Domain Adaptation for 2D/3D Change Detection in VHR Imagery via Calibration of Convolutional Neural Network under Prior Probability Shift

Hunsoo Song (1), Anjin Chang (2), Junho Yeom (3), Jinha Jung (4), Yongil Kim (1)

[1] Seoul National Univ., 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea
[2] Texas A&M Univ. Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 78412, USA
[3] Gyeongsang National Univ., 501 Jinju-daero, Jinju-si, Gyeongsangnam-do 52828, Korea
[4] Purdue Univ., 550 Stadium Mall Drive, West Lafayette, IN 47907-2051, USA
Email: songhunsoo92@gmail.com; anjin.chang@tamucc.edu; junho.yeom@gnu.ac.kr; jinha@purdue.edu; yik@snu.ac.kr

KEY WORDS: Domain Adaptation, Change Detection, Convolutional Neural Network, VHR, Prior Probability Shift

ABSTRACT: The success of deep learning algorithms applied on large-scale remotely sensed imagery rests on the representativeness of the samples used to train the model. When the source domain training samples do not fully represent the target domain samples where no prior information is available, the trained model may fail to adapt to the target domain, which makes it compulsory to repeat the time-consuming labeling process. Even though this issue could be a critical point to apply deep learning algorithms to change detection in remote sensing, the vast majority of change detection algorithms have not considered the discrepancy between the two domains, and in turn, failed to generalize their performance on a newly given area. To address this problem, we focus on the discrepancy where the class distributions of the two domains are not balanced, defined as the prior probability shift, associated with the supervised change detection. This research shows that the performance of a convolutional neural network (CNN) is largely impaired by the inconsistent changed ratio and proposes a novel approach to resolve the problem of prior probability shift. The proposed framework estimated the changed ratio of the target domain by patch-based change vector analysis (CVA) and calibrated the change threshold from the softmax output of the CNN to adapt well to the target domain. The framework was implemented in three sub-regions with various changed ratios of 20.0 %, 45.0 %, and 35.0 %, acquired from bi-temporal Unmanned Aerial Vehicle (UAV) imagery with 0.1-m resolution including RGB channels and digital surface model (DSM). A total of 1,000 and 2,000 training samples drawn from each of the three sub-regions (source domain), with the two different sampling scenarios (i.e., balanced sampling and imbalanced sampling), were trained by a CNN. Subsequently, the trained CNN was adapted using the patch-based CVA and tested on the two other sub-regions (target domains) of different change/non-change ratios, respectively. The proposed framework was separately applied to RGB imagery and RGB + DSM imagery. The results demonstrated an improvement in the change detection performance by evaluating the overall accuracy, recall, precision, and F1-score, thus confirming the effectiveness of the proposed framework to address the prior probability shift problem.

1. INTRODUCTION

With great advances in remote sensing technology, the amount of remotely sensed imagery produced by spaceborne and airborne sensors is growing enormously. To deal with rapidly increasing amounts of data, techniques to effectively interpret the earth's surface are developing. Change detection is one of the most important techniques in remote sensing, and is widely used to update changes to manage a large-scale area (Hansen, 2012; Tewkesbury, 2015). Significant knowledge for urban and environmental studies can be attained from accurate and timely change detection (Rahman, 2016; Sexton, 2015). Although there have been numerous efforts to develop change detection techniques, there is still no universal change detection algorithm with high accuracy that can keep the pace of rapidly accumulating remotely sensed data (Hussain, 2013; Qin, 2016). Therefore, development of a change detection is imperative.

Depending on the presence of prior knowledge in the investigation area, change detection algorithms are generally divided into supervised methods and unsupervised methods (Tewkesbury, 2015). Supervised methods learn the decision boundary of change/non-change by training the groundtruth data; however, unsupervised methods do not require prior knowledge of the data. Although unsupervised methods have the advantage that they do not require training samples, it is difficult to set a certain threshold of change, and the accuracy of the clustering method depends too much on the parameters chosen (Hussain, 2013). In addition, remote-sensing data are often obtained from multimodal sensors, namely, optical sensor, synthetic aperture radar, and LiDAR, so the properties and the dimensions of the data differ (Wegner, 2010; Bovolo, 2015; Zhang, 2015). Moreover, multimodal data consist of data with varying precisions and, in most cases, have different abilities for solving a given task (Byun, 2013; Jung, 2014; Chang, 2015). Therefore, an algorithm that can consider non-linear features, and assign a different weight to each of

the data therein, is desirable (Bovolo, 2015; Mou, 2018). However, most of the unsupervised algorithms lose information of the original data because most of them use simplified subspace of the original feature space to solve the problem, and cannot automatically identify the importance among features (Bovolo, 2015). Thus, the differences between bi-temporal images, intertwined with light scattering mechanisms and multimodal data sources, should be addressed using a learning-based method that can learn complex and non-linear behavior.

Among the supervised methods, deep neural networks have recently shown successful performance in change detection applications (Zhu, 2017; Song, 2018; Mou, 2018). Deep neural networks with multiple non-linear layers effectively learn high-level abstract features in the multi-temporal images, and outperform conventional machine learning algorithms (Heydari, 2019). However, to perform deep learning successfully, a large number of training samples are often required, and the success of deep learning algorithms largely depends on the representativeness of the samples used to train the model (LeCun, 2015). In many applications for remotely sensed imagery, however, collecting groundtruth information of a good representation is often expensive and labor-intensive. Thus, there have been studies reusing the trained model or the collected training samples to the scenes for which there is no existing knowledge (Persello, 2014); however, directly applying the trained model or the collected training samples to an unseen area often fails (Tuia, 2016).

