World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:5, 2022

A Deep Learning Framework for Polarimetric SAR Change Detection Using Capsule Network

Sanae Attioui, Said Najah

***Abstract*—**The Earth's surface is constantly changing through forces of nature and human activities. Reliable, accurate, and timely change detection is critical to environmental monitoring, resource management, and planning activities. Recently, interest in deep learning algorithms, especially convolutional neural networks, has increased in the field of image change detection due to their powerful ability to extract multi-level image features automatically. However, these networks are prone to drawbacks that limit their applications, which reside in their inability to capture spatial relationships between image instances, as this necessitates a large amount of training data. As an alternative, Capsule Network has been proposed to overcome these shortcomings. Although its effectiveness in remote sensing image analysis has been experimentally verified, its application in change detection tasks remains very sparse. Motivated by its greater robustness towards improved hierarchical object representation, this study aims to apply a capsule network for PolSAR image Change Detection. The experimental results demonstrate that the proposed change detection method can yield a significantly higher detection rate compared to methods based on convolutional neural networks.

***Keywords*—**Change detection, capsule network, deep network, Convolutional Neural Networks, polarimetric synthetic aperture radar images, PolSAR images.

I. Introduction

DUE to its ability to map independently all-time weather conditions and its capacity to provide high spatial image resolution, Synthetic Aperture Radar (SAR) holds a strong potential in remote sensing applications, such as urban studies [1] agricultural development [1], disaster evaluation [2], forest monitoring [3] and so on.

As a core task for SAR image interpretation, change detection has drawn increasing interest in remote sensing communities, focusing on identifying changed areas that resulted from the evolution of the earth's surface, through multi-temporal images acquired for the same geographical area at distinct observation dates.

In recent years, SAR images have received increasing attention and the interest in understanding them from a multidimensional and multilook perspective has likewise amplified significantly. Different from Single polarization SAR data, PolSAR data, multi-channel and multi-parameter imagery, obtained by transmitting and receiving electromagnetic waves in different polarimetric states, can provide enhanced physical and geometric information about the target by analyzing the behaviors of backscattered polarized electromagnetic waves, which leads to an improvement in the quality of the change detection map.

Various procedures have been proposed in change detection, such as hypothesis test [4], fuzzy clustering [5] statistic modeling [6]. Most of the previous procedures rely on handcrafted entity representations, require a lot of specialized knowledge and complex computation, and eventuate in a limited representational ability to describe complex and high-level change information. These methods cannot, however, meet the current needs and expectations of dynamic environmental monitoring.

Accurately and quickly extracting relevant information from massive datasets for the change detection task has become an urgent problem, especially with the rapid advances in remote sensing techniques where many new problems and challenges keep emerging.

The change detection methods can be performed in two different ways: supervised or unsupervised. Although supervised approaches perform better than unsupervised approaches, it is too difficult and time consuming to obtain a major volume of high-quality labeled samples and to ensure the availability of an appropriate multi-temporal ground truth. Consequently, the unsupervised methods are preferable over the supervised ones.

Unsupervised SAR image change detection techniques are usually based on three steps: (1) image preprocessing includes multi-looking, co-registration of images and noise diminution; 2) generating a difference image (DI) the use of various kinds of operators. In terms of simplicity and efficiency, image difference and image ratio are the most common techniques. However, these do not make full use of spatial information, and so the results can be noisy in some cases. A DI can also be achieved by using similarity measurements. The strength of these lies in the fact that the estimation of the change indicator considers the pixel and its neighborhood. 3) classification of the DI. It is critical step, as it aims to generate the final change map, whereby change detection is converted into a binary classification task where pixels are generally classified into changed and unchanged areas.

Thresholding methods and clustering methods have been commonly applied; however, these methods are likely to be affected by speckle noise, and the contextual information round each pixel is not considered. The critical step in the change detection algorithm is to develop a powerful classification model.

