Change Detection From Synthetic Aperture Radar Images via Dual Path Denoising Network

Junjie Wang [ID], Feng Gao [ID], *Member, IEEE*, Junyu Dong [ID], *Member, IEEE*, Qian Du [ID], *Fellow, IEEE*, and Heng-Chao Li [ID], *Senior Member, IEEE*

Abstract—Benefited from the rapid and sustainable development of synthetic aperture radar (SAR) sensors, change detection from SAR images has received increasing attentions over the past few years. Existing unsupervised deep learning-based methods have made great efforts to exploit robust feature representations, but they consume much time to optimize parameters. Besides, these methods use clustering to obtain pseudolabels for training, and the pseudolabeled samples often involve errors, which can be considered as "label noise." To address these issues, we propose a dual path denoising network (DPDNet) for SAR image change detection. In particular, we introduce the random label propagation to clean the label noise involved in preclassification. We also propose the distinctive patch convolution for feature representation learning to reduce the time consumption. Specifically, the attention mechanism is used to select distinctive pixels in the feature maps, and patches around these pixels are selected as convolution kernels. Consequently, the DPDNet does not require a great number of training samples for parameter optimization, and its computational efficiency is greatly enhanced. Extensive experiments have been conducted on five SAR datasets to verify the proposed DPDNet. The experimental results demonstrate that our method outperforms several state-of-the-art methods in change detection results.

Index Terms—Change detection, dual path denoising network (DPDNet), label noise, synthetic aperture radar (SAR).

I. INTRODUCTION

OWING to the sustainable development of remote sensing imaging science and technology, plenty of synthetic aperture radar (SAR) sensors are designed for aerial and spaceborne platforms. The ability to collect images of the same locations at higher frequency allows near real-time monitoring of the earth,

Junjie Wang, Feng Gao, and Junyu Dong are with the School of Computer Science and Technology, Ocean University of China, Qingdao 266100, China (e-mail: 1091416535@qq.com; gaofeng@ouc.edu.cn; dongjunyu@ouc.edu.cn).

Qian Du is with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762 USA (e-mail: du@ece.msstate.edu).

Heng-Chao Li is with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China (e-mail: hcli@swjtu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3159619

and therefore, increasingly SAR images of the same geographical area are available. These images provide insights to many applications, such as change detection [1]–[4], natural disaster monitoring [5]–[7], urban planning [8]–[10], and land cover data updating [11]. Among these applications, change detection aims to recognize the changed information of the same region by analyzing mutlitemporal images. It has attracted widespread interest in recent years [12].

SAR change detection is commonly considered as a classification task. In a general pipeline of SAR change detection, the difference image (DI) is first created, and the DI is classified into changed or unchanged pixels in a supervised or unsupervised manner [13]. However, the existence of speckle noise is a nonnegligible problem in SAR change detection task. Speckle noise in SAR images appears as a form of multiplicative noise and degrades the image quality. False alarms are usually produced due to the existence of speckle noise [14], [15]. Therefore, the DI classification for SAR images is still a challenging task [16], [17]. So, it is critical to design robust change detection techniques which are effective in speckle noise suppression [18]–[20].

In the past few years, researchers have devoted great efforts toward solving or alleviating the effect of speckle noise. There are two types of change detection techniques: supervised and unsupervised methods. Supervised methods require labeled training samples to learn the parameters of a classifier [21], [22]. Compared with it, unsupervised methods are more popular since they compare two multitemporal SAR images without any prior knowledge or manually labeled samples [23]–[25]. In this article, we mainly focus on developing unsupervised SAR image change detection method.

Existing unsupervised change detection methods mainly focus on DI generation and classification. In DI generation, log-ratio [26], Gauss-ratio [27], or neighborhood-based ratio [28] operators are used. Besides, the coefficient of variation based on time series SAR images is used to avoid the speckle noise [29]. For DI classification, many clustering methods have been employed, such as fuzzy c-means (FCM) [30], k-means clustering [31], multiple kernel clustering [32], and mean-shift algorithms [33]. In addition, the machine learning methods based on Markov random fields are also applied to the despeckle task [23].

To further enhance the DI classification performance, researchers incorporate deep learning-based classifiers into the traditional unsupervised model. Gong *et al.* [34] first assigned pseudolabels to the pixels in DI by an FCM-based joint classifier.

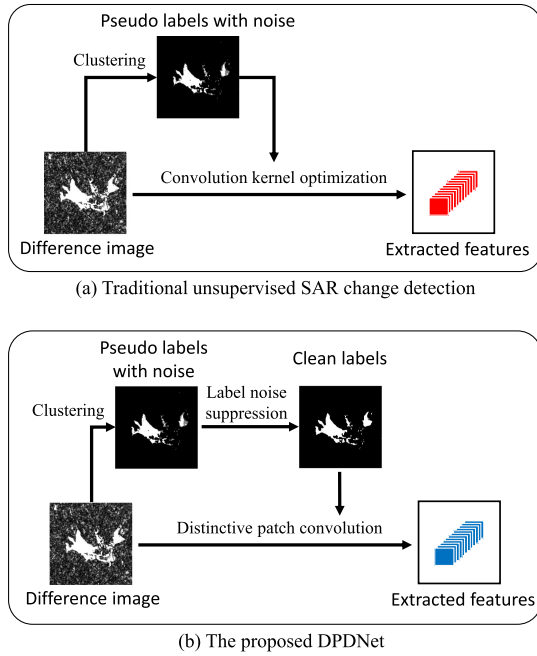(a) Traditional unsupervised SAR change detection

(b) The proposed DPDNet

Fig. 1. Comparisons of traditional methods with the proposed DPDNet. (a) General framework of existing unsupervised SAR image change detection method. These methods contain many noise labels and are commonly time-consuming in the training phase. (b) Proposed DPDNet. It generates clean labels and is more efficient in the training phase.

Then restricted Boltzmann machines (RBMs) were trained for change map generation. Hou *et al.* [35] presented a change detection method by combining saliency computation and low-rank algorithm. Zhan *et al.* [36] proposed a deep Siamese CNN model to extract discriminant features for change detection. Zhang *et al.* [37] presented a deep spatial-temporal gray-level co-occurrence CNN, which is capable of exploiting the spatial-temporal information of mutlitemporal images. Dong *et al.* [6] integrated unsupervised clustering with CNN to learn clustering-friendly feature representations from multitemporal SAR images. Qu *et al.* [38] proposed a dual-domain network. Features from both frequency and spatial domains are exploited to alleviate the speckle noise.

As illustrated in Fig. 1(a), existing deep learning-based unsupervised SAR change detection methods confront two challenges. 1) The obtained pseudolabels commonly involve errors. The existing methods generally use clustering to obtain pseudolabels, which contain some errors. This phenomenon is called "label noise," which can affect the subsequent network optimization. Accordingly, it is highly desired to devise a label noise tolerant method. 2) Deep learning-based methods are time-consuming in the training phase. To obtain satisfying results, the existing methods run many epochs during the training phase or utilize the pretraining strategy to obtain the desired features (Here, the pretraining time is also regarded as a part of time consumption). Both of them are time-consuming and require reliable training samples, which are quite limited in SAR change detection. Therefore, how to design a network that is less time-consuming and requires fewer training samples has become the key to the research.

To tackle the abovementioned problems, we propose a dual path denoising network (DPDNet) to alleviate the effect of speckle noise and label noise simultaneously [see Fig. 1(b)]. As illustrated in Fig. 2, one branch of DPDNet uses random label propagation to clean the label noise. Besides, the other branch of DPDNet is employed to extract shallow and deep features, and then combines them together for hierarchical feature representation. In this process, the designed distinctive patch convolution (DPConv) is utilized to simplify the network structure and reduce the training time. To validate the effectiveness of the proposed DPDNet, we conducted extensive experiments on five datasets. Moreover, we compared our method with six state-of-the-art models, and the experimental results demonstrate that our method can reduce the effect of speckle noise and label noise effectively. Besides, the DPDNet consumes less time than other models while achieving satisfying change detection results.

In summary, the main contributions of this article are as follows.

1) To the best of our knowledge, we make the first attempt to solve the problem of speckle noise and label noise simultaneously. Previous works mainly focus on the speckle noise. However, the label noise problem is detrimental to the change detection performance. The proposed DPDNet can alleviate two kinds of noise simultaneously, and generate more accurate change maps.

2) We present a novel distinctive patch convolution (DPConv), which simplifies the network structure and expedite the training phase. The patches taken from the original images are used as convolutional kernels, and therefore, it does not require many training samples for parameter optimization.

II. METHODOLOGY

As shown in Fig. 2, the proposed DPDNet comprises two branches: one is used to clean the label noise by propagating the correct labeled samples to the other unlabeled samples. In this process, the transformation weight is generated through homogeneous region generation and similarity measurement, followed by multiple random label propagation based on the weights and using voting to obtain the final labels. The other branch aims to improve the efficiency of the model while avoiding the speckle effect. To reduce the long training time required by conventional convolution, the designed DPConv is used to obtain hierarchical convolution features, so as to reduce the consumption of computational resources while improving feature expressiveness. Through the combination of these two branches, it not only enables the removal of noise from the data and improves the change detection performance, but also reduces the training time and improves the application potential.