Performance degradation is likely to occur when the trained model, through any machine learning algorithms, is applied to a new region (*target domain*) that is different from the region (*source domain*) where the training samples are collected. This is because of *the dataset shift* where the joint distribution of inputs and outputs differs between the training and test datasets (Moreno-Torres, 2012). To tackle the problem where the distribution in the source domain does not fully represent the distribution in the target domain, there have been studies to overcome the discrepancy between two domains; this is known as *domain adaptation* (Patel, 2015). Domain adaptation is a specific branch of the transfer learning problem with multitask learning, self-taught learning, sample selection bias, and covariate shift (Pan, 2009). Among the branches of transfer learning, the domain adaptation problem refers to the case where the joint probabilities distribution of the two domains describing the relationship between the input variable and the output variable are similar but not identical. Therefore, the goal of domain adaptation methods is to bridge the two distribution gaps so that the algorithms that work in the source domain work well in the target domain as well.

In remote sensing, the domain adaptation problem is very common and occurs in various forms in intertwined manners. This occurs when the imageries to be dealt with share common characteristics, but they were acquired from different areas, at different times, or from different sensors. Of course, domain adaptation problems also arise in performing change detection. However, the vast majority of change detection algorithms have not considered their performance in diverse domains, and in turn, failed to generalize their performance due to discrepancy among the domains. Among the dataset shifts, inducing domain adaptation problem, the prior probability shift specifically refers to changes in the distribution of the class variable (Moreno-Torres, 2012; Redko, 2019). The prior probability shift is prevalent in change detection because changes in remote sensing occur with the different speed regionally over a very large area. To elaborate, urban development does not change at the same rate over the entire range, but the rate of change in urban development varies considerably from region to region. Likewise, changes caused by disasters also vary greatly in each region. In this situation, obviously, the prior probability shift occurs and severely degrades the change detection performance. However, very few studies have considered the prior probability shift in change detection.

Figure 1 illustrates examples of the prior probability shifts in two-dimensional feature spaces. The samples of input variables X are represented as points in the feature spaces with the two different class Y_1 and Y_2. The goal of change detection algorithm is to find an optimal decision boundary to identify whether the incoming data change or not. The class conditional probability density of each class is the same in all domain, i.e., $P_s(X|Y_i) = P_{t1}(X|Y_i) = P_{t2}(X|Y_i)$, where $i = 1, 2$. However, the prior probability of each class is different in all domains. The prior probability ratio for each class is 50:50 in the source domain, whereas the prior probability ratio for each class is 20:80 in the target domain 1 and 80:20 in the target domain 2, i.e., $P_s(Y_i) \neq P_{t1}(Y_i)$ and $P_s(Y_i) \neq P_{t2}(Y_i)$, where $i = 1, 2$. As a result, the joint probabilities on the source and target domain is different, i.e., $P_s(X|Y_i)P_s(Y_i) \neq P_{t1}(X|Y_i)P_{t1}(Y_i)$ and $P_s(X|Y_i)P_s(Y_i) \neq P_{t2}(X|Y_i)P_{t2}(Y_i)$, where $i = 1, 2$. Therefore, the decision boundary in the source domain should be calibrated to attain optimal result in the target domain. This problem is of crucial importance since the changed ratio greatly differs in remotely sensed imagery.

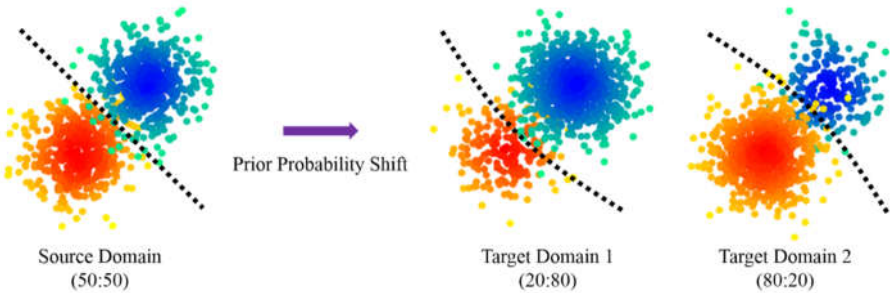

Figure 1. Examples of the Prior Probability Shift in Binary Classification

In this research, we focused on the prior probability shift problem in change detection. We firstly observed that the prior probability shift problem severely degrades the change detection performance. Then, an approach to resolve the problem of prior probability shift has been proposed. The proposed novel framework estimates the changed ratio of the target domain by the patch-based change vector analysis (CVA) and calibrates the decision boundary by ranking the softmax output of the convolutional neural network (CNN) to adapt well to the target domain. The proposed framework was implemented in three sub-regions with various changed ratios acquired from bi-temporal Unmanned Aerial Vehicle (UAV) imagery with 0.1-m resolution including RGB channels and digital surface model (DSM). The results from the six dataset shifts among the three sub-regions demonstrated the effectiveness of the proposed framework to address the prior probability shift problem.

The rest of this paper is organized as follows: Section 2 presents the proposed framework for domain adaptation under prior probability shift, Section 3 includes the datasets, experimental results, and discussion, and Section 4 concludes this paper.