Deep learning has proven to be an insightful new tool that could be the next trend in the change detection problem, attracting significant interest from the global Earth science and

Sanae Attioui and Said Najah (Prof.) are members of the Laboratory of Intelligent Systems and Application (LSIA), Faculty of Sciences and Technology, University Sidi Mohamed Ben Abdellah, Fes, Morocco (e-mail: sanae.attioui@usmba.ac.ma, said.najah@usmba.ac.ma).

remote sensing communities. Deep architectures, including convolutional neural networks (CNN), deep belief networks (DBN), and deep stacking networks, have all significantly improved the performance of many visual tasks, such as face recognition [7], image processing [8], [9], speech recognition [10], [11], change detection [12],[13].

Although CNNs have shown their superiority in the change detection field due to their powerful feature learning capabilities, there are still some shortcomings that limit their use. The most common deficiency is that CNNs typically require a large amount of training data with ground-truth labels, due to its difficulty to recognize the posture transformation of the image. In addition, during the network optimization, the influence of unbalanced samples can also lead to either underfitting or overfitting of CNN model training. CNNs architecture harnesses pooling layers to reduce the dimensionality of the representation, which could result in a loss lost some spatial information. In the CNNs structure, a single active neuron can only represent one entity, which greatly limits its capacity to represent multiple attributes of the same entity, knowing that the same land covers class can have more than one scattering mechanism.

More recently, a new architecture, called Capsule networks, has been proposed to break with the shortcomings of traditional deep learning networks. The capsule network (CapsNet) has shown a great performance in many image applications, and it has yielded more competitive results than CNNs, such as remote sensing image analysis [14], [15] and target recognition [16]. The CapsNet can contain more information about objects, which is advantageous for representing multiple scattering mechanisms of land covers, as a capsule is defined as a vector composed of a group of neurons to characterize the polarimetric attributes.

As far as we know, it has not been found that researchers have used CapsNet in the change detection of PolSAR images.

In this paper, we will investigate the potential ability of CapsNet in the change detection of PolSAR images, a CapsNet structure will be proposed. The rest of this paper is organized in 2 Sections. Section II provides details on the methodology followed. The case study and discussion are also processed in this section. The conclusion is presented in Section III.

II. Methodology

A. Representation of PolSAR images

The fundamental quantity measured by a polarimetric radar at each pixel and at each frequency is the scattering matrix [s] [17]. [s] represents all the information of the target scattering characteristics and can be expressed as:

$$S = \begin{bmatrix} s_{hh} & s_{hv} \\ s_{vh} & s_{vv} \end{bmatrix} \quad (1)$$

The scattering matrix [S] contains all the polarimetric information, with entries $s_{hh}, s_{hv}, s_{vh}, s_{vv}$ which designate the corresponding components of the backscattered electric fields. The first subscript refers to the polarization of the transmit antenna, and the second one to that of the receive antenna. The other elements have similar definitions.

For a monostatic SAR system, in which the same antenna performs transmission and reception, the off-diagonal terms of [S] are equal, $s_{hv} = s_{vh}$.

After multi-look processing, the covariance matrix [C] of the target can be represented as follows [18]:

$$C = \langle k_L \, . \, {k_L}^{*t} \rangle \quad (2)$$

$$k_L = \left[S_{XX}, \sqrt{2}\, S_{XY}, S_{YY} \right]^T \quad (3)$$

B. Capsule Network

CapsNets, a capsule structure-based neural network, is designed to overcome the loss of information that is seen in pooling operations and proposes the concept of a capsule. A capsule is a vector made up of a group of neurons each neuron of which represents various features of a specific entity in the images, e.g., position, size, orientation, hue, texture, etc., commonly called instantiation parameters. The length of the capsule can be interpreted as the probability that a particular entity being exists.

The structure of a simple CapsNet contains three types of layers: the convolutional layer, the primary capsule layer and the classification capsule layer.