A. Overview of the Proposed DPDNet

Most existing deep learning-based methods applied to SAR change detection tasks require multiple rounds of training for parameter optimization, which is very time-consuming. Besides, in the process of clustering, considerable noisy samples are inevitably introduced due to the limitation of clustering algorithm
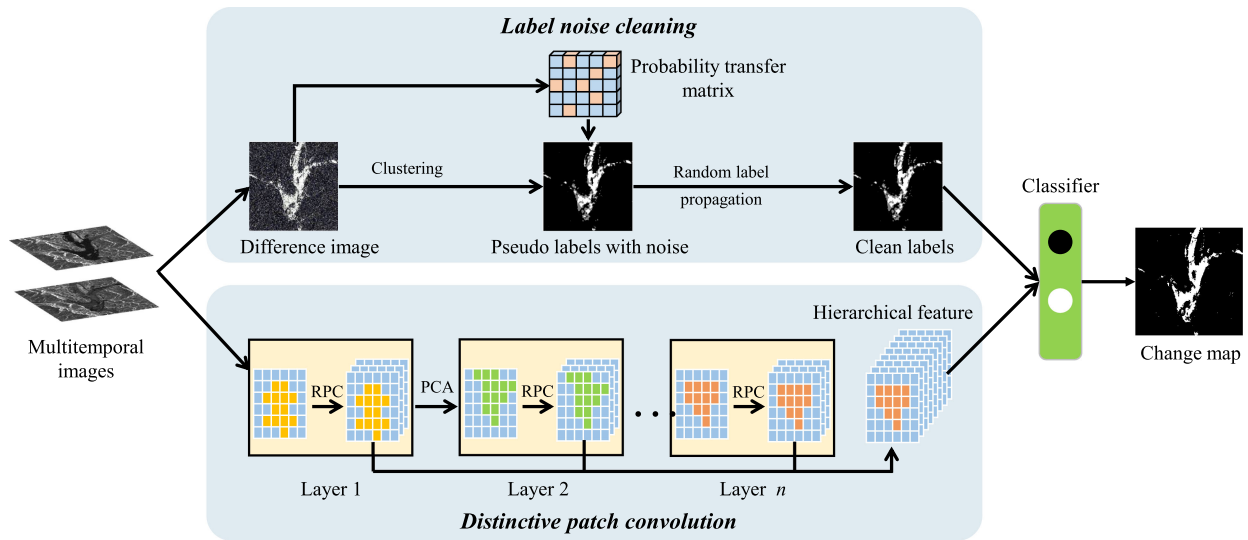
Fig. 2. Schematic illustration of the proposed DPDNet. One branch uses random label propagation to clean the label noise generated in preclassification. The other branch is utilized to extract shallow and deep features by stacking multiple convolution layers. Finally, the clean labels and stacked features are fed into the classifier to compute the final change detection results.
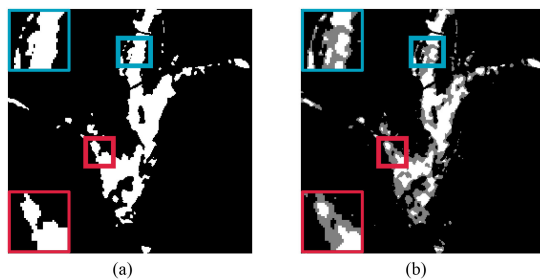

Fig. 3. Visualization of label noise. (a) Ground-truth change map. (b) Pseudolabeled change map generated by k-means [31] clustering. From the marked areas, it can be observed that there some differences between the ground-truth map with the generated pseudolabeled map. If these noisy labels are fed into the neural network, they will mislead the training process, and cause a decrease in the change detection performance.
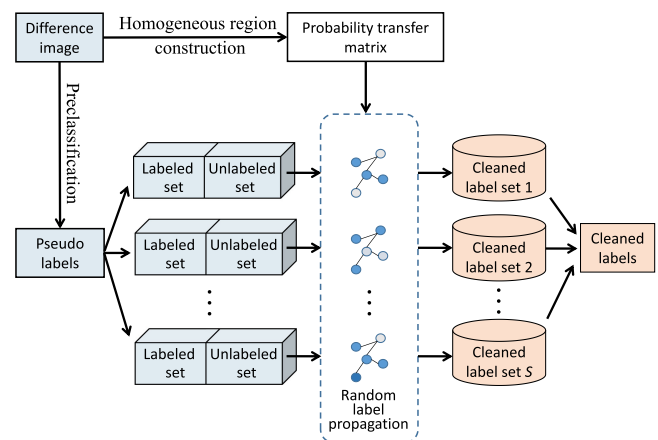

Fig. 4. Schematic of the proposed RLPA-based label noise cleaning for SAR change detection.

(see Fig. 3), which is rarely considered in previous works [25], [38], [39]. However, a robust change detection method should take label noise into account. To solve the above two problems, a dual path change detection model has been proposed, which is shown in Fig. 2. The first branch is used to guide the removal of label noise. Its core idea is to randomly remove the labels of some selected samples and repredict the labels of these selected samples based on the prior knowledge and spatial correlation. The other branch focuses on despeckle, and the proposed DP-Conv is used replacing the conventional convolution operation to reduce the time consumption while extracting meaningful features. The experimental results show that the combination of the two branches is an effective solution to the above problems and has a certain improvement both in terms of efficiency and evaluation indicators compared to other methods.

B. Label Noise Cleaning With Random Label Propagation

As illustrated in Fig. 3, there are some differences between the ground-truth change map with the generated pseudolabeled change map. These differences can be considered as the label noise. If these noisy labels are fed into the neural network, they will mislead the training process, and thus, cause a decrease in the final change detection performance. To effectively clean the label noise, we introduce the random label propagation algorithm (RLPA) [40] in our network, which is shown in Fig. 4. It mainly includes three phases—homogeneous region construction, generation of probabilistic transition matrix, and random label propagation.

1) Homogeneous Region Construction: We aim to propagate the label information among samples to clean the label noise. We assume that pixels from the same homogeneous region are more likely to belong to the same class. Thus, the construction

of homogeneous regions plays an important role in label noise cleaning. The entropy rate superpixel segmentation [41], which is an objective function of superpixel segmentation, is employed to construct homogeneous regions. Through this function, the whole image can be divided into several regions, and each region is recognized as a homogeneous region to facilitate subsequent label propagation.

2) Probabilistic Transition Matrix Generation [42]: Upon the acquired homogeneous regions, we can assign a weight to the pixels within a homogeneous region based on the spatial similarity, and let the weight between the pixels from different homogeneous regions be zero

$$W_{ij} = \begin{cases} \exp(-\frac{\text{sim}(p_i, p_j)^2}{2\sigma^2}), & p_i, p_j \in \chi_l \\ 0, & p_i \in \chi_l \text{ and } p_j \in \chi_p \end{cases} \quad (1)$$

where $\text{sim}(p_i, p_j)$ represents the spatial similarity, χ_l and χ_p are the lth and pth homogeneous region, and σ is the mean variance of all pixels in each homogeneous region.

In this article, the Euclidean distance is used to measure the similarity

$$\text{sim}(p_i, p_j) = \|p_i - p_j\|_2 \quad (2)$$

where $\| \cdot \|_2$ demotes the l_2 norm.

After that, the label information can propagate through the connected weight between different pixels. The larger the weight between two nodes, the easier it is likely to traverse. Hence, the two nodes are more likely to belong to the same class. We can define a probabilistic transition matrix T which can be seen as the probability of label conversion

$$T_{ij} = P(j \to i) = \frac{W_{ij}}{\sum_{k=1}^{N} W_{kj}} \quad (3)$$

where W_{ij} is the connected weight between different pixels. Here, N is the number of pixels in the homogeneous region, and k traverses all pixels in the homogeneous region, and T_{ij} can be considered as the probability to travel from node j to i.

3) Random Label Propagation: The training dataset X is divided into a labeled subset $X_L = \{x_1, x_2, \ldots, x_l\} \in R^l$ (X_L is the set of real vector of l elements) and an unlabeled subset $X_U = \{x_{l+1}, x_{l+2}, \ldots, x_N\} \in R^{N-l}$ (X_U denotes the set of unlabeled vector of $N - l$ elements). The label matrix of X_L is denoted as $\hat{Y}_L$, while the labels of X_U are discarded. In the experiments, we randomly choose two subsets from the total training samples. Afterward, the task turns to be predicting the labels $\hat{Y}_U$ of the unlabeled subset X_U based on the generated probabilistic transition matrix. To better propagate the label information, multiple iterations are adopted to propagate the labels from X_L to X_U. Hence, the label of x_i at time $t + 1$ is

$$y_i^{t+1} = \alpha \sum_{x_i, x_j \in X} T_{ij} y_j^t + (1 - \alpha) \widetilde{y}_i^{LU} \quad (4)$$

where α is a parameter to balance the contribution between the initial label information and the label information received from its neighbors, and $\widetilde{y}_i^{LU}$ is the ith column of $\hat{Y}_{LU} = [\hat{Y}_L \hat{Y}_U]^T$. It should be noted that the labels of the unlabeled subset are initialized to zero.

Since the initial labeled and unlabeled subsets are chosen randomly, the abovementioned process of random assignment of clean samples and unlabeled samples are repeated multitimes. Because each sample gets one label per round, each sample will obtain multiple labels after multiple rounds of random assignment. The final label can be computed by majority vote.

C. Alleviating Speckle Noise With Distinctive Patch Convolution

Speckle noise is a granular interference that inherently degrades the quality of SAR images, and it severely affects the change detection performance. Many deep learning-based classifiers have been used to suppress the speckle noise [43], [44]. However, these techniques are commonly time-consuming in the training period. Besides, these techniques generally require a great number of training samples for model training to achieve robust representation. To solve the mentioned issue, we propose a distinctive patch convolution (DPConv) for SAR image feature representation.