2. DOMAIN ADAPTATION FRAMEWORK

In remote sensing literature, various approaches have been presented to address the domain adaptation problem. Adaptation methods can be grouped into four categories (Tuia, 2016). The first category methods involve finding and utilizing invariant features during the shift between the two domains (Bruzzone, 2009). The second category methods involve aligning the distribution of the source domain with the distribution of the target domain and perform the adaptation using the same classifier (Yang, 2015). The third category methods involve applying the trained model from the source domain to the target domain in a semi-supervised manner leaving both distributions intact (Rajan, 2006; Bahirat, 2011). The last category methods involve assuming the presence of a few labeled samples in the target domain, and by utilizing these, making the classifier adapt to the target domain (Crawford, 2013; Yu, 2017).

To solve the domain adaptation problem, it is important to select an appropriate method considering the available data and the given type of dataset shift. In this study, we concentrated on the prior probability shift problem that occurs among spatially disjoint areas acquired in the same time zone. Since the changed ratio varies locally in large-scale remotely sensed imagery, we assumed that the different prior probabilities among the domains could be the decisive factor of the datasets shift in change detection tasks. In addition, we supposed the case where the training sample of the target domain is not available. Thus, among the aforementioned methods, it was determined that the selection of a methodology belonging to the third category, where the trained model is adapted to the target domain, is desirable.

In short, the proposed framework adapted the trained model to the target domain considering its prior probability estimated by unsupervised algorithm. We adopted CNN for training model and the patch-based CVA for estimating the prior probability of the target domain.

The flowchart of the proposed framework can be summarized as follows (Figure 2).

1) Let the CNN trained from the source domain infer the target domain and rank the probabilities of change for each input derived from the softmax output of the trained CNN.
2) Calculate the threshold that maximizes the overall accuracy in the source domain using the patch-based CVA on imagery.
3) Apply the threshold obtained in step 2) to the target domains and estimate the changed ratio of the target domain.
4) Assign the changed pixel in order of the highest probability derived from the CNN, assuming that the target domain change has been made by the changed ratio estimated using CVA.

A change detection map was generated by applying the proposed method for every pixel in the target scene.

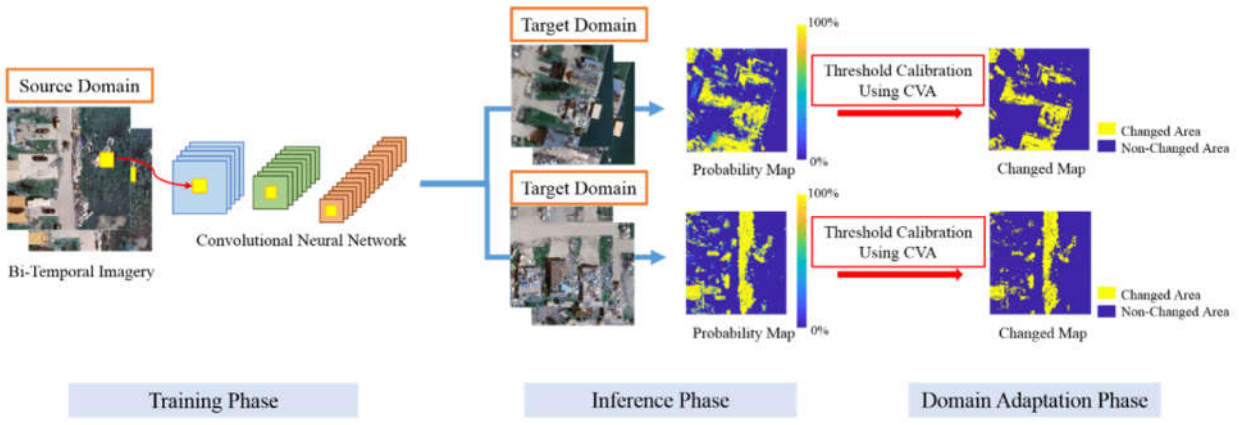

Figure 2. Overview of the Proposed Domain Adaptation Framework.

2.1 Estimation of the Changed Ratio via the patch-based CVA

CVA, a widely used unsupervised algorithm for change detection, is adopted to estimate the changed ratio of the target domain. This is because CVA does not require a training sample, and it is believed to be more insensitive to the change in the distribution than a CNN.

CVA subtracts corresponding bands of bi-temporal images to calculate the difference between two images and determine the changed pixel (Bovolo, 2006). In this experiment, a patch-wise performing CVA was adopted and only its magnitude to identify changed pixels was considered. Since a pixel-based CVA is assumed to be liable to illumination variations, which yield salt and pepper noises, we considered a patch-based CVA is better to estimate changed ratio accurately by averaging the noise effects. The patch size was determined to be 9 pixel by 9 pixel, approximately covering the area of 1-m by 1-m in this study area, in order to produce the optimal threshold. The magnitude $M(i,j)$ of the patch-based CVA was calculated by the following formula (1):

$$M(i,j) = \sum_m \sum_n \sum_b \left(X_b(i+m, j+n) - X'_b(i+m, j+n) \right) \quad (1)$$

where (i,j) is the pixel coordinate of the image, and m, n = -4, -3, -2, -1, 0, 1, 2, 3, 4. X_b and X'_b are the digital number of the b^{th} band of bi-temporal imagery.

