SAR Image Change Detection Based on Multiscale Capsule Network

Yunhao Gao, Feng Gao, Junyu Dong, and Heng-Chao Li

***Abstract*— Traditional synthetic-aperture radar (SAR) image change detection methods based on convolutional neural networks (CNNs) face the challenges of speckle noise and deformation sensitivity. To mitigate these issues, we proposed a multiscale capsule network (Ms-CapsNet) to extract the discriminative information between the changed and unchanged pixels. On the one hand, the multiscale capsule module is employed to exploit the spatial relationship of features. Therefore, equivariant properties can be achieved by aggregating the features from different positions. On the other hand, an adaptive fusion convolution (AFC) module is designed for the proposed Ms-CapsNet. The higher semantic features can be captured for the primary capsules. Feature extracted by the AFC module significantly improves the robustness to speckle noise. The effectiveness of the proposed Ms-CapsNet is verified on three real SAR data sets. The comparison experiments with four state-of-the-art methods demonstrate the efficiency of the proposed method. Our codes are available at *https://github.com/summitgao/SAR_CD_MS_CapsNet*.**

***Index Terms*— Change detection, deep learning, multiscale capsule network (Ms-CapsNet), synthetic-aperture radar (SAR).**

I. Introduction

SYNTHETIC-APERTURE radar (SAR) imaging acquisition technologies have been developed rapidly. A plenty of multitemporal SAR images are available to monitor the changed information of the Earth. Therefore, SAR image change detection has drawn an increasing attention recently. Researchers have designed a variety of SAR change detection methods for ecological surveillance, disaster monitoring [1], and urban planning [2].

Although a plenty of techniques have been proposed [3], the SAR image change detection is a still challenging task. An image quality is deteriorated by speckle noise which hinders the meticulous interpretation of SAR data. Many methods are implemented to address the issue of speckle noise. They are usually comprised of three steps: 1) image coregistration; 2) difference image (DI) generation; and 3) DI classification [4]. Image coregistration is a fundamental task to establish the spatial correspondences between multitemporal SAR images. In the second step, the DI is commonly generated by the log-ratio, Gauss-ratio [5], and neighborhood-ratio [6] operators. For the DI classification step, more research works are devoted to build a robust classifier. It is a nontrivial task since a powerful classifier directly determines the precision of change detection.

Many researchers are dedicated to developing powerful classifiers for change detection. Li *et al.* [7] designed two-level clustering algorithm for unsupervised change detection. In [8], local neighborhood information is embedded in the clustering objective function to improve the change detection performance. Gong *et al.* [9] developed an improved Markov random field (MRF) based on fuzzy *c*-means (FCM) clustering to suppress the speckle noise. In [4], the stacked restricted Boltzmann machines (RBMs) were employed for SAR image change detection. Although the above methods achieved a promising performance, the feature representation capabilities are still limited.

In recent years, the convolutional neural network (CNN) has greatly boosted the performance of many visual tasks. It is demonstrated to be rather effective for robust feature learning. CNN-based models have been successfully applied in remote sensing image change detection [10]. Wang *et al.* [11] proposed an end-to-end CNN framework to learn the discriminative features from mixed-affinity matrix for change detection. Later, the unsupervised deep noise modeling was developed for hyperspectral image change detection [12]. Liu *et al.* [13] proposed an elegant local restricted CNN (LR-CNN) for polarimetric SAR change detection. In [14], transferred deep learning was applied to sea ice SAR image change detection based on CNN. Although CNN-based methods have achieved an excellent performance in change detection, the accuracy sometimes deteriorates under the case of transformation, such as tilts and rotations. Specifically, CNN is incapable of modeling the positional relationship among ground objects.