Xu *et al.* [42] proposed a random patch network (RPNet), which utilizes the random patches taken from the image to build a random matrix, and then project the original vector into a proper subspace. This work reveals the potential value of extracting patches from the original images as convolution kernels, which has not been considered in SAR image change detection. However, there are certain problems with this approach, despite the advances it achieved. In [45], it points out that in a structured model, the randomness of sampling can risk the model's capability of focusing on nontrivial spatial contextual information. As we all know, SAR image is often composed of background and semantic objects. The texture and geometry information of semantic objects is distinctive, so these semantic objects (called distinctive region) commonly have higher activation values in the feature map. Therefore, when the convolutional kernels are selected from the background, they do not contain as much information as the semantic object regions, and this view is also demonstrated by comparative experiments in Section III-C. To solve the problem, we proposed DPConv, which selects patches from the distinctive regions as convolutional kernels to avoid the randomness of convolution kernel selection, so as to focus on important regions with salient texture and geometry information.

The implementation details of DPConv are shown in Fig. 5. The distinctive regions are obtained according to the activation values in the feature maps as

$$\hat{F}_{ij} = \frac{F_{ij} - F_{\min}}{F_{\max} - F_{\min}} \quad (5)$$

where F represents the input feature maps, $F_{\min}$ and $F_{\max}$ denote the minimum and maximum values, respectively. In our experiments, we set $\hat{F}_{ij} > 0.7$ as distinctive region. Then, we random choose m pixels in the distinctive region. Around these selected pixels, $k \times k$ patches are extracted. For pixels on the edge of the feature map, the vacant neighboring pixels are filled by mirroring the image. Then, these patches are considered as kernels to be convolved with the input data.
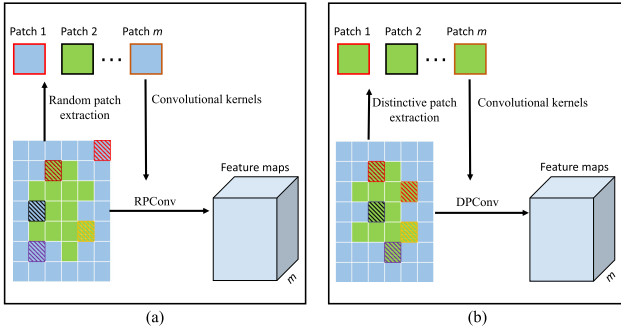
Fig. 5. Comparisons of RPConv with the proposed DPConv. (a) represents the feature extraction by RPConv. (b) represents the feature extraction by DPConv. The salient object is marked by green squares, the background is marked by blue square. The selected regions are marked by diagonal squares. By using RPConv, background region which contains little texture are chosen. In contrast, by using the proposed DPConv, convolutional kernels are selected from the distinctive region, and therefore, boosted classification performance can be achieved.
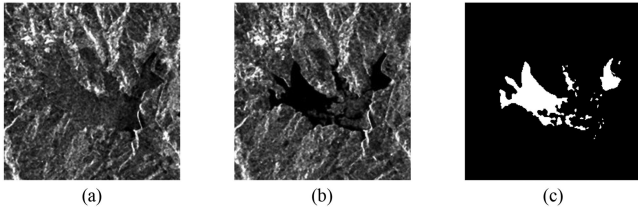

Fig. 6. Florence dataset. (a) Image captured on July 21, 2004. (b) Image captured on September 30, 2004. (c) Ground truth image.

Algorithm 1: The Workflow of Extracting Stacked Features by DPConv.

Input: Original data F, the number of patches m, and the depth of the network D.

Output: Stacked features F_{out}.

1:    **for** each layer d $(1 \le d \le D)$ **do**
2:      **if** $(d > 1)$ **then**
3:        PCA is employed to reserve the first three principal components of F.
4:      **end if**
5:      Find the distinctive region according to the activation value $\hat{F}_{ij}$:

$$\hat{F}_{ij} = \frac{F_{ij} - F_{\min}}{F_{\max} - F_{\min}}$$

6:      Extract m patches from the distinctive region:

$$W = Random(m, \hat{F}_{ij} > 0.7)$$

7:      Compute features by convolving F with patches extracted in the previous step:

$$DPConv = W \odot F$$

8:      Denote the features extracted in the dth layer as F^d
9:      If $d < D$, update the matrix F by F^d
10:    **end for**
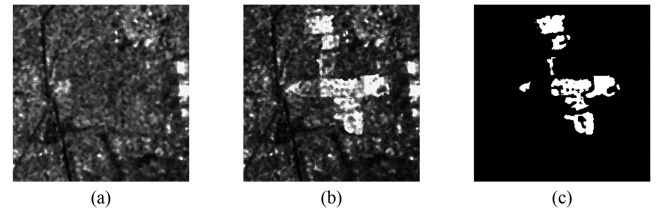11: $F^1, F^2, \ldots, F^D$ are concatenated to form the final output features F_{out}.

Fig. 7. Simulated dataset. (a) Image changed before. (b) Image changed after. (c) Ground truth image.

Fig. 2 shows that the network contains multiple convolution layers. After each convolution layer, principal component analysis (PCA) is used to retain the first three principal components. The philosophy behind this behavior is that we can obtain better change detection results by fusing the multilevel features. Next, we combine the deep and shallow features which are normalized to alleviate the speckle noise. Finally, all features and the original input data are utilized to predict the changed and unchanged area via a SVM classifier. The detailed implementation steps are shown in Algorithm 1.

III. EXPERIMENTS AND ANALYSIS

In this section, five multitemporal SAR datasets are used to verify the effectiveness of the proposed DPDNet. An extensive study of critical parameters are presented. Next, we compare DPDNet with its variants and make a series of ablation studies. Then, the proposed DPDNet is compared with several closely related methods. Finally, we analyze the computational complexity of DPDNet to show its efficiency.

A. Experimental Datasets

To validate the effectiveness of the proposed DPDNet, we carried out experiments on five multitemporal SAR datasets. The ground truth change maps are manually annotated carefully by experts for accuracy assessment.

The first dataset is the Florence dataset, as shown in Fig. 6, and it is obtained over the city of Florence, Italy. Some parts

of the river have changed between the two images. The images were captured on July 21 and September 30, 2004, respectively.

The second dataset is the Simulated dataset, which is a simulated dataset [46]. The size of the image is 350×250 pixels. The speckle noise on the Simulated dataset is rather severe. The dataset is shown in Fig. 7.

The third dataset is the Sulzberger dataset. It is cropped from a large SAR image from Sulzberger Ice Shelf. The dataset was taken by the Envisat satellite of the European Space Agency. Both images show the sea ice breakup. In March 2011, the ocean waves generated by the Tohoku Tsunami caused the Sulzberger Ice Shelf to break. To better reveal the change information, we chose one typical region. The ground truth change map is generated by experts with rich prior knowledge. The dataset is illustrated in Fig. 8.

The next dataset is the Havana dataset, which is centred over the city of Havana, Cuba. Two images show the landscape
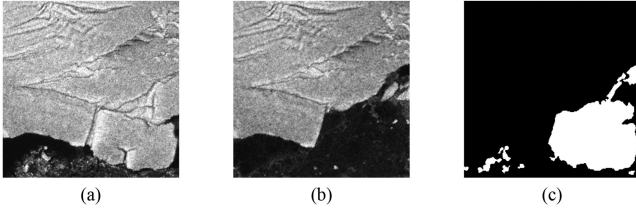
Fig. 8. Sulzberger dataset. (a) Image captured on March 11, 2011. (b) Image captured on March 16, 2011. (c) Ground truth image.
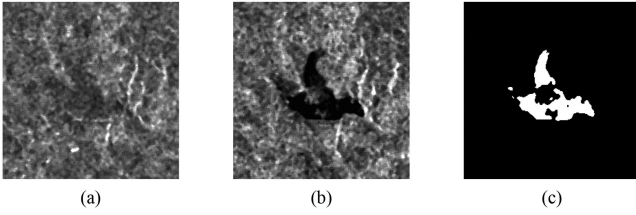

Fig. 9. Havana dataset. (a) Image captured in May 1997. (b) Image captured in July 1997. (c) Ground truth image.
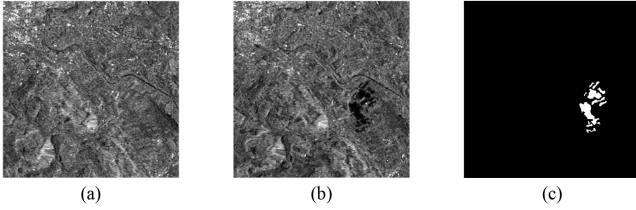

Fig. 10. Bern dataset. (a) Image captured in April 1997. (b) Image captured in May 1997. (c) Ground truth image.

changes from dry season to wet season in Havana. Two SAR images were obtained by ERS-2 satellite in May and July 1997, respectively. Some regions are flooded during the period. The Havana dataset is shown in Fig. 9.

The last dataset is the Bern dataset, which is captured in the area near the city of Bern, Switzerland. Two images show the geomorphic changes after the Aare River flooded parts of the cities of Thun and Bern and the airport of Bern entirely [47]. Two images were obtained by ERS-2 satellite in April and May 1997, which is illustrated in Fig. 10.

A critical issue in change detection is the selection of evaluation criteria [25], [38], [48]. In this article, false positives (FP), false negatives (FN), percentage correct classification (PCC), Kappa coefficient (KC), and F1 score (F1) are used to evaluate the change detection performance.

1) FP: the number of unchanged pixels in the ground truth image but falsely classified as changed.

2) FN: the number of changed pixels in the ground truth image but falsely classified as unchanged.