In addition, for RGB-D imagery, a weight of three was assigned to DSM compared to the weight of one assigned to each RGB band. This is because the DSM information is more insensitive to the illumination variations than the RGB channel's spectral information in general (Tian, 2013; Qin, 2014). For RGB imagery, the changed ratio was estimated with the patch-based CVA using only RGB imagery. This is firstly for those cases where only RGB is available, and secondly, we tried to test whether the CVA estimates the changed ratio soundly using the optic image alone provided that most of the changes involve the changes in RGB. If the estimation is acceptable, even with RGB alone, it is also possible to estimate the change/non-change ratio when additional data, other than RGB and requiring non-linear operations, should be considered.

The estimation for the changed ratio of the target domain via the patch-based CVA is as follows. First, find the optimal magnitude of the patch-based CVA that maximizes the overall accuracy in the source domain. In this process, find the n^{th} percentile of the magnitude that yields the highest overall accuracy in the source domain. Then, the threshold value maximizing overall accuracy in the source domain is directly applied to the target domain to estimate its changed ratio. Under the assumption that the optimal threshold obtained from patch-based CVA does not vary greatly according to the training sample, the threshold was determined once at the start by examining the entire pixel of each source domain.

2.2 CNN Architecture

In this study, the proposed framework adopted a CNN as a training model. This is because CNN is a widely used supervised method for change detection and has demonstrated successful performance. A CNN consists of a large number of trainable parameters with consecutive layers that allow the network to learn high-level abstract features from the massive input data (LeCun, 2015). However, a CNN learns its parameter to minimize the loss with the given distribution of the training samples similar to other supervised algorithms. Thus, the trained model is prone to failure in classification when the test samples lie on different distributions. Therefore, adaptation of a CNN to the target domain is crucial to maintain its good performance.

The adopted CNN consists of three convolutional layers, and each layer having 10, 20, and 40 filters, respectively. The filter size of the first convolutional layer was 3 by 3, and the filter size of the second and third convolutional layers was 2 by 2. The CNN takes three-dimensional patch as inputs. The patch size of the CNN was also determined to be 9 pixel by 9 pixel, same as the patch-based CVA. After taking this input, the trained CNN can produce the probability whether the center pixel of the patch has changed or not with the softmax output (Alshehhi, 2017; Xu, 2018). As two sets of stacked bi-temporal RGB imagery and RGB-D imagery were used, the CNN takes the input either with a size of 9 by 9 by 6 for stacked RGB imagery or with a size of 9 by 9 by 8 for stacked RGB-D imagery. Then, each CNN learns the decision boundary to solve a binary problem. As a small size of patch is used as an input, the pooling layer is not used. Also, the fully-connected layer is excluded to lower the complexity. ReLU (Rectified Linear Unit) and the Adam optimizer (Kingma, 2014) with a learning rate of 0.001 were used.

Most CNNs have a deep and complex network with a large number of trainable parameters because the distinctive advantage of the CNN is its ability to deal with big and intricate data (Canziani, 2016). However, the CNN used in this experiment has considerably low complexity. This is because a CNN with high complexity increases the computational burden and necessitates a large number of training samples for the network to perform well without overfitting (He, 2016; Song, 2019). Considering the given task of performing a binary classification with a small size of patch as an input, a lightweight CNN was used as adopted in (Song, 2019), but the proposed framework is not limited to the CNN used here.

2.3 Calibration of the Prior Probability Shift

Although CNN fails to produce accurate posterior probability, the rank of the probabilities of change, inferred from the softmax output in the CNN, is still valid provided only a prior probability shift has occurred. Therefore, it was assumed that the dataset shift problem can be alleviated by estimating the prior probability in the target domain from the information produced from the source domain.

From the above observations, the proposed framework effectively calibrates the prior probability shift problem. First, the changed ratio of the target domain is estimated as described in subsection 2.1. Then, the trained CNN, illustrated in subsection 2.2., infers the target domain and produces the probabilities of change using the softmax output. Subsequently, it is determined whether the sample has changed in the order of the highest probability under the assumption that the target domain has changed by the estimated changed ratio from the patch-based CVA. With this process, it is possible to calibrate the bias of the CNN created by the different prior probability of the classes.

3. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the study area and the changed ratios predicted by the CVAs are presented. Afterward, the performance of the proposed framework is assessed. The results from the six dataset shifts among three sub-regions (i.e., from A to B and C, from B to A and C, and from C to A and B) with two different sampling scenarios and the two different training sample sizes are illustrated.

3.1 Datasets and Preprocessing

The proposed domain adaptation framework was evaluated in the study area where the hurricane caused many changes in buildings and vegetation. The study area, a Holiday Beach in Rockport in southern Texas, had experienced significant changes during the recovery of Hurricane Harvey (Figure 3). Bi-temporal images were acquired from UAV and processed using Agisoft Photoscan Pro Software to generate the orthomosaic RGB imagery and DSM. Since the UAV imagery covers a large area, we clipped three sub-regions and down-sampled them to 0.1-m resolution. The three sub-regions (A, B, and C) consisted of 500 by 500 pixels each and had the changed ratio of 20.0 %, 45.0 %, and 35.0 %, respectively. Water bodies were excluded from the experiment because orthomosaic images and DSM over the water bodies are generally of poor quality and not reliable. We also excluded areas scattered with small, insignificant objects because it is difficult to identify whether they change.