More recently, Sabour and Hinton proposed the capsule network (CapsNet) to provide solutions to problems where CNN models are inadequate [15]. In CapsNet, an activity vector from capsules represents the entity instantiation parameters such as pose, texture, and deformation. The existence of entities is expressed by the length of instantiation parameters. The dynamic routing mechanism is utilized for information propagation. It is empirically verified that the CapsNet is

Manuscript received November 22, 2019; revised February 10, 2020; accepted February 23, 2020. This work was supported in part by the Next Generation AI Project of China under Grant 2018AAA0100602 and in part by the National Natural Science Foundation of China under Grant U1706218 and Grant 41927805. *(Corresponding author: Feng Gao.)*

Yunhao Gao, Feng Gao, and Junyu Dong are with the Qingdao Key Laboratory of Mixed Reality and Virtual Ocean, School of Information Science and Engineering, Ocean University of China, Qingdao 266100, China (e-mail: gaofeng@ouc.edu.cn).

Heng-Chao Li is with the Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2020.2977838

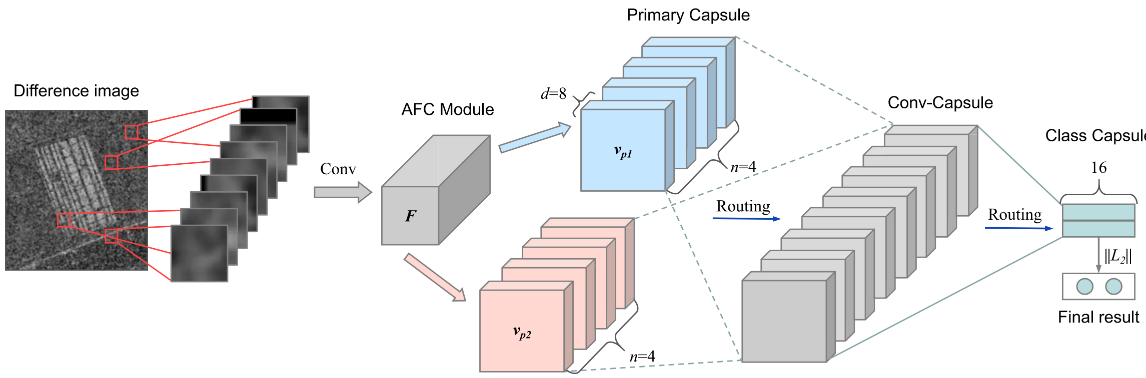

Fig. 1. Illustration of the proposed change detection method based on Ms-CapsNet. First, image patches are fed into AFC module for higher semantic features. Then, multiscale primary capsule layers are adopted with kernel 3×3 and 5×5 to obtain the primary capsules υ_{p1} and υ_{p2}. Later, υ_{p1} and υ_{p2} are input to conv-capsule layer and class capsule layer, respectively. Finally, fuse the output of class capsule layers to calculate the final result.

effective for remote sensing image analysis [16], [17]. As far as we know, the literature on the CapsNet-based SAR change detection is very sparse.

We argue that the weakness of the existing SAR image change detection approaches mainly comes from two aspects: One is the correlation of features from different positions, which fails to be modeled effectively and the other one lies in the intrinsic speckle noise in SAR images. To tackle the aforementioned issues, a multiscale CapsNet (Ms-CapsNet) is proposed to extract the discriminative information between multitemporal SAR images. The proposed Ms-CapsNet has a similar structure with the CapsNet [15] without the multiscale operator and the adaptive fusion convolution (AFC) module. The Ms-CapsNet provides a group of instantiation parameters to capture the features from different positions. To tackle the problem of speckle noise, the AFC module is designed to convert the pixel intensities to the activities of the local features. Accordingly, local features become noise robustness. Extensive experiments on three real data sets are conducted to show the superiority of the proposed method over four state-of-the-art works.

For clarity, the main contributions are summarized as follows.