By multiplying a spatial transformation matrix W, the vector U_i of each the low-level feature capsule i will generate a prediction vector of the high-level feature, also called a vote, i.e.,

$$\hat{\mathrm{U}}_{j/i} = W_{ij} U_i \quad (4)$$

Prediction vector $\hat{\mathrm{U}}_{j/i}$ is multiplied by the corresponding coupling coefficient C_{ij}, which is updated by the dynamic routing process, and passed to the capsule unit of the advanced capsule layer. Assuming that there are m capsules in the high-level capsule layer, the calculation method of the coupling coefficient C_{ij} between the j-th capsule of this level and the i-th capsule of the low-level capsule layer is as follows:

$$C_{ij} = \frac{\exp(b_{ij})}{\sum_{j=1}^{m} \exp\,(b_{ij})} \quad (5)$$

where b_{ij} is the log probability and it indicates whether lower level capsule i should be coupled with with higher level capsule j, at the beginning of the routing by agreement process, its set to be 0 for initialization, and later updated. All votes from different child capsules are then clustered to obtain the input vector of each parent capsule, i.e.,

$$S_j = \sum_i C_{ij} \hat{\mathrm{U}}_{j/i} \quad (6)$$

The output vector in the CapsNet is obtained using a nonlinear squashing function:

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|} \quad (7)$$

The Squash function also has the function of making the length of the vector no more than 1, and keeping v_j and S_j in the same direction.

The network needs to be improved by the margin loss function, which is defined as follows:

$$L = T_k \max(0, m^+ - \|v_0\|)^2 + \lambda(1 - T_k)\max(0, \|v_0\| - m^-)^2 \quad (8)$$

m^+ and m^- denote maximum margin and minimum margin respectively. T_k represents whether the k appears, T_k= 1if only the label k is presented (k = 0 means the unchanged class, and k = 1 means the changed class). λ = 0.5 is down-weighting factor is used to guarantee the final convergence.

C. Overview of the Proposed Method

Given two polarimetric SAR images acquired at different times t_1 and t_2 over the same geographical area; our goal is, then, to generate a binary change map: I (i, j) ∈ {0,1}, which represents the changed information, where 1 indicates that the position (i,j) is changed, and 0 indicates that the position (i,j) is unchanged.

The proposed method works in an unsupervised manner which involves the following steps:

Step1. Image preprocessing.

Step2. A preclassification step based on the similarity measurement and generalized histogram thresholding. The step aims to divide DI into changed and unchanged classes.

Step3. Training CapsNet Model: the training samples are randomly selected from DI.

Step4. Classification of changed and unchanged pixels: pixels in the DI are classified by the trained CapsNet to achieve the final change map.

The first step is significantly important as it mainly includes multi-looking, coregistration of images and speckle filtering. As several methods for speckle filtering are available, we opted for the refined Lee filter with a kernel size of 5, that can preserve edge details in low as well as in high contrast. Furthermore, geometric correction is applied for the co-registration of images for comparative purposes and matching.

The next step is generating a DI. To obtain DI, the similarity between two pixels i and j should be measured. Thanks to its high sensitivity to different types of changes and its computational simplicity, the symmetric revised Wishart (SRW) distance is exploit as the similarity measurement, more detail of the SRW distance can be found in [19].

For two pixels i and j, which can be characterized by covariance matrix C_i and C_j, the SRW distance is like as:

$$d_{SRW}(C_i, C_j) = \frac{1}{2}\left(tr\left(C_i^{-1}C_j + C_j^{-1}C_i\right)\right) - p \quad (9)$$

After obtaining the DI, considering the efficiency the Generalized Histogram Thresholding (GHT) technique for histogram-based image thresholding [20], the threshold method is utilized to partition pixels in DI into two groups: the changed class Ω_c and unchanged class Ω_u. where the pixels belonging to the two classes have a high probability of being modified or unchanged.