The CNN was trained using the training samples drawn from each of the three sub-regions (source domain). Subsequently, the trained CNN was adapted and tested to the two other sub-regions (target domain) that have different change/non-change ratios such that the proposed framework was assessed on six dataset shifts (i.e., from A to B and C, from B to A and C, and from C to A and B). In addition, a total of 1,000 and 2,000 training samples were tested with balanced sampling scenario and imbalanced scenario. The balanced sampling scenario has the same size of training samples for each class (i.e., 500 samples for non-changed class and 500 samples for changed class out of a total of 1,000 training samples). On the contrary, the imbalanced sampling scenario sampled the training samples for each class with the same proportion to the class ratio of each scene. To sum up, with the two

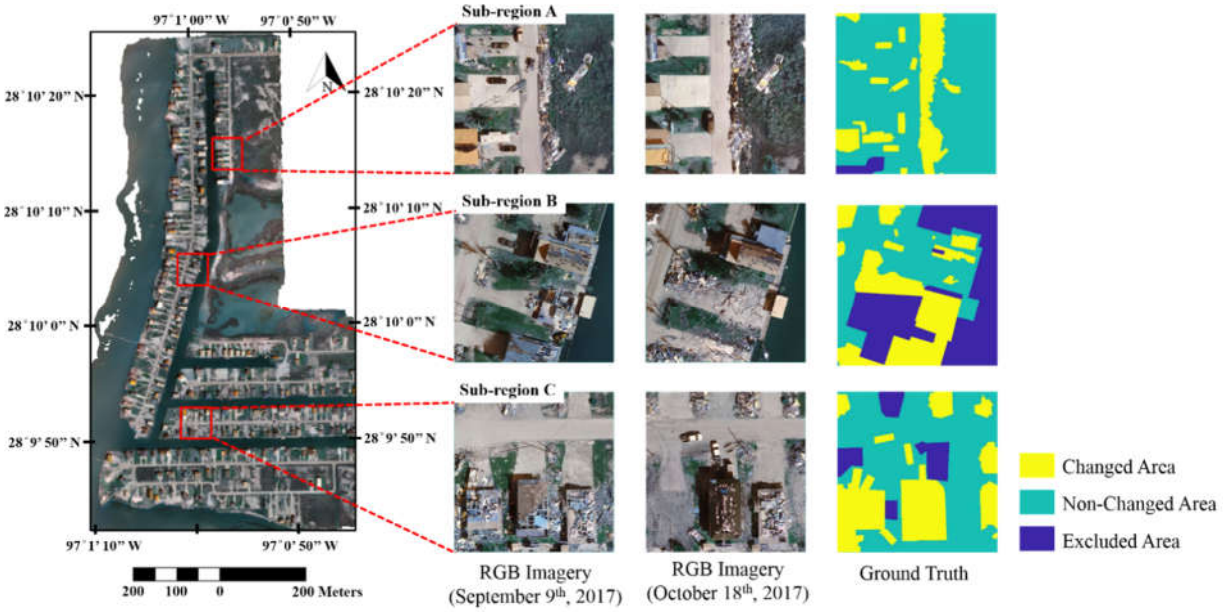

Figure 3. Experimental Areas

sampling scenarios and the two different training sample sizes, the performance of the proposed framework was evaluated. Additionally, the proposed framework was applied separately to RGB and RGB-D imagery. For the CNN, bi-temporal images were stacked to conserve the full information in original images and binary classification was performed as adopted in (Volpi, 2013). All values in RGB and DSM were rescaled to the range of [0, 1] by min-max normalization before the operation of the CNN and CVA.

3.2 Estimated Changed Ratio through CVA

The estimated changed ratios and their errors are reported in Table 1. The results of the pixel-based CVA and the patch-based CVA from the two types of imageries are illustrated. In addition, the mean absolute error (MAE) for each case is calculated. As a result, the patch-based CVAs estimated the changed ratio more accurately by 4.1 % for RGB and by 4.4 % for RGB-D than the pixel-based CVAs. Therefore, it was confirmed that the use of patch-based CVA is desirable.

Another notable point is that the error varies depending on the source domain. When the source domain is A, the changed ratios were underestimated in all cases. This is presumably due to the fact that the source domain A has a relatively small changed ratio, hence the calculated threshold maximizing the overall accuracy (OA) might be low.

Table 1. Estimation of the Changed Ratio Using the Pixel-based CVA and the Patch-based CVA (The values in parenthesis indicate the error.)

	Threshold from the Pixel-Based CVA				Threshold from the Patch-Based CVA			
RGB Imagery	Target / Source	A	B	C	Target / Source	A	B	C
	A	-	14.2% (-30.8%)	12.0% (-23.0%)	A	-	28.5% (-16.5%)	24.5% (-10.5%)
	B	36.5% (+16.5%)	-	39.7% (+4.7%)	B	29.9% (+9.9%)	-	44.3% (+9.3%)
	C	30.0% (+10.0%)	37.7% (-7.3%)	-	C	28.7% (+8.7%)	46.5% (+1.5%)	-
	MAE (%)	11.6%			MAE (%)	7.5%		
RGB-D Imagery	Target / Source	A	B	C	Target / Source	A	B	C
	A	-	21.5% (-23.5%)	14.7% (-20.3%)	A	-	36.2% (-8.8%)	29.5% (-5.5%)
	B	32.4% (+12.4%)	-	39.2% (+4.2%)	B	27.5% (+7.5%)	-	40.3% (+5.3%)
	C	29.2% (+9.2%)	43.6% (-1.4%)	-	C	26.8% (+6.8%)	46.7% (+1.7%)	-
	MAE (%)	9.2%			MAE (%)	4.8%		

3.3 Domain Adaptation Results

In this subsection, the domain adaptation performance of the proposed framework in RGB and RGB-D imageries is evaluated. The proposed network trained a total of 1,000 and 2,000 training sample sizes from the source domain with the two different sampling scenarios (i.e., balanced sampling and imbalanced sampling), and then predicted all pixels in the target domains. The estimated changed ratios produced from the patch-based CVA (in Table 1) were applied to RGB and RGB-D imagery. Random samplings for every case were replicated five times to increase the statistical confidence of the results.