1) The proposed Ms-CapsNet has the capability to extract the robust features from different positions. The equivariant properties can be achieved by the capsule module. Therefore, the demand for a large amount of training samples is reduced by the correlative and complete information.
2) A simple yet effective AFC module is designed, which can effectively convert pixel intensities to the activities of local features. The AFC module extracts the higher semantic features and emphasizes the meaningful one through attention-based strategy. Therefore, the activity of local features becomes more noise robustness and immediately accepted as the input of the primary capsule.
3) Extensive experiments have been implemented on three SAR data sets to validate the effectiveness of the proposed method. Moreover, we have released the codes and setting to facilitate future research in multitemporal remote sensing image analysis.

II. Methodology

The proposed method is shown in Fig. 1. A DI is first generated by the log-ratio operator. Then, the training samples are selected randomly from DI for Ms-CapsNet. Finally, pixels in the DI are classified by the trained Ms-CapsNet to obtain the final change map.

In our implementations, the Ms-CapsNet is comprised of AFC and capsule modules. The AFC module is used to convert pixel intensities to high semantic features, through which the speckle noise is suppressed to some extent. The capsule module is utilized to activate high semantic features. In Sections II-A and II-B, we will describe both the modules in detail.

A. AFC Module

As shown in Fig. 2, the proposed AFC module is utilized to encode the input. Recently, some studies suggest that the long-range feature dependencies can be captured by the self-attention mechanism. Hu *et al.* [18] demonstrated it in large-scale image recognition task. In this letter, we introduce the self-attention mechanism into the SAR image change detection task and designed a simple yet effective AFC module. First, a set of convolutions (Conv 1–1, Conv 1–2, and Conv 1–3 with kernel size 3×3) are employed with different dilation rates, which are set to 1, 2, and 3, respectively, to capture multiscale features. Then, the multiscale features are aggregated by feature fusion based on channel-wise attention (CA).

The input features $\mathbf{F}_{\text{in}} \in \mathbb{R}^{w_0 \times w_0 \times c_0}$ from atrous convolution are fed into CA. Then, global average pooling (GAP) squeezes $\mathbf{F}_{\text{in}}$ in the spatial domain to obtain $\mathbf{F}_{\text{avg}} \in \mathbb{R}^{1 \times 1 \times c_0}$. Then, 1-D convolution (1D-Conv) is employed to explore the channel relationship of $\mathbf{F}_{\text{avg}}$. After the $Sigmoid$ function, a channel weighting-based vector M can be obtained. Finally, the channel weighting-based feature $\mathbf{F}_{\text{out}}$ can be computed as $\mathbf{F}_{\text{out}} = M \otimes F_{\text{in}}$, where $\otimes$ denotes channel-wise multiplication. Therefore, the channel weighting-based features from Conv 1–1, Conv 1–2, and Conv 1–3 are $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{F}_3$, respectively. We fused the features by pixel-wise summation as

$$\mathbf{F} = D_1(\mathbf{F}_1) + D_2(\mathbf{F}_2) + D_3(\mathbf{F}_3) \tag{1}$$

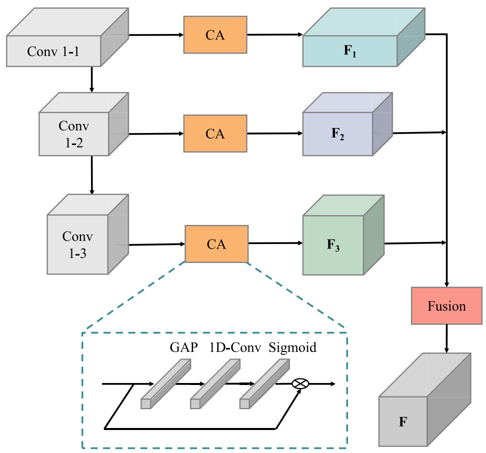

Fig. 2. Illustration of the AFC module.

where $\mathbf{F}$ represents the fused features, and D_1, D_2, and D_3 are dimension matching functions which are operated by 1×1 convolution.

B. Capsule Module

The capsule module is a neural network comprised of the primary capsule layer, conv-capsule layer, and fully connected layer, as shown in Fig. 1.