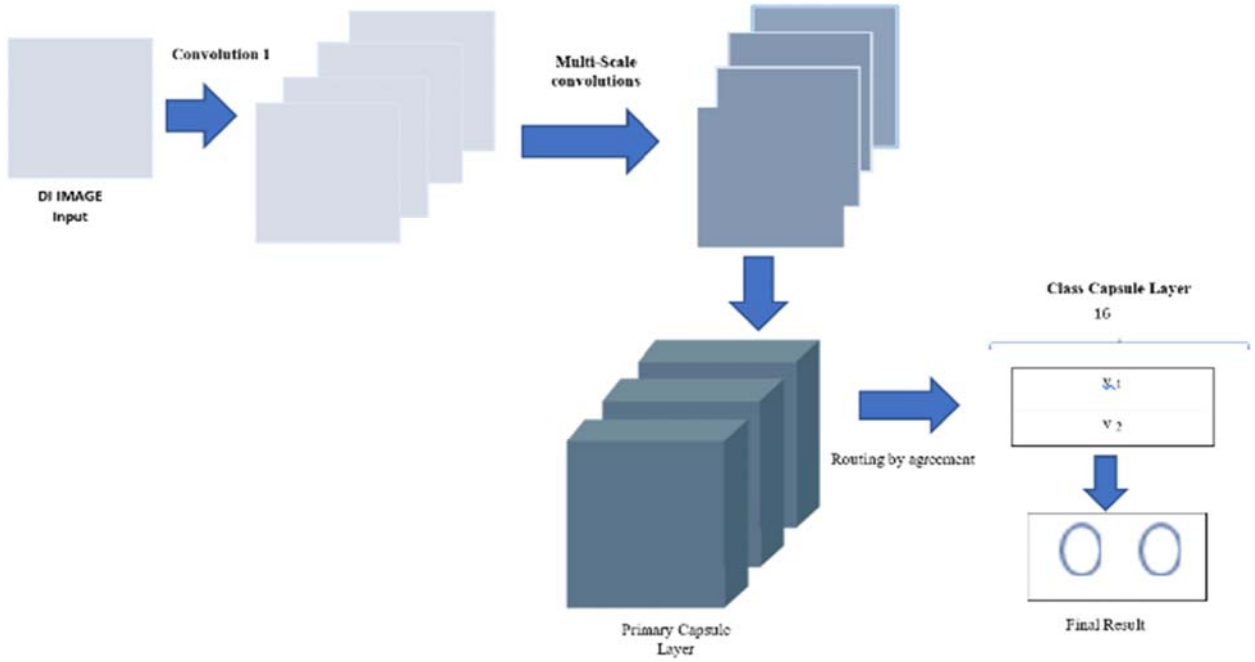


Fig. 1 The proposed change detection method

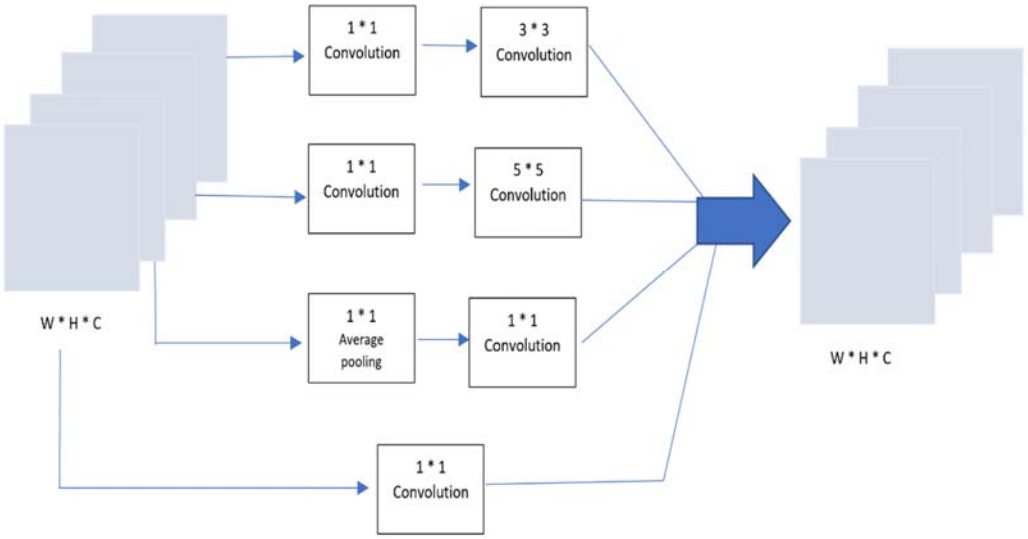
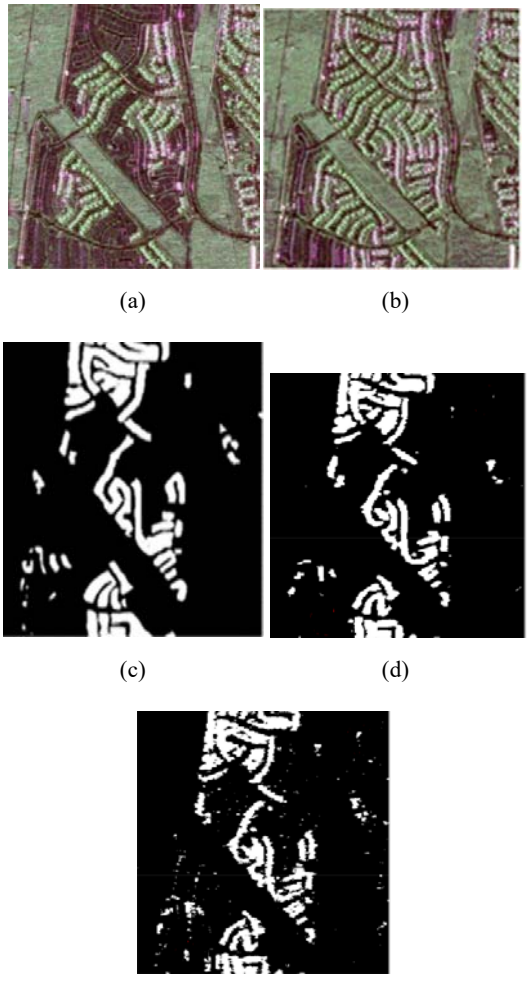