3) PCC: Classification accuracy in change detection. The PCC is calculated by

$$PCC = 1 - \frac{FP + FN}{N_c + N_{uc}} \qquad (6)$$

where N_c denotes the number of changed pixels in the ground truth image, and N_{uc} is the number of unchanged pixels.
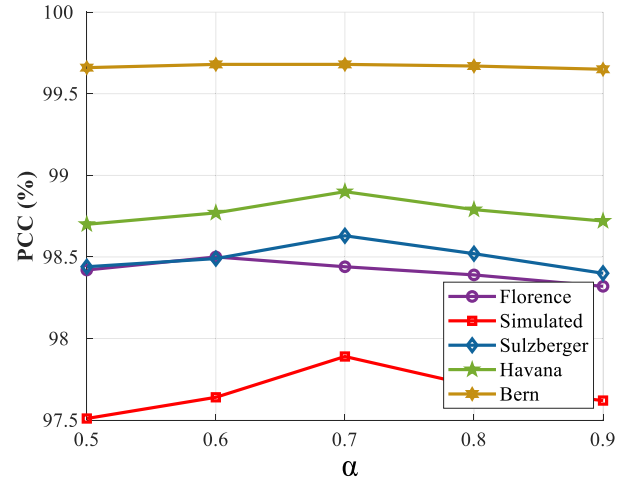

Fig. 11. Relationship between the classification accuracy and parameter α.

4) KC: Kappa coefficient is used for reliability evaluation, and it is computed as:

$$KC = \frac{PCC - PRE}{1 - PRE} \qquad (7)$$

$$PRE = \frac{(N_c + FP - FN) \times N_c + (N_{uc} + FN - FP) \times N_{uc}}{(N_c + N_{uc}) \times (N_c + N_{uc})}. \qquad (8)$$

5) F1: F1 score is the harmonic mean of precision and sensitivity, and it is computed as

$$F1 = \frac{2TP}{2TP + FP + FN} \qquad (9)$$

where TP is the number of changed pixels in the ground truth image and truly classified as changed.

B. Analysis of the Important Parameters of DPDNet

Analysis of the parameter α. In the proposed DPDNet, an important branch is used to clean the label noise. The critical parameter α is used to balance the contribution between the initial label itself and the labels received from its neighbors. We evaluate the change detection performance by taking $\alpha = 0.5$, 0.6, 0.7, 0.8, and 0.9. Fig. 11 shows the quantitative analysis result of the parameter α. It can be observed that when $\alpha = 0.6$, the PCC obtains the best value on the Florence datasets, while on the Simulated, Sulzberger, Havana, and Bern datasets, the result achieves the best when $\alpha = 0.7$. Since different datasets have different spatial structures in the label propagation, the contributions between the initial label and the label received from its neighbors are different. Hence, we set α to 0.6 on the Florence datasets, and set α to 0.7 on the Simulated, Sulzberger, Havana, and Bern datasets in our following experiments.

1) Analysis of the Image Patch Size: As we all know, the contextual information plays an important role in alleviating the speckle noise. However, the patch size controls the contextual information contained in the training samples, and therefore, we design experiments to test the parameter w. Here, w denotes
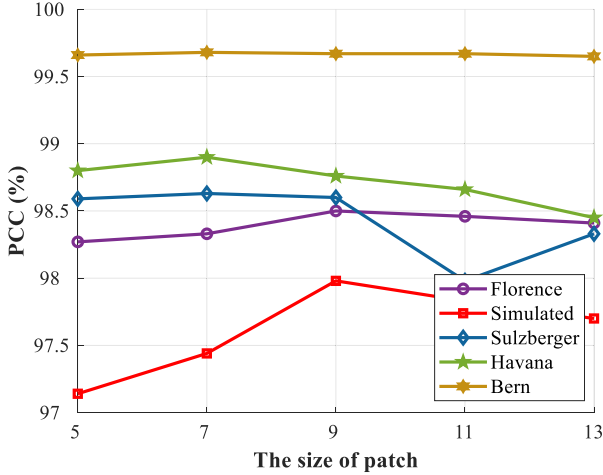
Fig. 12. Relationship between PCC and the size of patch.

Fig. 13. Relationship between the change detection accuracy and the number of training samples.

the size of the patches used as the training sample. We evaluate the change detection performance by taking $w = 5, 7, 9, 11,$ and 13. As shown in Fig. 12. We can see that the PCC value is not satisfying when $w = 5$, because such small patches can hardly extract the contextual information effectively. On the Sulzberger, Havana, and Bern datasets, the DPDNet achieves the best result when $w = 7$. On the Florence and Simulated datasets, $w = 9$ obtains the best PCC values. It indicates that different datasets contain different contextual information and require different size of receptive fields for feature extraction. Therefore, in the following implementation, we set $w = 7$ on the Sulzberger, Havana, and Bern datasets and set $w = 9$ on the Florence and Simulated datasets.

2) Analysis of the Training Sample Numbers: It is evident that the number of training samples has a great impact on the change detection performance. Hence, we investigate the relationship between PCC value and the training sample numbers on different datasets. As mentioned before, the training samples are randomly selected from the preclassification results by the clustering algorithm. We randomly selected 4%, 6%, 8%, 10%, 12%, and 14% pixels as training samples. The change detection performance on different datasets by using different training sample numbers are shown in Fig. 13. From the curves in the figure, it can be seen that there is a sharp increase in PCC value when the training sample number ranges from 4% to 10%. On Florence, Sulzberger, Havana, and Bern datasets, the best result is achieved when the training sample ratio is 12%. After that, the PCC value tends to be stable even the number of training samples increases. Hence, we select 12% pixels from the preclassification result as training samples.

3) Analysis of the Stacked Layer Numbers: The main objective of this study is to use distinctive patches extracted from the original images as convolution kernels for feature extraction. In this article, we use multiple convolutional layers to extract shallow and deep features. Therefore, the number of convolutional layers affects the number of stacked features, which has a great impact on the change detection results. In this experiment, we
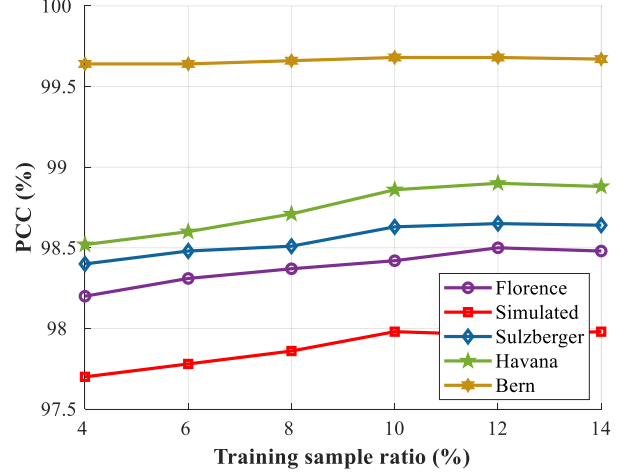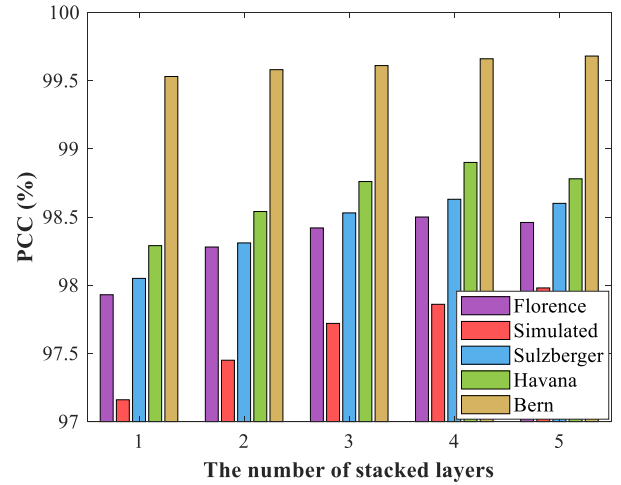

Fig. 14. Relationship between PCC and the stacked layer number.

explore the relationship between the number of stacked layers D and the change detection results. The experimental results are illustrated in Fig. 14. As can be observed, when $D = 4$, the PCC values reach the best on the Florence, Sulzberg, and Havana datasets. On the Simulated and Bern dataset, the best performance is achieved when $D = 5$. It is owing to the reason that different datasets have different texture details. With the increase of stacked layers, better feature representations can be captured on different regions. The visualized results with different number of stacked layers on the Florence and Havana datasets are shown in Figs. 15 and 16, respectively. As can be seen, with the increase of stacked layers, the change map contains less noise regions. Therefore, $D = 5$ is used on the Simulated and Bern dataset, $D = 4$ is employed on other datasets. In real applications, the value of D should be adjusted according to the texture or geometry characteristics of each dataset.
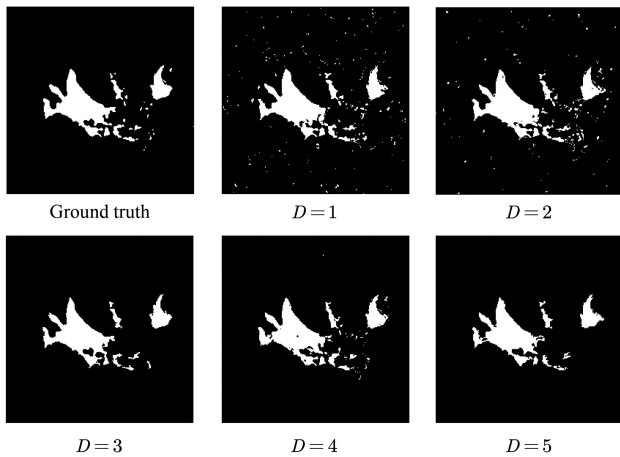
Fig. 15.   Change detection results on Florence dataset by different stacked features.
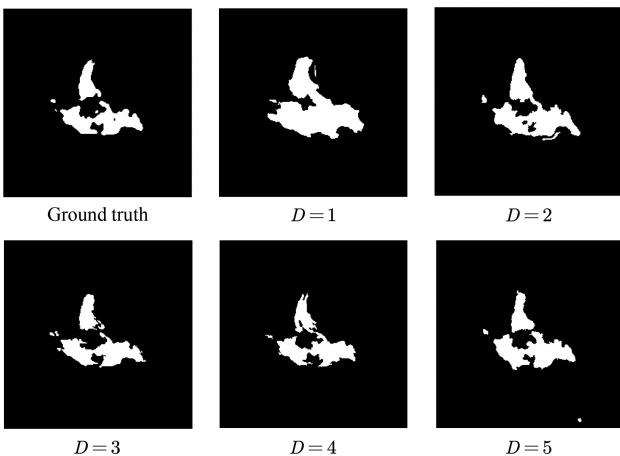