1) Primary Capsule Layer: This layer is employed to extract the low-level features from multidimensional entities through convolutional-like operation with kernel size $k \times k$. Different from traditional convolution, multiple feature maps will be obtained instead of one. The primary capsule layer first receives the feature map $F \in \mathbb{R}^{w \times w \times c}$ from the AFC module. Then, convolutional-like operation and squashing activation function are employed to obtain the output capsules v_p. The squashing activity function is denoted as

$$v = \frac{\|s\|^2}{1 + \|s\|^2} \frac{s}{\|s\|} \tag{2}$$

where s is the total input, and v is the vector output of capsule. In the primary capsule layer, the size of the output capsules v_p is $w_1 \times w_1 \times n \times d$, where n is the number of feature maps, $n \times d = c$, and $d = 8$. The $[w_1 \times w_1]$ grid is the shared weights. In other words, we obtain $[w_1 \times w_1 \times n]$ 8D vectors in total primary capsules. In our implementations, multiscale information is taken into account. Two primary capsule layers are employed with kernel size $k = 3$ and $k = 5$, respectively. Therefore, multiscale feature representation can be obtained. Feature vectors from two scales are denoted by v_{p1} and v_p, respectively.

2) Conv-Capsule Layer: This layer uses local connections and the shared transformation matrix to reduce the number of parameters to some extent [17]. Conv-capsule layer uses the dynamic routing strategy to update the coupling coefficient c. The connection (transformation matrix) between the primary capsule layer and the conv-capsule layer is W, and the transformation matrix W is also shared in each grid. Therefore, the output v_c of the conv-capsule layer can be expressed as

$$v_c = \text{squashing}\left(\sum c \cdot u\right) \tag{3}$$

where c is the coupling coefficient, and $u = W \cdot v_p$. v_p is the output of the primary capsule layer. For dynamic routing, we first set the agreement b to 0. The coupling coefficient c can be calculated by $c = softmax(b)$. That is to say, we update b to calculate the latest coupling coefficient c. In addition, the update process of b can be expressed as $b \leftarrow b + u \cdot v_c$. The detailed descriptions of the dynamic routing can be found in [15].

3) Class Capsule Layer: The class capsule layer can be considered as a fully connected layer. Dynamic routing mechanism is still used for coupling coefficient updating. In this layer, multiscale activity vectors $v_{o1} \in \mathbb{R}^{2\times 16}$ and $v_{o2} \in \mathbb{R}^{2\times 16}$ from class capsule layer are fused by summation $v_o = v_{o1} \oplus v_{o2}$. Then, the vector norm is calculated to measure the probability of classes. The loss function of Ms-CapsNet can be defined as

$$\begin{aligned} L = T_k \max(0, m^+ - \|v_o\|)^2 \\ + \lambda(1 - T_k) \max(0, \|v_o\| - m^-)^2 \end{aligned} \tag{4}$$

where $T_k = 1$ when the label k is presented ($k = 0$ means the unchanged class, and $k = 1$ means the changed class). $\lambda = 0.5$ is used to constrain the length of the activity vector of the initial class capsule. If there is a changed class object in the image, the class capsule of the changed class should output a vector with a length of at least $m^+ = 0.9$. On the contrary, if there is no object of the changed class, a vector with a length less than $m^- = 0.1$ will be output from the class capsule. Then, the final change map can be calculated by pixel-wise classification.

III. Experimental Results and Analysis

In this section, we first describe the data sets and evaluation criteria in our experiments. Next, an exhaustive investigation of several vital parameters on the change detection performance is presented. Finally, we conduct extensive experiments to verify the effectiveness of the proposed method.

