# Cross-Sensor Image Change Detection Based on Deep Canonically Correlated Autoencoders

Yuan Zhou[1,2(✉)], Hui Liu[1], Dan Li[1], Hai Cao[1], Jing Yang[2], and Zizi Li[2]

[1] National Ocean Technology Center, Tianjin 300072, China
zhouyuan@tju.edu.cn
[2] Tianjin University, Tianjin 300072, China

**Abstract.** Change detection for cross-sensor remote sensing images is an important research topic with a wide range of applications in disaster treatment, environmental monitoring and so on. It is a challenging problem as images from various acquisitions have difference in the spatial and spectral domains. Change detection models need effective feature representations to estimate interesting changes, but sometimes the hand-crafted low-level features affect the detection result. In this paper, we propose a novel cross-sensor remote sensing image change detection method based on deep canonically correlated autoencoders (DCCAE). The method extracts abstract and robust features of two multi-spectral images through two autoencoders, and then project them into a common latent space, in which any change detection models can be applied. Our experimental results on real datasets demonstrate the promising performance of the proposed network compared to several existing approaches.

**Keywords:** Change detection · Cross-sensor · Deep canonically correlated autoencoders (DCCAE)

## 1 Introduction

With the development of satellite technology, images shot by various sensors covering the same geographic area are available now. Accordingly, the considerable volume of remote sensing images makes it possible to study changes taking place on the surface of the earth [10]. Cross-sensor remote sensing image change detection aims to detect changes of interest between two images shot by two types of sensors of the same area graphical area at different times. It is an effective method to solve emergencies that require quickly responses, such as those natural disaster management and assessment [2,4]. Therefore, cross-sensor change detection has become an important issue for the remote sensing community [11].

There exist some works performing general cross-sensor multi-spectral change detection. Post-classification comparison (PCC) derives the classification map of each image independently and compare the maps pixel by pixel to detect changes. Object-based image analysis is considered as an ensemble approach

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2019
Published by Springer Nature Switzerland AG 2019. All Rights Reserved
S. Han et al. (Eds.): AICON 2019, LNICST 286, pp. 251–257, 2019.
https://doi.org/10.1007/978-3-030-22968-9_22

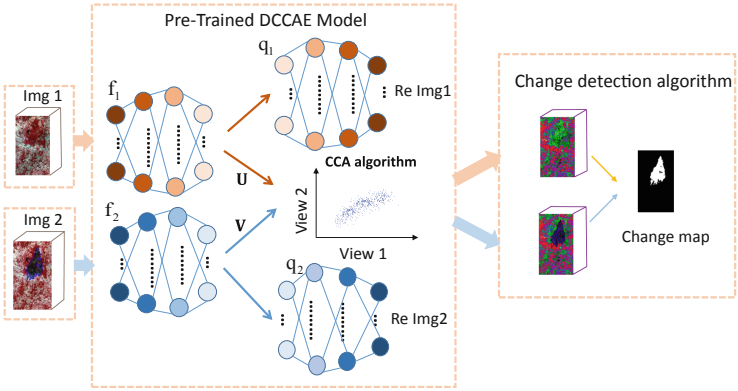


**Fig. 1.** Schematic diagram of DCCAE-based cross-sensor change detection.

which combines image segmentation and classification in image analysis [1]. The main advantage of object-based change detection is that the change map can avoid the error caused by noise. However, object-based methods heavily relies on the segmentation accuracy of the generations of objects. Michele et al. [7] first used kernel canonical correlation analysis (KCCA)[3] to learn nonlinear spectral feature transformations to enhance the accuracy of the change detection. Yang [9] uses a DNN extension of canonical correlation analysis termed DCCA to perform the spectral alignment. Other methods such as manifold learning and Bayesian nonparametric model are studied on the cross-sensor remote sensing image change detection. However, these methods only learn the low-level hand-crafted features of the images, which may affect the change detection accuracy and limit the applications.

In this paper, we propose a novel cross-sensor remote sensing image change detection method based on deep canonically correlated autoencoders (DCCAE). The method learns abstract and high-level features by exploiting a set of pixel samples belonging to the unchanged areas to train the DCCAE network. Feature transformation is made simultaneously from both sides of the network by maximizing the canonical correlation analysis function and minimizing the reconstruction errors. The two cross-sensor images are projected into a common latent space. In the space, any change detection models such as change vector analysis (CVA) can be used. As far as we know, this is the first time that DCCAE models has been successfully applied for cross-sensor remote sensing image change detection.