Fig. 2 Multi-Scale feature extraction module



(a) (b) (c) (d) (e)

Fig. 3 Change detection results

After preclassification, neighborhood features of sample pixels (selected randomly) belonging to two kinds of classes are generated and CapsNet classifier is trained based on these features.

The change detection performance of CapsNet is directly determined by the network structure. In this study, the proposed CapsNet contains a multi-scale feature extraction module before the primary capsule layer which is designed to guarantees the extraction of richer features. The model includes 1 * 1 convolution, two different size convolution kernels, and average pooling module. The purpose of using 1 * 1 to reduce dimensionality, decreasing the number of feature maps while retaining their salient features.

Finally, by using the trained CapsNet the pixels in the DI are classified so as to acquire the final change map.

D.Experimental Results

To assess the effectiveness of the proposed method, we use a set of two co-registered L-band UAVSAR full polarimetric images belonging to the city of Los Angeles, California, by the Jet Propulsion Laboratory/National Aeronautics and Space Administration UAVSAR (1.26 GHz), which were taken on 23 April 2009 (Fig. 3 (a)) and 11 May 2015 (Fig. 3 (b)), respectively.

Figs. 3 (a) and (b) show the Pauli decomposition images of these two PolSAR images (Red: |HH–VV|; Green: 2|HV|; Blue: |HH+VV|). The ground-truth map is represented in Fig. 3 (c) and includes nochange test pixels with white and change test pixels with gray.

In order to evaluate the change detection algorithm, a ground truth change image was compared to the final change detection map. False positives (FPs), false negatives (FNs), percentage correct classification (PCC), overall errors (OEs), and Kappa coefficient (KC) have been adopted as evaluation criteria.

TABLE I
CHANGE DETECTION RESULTS OF LOS ANGELES DATA SET

Method	FN	FP	OE	PCC (%)	KC (%)
CNN	1645	1768	3413	93.36	81.28
Proposed Method	**800**	**1157**	**1957**	**97,08**	**88,90**

The results are shown in Fig. 3 and listed in Table I. As seen, Fig. 3 (e) shows the change map generated by CNN that has many white spots which are listed as high values of FP in Table I. From a visual point of view, we can see that the final change map generated by our method is much closer to the ground-truth image than the CNN processes.

The proposed method is better and reveals more expressive results.