Fig. 16.   Change detection results on Havana dataset by different the number of stacked features.
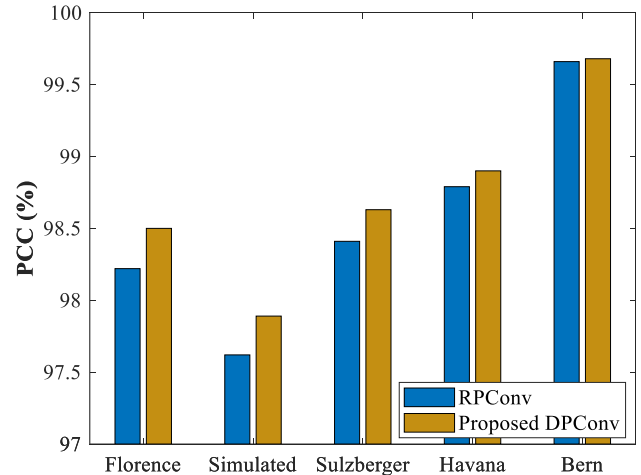

Fig. 17.   Comparison between distinctive patch convolution with random patch convolution.
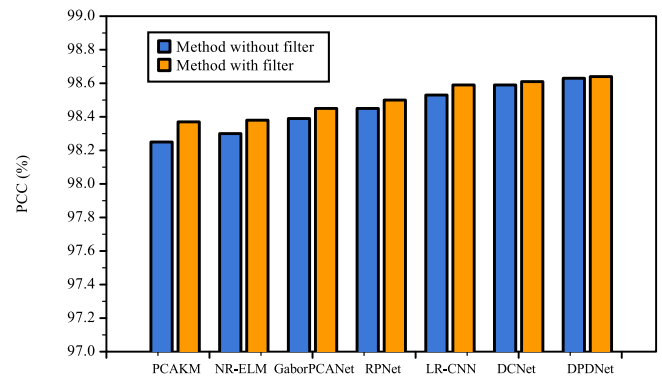
Fig. 18.   Effect of filtering operations on Sulzberger dataset.

C. Verification of Distinctive Patch Convolution Effectiveness

1) Distinctive Patch Convolution Versus Random Patch Convolution: To alleviate the speckle noise, our method introduces the distinctive patch convolution (DPConv) to extract the desired features. In most existing works, a great number of training samples are needed to optimize the parameters of the network, which is time-consuming. To make up this shortcoming, Xu *et al.* [42] proposed random patch convolution (RPConv) in which random patches taken from the image are used as convolution kernels. However, the randomness of RPConv risks the model's stability. Compared with RPNet, the proposed DPConv pinpoints the most nontrivial units of the image and offers better feature representation upon sampling position. We compare the models using DPConv with RPConv for feature extraction. As illustrated in Fig. 17, in most cases, the standard DPDNet offers a better performance. Meanwhile, the performance of RPConv is not stable, and it achieves relatively good change detection results when it takes samples from representative locations.

2) Distinctive Patch Convolution Versus Nonlocal Filter: To verify the performance of the proposed DPConv for speckle

noise suppression, the nonlocal means filter [49] was used for data preprocessing for all methods. As shown in Fig. 18, the filtering operation is performed on the Sulzberger dataset for these methods, and PCC value is utilized as the evaluation metric. It is obvious that nonlocal filter reduces the speckle noise to some extent and improves the change detection performance for most of the methods. As for DPDNet, the performance is also improved after using nonlocal filter for data preprocessing, but the improvement was smaller compared to the other methods. This is because it provides an end-to-end way to reduce the speckle noise via the strong learning capability of the network. Furthermore, the evaluation metric shows that DPDNet still outperforms the other methods without the addition of filter, which also proves that it is beneficial in speckle noise suppression.

D. Ablation Study

We further verify the effectiveness of both branches of the proposed DPDNet on five datasets. Note that to make a fair comparison, we have designed a Baseline (#1), which is completely consistent with the structure of the branch using DPConv in DPDNet, including the size of the convolution kernel and

TABLE I
ABLATION STUDIES ON FOUR DATASETS

#	DPConv	RPConv	RLPA	PCC (%) on different datasets				
				Florence	Simulated	Sulzberger	Havana	Bern
1	✗	✗	✗	98.02	97.43	98.22	98.37	99.60
2	✓	✗	✗	98.39	97.71	98.48	98.72	99.65
3	✗	✓	✗	98.14	97.57	98.34	98.56	99.63
4	✗	✗	✓	98.13	97.54	98.39	98.50	98.62
5	✗	✓	✓	98.21	97.66	98.47	98.71	99.65
6	✓	✗	✓	**98.50**	**97.98**	**98.63**	**98.90**	**99.68**

the number of layers, except that DPConv is replaced by the traditional convolution.

The Baseline (#1) has the same number of training epochs and training samples as DPDNet. However, as reported in Table I, the performance of *Baseline* is unsatisfying since the number of training samples is quite limited. When we use the distinctive patch extracted from the image as convolution kernels, it reduces the dependence of the network parameters and achieves better results (#1 vs #3). Meanwhile, the RPConv will risk the stability of the network. For instance, when the kernels are selected from the smooth-textured background, the parameters of the convolution kernels are similar, and the extracted features cannot be representative. Hence, the proposed DPConv can effectively select kernels from the distinctive region, so as to reduce the influence of randomness. The results of #2 and #3 in Table I demonstrate the effectiveness of the proposed DPConv in feature representation.

Furthermore, label noise is introduced into change detection. To handle the problem, we use RLPA to clean the label noise hidden in the preclassification results. To evaluate the effectiveness of label noise cleaning, we make extensive experiments (#4 vs #1, #6 vs #2, #5 vs #3). It is evident that when we introduce a label noise cleaning branch, the change detection performance always significantly improved. Moreover, it can be observed that the proposed DPDNet (#6) not only enhances the feature representation by using DPConv, but also improves the classification performance by alleviating the label noise in preclassification.

E. Change Detection Results on Five SAR Datasets

To validate the performance of the proposed DPDNet, six state-of-the-art methods, including PCAKM [31], NR-ELM [46], GaborPCANet [50], RPNet [42], LR-CNN [51], and DCNet [52] are used as the comparison methods. For PCAKM, the contextual information is analyzed by PCA, and the features are handled k-means clustering. NR-ELM utilizes extreme learning machine (ELM) [53] as the classifier. Gabor-PCANet is composed of cascaded PCA layer, binary hashing layer, and block-wise histogram generation. RPNet uses random patches taken from the original image as convolution kernels to extract shallow and deep concolutional features. LR-CNN is a deep learning based method with local spatial restrictions.
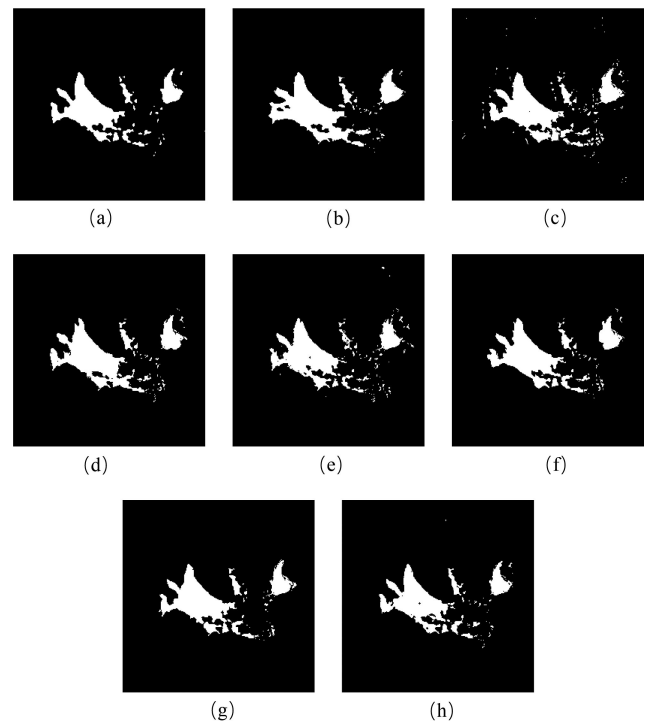

Fig. 19. Comparisons of the change maps generated by different methods on the Florence dataset. (a) Ground truth image. (b) Result by PCAKM [31]. (c) Result by NR-ELM [46]. (d) Result by GaborPCANet [50]. (e) Result by RPNet [42]. (f) Result by LR-CNN [51]. (g) Result by DCNet [52]. (h) Result by the proposed DPDNet.

In DCNet, a channel weighting-based module is designed for feature enhancement. Among these methods, RPNet, LR-CNN, and DCNet are deep learning-based methods. It should be noted that the RPNet is originally designed for hyperspectral image classification, and we made some adjustments so that it can be applied to change detection tasks.

