



HAL
open science

Feature Learning and Change Feature Classification Based on Variational Auto-encoder for SAR Change Detection

Huan Chen, Hongming Zhang, Zhixi Feng

► **To cite this version:**

Huan Chen, Hongming Zhang, Zhixi Feng. Feature Learning and Change Feature Classification Based on Variational Auto-encoder for SAR Change Detection. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.409-416, 10.1007/978-3-031-14903-0_44 . hal-04666449

HAL Id: hal-04666449

<https://hal.science/hal-04666449v1>

Submitted on 1 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Feature Learning and Change Feature Classification based on Variational Auto-encoder for SAR Change Detection

Huan Chen¹, Hongming Zhang¹ and Zhixi Feng²

¹ College of Information Engineering, North West A&F University, Yang Ling 712100, China

² School of Artificial Intelligence, Xidian University, Xi'an 710071, China

huanchen@nwfau.edu.cn

Abstract. As a special auto-encoder, variational auto-encoder (VAE) is not only known as a branch of generation model, but also plays an important role in image feature extraction. Since VAE can get Gaussian distribution with different parameters from different types of inputs or with approximately the same parameters from the same type of input, the latent variables have obvious differences between different categories. This paper put forward a supervised variational auto-encoder (SVAE) to study the representation ability for synthetic aperture radar change detection. Firstly, the difference image is obtained by log ratio method from two original images preprocessed. Then fuzzy c-means (FCM) clustering is used to analyze difference image with the aim of acquiring pseudo labels. As for the inputs of the SVAE, they are selected directly from two SAR images instead of sampling from difference image (DI). Having the inputs and pseudo labels, SVAE can learn latent Gaussian representations according to which SVAE can make a classification for change detection (CD). Experiments on four data set demonstrate that SVAE can obtain discriminative features for CD and outperforms some related approaches.

Keywords: Generation model, Variational auto-encoder, Change detection, Supervised variational auto-encoder, Representation ability.

1 Introduction

Change detection (CD) is a meaningful process which refers to recognize changes by analyzing two or more images over the same area acquired at different times. It has been widely used in various applications, such as urban change analysis [1], agricultural survey [2], environmental monitoring [3], and hazard assessment [4]. Due to the coherent principle of SAR imaging system, there would generate some speckle noises, which makes it more difficult to detect changes than optical ones. It is essential to develop a robust SAR image change detection technique against the speckle noise.

Generally, CD approaches are often comprised of three steps: (1) image preprocessing, (2) DI generation, and (3) unsupervised or supervised classification of the DI into changed and unchanged classes. Recently, deep learning has become a powerful

tool in the field of feature learning. Inspired by the architectural depth of the brain, deep learning has the ability to transform the raw pixels into a higher feature space to capture the discrimination information, and has become a popular way for automatically classification. Geng Jie [5] developed a supervised contractive auto-encoder (SCAE) combined with FCM for SAR CD. Changed features were extracted from two original images and DI based on SCAE, which has been proved to optimize feature representations and enhance classification accuracies. Gong [6] applied a deep belief network (DBN) to SAR images CD without the generation of DI, where CD problem is transformed into a classical classification task. In our previous work [7], a fast unsupervised deep fusion framework based on SAE was proposed for SAR CD. However, DI was generated based on the extracted feature in this work, resulting in redundancy. In addition, unsupervised clustering analysis of DI may lead to low accuracy. However, as mentioned in [5], methods conducted on DI may lose useful information during the DI generation. If representative features are directly extracted from two original images and a more accurate classification model is utilized, there would be a considerable room to improve.

In this paper, supervised variational auto-encoder (SVAE) combined with FCM are developed for SAR CD. The SVAE consists of the encoder of a variational auto-encoders [8] (VAE) and a softmax layer following. The purpose of SVAE is to learn a latent variable which obeys Gaussian distribution and to classify it by the following softmax layer. Since VAE can get Gaussian distribution with different parameters from different types of inputs or with approximately the same parameters from the same type of inputs, the latent variables have obvious differences between different categories. Better results can be achieved in classification and retrieval tasks. To explore the representation ability of VAE in SAR CD, image pairs from two original image are fed to VAE and the learned latent variables are put into a softmax layer. In addition, the training of SVAE are guided by pseudo labels obtained by FCM. Once the SVAE training is done, all image pairs from the two raw image are fed into the learned SVAE, and the final change detection results can be acquired.

