SCIENCE CHINA
Information Sciences

• LETTER •

April 2023, Vol. 66 140308:1–140308:2
https://doi.org/10.1007/s11432-022-3693-0

Special Topic: Artificial Intelligence Innovation in Remote Sensing

Multi-layer composite autoencoders for semi-supervised change detection in heterogeneous remote sensing images

Jiao SHI[1], Tiancheng WU[1], Hanwen YU[2], A. K. QIN[3], Gwanggil JEON[4] & Yu LEI[1*]

[1]*School of Electronics and Information, Northwestern Polytechnical University, Xi'an* 710129*, China;*
[2]*School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu* 611731*, China;*
[3]*Department of Computer Science and Software Engineering, Swinburne University of Technology, Melbourne* 3122*, Australia;*
[4]*Department of Embedded Systems Engineering, Incheon National University, Incheon* 22012*, Republic of Korea*

Received 2 November 2022/Revised 8 January 2023/Accepted 18 January 2023/Published online 17 March 2023

Citation Shi J, Wu T C, Yu H W, et al. Multi-layer composite autoencoders for semi-supervised change detection in heterogeneous remote sensing images. Sci China Inf Sci, 2023, 66(4): 140308, https://doi.org/10.1007/s11432-022-3693-0

With the increasing complexity of application scenarios, the fusion of different remote sensing data types has gradually become a trend, which can greatly improve the utilization of massive remote sensing data.

While the problem of change detection for heterogeneous remote images can be much more complicated than the traditional change detection for homologous remote sensing images, there are huge differences between heterogeneous images caused by factors such as the light sensitivity and object reflection properties [1, 2]. So the common methods are meant to align the images from two different domains and then compare the original data in the common domain to highlight the difference [3].

However, most traditional heterogeneous change detection frameworks have complex alignment tasks due to their unique design [4, 5]. However, these additional auxiliary tasks will deepen the network's complexity, and it is difficult to balance multiple tasks such as the mentioned loss learning of cycle consistency, weighted translation, etc.

Therefore, it is hoped to propose a concise network framework. One method is to use a small amount of label information to spread its guiding role in network learning. Thus, it is not required to complete the additional tasks of transformation and alignment of two images in the model. On the other hand, by making more use of the discriminant difference information extracted from the traditional framework that has not been fully utilized, it can complete the difference learning only by mapping the labeled samples, avoiding complex additional operations.

Method. Figure 1 shows the structure of proposed multi-layer composite autoencoders (MLCAE) for change detection.

Step 1. For heterogeneous images, two AE networks with same set are used to extract discriminative features of $I1$ and $I2$. It completes the reconstruction of unsupervised data. In this way, their identification information will be more prominent in hidden higher-order space. So the unsupervised loss is shown as follows:

$$L_u(\theta) = N_{I1}(I1, \dot{I}1) + N_{I2}(I2, \dot{I}2), \tag{1}$$

where $N_{I1}(\cdot)$ and $N_{I2}(\cdot)$ represent the operation of network training to $I1$ and $I2$ respectively, which are CAE1 and CAE2 in step 1 of Figure 1. $\dot{I}1$ and $\dot{I}2$ are the output of the decoder, which are meant for reconstruction. It represents the unsupervised part of MLCAE.

Step 2. The combination of multi-layer composite features is carried out in the multi-layer outputs, which are CM_1, CM_2, CM_3, CM_4. To obtain independent change detection results, each corresponding layer of the two AE networks will be combined to connect a classification layer:

$$L_s(\theta) = -I_{\text{label}} \log(Z) + (1 - I_{\text{label}}) \log(1 - Z), \tag{2}$$

where Z is the extracted high-order features in the process of unsupervised learning in step 2 of Figure 1 and I_{label} are the true-labels, being used for the supervised learning in comparison to Z. The supervised loss is the classical cross entropy loss.

So the total training loss can be expressed as follows:

$$\min L(\theta) = L_u(\theta) + \sum_{i=1}^{4} L_s^i(\theta), \tag{3}$$

where $L_s^1(\theta)$, $L_s^2(\theta)$, $L_s^3(\theta)$, $L_s^4(\theta)$ represent the losses of the classification layer connected to each layer of the encoders. Although network training appears to have many losses, it does not require any additional manual parameter settings because it is a mapping of label information.