1) Results on the Florence Dataset: Fig. 19 shows the experimental results on the Florence dataset, while Table II lists the corresponding quantitative metrics. From Table II, we can see that, compared with DPDNet, other methods suffer from high FP values resulting in many noisy spots in the change map as shown in Fig. 19. Since other learning-based methods do not take the label noise into account, and the training samples with errors mislead the learning process. In addition, the FN

TABLE II
CHANGE DETECTION RESULTS OF DIFFERENT METHODS ON THE FLORENCE DATASET

Methods	FP	FN	PCC(%)	KC(%)	F1(%)
PCAKM [31]	1289	672	97.01	78.59	82.20
NR-ELM [50]	1203	590	97.26	86.03	83.34
GaborPCANet [51]	1055	564	97.53	86.19	84.86
RPNet [42]	987	536	97.68	86.24	85.07
LR-CNN [52]	882	491	97.90	**86.32**	86.78
DCNet [53]	805	410	98.15	86.31	88.06
Proposed DPDNet	**686**	**298**	**98.50**	**86.32**	**89.88**

TABLE III
CHANGE DETECTION RESULTS OF DIFFERENT METHODS ON THE SIMULATED DATASET

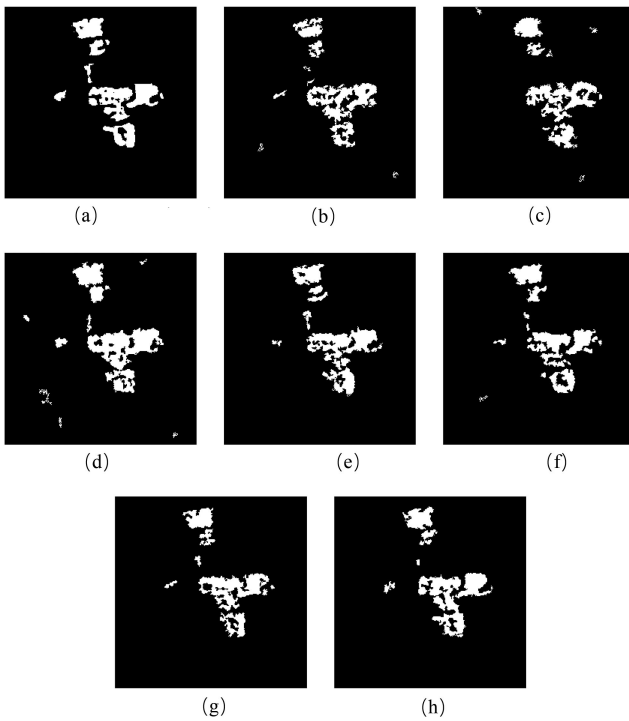| Methods	FP	FN	PCC(%)	KC(%)	F1(%)
PCAKM [31]	680	1247	97.06	74.63	76.19
NR-ELM [50]	903	1286	96.66	71.77	73.55
GaborPCANet [51]	1206	**507**	97.39	80.30	81.70
RPNet [42]	641	831	97.75	81.42	82.62
LR-CNN [52]	523	906	97.82	81.57	82.74
DCNet [53]	**420**	979	97.87	81.60	82.73
Proposed DPDNet	551	771	**97.98**	**83.26**	**84.34**

Fig. 20. Comparisons of the change maps generated by different methods on the Simulated dataset. (a) Ground truth image. (b) Result by PCAKM [31]. (c) Result by NR-ELM [46]. (d) Result by GaborPCANet [50]. (e) Result by RPNet [42]. (f) Result by LR-CNN [51]. (g) Result by DCNet [52]. (h) Result by the proposed DPDNet.
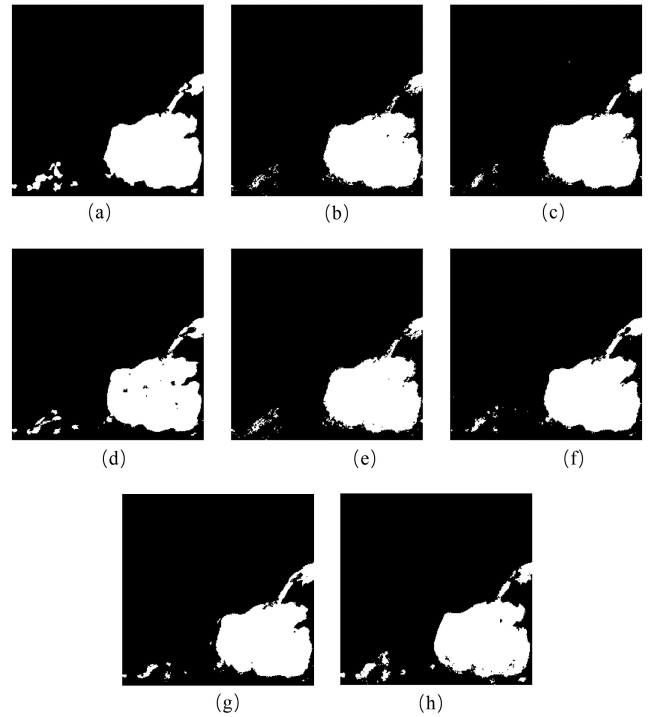

Fig. 21. Comparisons of the change maps generated by different methods on the Sulzberger dataset. (a) Ground truth image. (b) Result by PCAKM [31]. (c) Result by NR-ELM [46]. (d) Result by GaborPCANet [50]. (e) Result by RPNet [42]. (f) Result by LR-CNN [51]. (g) Result by DCNet [52]. (h) Result by the proposed DPDNet.

value of DPDNet is the lowest, which means that the proposed model does not ignore some subtle change pixels. Among these methods, the proposed DPDNet achieved the best PCC, KC, and F1 values. The results on the Florence dataset demonstrate that label cleaning is necessary. The comparisons demonstrate the superior performance of the proposed DPDNet on the Florence dataset.

2) Results on the Simulated Dataset: Fig. 20 presents the change maps on the simulated dataset. The corresponding quantitative metrics are shown in Table III. In addition to the complex spatial structure, the Simulated dataset is severely interfered by speckle noise. Thus, it is rather challenging to identify changed pixels accurately on this dataset. The results of NR-ELM and

GaborPCANet are polluted with noisy regions. Therefore, NR-ELM and GaborPCANet suffer from high FP values. Generally speaking, the deep learning-based models improve the change detection performance with a clear margin. In addition, the proposed DPDNet performs favorably against other methods on this dataset. The comparison shows that the DPDNet is powerful in noise suppression on the Simulated dataset.

3) Results on the Sulzberger Dataset: Fig. 21 shows the change detection results on the Sulzberger dataset. From visual inspection, it can be seen that most methods missed some subtle changed areas, which generate higher FN values in Table IV. Moreover, deep learning-based methods (LR-CNN, RPNet, and DCNet) generally obtain better performance. Furthermore, because of the stacked DPConv features and label cleaning, the

TABLE IV
CHANGE DETECTION RESULTS OF DIFFERENT METHODS
ON THE SULZBERGER DATASET

Methods	FP	FN	PCC(%)	KC(%)	F1(%)
PCAKM [31]	144	998	98.25	93.91	94.97
NR-ELM [50]	**74**	1037	98.30	94.07	95.09
GaborPCANet [51]	203	850	98.39	94.44	95.41
RPNet [42]	184	831	98.45	94.64	95.58
LR-CNN [52]	151	810	98.53	94.92	95.81
DCNet [53]	167	756	98.59	95.13	95.99
Proposed DPDNet	240	**659**	**98.63**	**95.29**	**96.12**

TABLE V
CHANGE DETECTION RESULTS OF DIFFERENT METHODS
ON THE HAVANA DATASET

Methods	FP	FN	PCC(%)	KC(%)	F1(%)
PCAKM [31]	1077	119	98.18	83.98	84.93
NR-ELM [50]	1542	77	97.53	79.55	80.83
GaborPCANet [51]	2387	**42**	96.29	64.85	73.95
RPNet [42]	795	258	98.39	85.11	86.35
LR-CNN [52]	553	383	98.57	86.14	87.22
DCNet [53]	447	392	98.72	87.40	89.04
Proposed DPDNet	**295**	428	**98.90**	**88.86**	**89.44**

Fig. 22. Comparisons of the change maps generated by different methods on the Havana dataset. (a) Ground truth image. (b) Result by PCAKM [31]. (c) Result by NR-ELM [46]. (d) Result by GaborPCANet [50]. (e) Result by RPNet [42]. (f) Result by LR-CNN [51]. (g) Result by DCNet [52]. (h) Result by the proposed DPDNet.
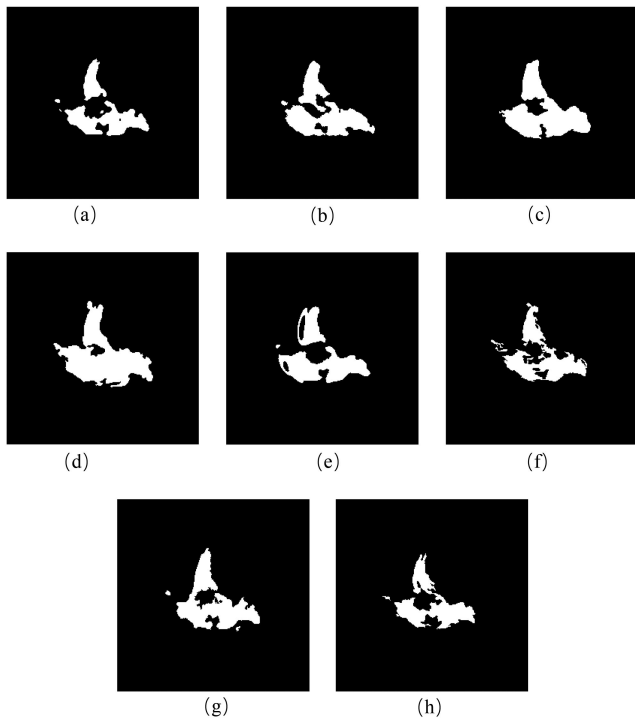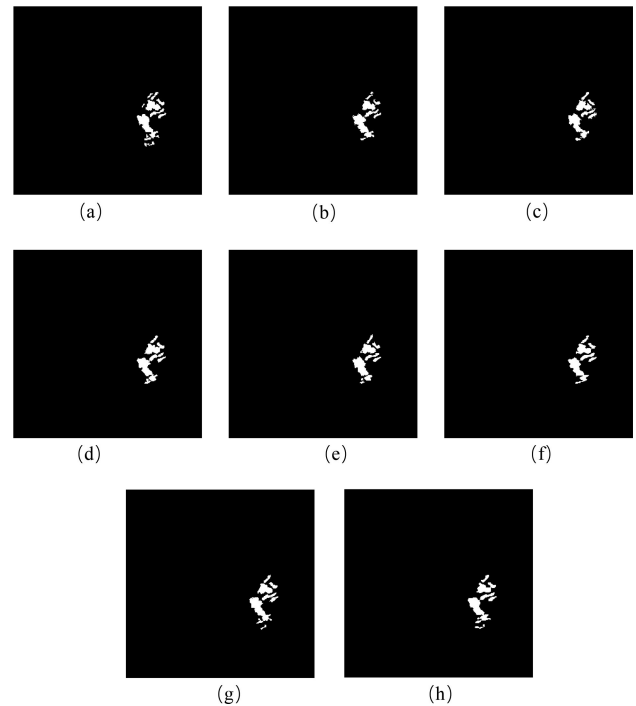