A. Data Set and Evaluation Criteria

To verify the effectiveness of the proposed method, we employed Ms-CapsNet on three multitemporal SAR data sets acquired by different sensors. The first data set is the Sulzberger data set. It is captured at Sulzberger Ice Shelf by ENVISAT satellite of the European Space Agency on March 11 and 16, 2011, respectively. The size of the data set is 256×256 pixels, as shown in the first row of Fig. 3(a)–(c). The other two data sets named Yellow River I and Yellow River II are captured at the Yellow River Estuary by Radarsat-2 in June 2008 and June 2009, respectively. Their sizes are 257×289 and 306×291 pixels, respectively. It is very challenging to perform the change detection on the Yellow River data set since the speckle noise is much stronger. The geometric corrections have been performed on these data sets, and the ground truth images were manually annotated carefully with expert knowledge.

In the following experiments, the proposed Ms-CapsNet is compared with four closely related methods, including the PCA-based neural networks (PCANet) [19], the transferred

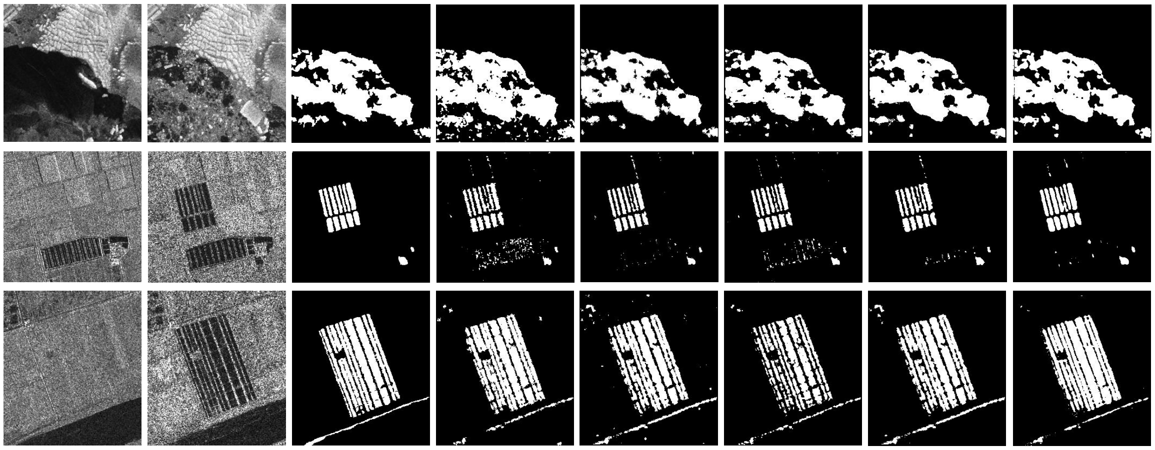

Fig. 3. Visualized results of different change detection methods on three data sets. (a) Image captured at t_1. (b) Image captured at t_2. (c) Ground truth image. (d) Result by PCANet. (e) Result by MLFN. (f) Result by DCNN. (g) Result by LR-CNN. (h) Result by the proposed Ms-CapsNet.

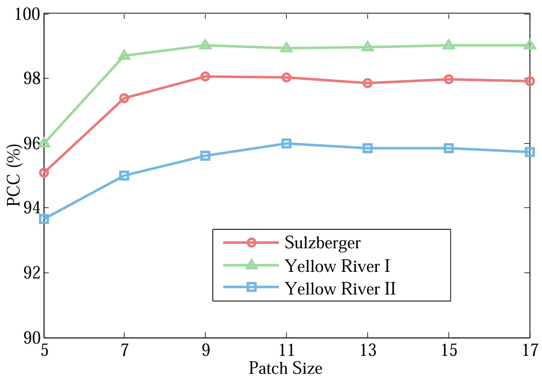

Fig. 4. Relationship between the PCC values and patch size.

multilevel fusion network (MLFN) [14], the deep CNNs (DCNN) [20], and the CNN with local spatial restrictions (LR-CNN) [13]. To verify the effectiveness of the proposed Ms-CapsNet, false positives (FPs), false negatives (FNs), percentage correct classification (PCC), overall errors (OEs), and Kappa coefficient (KC) are adopted as the evaluation criteria.