The rest of this paper is organized as follows: Section 2 describes the proposed method in details. The experimental results and the corresponding analysis are illustrated in Section 3. Finally, a conclusive remark on this study are summarized briefly in Section 4.

2 The proposed method

In this section, the proposed method is illustrated in detail. The general flowchart of the proposed framework is shown in Fig. 1. As we can see, given two coregistered SAR images, DI can be obtained by log ratio method, and then FCM is used to analyse DI in order to acquire the pseudo labels. After that, the corresponding neighborhood pixels of two SAR images are selected and transformed into vectors as the input of SVAE. Having the input and the corresponding pseudo labels, SVAE can be trained to learn the concept of changes and unchanges. At the testing stage, all sam-

ples from the two SAR images are fed into SVAE and the final change results can be obtained.

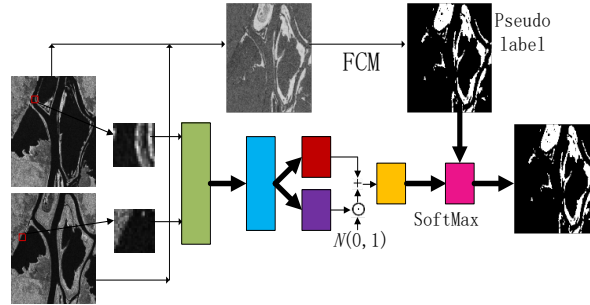


Fig. 1. Illustration of the architecture of SVAE. The green box is the first hidden layer and the blue box is the second hidden layer. The final red box is the latent variable constrained by a Gaussian variation (red box and purple box).

2.1 Preprocessing and FCM

In order to obtain the pseudo labels, we first preprocess two SAR images by mean filter. This process can remove some speckle noises of SAR images. Next, log ratio method is used to acquire DI, and FCM is adopted to analyse DI in order to obtain pseudo labels. Specially, FCM is applied to detect changed and unchanged regions by clustering with high probability, and percent of changed and unchanged are used as pseudo labels. As for the input of SVAE, we flatten each pixel combined with its neighbor pixels of two SAR images and concatenate them into pixel vectors. Thus, pixel vectors and pseudo labels are fed into SVAE for training.

2.2 SVAE for feature learning and classification

Unlike the traditional methods that analyze DI to obtain change results, we desire to learn the representation from two SAR images directly instead of DI, considering useful information would be lost during generating DI. Neighborhood pixel features of two SAR images are input into SVAE to extract effective representations due to the powerful ability of VAE. Traditional AE learn the representation of data by reconstructing the input. While the VAE adds a variation constraint to latent variable, which obeys a Gaussian distribution, and can generate a new sample through the decoder network by sampling from a Gaussian distribution. Thus, VAE not only compresses the input but also studies the distribution of data. The parameters of Gaussian distribution of the same type are approximately the same, while the input belongs to different types have quite different parameters of Gaussian distribution. So VAE can often achieve better results in classification and retrieval tasks [9].

Suppose the pixel vectors $s^i = [s_1^i, s_2^i]$, then in the encoding step, the vector s^i is transformed into hidden representation h by the following expression:

$$h = \tanh(W_1 \cdot s^i + b_1) \quad (1)$$

The mean and variance of the approximation Gaussian distribution is computed as:

$$\mu = W_2 \cdot h + b_2 \quad (2)$$

$$\sigma = W_3 \cdot h + b_3 \quad (3)$$

Then the final latent expression is:

$$z = \mu + \sigma \odot \varepsilon \quad (4)$$

As for the decoding step, the expression is as follows:

$$\hat{h} = \tanh(W_4 \cdot z + b_4) \quad (5)$$

$$\hat{s}^l = \tanh(W_5 \cdot \hat{h} + b_5) \quad (6)$$

Then the last optimization can be written as:

$$L = \frac{1}{2} \|\hat{s}^l - s^l\|^2 + \frac{1}{2} (-\log \sigma^2 + \mu^2 + \sigma^2 - 1) \quad (7)$$