Fig. 23. Comparisons of the change maps generated by different methods on the Bern dataset. (a) Ground truth image. (b) Result by PCAKM [31]. (c) Result by NR-ELM [46]. (d) Result by GaborPCANet [50]. (e) Result by RPNet [42]. (f) Result by LR-CNN [51]. (g) Result by DCNet [52]. (h) Result by the proposed DPDNet.

proposed DPDNet obtains the best PCC, KC, and F1 values. Overall, the proposed DPDNet outperforms in terms of PCC, KC, and F1 values on the Sulzberger dataset.

4) Results on the Havana Dataset: Fig. 22 shows the change maps on the Havana dataset. The corresponding evaluation metrics are listed in Table V. As we can see, PCAKM, NR-ELM, and GaborPCANet do not perform well since many unchanged regions are falsely classified into changed regions with high FP values. Moreover, the deep learning-based methods (LR-CNN, RPNet, and DCNet) can suppress the speckle noise to some extent by integrating spatial information. Therefore, the PCC, KC, and F1 values of these deep learning-based models have been greatly improved. However, some changed pixels are missed, and thus the FN values are relatively high. The proposed

DPDNet obtains the best PCC, KC, and F1 values compared with other methods. Generally speaking, the comparisons demonstrate the effectiveness of the proposed method on the Havana dataset.

5) Results on the Bern Dataset: The visualized change detection results are shown in Fig. 23, and the evaluation metrics obtained by different methods are listed in Table VI. It can be observed that the proposed DPDNet achieves the best performance in most metrics with the assistance of cleaned labels and multilevel features. Meanwhile, the result map of DPDNet is the most similar to the ground truth map. The satisfying performance on this open dataset proves the effectiveness of the proposed method in change detection task.
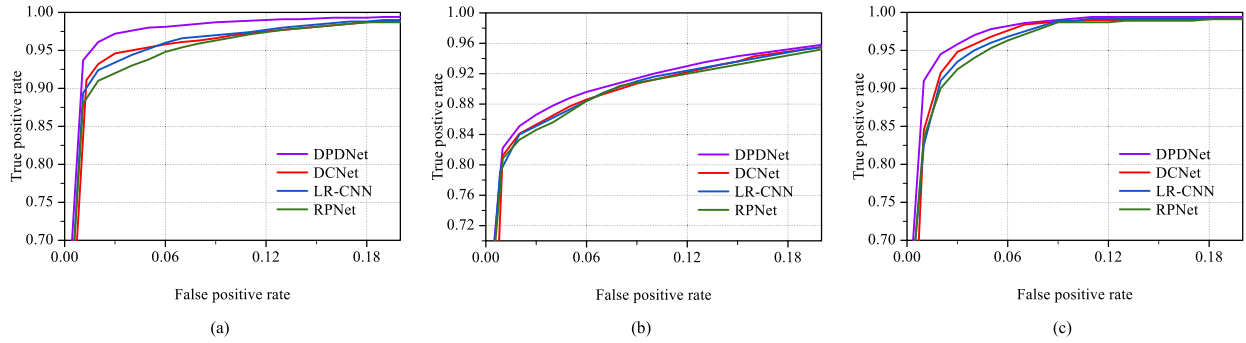
Fig. 24. ROC curves of the change detection results on different datasets. (a) ROC curves on the Florence dataset. (b) ROC curves on the Simulated dataset. (c) ROC curves on the Sulzberger dataset.

TABLE VI
CHANGE DETECTION RESULTS OF DIFFERENT METHODS
ON THE BERN DATASET

Methods	FP	FN	PCC(%)	KC(%)	F1(%)
PCAKM [31]	**95**	258	99.61	83.36	83.56
NR-ELM [50]	104	228	99.63	84.63	84.81
GaborPCANet [51]	131	193	99.64	85.41	85.59
RPNet [42]	103	226	99.64	84.77	84.96
LR-CNN [52]	97	216	99.65	85.54	85.71
DCNet [53]	146	161	99.66	86.45	86.62
Proposed DPDNet	131	**160**	**99.68**	**87.08**	**87.24**

TABLE VII
AUCs OF DIFFERENT METHODS ON THE FLORENCE, SIMULATED,
AND SULZBERGER DATASETS

Methods	Florence	Simulated	Sulzberger
RPNet [42]	0.9831	0.9585	0.9867
LR-CNN [52]	0.9847	0.9603	0.9878
DCNet [53]	0.9852	0.9602	0.9885
Proposed DPDNet	0.9910	0.9632	0.9912

TABLE VIII
COMPARISON AMONG DIFFERENT CHANGE DETECTION METHODS
OF THE EVALUATION TIME (IN SECONDS)

Methods	Florence	Simulated	Sulzberger
PCAKM [31]	3.52	3.41	3.26
NR-ELM [50]	24.62	23.81	22.75
GaborPCANet [51]	441.68	435.80	431.76
RPNet [42]	158.62	155.78	153.30
LR-CNN [52]	482.53	478.33	465.90
DCNet [53]	492.01	488.62	481.80
Proposed DPDNet	165.15	163.24	162.87

F. ROC Curves of the Change Detection Results

Fig. 24 shows the receiver operating characteristic (ROC) curves of the change maps on the Florence, Simulated, and Sulzberger datasets. It should be noted that the PCAKM, NR-ELM, GaborPCANet do not generate results with probability. Therefore, we only provide ROC curves of DPDNet, DCNet, LR-CNN, and RPNet.

The ROC curve of one good change detection method should be close to the top-left corner. On the three datasets, the ROC curves of the proposed DPDNet perform better than the other methods. The ROC curves of RPNet are closest to the bottom-right corner and have the worst performance. It is evident that the proposed DPDNet effectively improves the change detection performance by using the distinctive patch convolution. The corresponding area under the ROC curve (AUC) is provided in Table VII. The higher the AUC, the better the performance of the model at distinguishing between the changed and unchanged pixels. We can observe that the DPDNet obtains the best AUC, which also demonstrates the effectiveness of the proposed method.

G. Analysis of the Computation Time Requirement

Table VIII shows the comparison among different change detection methods of the evaluation time on the Florence, Simulated and Sulzberger datasets. The traditional methods such as PCAKM and NR-ELM, cost less time, due to its relatively simple workflow. However, their performances are not as good as deep learning-based models. By contrast, the time consumptions of GaborPCANet, LR-CNN, and DCNet are about more than twice higher than the proposed DPDNet. This demonstrates the computational efficiency of the DPDNet. The DPDNet requires much less computation time owing to the prefixed convolution kernels. Compared with RPNet, DPDNet costs more time to clean the label noise, while this further improves the experimental results. All experiments are carried out on a desktop PC with an Intel Xeon E5-2620 processor and an NVIDIA GTX 1080Ti GPU.

IV. CONCLUSION

In this article, we present a novel SAR image change detection model called DPDNet, which can suppress the speckle noise and label noise simultaneously. The DPDNet is composed of

two branches. One branch cleans the label noise via random label propagation, and the other branch combines shallow and deep features by using distincitve path convolution. In particular, the attention mechanism is used to select distinctive pixels in the feature map. Patches around these pixels are selected as convolution kernels. It does not require many training samples for parameter optimization, and therefore, the proposed DPDNet is of high computational efficiency. The experimental results on five multitemporal SAR datasets demonstrate that the DPDNet can achieve substantially higher accuracy over state-of-the-art methods.

REFERENCES

[1] 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.

[2] A. B. Campos, M. I. Pettersson, V. T. Vu, and R. Machado, "False alarm reduction in wavelength-resolution SAR change detection schemes by using a convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4004805.

[3] J. G. Vinholi, D. Silva, R. Machado, and M. I. Pettersson, "CNN-based change detection algorithm for wavelength-resolution SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4003005.

[4] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5530–5538, Dec. 2018.

[5] H. Zhuang, M. Hao, K. Deng, K. Zhang, X. Wang, and G. Yao, "Change detection in SAR images via ratio-based Gaussian kernel and nonlocal theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5210215.

[6] H. Dong, W. Ma, L. Jiao, F. Liu, and L. Li, "A multiscale self-attention deep clustering for change detection in SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5207016.

[7] D. Brunner, G. Lemonie, 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.

[8] 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.

[9] B. Wang, C. Zhao, Q. Zhang, Z. Lu, and A. Pepe, "Long-term continuously updated deformation time series from multisensor InSAR in Xi'an, China from 2007 to 2021," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7297–7309, 2021.

[10] B. Zou, W. Li, and L. Zhang, "Built-up area extraction using high-resolution SAR images based on spectral reconfiguration," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 8, pp. 1391–1395, Aug. 2021.