B. Parameter Analysis of the Proposed Ms-CapsNet

1) Analysis of the Patch Size: The patch size r represents the scale of spatial neighborhood information. Fig. 4 shows the relationship between r and PCC, where r is changed from 5 to 17. As shown in Fig. 4, the PCC values increase first and then tend to be stable. It is evident that the contextual information is important for change detection. However, a large patch size will increase the computational cost. Therefore, we choose $r = 9$ for the Sulzberger and Yellow River I data sets, and $r = 11$ for the Yellow River II data set.

2) Analysis of the Training Sample Numbers: Table I shows the comparison of the Ms-Capsule with other methods on the Yellow River II data set when considering a different number of training samples, i.e., 200, 400, 600, 800, and 1000. We can observe that the accuracy of other methods drops sharply when the number of samples is less than 600. Especially, DCNN and LR-CNN depend heavily on large volumes of training data, and few training samples will lead to overfitting which degrades the performance. In summary, the PCC values of the proposed method are less afflicted with the training sample numbers. It is because the feature spatial correlations can reduce the dependence on training samples to some extent.

TABLE I

RELATIONSHIP BETWEEN THE PCC VALUES AND THE NUMBER OF TRAINING SAMPLES

Method	PCC of different training samples number (%)				
	200	400	600	800	1000
PCANet [19]	91.44	92.12	93.34	93.88	94.26
MLFN [14]	94.20	94.72	95.01	95.24	95.48
DCNN [20]	92.51	92.58	92.66	93.07	93.81
LR-CNN [13]	93.81	94.24	94.68	95.00	95.32
Ms-CapsNet	94.54	94.98	95.28	95.65	96.00

TABLE II

ABLATION STUDIES (PCC) OF THE PROPOSED MS-CAPSNET

Method	Sulzberger	Yellow River I	Yellow River II
CapsNet	97.54	98.45	94.95
w/o AFC	97.58	98.68	95.03
w/o multiscale	97.98	98.91	95.51
Full model	98.16	99.02	96.00

3) Ablation Studies: We conduct experiments to compare the performance of several variants of our method for ablation studies. The qualitative results are shown in Table II. The full model represents the proposed Ms-CapsNet. CapsNet denotes the traditional CapsNet [15] without AFC module and multiscale operator. In addition, we implement our model without AFC module (w/o AFC) and without multiscale operator (w/o multiscale). It can be observed that both multiscale operator and the AFC module can boost the change detection performance. The PCC values improve 0.14, 0.23, and 0.08 by the multiscale operator on three data sets, respectively. This is because the multiscale operator is beneficial to enrich the feature representations. In addition, the PCC values improve 0.36, 0.46, and 0.46 by the AFC module on three data sets, respectively. It is evident that local features can be effectively converted for primary capsules.

TABLE III
CHANGE DETECTION RESULTS ON THREE DATA SETS

Method	Sulzberger dataset					Yellow River I dataset					Yellow River II dataset				
	FP	FN	OE	PCC(%)	KC(%)	FP	FN	OE	PCC(%)	KC(%)	FP	FN	OE	PCC(%)	KC(%)
PCANet [19]	1410	1437	2847	95.66	88.63	2942	493	3435	96.14	71.55	2435	1533	3968	94.66	82.43
MLFN [14]	616	664	1280	98.05	94.89	721	863	1584	98.22	83.82	1544	1972	3516	95.27	83.82
DCNN [20]	312	1467	1779	97.29	92.74	1468	712	2180	97.55	79.40	1231	2370	3601	95.15	83.08
LR-CNN [13]	1198	680	1878	97.13	92.58	1118	423	1541	98.27	85.36	1923	1460	3383	95.45	84.83
Ms-CapsNet	425	779	1204	98.16	95.16	468	407	875	99.02	91.22	1173	1798	2971	96.00	86.25

C. Change Detection Results on Three Data Sets

In this section, the proposed method is compared with four closely related methods. The quantitative results and visual results with all competitors are shown in Table III and Fig. 3, respectively.