Step 3. After voting of multiple change detection results from multi-layer composite features, the pseudo-labeled samples with higher confidence are selected, which is shown in step 3 of Figure 1. And it is re-added to the

* Corresponding author (email: leiy@nwpu.edu.cn)

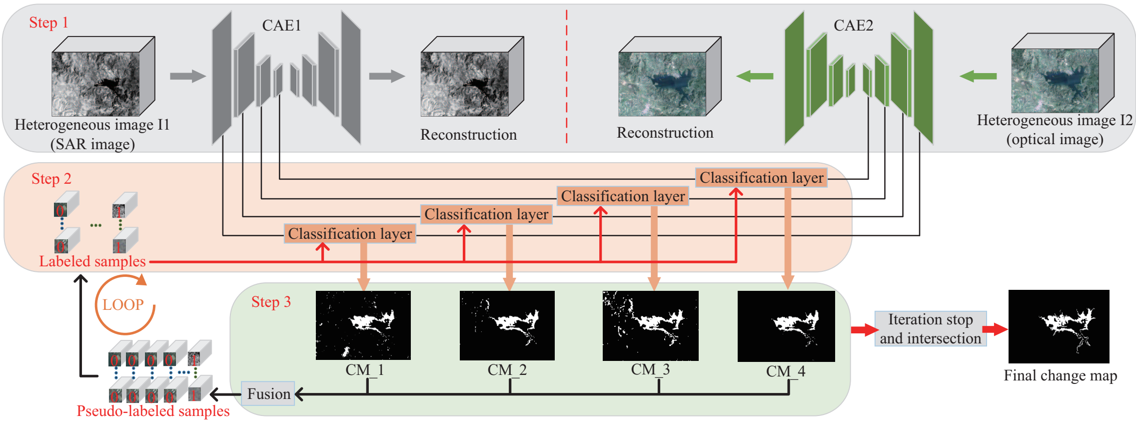

Figure 1 (Color online) Overview of the proposed MLCAE.

0.1% true-labeled training samples to obtain a new training data set to re-train the network. By adding high-confidence pseudo-labeled samples, it can not only avoid the over-fitting problem caused by a small number of true-labeled samples but also mine the information of unlabeled samples more effectively, optimizing the process of change detection.

Finally, when the change degree of the pseudo-labeled training sample is less than a fixed value, the iteration can be stopped. From the experimental results, generally, the global label pool can be diffused after one iteration of the pseudo-labels, and the final change map can be output.

Experiments and results. For change detection in heterogenous remote sensing images, the proposed framework is compared to both unsupervised and semi-supervised methods. Because of space limitations, the detailed results are presented in Appendixes A and B.

Conclusion. We proposed MLCAE for change detection in heterogeneous remote sensing images, which avoids additional tasks such as complex alignment or transformation in the traditional change detection framework for heterogeneous remote sensing images. It only requires 0.1% amount of true labels at the approaching cost of unsupervised models.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 62076204).

Supporting information Appendixes A and B. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Wu Y, Li J, Yuan Y, et al. Commonality autoencoder: learning common features for change detection from heterogeneous images. IEEE Trans Neural Netw Learn Syst, 2022, 33: 4257–4270
2 Luppino L T, Bianchi F M, Moser G, et al. Unsupervised image regression for heterogeneous change detection. IEEE Trans Geosci Remote Sens, 2019, 57: 9960–9975
3 Su L, Gong M, Zhang P, et al. Deep learning and mapping based ternary change detection for information unbalanced images. Pattern Recogn, 2017, 66: 213–228
4 Liu Z G, Zhang Z W, Pan Q, et al. Unsupervised change detection from heterogeneous data based on image translation. IEEE Trans Geosci Remote Sens, 2022, 60: 1–13
5 Luppino L T, Hansen M A, Kampffmeyer M, et al. Code-aligned autoencoders for unsupervised change detection in multimodal remote sensing images. IEEE Trans Neural Netw Learn Syst, 2022, doi: 10.1109/TNNLS.2022.3172183