## 2 DCCAE-Based Cross-Sensor Change Detection Method

The framework of the proposed DCCAE-based method, as shown in Fig. 1, is made up of two stages, including joint transformation of spectral domains and change detection methods.

Let $X$ and $Y$ represent a pair of coregistered images acquired by two types of sensors. Here, we assume the images are coregistered geometrically. We select $N$ unchanged pixels corresponding to each position from both images as the training set. Let $\mathbf{X} = [x_1, x_2, ..., x_N] \in \mathbb{R}^{N \times D_x}$ and $\mathbf{Y} = [y_1, y_2, ..., y_N] \in \mathbb{R}^{N \times D_y}$ be the training data matrices, where $D_x$ and $D_y$ represent the spectral dimension of each image, respectively.

DCCAE proposed by Wang et al. [8] is a model consisting of two autoencoders and the CCA algorithm. The encoder network $f_1,\ f_2$ and the corresponding decoder network $q_1,\ q_2$ are trained to minimize reconstruction errors, which amounts to maximizing the distribution characteristics of the input images [5]. Meanwhile, we project the features to a common latent space by CCA. Projection matrices $\mathbf{U}$ and $\mathbf{V}$ are determined by maximizing the correlation between the learning bottleneck representations $f_1(\mathbf{X}, \theta_1)$ and $f_2(\mathbf{Y}, \theta_2)$, as indicated in Fig. 1. $\{\theta_1, \theta_2\}$ indicate parameters of encoder network and $\{\theta_3, \theta_4\}$ indicate the corresponding parameters of decoder network. The distance between paired feature vectors transformed from unchanged positions is shrunk, and the distance between paired feature vectors transformed from changed positions is enlarged. Mathematically, the DCCAE model optimizes the combination of canonical correlation and the reconstruction errors of the autoencoders:

$$\begin{aligned} \underset{\theta_1,\theta_2,\theta_3,\theta_4,\mathbf{U},\mathbf{V}}{\arg\min} & -\frac{1}{N}\mathrm{tr}(\mathbf{U}^T f_1(\mathbf{X}) f_2(\mathbf{Y})^T \mathbf{V}) \\ & + \frac{\lambda}{N} \sum_{i=1}^{N} (||x_i - q_1(f_1(x_i))||^2 + ||y_i - q_2(f_2(y_i))||^2) \\ & s.t.\ \mathbf{U}^T (\frac{1}{N} f_1(\mathbf{X})^T f_1(\mathbf{X}) + r_1 \mathbf{I}) \mathbf{U} = \mathbf{I}, \\ & \mathbf{V}^T (\frac{1}{N} f_2(\mathbf{Y})^T f_2(\mathbf{Y}) + r_2 \mathbf{I}) \mathbf{V} = \mathbf{I}, \\ & \mathbf{u}_i^T f_1(\mathbf{X}) f_2(\mathbf{Y})^T \mathbf{v}_j = 0, \text{ for } i \neq j, \end{aligned} \tag{1}$$

where $\mathbf{U} \in \mathbb{R}^{d_1 \times l}$ and $\mathbf{V} \in \mathbb{R}^{d_2 \times l}$ are the projection matrices, $l$ is the number of projection vectors. $\lambda$ is a weight that trades off the correlation with reconstruction errors. We use SGD for optimization. The gradient of the DCCAE model is the sum of the gradient for CCA term and the gradient for reconstruction term.

## 3 Experiments and Results

To validate the effectiveness of the proposed method, it is compared with three methods (CCA [7], KCCA [6] and DCCA [9]).

### 3.1 Dataset

The performance of the proposed method is evaluated on the Bastrop country complex fire datasets, which is made up four images acquired by three types of

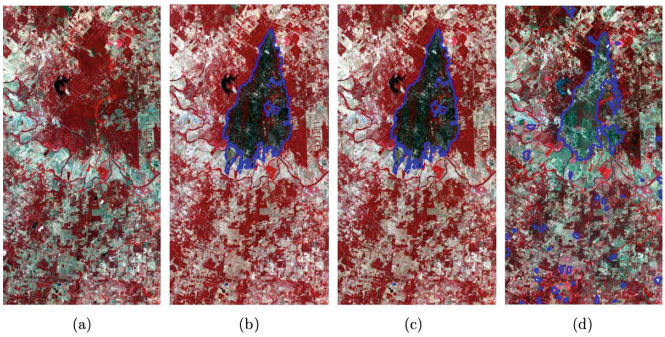


**Fig. 2.** The Bastrop country complex fire datasets. (a) T1 image (landsat 5, Aug.2011) (b) T2 image (landsat 5, Sep.2011) (c) T2 image (EO-1 ALI, Sep.2011) (d) T2 image (Landsat 8, June.2013).