After that, we can optimize the VAE with SGD or Adam. Once the VAE has been trained well, the decoder part is removed and the weightings/biases in the encoder are fixed. Then a softmax classifier is added on the top of encoder. At this moment, the pseudo labels of the corresponding pixel vectors are adopted to train the classifier and fine tune the whole SVAE network. At last, the final change detection results can be acquired directly by feeding all pixel vectors one by one.

3 Experimental settings and results analysis

3.1 Data Description

To confirm the advantages of SVAE, extensive experiments have been performed on the following data sets:

(1) Ottawa: This data set contains a section (290350 pixels) of two SAR images (C-band and HH polarization) with 10-m resolution acquired in May and August 1997, respectively. They are over the city of Ottawa acquired by RADARSAT SAR sensor and provided by the Defence Research and Development Canada, Ottawa. The available ground truth (i.e., reference image) is shown in Fig. 2 (c), and was created by integrating prior information and photointerpretation based on the input images Fig. 2 (a) and (b). The two images present the areas where they were once afflicted with floods. The changed areas represent the affected areas. The experiment on Ottawa dataset is an instance of disaster evaluation.

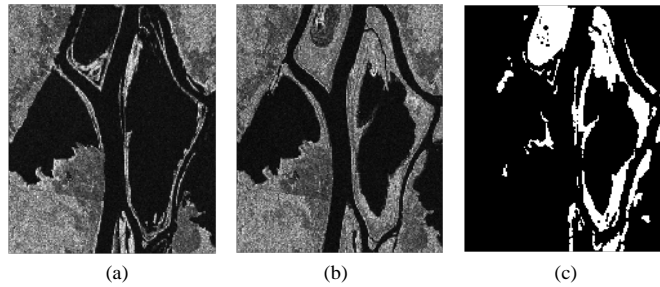


Fig. 2. Illustration of Ottawa data set. (a) Image acquired in July 1997. (b) Image acquired in August 1997. (c) Ground truth.

(2) San Francisco: The second dataset, as shown in Fig. 3, is called the San Francisco dataset. It presents a section (256 256 pixels) of two SAR images acquired by the ERS-2 SAR sensor. The original images are 77497713 pixels. The images were captured in August 2003 and May 2004, respectively. The ground truth change map shown in Fig. 3 (c) is generated by integrating prior information with photo interpretation based on the input images in Fig. 3 (a) and (b).

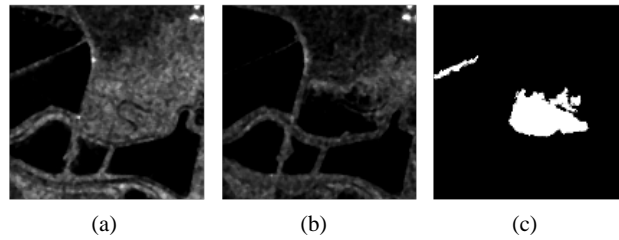


Fig. 3. Illustration of San Francisco dataset. (a) Image acquired in August 2003. (b) Image acquired in May 2004. (c) Ground-truth image.

3.2 General Information

Having change detection results, it is necessary to use some criteria to evaluate the results quantitatively. FP: This index is the number of pixels falsely classified as changed class. FN: This index is the number of pixels falsely classified as unchanged class. OE: The overall error. It is the sum of the FP and FN. KC: Kappa coefficient. It is used to evaluate the effect of the result in the domain of image segmentation. The higher the value of Kappa is, the better the segmentation result is. PCC: Percentage correct classification. It represents the proportion of observed agreement in Kappa. AA: The averaged accuracy, which represents the average percentage of correctly classified pixels for each class. In order to demonstrate the effectiveness of the proposed approach, principal component analysis (PCA), AE and Sparse AE (SpAE) are used to compare with our method. All methods, extracting discriminative features, are followed by FCM to generate the final change detection results.