Fig. 3(d)–(h) shows the change maps corresponding to the experiments as presented in Table III. On the Sulzberger data set (the first row of Fig. 3), the result of PCANet tends to be rather noisy, and it is afflicted with high FP value. Although other methods generally performed well, the results are deteriorated due to higher OE values. The proposed Ms-CapsNet exhibits less misclassified pixels and obtains the best PCC and KC values.

On the Yellow River I and II data sets (the second and third rows of Fig. 3), we can observe that the proposed Ms-CapsNet achieves at least 0.5% improvement over other compared methods. Considering that the interference of different characteristics of speckle noise weakens the model performance, the proposed method is relatively noise robust. The PCANet suffers from high FP value, and there are many noisy regions in the generated change maps. LR-CNN performs better since local spatial restrictions can balance the influence of local noise. CNN-based methods can suppress noise interference to some extent through deep feature representation. However, relatively high OE values are still obtained. In general, the proposed Ms-CapsNet exhibits the best performance as shown in Table III and Fig. 3. It reveals that the proposed Ms-CapsNet benefits from the spatial relation exploration.

IV. CONCLUSION

In this letter, the Ms-CapsNet is proposed for SAR image change detection. The Ms-CapsNet benefits from two aspects: first, to enhance the spatial feature correlations, a multiscale capsule module is utilized to model the spatial relationship of features between one object and another. Equivariant properties can be achieved by aggregating the feature from different positions. Furthermore, we design an AFC module to alleviate the interference of speckle noise. The module can effectively convert the pixel-wise intensities to the activity of local features. Extensive experiments are conducted on three SAR data sets, and the experimental results demonstrate the superior performance of the proposed method.

REFERENCES

[1] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, May 2010.

[2] S. Quan, B. Xiong, D. Xiang, L. Zhao, S. Zhang, and G. Kuang, "Eigenvalue-based urban area extraction using polarimetric SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 2, pp. 458–471, Feb. 2018.

[3] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[4] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, "Change detection in synthetic aperture radar images based on deep neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 125–138, Jan. 2016.

[5] B. Hou, Q. Wei, Y. Zheng, and S. Wang, "Unsupervised change detection in SAR image based on gauss-log ratio image fusion and compressed projection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 8, pp. 3297–3317, Aug. 2014.

[6] M. Gong, Y. Cao, and Q. Wu, "A neighborhood-based ratio approach for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 307–311, Mar. 2012.

[7] H.-C. Li, T. Celik, N. Longbotham, and W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal SAR images based on two-level clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2458–2462, Dec. 2015.

[8] L. Jia, M. Li, P. Zhang, Y. Wu, and H. Zhu, "SAR image change detection based on multiple kernel K-Means clustering with local-neighborhood information," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 6, pp. 856–860, Jun. 2016.

[9] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.

[10] 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.

[11] Q. Wang, Z. Yuan, Q. Du, and X. Li, "GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, Jan. 2019.

[12] 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.

[13] F. Liu, L. Jiao, X. Tang, S. Yang, W. Ma, and B. Hou, "Local restricted convolutional neural network for change detection in polarimetric SAR images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 818–833, Mar. 2019.

[14] Y. Gao, F. Gao, J. Dong, and S. Wang, "Transferred deep learning for sea ice change detection from synthetic-aperture radar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 10, pp. 1655–1659, Oct. 2019.

[15] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3859–3869.

[16] 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.

[17] 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.

[18] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 7132–7141.

[19] F. Gao, J. Dong, B. Li, and Q. Xu, "Automatic change detection in synthetic aperture radar images based on PCANet," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1792–1796, Dec. 2016.

[20] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3173–3184, Jun. 2018.