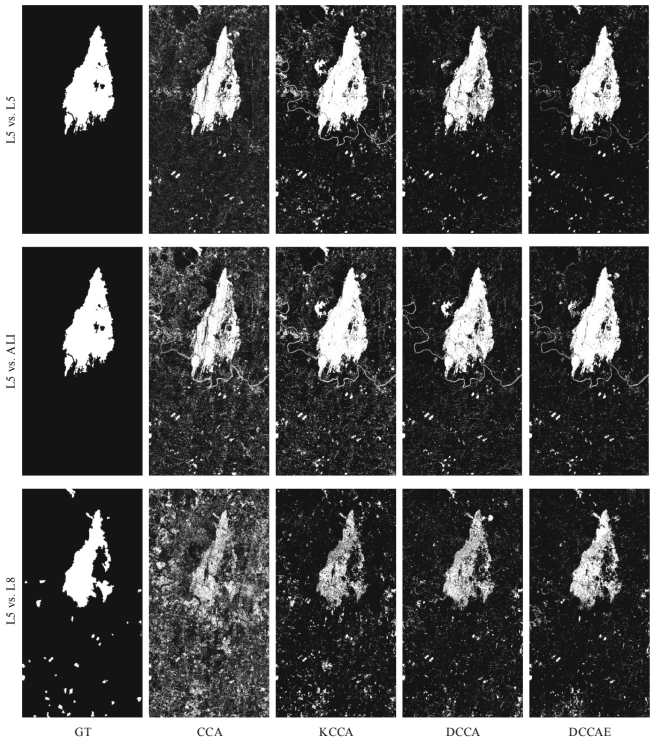


**Fig. 3.** The ground truth map and change maps generated by the four methods.

sensors. The main cause of surface changes was the fire on September 4, 2011. Figure 2 shows the specific information of the datasets. The size of all the images are $1534 \times 808$ pixels with different spectrum channels.

### 3.2 Experimental Settings

(1) Training Set: To study the impact of samples on the results, the number of training samples $N$ is selected from $\{50, 100, 250, 500, 1000\}$. For fair comparison, the same samples are used for the four methods.
(2) DCCAE Network: The networks $f_1$, $f_2$ are built with 3 hidden layers, each of 100 sigmoid units, and an output layer for linear CCA with 7 units. The networks $q_1$, $q_2$ are implemented by 3 hidden layers, each of 100 sigmoid units, and an output layer with the same units as the width of input dataset. In the training procedure, the hyper-parameter of DCCAE model with SGD optimization is set as follows: the learning rate is 0.001 and momentum is 0.99, respectively.
(3) Change Detection Algorithm: We use the change vector analysis (CVA) as the change detection algorithm. We randomly selected 40 validation pixels from the changed or unchanged areas to tune the threshold of each binary CVA model in a supervised wayin order to select a proper threshold.