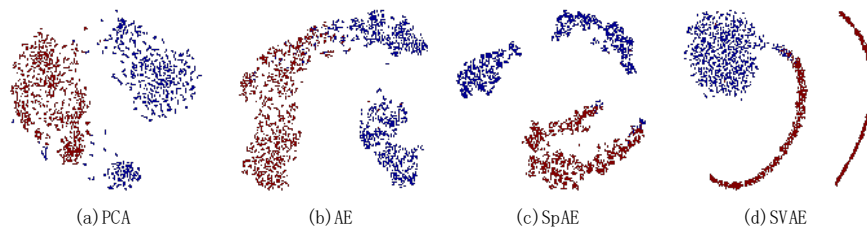


Fig. 4. Latent representation visualization of different methods. (a) PCA. (b) AE. (c) SpAE. (d) SVAE.

3.3 Analysis of the representation ability

In this subsection, we evaluate the representation ability of SVAE based on Ottawa

dataset. As a comparison, PCA, AE and SpAE are selected. We use the t-SNE [10] to visualize the latent representations learned by different methods. The visualization is shown in Fig. 4. It can be seen that the intra class separation distance in Fig. 4 (d) is the minimum, while the intra class separation distance of PCA and AE in Fig. 4 (a) and (b) are larger. This proves that SVAE can learn better latent representations than PCA, AE and SpAE. The following comparison will also illustrate this point.

3.4 Results comparison with other methods on Ottawa Data Set

The results of experiment on the Ottawa data set are shown in Fig. 5 and listed in Table 1. As shown in Fig. 5 (a), it can be found that there exists many white noise spots in the change detection results obtained by PCA. Comparing with the PCA method, AE approach can gain a much better result, as shown in Fig. 5 (b). Due to the additional constraint condition, SpAE can learn the useful information from the samples. As for the result of SVAE shown in Fig. 5 (d), it looks most like the Ground truth. In addition, Table 1 also proof that SVAE can acquire better change detection results than other methods.

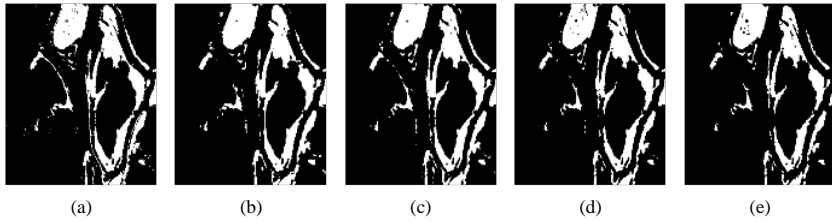


Fig. 5. Change detection results for Ottawa data set obtained by different methods. (a) PCA. (b) AE. (c) SpAE. (d) SVAE. (e) Ground truth.

Table 1. Results obtained by different methods on Ottawa data sets.

Method	FN	FP	OE	PCC(%)	KC	AA
PCA	694	2976	3670	96.38	0.87	96.09
AE	735	3006	3741	96.31	0.86	95.95
SpAE	137	2663	2800	97.24	0.9	98.01
SVAE	312	1092	1404	98.61	0.94	98.38

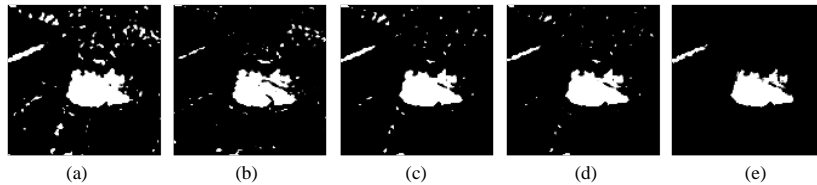


Fig. 6. Change detection results for San Francisco data set obtained by different methods. (a) PCA. (b) AE. (c) SpAE. (d) SVAE. (e) Ground truth.