The results of OA from the six dataset shifts (i.e., from A to B and C, from B to A and C, and from C to A and B) with the two sampling scenarios using RGB and RGB-D imagery are shown in Table 2 and Table 3, respectively. All values in Table 2 and Table 3 are in percent, and the results from using the training sample size of 2,000 was excluded because it showed a similar trend to that of 1,000. In both scenarios, the accuracy increased after domain adaptation except for the few cases.

In particular, when the target domain is scene A, the proposed framework greatly increased the accuracy. This is because the actual change rate of scene A is 20.0 %, which is quite different from the balanced sampling scenario's assumption of 50.0 %. Overall, including the two sampling scenarios, the average of OA increased by 3.2 % and 4.6 % after the domain adaptation for RGB and RGB-D imagery, respectively. Without the domain adaptation, imbalanced sampling was more accurate than the balanced sampling, but after the domain adaptation, the two accuracies were similarly high regardless of the sampling method. The difference in accuracy between the balanced

and imbalanced sampling scenarios was less than 0.5 % in both RGB and RGB-D after the domain adaptation. This result indicates that the proposed framework is less sensitive to sampling strategy, which can later relieve a burden of determining a sampling ratio. Additionally, when comparing RGB and RGB-D, the use of RGB-D is 3.2 % more accurate than the use of RGB imagery alone.

Table 4 shows the results of the domain adaptation with the proposed framework for each sub-region. The results from the two sample sizes with the balanced sampling are shown. The change detection results without dataset shift (the training and the test were performed in the same sub-region) are also shown in Table 4. The recall, precision and F1-score were reported together. Overall, the proposed framework increased the accuracy regardless of imagery set and training sample size. The increase in accuracy was particularly significant when the target scene was A. The results can be attributed to the relatively small changed ratio of the scene A compared to the balanced sampling scenario's assumption of 50.0 %. Additionally, when the domain adaptation was performed on the scene A, the recall decreased but precision increased, resulting in an increase in F1-score. This is because the CNN trained with the balanced sample overestimated the changed ratio of A as it assumed the prior probability of change was half. The larger the difference between the actual change rate and the 50.0 % change rate, the assumption of balanced sampling, the greater the effect of domain adaptation. This confirms that the prior probability shift can significantly degrade the performance of supervised algorithm. Furthermore, it suggests that the proposed framework can be more effective especially where the variation of changed ratios among domains is large.

However, the accuracies after domain adaptation are still lower than the accuracies when the dataset shift does not occur. This discrepancy can be attributed to the wrong estimation of the changed ratio. Another reason is due to the fact that the proposed framework is developed based on the fact that only the prior probability shift may occur. The changed patterns might have been different for each sub-region, and thus, the CNN could not have learned some patterns of change that occurred in the target domain. Further research is required to address other kinds of dataset shifts in change detection.

Another noteworthy observation is that, when the domain adaptation was not performed, the accuracy with the training sample size of 2,000 is lower than the accuracy after the domain adaptation was performed with training sample size of 1,000. The results after the domain adaptation using training sample size of 1,000 is 0.4 % and 3.5 % more accurate than that of using training sample size of 2,000 without domain adaptation in RGB and RGB-D imagery, respectively. This finding confirms that domain adaptation is crucial when the class imbalance among domains is severe and that the proposed framework can improve change detection performance.

Table 2. Domain Adaptation Results with RGB Imagery (training sample size = 1,000, and all values are in percent.)

RGB Imagery	Balanced Sampling Scenario						Imbalanced Sampling Scenario					
	w/o Domain Adaptation			w/ Domain Adaptation			w/o Domain Adaptation			w/ Domain Adaptation		
Source \ Target	A	B	C	A	B	C	A	B	C	A	B	C
A	-	76.3	73.0	-	74.0	74.2	-	71.2	73.4	-	72.4	73.9
B	71.3	-	74.6	81.2	-	74.6	73.1	-	75.4	79.4	-	75.3
C	64.9	72.5	-	80.1	73.7	-	74.9	76.1	-	79.6	76.3	-
Avg. OA	**72.1**			**76.3**			**74.0**			**76.1**		

Table 3. Domain Adaptation Results with RGB-D Imagery (training sample size = 1,000 and all values are in percent.)

RGB-D Imagery	Balanced Sampling Scenario						Imbalanced Sampling Scenario					
	w/o Domain Adaptation			w/ Domain Adaptation			w/o Domain Adaptation			w/ Domain Adaptation		
Source \ Target	A	B	C	A	B	C	A	B	C	A	B	C
A	-	80.0	76.3	-	79.7	77.2	-	71.7	73.7	-	78.3	76.0
B	66.2	-	75.8	83.2	-	76.9	76.0	-	77.0	82.9	-	77.2
C	62.7	78.8	-	82.1	79.0	-	84.2	76.4	-	82.4	78.5	-
Avg. OA	**73.3**			**79.7**			**76.5**			**79.2**		

Table 4. Domain Adaptation Results with the Two Different Sample Sizes (All values are in percent.)