**Table 1.** OA (standard deviation) comparison of change detection methods.

| | CCA-based | KCCA-based | DCCA-based | DCCAE-based |
|---|---|---|---|---|
| *L*5 vs. *L*5 | | | | |
| 50 | 84.1(6.3) | 91.2(4.7) | 90.7(3.8) | **93.4(2.1)** |
| 100 | 86.2(3.6) | 88.9(6.5) | 91.6(3.2) | **94.3(1.7)** |
| 250 | 82.9(7.7) | 90.7(3.1) | 93.2(1.7) | **93.7(2.9)** |
| 500 | 85.2(4.3) | 90.9(3.1) | 93.6(4.2) | **94.4(1.6)** |
| 1000 | 87.2(3.6) | 92.6(3.0) | 95.6(1.1) | **96.6(1.8)** |
| *L*5 vs. *ALI* | | | | |
| 50 | 80.0(7.0) | 90.2(4.3) | 89.4(4.5) | 89.9(3.8) |
| 100 | 78.4(4.8) | 91.5(2.2) | 92.2(2.6) | **92.9(2.7)** |
| 250 | 80.0(3.3) | 91.2(2.3) | 92.6(3.4) | **93.5(2.9)** |
| 500 | 81.2(5.8) | 90.3(4.0) | 93.4(2.8) | **94.3(2.4)** |
| 1000 | 79.3(5.6) | 89.8(2.3) | 94.6(1.6) | **95.7(2.3)** |
| *L*5 vs. *L*8 | | | | |
| 50 | 75.0(3.2) | 86.9(5.0) | 81.1(7.1) | 84.5(3.3) |
| 100 | 76.8(3.0) | 88.5(4.3) | 85.9(4.8) | **88.9(2.7)** |
| 250 | 77.2(3.1) | 88.9(3.7) | 90.2(1.4) | **90.6(2.8)** |
| 500 | 76.2(1.9) | 91.3(1.7) | 92.3(1.1) | **92.9(1.5)** |
| 1000 | 77.1(3.0) | 91.7(1.2) | 93.9(2.2) | **94.3(1.2)** |

### 3.3 Evaluation Results

Table 1 presents the OAs of all competitors and the proposed method under different number of training sets. When the number reaches 100, DCCAE-based method can get more accurate change detection results. Figure 3 lists the change detection results of the Bastrop datasets obtained by the CVA algorithm after the projection of the datasets by CCA, KCCA, DCCA and DCCAE. As can be seen, for the L5T1 vs. L5T2 experiment, the proposed DCCAE-based change maps can detect correct burned area effectively. For the L5T1 vs. ALIT2 experiment, DCCAE-based change map shows the least additional other changes. Other change maps include more small water basins. For the L5T1 vs. L5T8 experiment, DCCAE-based change maps miss less internal pixels than other maps.

## 4 Conclusion

In this paper, we propose a novel change detection method based on deep canonically correlated autoencoders (DCCAE) for cross-sensor change detection. Our experimental results show that the proposed method significantly can effectively detect the interesting changes. Considering the practical application scenarios, future work should address the cross-sensor change detection task in an unsupervised way.

## References

1. Gong, M., Zhan, T., Zhang, P., Miao, Q.: Superpixel-based difference representation learning for change detection in multispectral remote sensing images. IEEE Trans. Geosci. Remote Sens. **55**(5), 2658–2673 (2017)
2. Gueguen, L., Hamid, R.: Toward a generalizable image representation for large-scale change detection: application to generic damage analysis. IEEE Trans. Geosci. Remote Sens. **54**(6), 3378–3387 (2016)
3. Lai, P.L., Fyfe, C.: Kernel and nonlinear canonical correlation analysis. Int. J. Neural Syst. **10**(05), 365–377 (2000)
4. Roemer, H., Kaiser, G., Sterr, H., Ludwig, R.: Using remote sensing to assess tsunami-induced impacts on coastal forest ecosystems at the andaman sea coast of thailand. Nat. Hazards Earth Syst. Sci. **10**(4), 729 (2010)
5. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.A.: Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J. Mach. Learn. Res. **11**(Dec), 3371–3408 (2010)
6. Volpi, M., Camps-Valls, G., Tuia, D.: Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis. ISPRS J. Photogramm. Remote. Sens. **107**, 50–63 (2015)
7. Volpi, M., de Morsier, F., Camps-Valls, G., Kanevski, M., Tuia, D.: Multi-sensor change detection based on nonlinear canonical correlations. In: 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1944–1947. IEEE (2013)

8. Wang, W., Arora, R., Livescu, K., Bilmes, J.: On deep multi-view representation learning: objectives and optimization. arXiv preprint arXiv:1602.01024 (2016)
9. Yang, J., Zhou, Y., Cao, Y., Feng, L.: Heterogeneous image change detection using deep canonical correlation analysis. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp. 2917–2922. IEEE (2018)
10. Zhang, Z., Vosselman, G., Gerke, M., Tuia, D., Yang, M.Y.: Change detection between multimodal remote sensing data using siamese CNN. arXiv preprint arXiv:1807.09562 (2018)
11. Zhao, W., Wang, Z., Gong, M., Liu, J.: Discriminative feature learning for unsupervised change detection in heterogeneous images based on a coupled neural network. IEEE Trans. Geosci. Remote Sens. **55**(12), 7066–7080 (2017)