Table 2. Results obtained by different methods on San Francisco data sets.

Method	FN	FP	OE	PCC(%)	KC	AA
PCA	10	2906	2916	95.55	0.73	97.5
AE	416	1400	1816	97.22	0.81	94.4
SpAE	68	1061	1129	98.27	0.88	98.4
SVAE	126	744	870	98.67	0.91	98.04

3.5 Results comparison with other methods on San Francisco Set

The second experiment is carried on the San Francisco data set. The change detection results of different methods are shown in Fig. 6. It can be seen that there are so many isolated white noisy points in Fig. 6 (a), which reflects higher FP in Table 2. Moreover, the FP values of AE and SpAE are smaller than that of PCA, this can also be seen from Fig. 6 (b) and (c) where exist less white noisy points. Comparing with the other three methods, SVAE can obtain the best result as shown in Fig. 6 (d). The FP, OE, PCC and KC values of SVAE in Table 2 are the best.

4 Conclusion

In this paper, SVAE combined with FCM are developed for SAR CD with the aim of studying the representation ability. SVAE can learn a latent variable which obeys Gaussian distribution and to classify it by the following softmax layer. What is more, SVAE can get Gaussian distribution with different parameters from different types of inputs or with approximately the same parameters from the same type of input, the latent variables have obvious differences between different categories. To explore the representation ability of SVAE in SAR CD, image pairs from two original image are fed to SVAE and the learned latent variables are put into a softmax layer. In addition, the training of SVAE are guided by pseudo labels obtained by FCM. Once the SVAE training is done, all image pairs from the two raw image are fed into the learned SVAE, and the final change detection results can be acquired. Experiments on two data set demonstrate that SVAE can obtain discriminative features for CD and outperforms some related approaches.

References

1. Ban, Y., Yousif, O. A.: Multitemporal spaceborne sar data for urban change detection in china. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(4), 1087-1094(2012).
2. P. Lv, Y. Zhong, Z. Ji, H. Jiao, L. Zhang.: Change detection based on a multifeature probabilistic ensemble conditional random field model for high spatial resolution remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters* 13 (12), 1965-1969 (2017).
3. J. F. Mas.: Monitoring land-cover changes: A comparison of change detection

- techniques. *International Journal of Remote Sensing* 20 (1), 139-152 (1999).
4. L. Giustarini, R. Hostache, P. Matgen, J. P. Schumann, P. D. Bates, D. C. Mason.: A change detection approach to ood mapping in urban areas using terrasar-x. *IEEE Transactions on Geoscience and Remote Sensing* 51 (4), 2417-2430(2013).
 5. G. Jie, H. Wang, J. Fan, X. Ma.: Change detection of sar images based on supervised contractive autoencoders and fuzzy clustering. In: *International Workshop on Remote Sensing with Intelligent Processing*, pp. 1-3. IEEE, Shanghai (2017).
 6. M. Gong, J. Zhao, J. Liu, Q. Miao, L. Jiao.: Change detection in synthetic aperture radar images based on deep neural networks. *IEEE Transactions on Neural Networks and Learning Systems* 27 (1), 125-138 (2017).
 7. H. Chen, L. Jiao, M. Liang, F. Liu, S. Yang, B. Hou.: Fast unsupervised deep fusion network for change detection of multitemporal sar images. *Neurocomputing* 332, 56-70 (2019).
 8. D. P. Kingma, M. Welling.: Auto-encoding variational bayes. *arXiv.org*, (2014).
 9. X. Fu, Y. Wei, F. Xu, T. Wang, Y. Lu, J. Li, J. Z. Huang.: Semi-supervised aspect-level sentiment classification model based on variational autoencoder. *Knowledge Based Systems* 171, (2019).
 10. L. v. d. Maaten, G. Hinton.: Visualizing data using t-sne. *Journal of machine learning research* 9 (11), 2579-2605 (2008).