Balanced Sampling		Training Sample Size = 1,000									Training Sample Size = 2,000								
		w/o Dataset shift (Source = Target)			w/ Dataset shift						w/o Dataset shift (Source = Target)			w/ Dataset shift					
					w/o Domain Adaptation			w/ Domain Adaptation						w/o Domain Adaptation			w/ Domain Adaptation		
Imagery \ Target		A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
RGB	OA	88.7	86.3	78.9	68.1	74.4	73.8	80.7	73.9	74.4	89.5	88.3	81.8	74.5	77.7	75.6	82.0	76.2	75.2
	Recall	89.9	86.9	87.4	84.8	75.2	72.7	75.0	62.7	62.5	93.0	87.7	84.2	83.5	74.8	75.5	78.2	65.2	63.6
	Precision	66.4	83.5	65.2	38.0	71.5	60.9	51.1	76.7	64.8	67.4	86.8	70.4	44.1	75.8	62.7	53.4	79.5	66.0
	F1-Score	76.2	85.1	74.4	52.1	72.4	65.7	60.8	68.0	62.3	78.1	87.1	76.4	57.3	75.0	68.4	63.5	70.6	63.4
	Avg. OA	**84.6**			**72.1**			**76.3**			**86.5**			**75.9**			**77.8**		
RGB-D	OA	90.1	87.0	89.3	64.4	79.4	76.1	82.7	79.4	77.0	91.1	90.7	86.1	69.0	81.6	78.0	83.3	81.2	77.8
	Recall	91.7	87.0	88.2	86.8	76.7	71.7	74.6	73.1	67.1	93.9	88.0	90.2	86.3	80.3	73.0	76.1	75.2	68.2
	Precision	69.7	84.5	82.6	37.2	78.0	64.9	55.0	79.9	67.7	71.2	91.0	76.4	39.1	79.6	67.2	56.1	81.8	68.9
	F1-Score	78.9	85.6	85.3	50.9	76.9	67.3	63.3	76.1	67.0	80.9	89.5	82.4	53.4	79.7	69.8	64.6	78.0	68.1
	Avg. OA	**88.8**			**73.3**			**79.7**			**89.3**			**76.2**			**80.8**		

4. CONCLUSION

In this paper, we proposed a novel framework for domain adaptation in change detection working under prior probability shift. The proposed framework estimated the changed ratio of the target domain by the patch-based CVA and adjusted the change threshold for a CNN to adapt well to the target domain.

The experiments on the two sets of bi-temporal imagery, RGB and RGB-D with diverse changed ratios, demonstrate that the proposed framework successfully rectifies the biased threshold and improves the change detection performance. The proposed framework was evaluated with the six different dataset shifts from three sub-regions, including two different sampling scenarios and two different training sample sizes, thus the effectiveness of the proposed method was validated for various conditions. In addition, it is confirmed that the accuracy of the estimated changed ratio is vital where the prior probability shift is dominant and the patch-based CVA is more accurate in estimating the changed ratio than the pixel-based CVA. Furthermore, experiments on the two different sample sizes showed that the domain adaptation was essential to improve the change detection performance in the target domain.

With the increase in the amount of remotely sensed data, previous studies excessively focused on deep complex networks trained with massive data. However, the results of this experimental study indicate that deep learning with a large amount of training data has limitations and its performance can be improved with domain adaptation. This study showed that domain adaptation is compulsory to utilize the training sample appropriately and to process a large amount of data effectively. In addition, changes in remotely sensed imagery occur globally and differ regionally. Thus, the proposed framework being capable of solving the prevailing prior probability shift occurring over a large area has advantages. Last but not the least, the proposed framework can be applied to a wide variety of classifiers. The use of this framework with the combination of other supervised algorithms providing classification results as a probability and other unsupervised algorithms estimating the changed ratio can further increase the change detection accuracy.

Further study will focus on the improvement of changed ratio estimation using different unsupervised algorithms and address other kinds of dataset shifts with different types of data.

ACKNOWLEDGMENT

This research was supported by a grant (2019-MOIS32-015) of Disaster-Safety Industry Promotion Program funded by Ministry of Interior and Safety (MOIS, Korea).

REFERENCES

Alshehhi, R., Marpu, P. R., Woon, W. L., & Dalla Mura, M., 2017. Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing, 130, pp. 139-149.

Bahirat, K., Bovolo, F., Bruzzone, L., & Chaudhuri, S., 2011. A novel domain adaptation Bayesian classifier for updating land-cover maps with class differences in source and target domains. IEEE Transactions on Geoscience and Remote Sensing, 50(7), pp. 2810-2826.

Bovolo, F., & Bruzzone, L., 2006. A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain. IEEE Transactions on Geoscience and Remote Sensing, 45(1), pp. 218-236.

Bovolo, F., & Bruzzone, L., 2015. The time variable in data fusion: A change detection perspective. IEEE Geoscience and Remote Sensing Magazine, 3(3), pp. 8-26.

Bruzzone, L., & Persello, C., 2009. A novel approach to the selection of spatially invariant features for the classification of hyperspectral images with improved generalization capability. IEEE transactions on geoscience and remote sensing, 47(9), pp. 3180-3191.

Byun, Y., Choi, J., & Han, Y., 2013. An area-based image fusion scheme for the integration of SAR and optical satellite imagery. IEEE journal of selected topics in applied earth observations and remote sensing, 6(5), pp. 2212-2220.

Canziani, A., Paszke, A., & Culurciello, E., 2016. An analysis of deep neural network models for practical applications. arXiv preprint arXiv:1605.07678.