III. Conclusion

CapsNet based models have been recently discussed to solve the challenges of PolSAR image change detection (i.e., manual feature extraction methods and scarce and very small open labeled data sets). In this work, we have presented a multi-scale feature extraction module into the CapsNet which ensures that richer features are transmitted into the capsule. Compared with CNN methods, the proposed method exhibits good performance. In the future, we are hoping to improve the performance of the network and further apply to the PolSAR image datasets.

References

[1] M. K. Ridd and J. Liu, "A comparison of four algorithms for change detection in an urban environment," Remote Sens. Environ., vol. 63, no. 2, pp. 95–100, 1998.

[2] B. Brisco, A. Schmitt, K. Murnaghan, S. Kaya, and A. Roth, "SAR polarimetric change detection for flooded vegetation," Int. J. Digit. Earth, vol. 6, no. 2, pp. 103–114, 2013.

[3] T. Hame, I. Heiler, and J. S. Miguel-Ayanz, "An unsupervised change detection and recognition system for forestry," Int. J. Remote Sens., vol. 19, no. 6, pp. 1079–1099, 1998.

[4] G. Liu, L. Jiao, F. Liu, H. Zhong, and S. Wang, "A new patch-based change detector for polarimetric SAR data," Pattern Recognit., vol. 48, no. 3, pp. 685–695, 2015.

[5] A.Ghosh, N. S. Mishra, and S. Ghosh, "Fuzzy clustering algorithms for unsupervised change detection in remote sensing images," Inf. Sci., vol. 181, no. 4, pp. 699–715, 2011.

[6] M. Liu, H. Zhang, C. Wang, and F. Wu, "Change detection of multilook polarimetric SAR images using heterogeneous clutter models," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 12, pp. 7483–7494, Dec. 2014

[7] C. Ding and D. Tao, "Robust face recognition via multimodal deep face representation," IEEE Trans. Multimedia, vol. 17, no. 11, pp. 2049–2058, Nov. 2015.

[8] J. Li, H. Chang, and J. Yang. "Sparse deep stacking network for image classification." 2015

[9] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015.

[10] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," IEEE Trans. Audio Speech Lang. Process., vol. 20, no. 1, pp. 30–42, Jan. 2012.

[11] L. Deng, D. Yu, and J. Platt, "Scalable stacking and learning for building deep architectures," in Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Mar. 2012, pp. 2133–2136.

[12] Q. Liu, R. Hang, H. Song, and Z. Li, "Learning multiscale deep features for high-resolution satellite image scene classification," IEEE Trans. Geosci. Remote Sens., vol. 56, no. 1, pp. 117–126, Jan. 2018.

[13] X. Li, Z. Yuan, and Q. Wang, "Unsupervised deep noise modeling for hyperspectral image change detection," Remote Sens., vol. 11, no. 3, p. 258, Jan. 2019.

[14] M. E. Paoletti et al., "Capsule networks for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 57, no. 4, pp. 2145–2160, Apr. 2019.

[15] K. Zhu, Y. Chen, P. Ghamisi, X. Jia, and J. A. Benediktsson, "Deep convolutional capsule network for hyperspectral image spectral and spectral-spatial classification," Remote Sens., vol. 11, no. 3, p. 223, Jan. 2019.

[16] Guo, Y.; Pan, Z.; Wang, M.; Wang, J.; Yang, W. Learning Capsules for SAR Target Recognition. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 4663–4673.

[17] Lee, J.S.; Pottier, E. Polarimetric Radar Imaging: From Basics to Applications; CRS Press: Boca Raton, FL, USA, 2009.

[18] Ziegler, V.; Lüneburg, E.; Schroth, A. Mean backscattering properties of random radar targets-A polarimetric covariance matrix concept. In Proceedings of the IGARSS'92; Proceedings of the 12th Annual Int Geo and Remote Sensing Symposium, Houston, TX, USA, 26–29 May 1992; Volume 1, pp. 266–268.

[19] N. Anfinisen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric SAR data with Wishart-derived distance measures," in Proc. POLInSAR 2007, Esrin

[20] Barron, J.T.: A generalization of otsu's method and minimum error thresholding. ECCV (2020).