[11] D. C. Zanotta and V. Haertel, "Gradual land cover change detection based on multitemporal fraction images," *Pattern Recognit.*, vol. 45, no. 2, pp. 2927–2937, 2012.

[12] D. Amitrano, R. Guida, and P. Iervolino, "Semantic unsupervised change detection of natural land cover with multitemporal object-based analysis on SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5494–5514, Jul. 2021.

[13] M. N. Sumaiya and R. Shantha SelvaKumari, "Logarithmic mean-based thresholding for SAR image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 11, pp. 1726–1728, Nov. 2016.

[14] G. Yang, H. Li, W. Wang, W. Yang, and W. J. Emery, "Unsupervised change detection based on a unified framework for weighted collaborative representation with RDDL and fuzzy clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8890–8903, Nov. 2019.

[15] H. Wu, H. Li, G. Yang, and W. Yang, "Change detection of remote sensing images based on weighted nonnegative matrix factorization," in *Proc. Int. Workshop Anal. Multitemporal Remote Sens. Images*, 2019, pp. 1–4.

[16] M. Li, M. Li, P. Zhang, Y. Wu, W. Song, and L. An, "SAR image change detection using PCANet guided by saliency detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 3, pp. 402–406, Mar. 2019.

[17] M. Pham, G. Mercier, and J. Michel, "Change detection between SAR images using a pointwise approach and graph theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2020–2032, Apr. 2016.

[18] S. Saha, F. Bovolo, and L. Bruzzone, "Building change detection in VHR SAR images via unsupervised deep transcoding," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 1917–1929, Mar. 2021.

[19] C. Marin, F. Bovolo, and L. Bruzzone, "Building change detection in multitemporal very high resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2664–2682, May 2015.

[20] L. An, M. Li, P. Zhang, Y. Wu, L. Jia, and W. Song, "Multicontextual mutual information data for SAR image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1863–1867, Sep. 2015.

[21] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone, "Supervised change detection in VHR images using contextual information and support vector machines," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 20, pp. 77–85, Feb. 2013.

[22] S. Wang, S. Yang, and L. Jiao, "Saliency-guided change detection for SAR imagery using a semi-supervised Laplacian SVM," *Remote Sens. Lett.*, vol 7, no. 11, pp. 1043–1052, Nov. 2016.

[23] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.

[24] Z. Yetgin, "Unsupervised change detection of satellite images using local gradual descent," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1919–1929, May 2012.

[25] M. Gong, H. Yang, and P. Zhang, "Feature learning and change feature classification based on deep learning for ternary change detection in SAR images," *ISPRS J. Photogramm. Remote Sens.*, vol. 129, pp. 212–225, 2017.

[26] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.

[27] 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.

[28] 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, Feb. 2012.

[29] M. Jiang et al., "Delineation of built-up land change from SAR stack by analysing the coefficient of variation," *ISPRS J. Photogramm. Remote Sens.*, vol. 169, pp. 93–108, 2020.

[30] N. S. Mishra, S. Ghosh, and A. Ghosh, "Fuzzy clustering algorithms incorporating local information for change detection in remotely sensed images," *Appl. Soft Comput.*, vol. 12, no. 8, pp. 2683–2692, 2012.

[31] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Apr. 2009.

[32] L. Jia, M. Li, P. Zhang, Y. Wu, 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.

[33] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and C. Zoppetti, "Non-parametric change detection in multitemporal SAR images based on mean-shift clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2022–2031, Apr. 2013.

[34] 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.

[35] B. Hou, Y. Wang, and Q. Liu, "Change detection based on deep features and low rank," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2418–2422, Dec. 2017.

[36] Y. Zhan, K. Fu, M. Yan, X. Sun, H. Wang, and X. Qiu, "Change detection based on deep siamese convolutional network for optical aerial images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1845–1849, Oct. 2017.

[37] X. Zhang, X. Su, Q. Yuan, and Q. Wang, "Spatial-temporal gray-level co-occurrence aware CNN for SAR image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4018605.

[38] X. Qu, F. Gao, J. Dong, Q. Du, and H. -C. Li, "Change detection in synthetic aperture radar images using a dual-domain network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4013405.

[39] Y. Gao, F. Gao, J. Dong, and H. -C. Li, "SAR image change detection based on multiscale capsule network," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 484–488, Mar. 2021.

[40] J. Jiang, J. Ma, Z. Wang, C. Chen, and X. Liu, "Hyperspectral image classification in the presence of noisy labels," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 851–865, Feb. 2019.

[41] M. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate super-pixel segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 2097–2104.

[42] Y. Xu, F. Zhang, and L. Zhang, "Hyperspectral image classification via a random patches network," *ISPRS J. Photogramm. Remote Sens.*, vol. 142, pp. 344–357, May 2018.

[43] H. Shen, C. Zhou, J. Li, and Q. Yuan, "SAR image despeckling employing a recursive deep CNN prior," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 273–286, Jan. 2021.

[44] X. Ma, C. Wang, Z. Yin, and P. Wu, "SAR despeckling by noisy reference-based deep learning method," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8807–8818, Dec. 2020.

[45] Z. Ouyang, Y. Feng, Z. He, T. Hao, T. Dai, and S. -T. Xia, "Attentiondrop for convolutional neural networks," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2019, pp. 1342–1347.

[46] F. Gao, J. Dong, B. Li, Q. Xu, and C. Xie, "Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine," *J. Appl. Remote Sens.*, vol. 10, no. 4, 2016, Art. no. 046019.

[47] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.

[48] Y. Gao, F. Gao, J. Dong, Q. Du, and H. -C. Li, "Synthetic aperture radar image change detection via siamese adaptive fusion network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10748–10760, Oct. 2021.

[49] X. Su, C. A. Deledalle, F. Tupin, and H. Sun, "NORCAMA: Change analysis in SAR time series by likelihood ratio change matrix clustering," *ISPRS J. Photogramm. Remote Sens.*, vol. 101, pp. 247–261, 2015.

[50] 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.

[51] 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.

[52] Y. Gao, F. Gao, J. Dong, and S. Wang, "Change detection from synthetic aperture radar images based on channel weighting-based deep cascade network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 11, pp. 4517–4529, Nov. 2019.

[53] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2004, pp. 985–990.

Junyu Dong (Member, IEEE) received the B.Sc. and M.Sc. degrees in applied mathematics from Ocean University of China, Qingdao, China, in 1993 and 1999, respectively, and the Ph.D. degree in image processing from the Department of Computer Science, Heriot-Watt University, Edinburgh, U.K., in 2003.
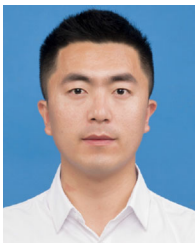
He is currently a Professor and Dean with the School of Computer Science and Technology, Ocean University of China. His research interests include visual information analysis and understanding, machine learning, and underwater image processing.

Qian Du (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland at Baltimore, Baltimore, MD, USA, in 2000.

She is currently the Bobby Shackouls Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. Her research interests include hyperspectral remote-sensing image analysis and applications, and machine learning.

Dr. Du was the recipient of the 2010 Best Reviewer Award from the *IEEE Geoscience and Remote Sensing Society (GRSS)*. She was the Cochair for the Data Fusion Technical Committee of the IEEE GRSS from 2009 to 2013, Chair for the Remote Sensing and Mapping Technical Committee of International Association for Pattern Recognition from 2010 to 2014, and General Chair for the Fourth IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing held at Shanghai, China, in 2012. She was an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATION AND REMOTE SENSING, *Journal of Applied Remote Sensing*, and IEEE SIGNAL PROCESSING LETTERS. From 2016 to 2020, she was the Editor-in-Chief of the *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*. She is a Fellow of SPIE-International Society for Optics and Photonics (SPIE). She is currently a Member of the IEEE Periodicals Review and Advisory Committee and SPIE Publications Committee.

Junjie Wang received the B.Sc. degree in computer science in 2018 from the Ocean University of China, Qingdao, China, where he is currently working toward the M.Sc. degree in computer science and applied remote sensing with the School of Information Science and Technology, Ocean University of China.

His research interests include computer vision and remote sensing image processing.

Feng Gao (Member, IEEE) received the B.Sc. degree in software engineering from Chongqing University, Chongqing, China, in 2008, and the Ph.D. degree in computer science and technology from Beihang University, Beijing, China, in 2015.

He is currently an Associate Professor with the School of Information Science and Engineering, Ocean University of China, Qingdao, China. His research interests include remote sensing image analysis, pattern recognition, and machine learning.

Heng-Chao Li (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from Southwest Jiaotong University, Chengdu, China, in 2001 and 2004, respectively, and the Ph.D. degree from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2008, all in information and communication engineering.

He is currently a Full Professor with the School of Information Science and Technology, Southwest Jiaotong University. His research interests include statistical analysis of synthetic aperture radar (SAR) images, remote sensing image processing, and pattern recognition.

Dr. Li is an Editorial Board Member of the *Journal of Southwest Jiaotong University* and *Journal of Radars*. He was the recipient of the 2018 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society for his service to the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. Moreover, he was the Guest Editor for the Special Issues of the *Journal of Real-Time Image Processing*, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING PUBLICATION INFORMATION, and the IEEE JOURNAL ON MINIATURIZATION FOR AIR AND SPACE SYSTEMS, a Program Committee Member for the 26th International Joint Conference on Artificial Intelligence (IJCAI-2017) and 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MULTITEMP-2019), and the Session Chair for the 2017 International Geoscience and Remote Sensing Symposium (IGARSS-2017) and 2019 Asia-Pacific Conference on Synthetic Aperture Radar (APSAR-2019). He is an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING PUBLICATION INFORMATION.