Chang, A., Jung, J., & Kim, Y., 2015. Estimation of forest stand diameter class using airborne lidar and field data. Remote sensing letters, 6(6), pp. 419-428.

Crawford, M. M., Tuia, D., & Yang, H. L., 2013. Active learning: Any value for classification of remotely sensed data?. Proceedings of the IEEE, 101(3), pp. 593-608.

Hansen, M. C., & Loveland, T. R., 2012. A review of large area monitoring of land cover change using Landsat data. Remote sensing of Environment, 122, pp. 66-74.

He, K., Zhang, X., Ren, S., & Sun, J., 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 770-778.

Heydari, S. S., & Mountrakis, G., 2019. Meta-analysis of deep neural networks in remote sensing: A comparative study of mono-temporal classification to support vector machines. *ISPRS Journal of Photogrammetry and Remote*

Sensing, *152*, pp. 192-210.
Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS Journal of photogrammetry and remote sensing, 80, pp. 91-106.
Jung, J., Pasolli, E., Prasad, S., Tilton, J. C., & Crawford, M. M., 2014. A framework for land cover classification using discrete return LiDAR data: Adopting pseudo-waveform and hierarchical segmentation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(2), pp. 491-502.
Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
LeCun, Y., Bengio, Y., & Hinton, G., 2015. Deep learning. nature, 521(7553), pp. 436.
Moreno-Torres, J. G., Raeder, T., Alaiz-RodríGuez, R., Chawla, N. V., & Herrera, F. (2012). A unifying view on dataset shift in classification. Pattern Recognition, 45(1), 521-530.
Mou, L., Bruzzone, L., & Zhu, X. X., 2018. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 57(2), pp. 924-935.
Pan, S. J., & Yang, Q., 2009. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*(10), pp. 1345-1359.
Patel, V. M., Gopalan, R., Li, R., & Chellappa, R., 2015. Visual domain adaptation: A survey of recent advances. IEEE signal processing magazine, 32(3), pp. 53-69.
Persello, C., & Bruzzone, L., 2014. Active and semisupervised learning for the classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 52(11), pp. 6937-6956.
Qin, R., 2014. Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery. ISPRS journal of photogrammetry and remote sensing, 96, pp. 179-192.
Qin, R., Tian, J., & Reinartz, P., 2016. 3D change detection–approaches and applications. ISPRS Journal of Photogrammetry and Remote Sensing, 122, pp. 41-56.
Rahman, M., 2016. Detection of land use/land cover changes and urban sprawl in Al-Khobar, Saudi Arabia: An analysis of multi-temporal remote sensing data. ISPRS International Journal of Geo-Information, 5(2), pp. 15.
Rajan, S., Ghosh, J., & Crawford, M. M., 2006. Exploiting class hierarchies for knowledge transfer in hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 44(11), pp. 3408-3417.
Redko, I., Courty, N., Flamary, R., & Tuia, D., 2018. Optimal transport for multi-source domain adaptation under target shift. arXiv preprint arXiv:1803.04899.
Sexton, J. O., Noojipady, P., Anand, A., Song, X. P., McMahon, S., Huang, C. & Townshend, J. R., 2015. A model for the propagation of uncertainty from continuous estimates of tree cover to categorical forest cover and change. Remote Sensing of Environment, 156, pp. 418-425.
Song, A., Choi, J., Han, Y., & Kim, Y., 2018. Change detection in hyperspectral images using recurrent 3d fully convolutional networks. Remote Sensing, 10(11), pp. 1827.
Song, H., Kim, Y., & Kim, Y., 2019. A Patch-Based Light Convolutional Neural Network for Land-Cover Mapping Using Landsat-8 Images. *Remote Sensing*, *11*(2), pp. 114.
Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., & Fisher, P. F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. Remote Sensing of Environment, 160, pp. 1-14.
Tian, J., Reinartz, P., d'Angelo, P., & Ehlers, M., 2013. Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 79, pp. 226-239.
Tuia, D., Persello, C., & Bruzzone, L., 2016. Domain adaptation for the classification of remote sensing data: An overview of recent advances. IEEE geoscience and remote sensing magazine, 4(2), pp. 41-57.
Volpi, M., Tuia, D., Bovolo, F., Kanevski, M., & Bruzzone, L., 2013. Supervised change detection in VHR images using contextual information and support vector machines. International Journal of Applied Earth Observation and Geoinformation, 20, pp. 77-85.
Wegner, J. D., Hansch, R., Thiele, A., & Soergel, U., 2010. Building detection from one orthophoto and high-resolution InSAR data using conditional random fields. IEEE Journal of selected topics in applied Earth Observations and Remote Sensing, 4(1), pp. 83-91.
Xu, Y., Wu, L., Xie, Z., & Chen, Z., 2018. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. Remote Sensing, 10(1), pp. 144.
Yang, H. L., & Crawford, M. M., 2015. Domain adaptation with preservation of manifold geometry for hyperspectral image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(2), pp. 543-555.
Yu, H., Yang, W., Hua, G., Ru, H., & Huang, P., 2017. Change detection using high resolution remote sensing images based on active learning and Markov random fields. *Remote Sensing*, *9*(12), pp. 1233.
Zhang, Y., Yang, H. L., Prasad, S., Pasolli, E., Jung, J., & Crawford, M., 2014. Ensemble multiple kernel active learning for classification of multisource remote sensing data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(2), pp. 845-858.
Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F., 2017. Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), pp. 8-36.