt from the source labelled data to target data so that the effort on annotating labelled data can be decreased. In order to investigate the transferable ability of source labelled data and

reliable target pseudo-labeled data, we propose to select source labelled data that are close to the centers of reliable target data for each class based on the assumption that the nearest neighbours of a correctly predicted target data are likely to be classified. Experimental verification has been performed on the target UC Merced dataset with different training ratios under three CNN architectures to demonstrate the superiority of the approach proposed in this paper when NWPU-RESISC45 dataset is chosen as the source samples. The proposed source sample selection can produce 91.2% domain adaptation accuracy when only 5% target labelled data are available and Resnet-101 is employed as feature extractor. The selected source samples and reliable target pseudo-labeled data can increase the domain adaptation performance and reliable target pseudo-labeled data may play a more important role compared with selected source samples.

However, the proposed method is only experimented when the spectral shift between images from diverse domains is small. We will extend the proposed method to the large spectral shift in the future direction.

ACKNOWLEDGEMENTS

We would like to thank anonymous editors and reviewers for their kind suggestions. This study is supported by The National Key Research and Development Program of China under grant no. 2017YFC0803802.

REFERENCES

- Bahirat K, Bovolo F, Bruzzone L, 2011. A novel domain adaptation Bayesian classifier for updating land-cover maps with class differences in source and target domains. *IEEE Trans Geosci Remote Sens*, 50(7), 2810-2826. [Doi.org/10.1109/TGRS.2011.2174154](https://doi.org/10.1109/TGRS.2011.2174154).
- Cheng G, Han J, Lu X, 2017. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10), 1865-1883. doi.org/10.1109/JPROC.2017.2675998
- Fernando B, Habrard A, Sebban M, 2013. Unsupervised visual domain adaptation using subspace alignment. *Proceedings of the IEEE international conference on computer vision*, 2960-2967. doi.org/10.1109/ICCV.2013.368.
- Ham J, Lee D D, Saul L K, 2005. Semi-supervised alignment of manifolds. *Proceedings of the Annual Conference on Uncertainty in Artificial Intelligence*. 120, 27.
- He K, Zhang X, Ren S, 2016. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778. doi.org/10.1109/CVPR.2016.90
- Jiang W, Zavesky E, Chang S F, 2008. Cross-domain learning methods for high-level visual concept classification. *IEEE International Conference on Image Processing*, 161-164, doi.org/10.1109/ICIP.2008.4711716.
- Liu Y, Li X, 2014. Domain adaptation for land use classification: A spatio-temporal knowledge reusing method. *ISPRS J Photogramm Remote Sens*, 98, 133-144, doi.org/10.1016/j.isprsjprs.2014.09.013
- Ma L, Crawford M M, Zhu L, 2018. Centroid and covariance alignment-based domain adaptation for unsupervised classification of remote sensing images. *IEEE Trans Geosci Remote Sens*, 57(4), 2305-2323, doi.org/10.1109/TGRS.2018.2872850.
- Mora C, Vieira G, Pina P, 2015. Land cover classification using high-resolution aerial photography in adventdalen, svalbard. *Geografiska Annaler: Series A, Physical Geography*, 97(3), 473-488, doi.org/10.1111/geoa.12088.
- Pan S J, Tsang I W, Kwok J T, 2010. Domain adaptation via transfer component analysis. *IEEE Trans Neural Netw*, 22(2), 199-210, doi.org/10.1109/TNN.2010.2091281.
- Simonyan K, Zisserman A, 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sun B, Feng J, Saenko K, 2017. Correlation alignment for unsupervised domain adaptation. *Domain Adaptation in Computer Vision Applications*, 153-171, doi.org/10.1007/978-3-319-58347-1_8
- Tuia D, Persello C, Bruzzone L, 2016. Domain adaptation for the classification of remote sensing data: An overview of recent advances. *IEEE Geosci Remote Sens Mag*, 4(2), 41-57, doi.org/10.1109/MGRS.2016.2548504.
- Yang J, Yan R, Hauptmann A G, 2007. Cross-domain video concept detection using adaptive svms. *Proceedings of the 15th ACM international conference on Multimedia*, 188-197, doi.org/10.1145/1291233.1291276.
- Yang Y, Newsam S, 2010. Bag-of-visual-words and spatial extensions for land-use classification. *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, 270-279, doi.org/10.1145/1869790.1869829
- Yeh Y R, Huang C H, Wang Y C F, 2014. Heterogeneous domain adaptation and classification by exploiting the correlation subspace. *IEEE Trans Image Process*, 23(5), 2009-2018, doi.org/10.1109/TIP.2014.2310992.
- Zhong E, Fan W, Peng J, 2009. Cross domain distribution adaptation via kernel mapping. *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1027-1036, doi.org/10.1145/1557019.1557130.
- Zhu R, Yan L, Mo N, 2019 Semi-supervised center-based discriminative adversarial learning for cross-domain scene-level land-cover classification of aerial images. *ISPRS J Photogramm Remote Sens*, 155, 72-89, doi.org/10.1016/j.isprsjprs.2019.07.001.