

A Deep Learning Method for Change Detection in Synthetic Aperture Radar Images

Yangyang Li^{ID}, Senior Member, IEEE, Cheng Peng, Yanqiao Chen^{ID}, Student Member, IEEE,
Licheng Jiao, Fellow, IEEE, Linhao Zhou, and Ronghua Shang^{ID}, Member, IEEE

Abstract—With the rapid development of various technologies of satellite sensor, synthetic aperture radar (SAR) image has been an import source of data in the application of change detection. In this paper, a novel method based on a convolutional neural network (CNN) for SAR image change detection is proposed. The main idea of our method is to generate the classification results directly from the original two SAR images through a CNN without any preprocessing operations, which also eliminate the process of generating the difference image (DI), thus reducing the influence of the DI on the final classification result. In CNN, the spatial characteristics of the raw image can be extracted and captured by automatic learning and the results with stronger robustness can be obtained. The basic idea of the proposed method includes three steps: it first produces false labels through unsupervised spatial fuzzy clustering. Then we train the CNN through proper samples that are selected from the samples with false labels. Finally, the final detection results are obtained by the trained convolutional network. Although training the convolutional network is a supervised learning fashion, the whole process of the algorithm is an unsupervised process without priori knowledge. The theoretical analysis and experimental results demonstrate the validity, robustness, and potential of our algorithm in simulated and real data sets. In addition, we try to apply our algorithm to the change detection of heterogeneous images, which also achieves satisfactory results.

Index Terms—Change detection, convolutional neural network (CNN), spatial fuzzy clustering, synthetic aperture radar (SAR) images.

I. INTRODUCTION

IN the field of remote sensing, the change detection is used to detect changes in the same place on the surface of the Earth at different times. Synthetic aperture radar (SAR) images

Manuscript received August 3, 2017; revised February 9, 2018, June 18, 2018 and December 19, 2018; accepted February 3, 2019. This work was supported by the National Natural Science Foundation of China under Grant 61772399, Grant U1701267, Grant 61773304, Grant 61672405, and Grant 61772400, in part by the Program for Cheung Kong Scholars and Innovative Research Team in University under Grant IRT_15R53, in part by the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) under Grant B07048, in part by the Major Research Plan of the National Natural Science Foundation of China under Grant 91438201, and in part by the Technology Foundation for Selected Overseas Chinese Scholar in Shaanxi under Grant 2017021 and Grant 2018021. (*Corresponding authors:* Cheng Peng; Yanqiao Chen)

The authors are with the Key Laboratory of Intelligent Perception and Image Understanding, Ministry of Education, International Research Center for Intelligent Perception and Computation, Joint International Research Laboratory of Intelligent Perception and Computation, School of Artificial Intelligence, Xidian University, Xi'an 710071, China (e-mail: pengchengmaster@163.com; chenyanqiao2016@163.com).

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

Digital Object Identifier 10.1109/TGRS.2019.2901945

have an important research value because of their independence on atmospheric and sunlight conditions. The increasing availability of SAR images offers unprecedented convenience and opportunities for the following research directions such as environmental monitoring [1], [2], urban research [3], [4], disaster assessment [5], forest resource monitoring [6], and so on. However, the microwave imaging mechanism of SAR images makes the background information mostly complicated, and the features of the region are mixed together. These difficulties are reflected in the structural sensitivity, imaging geometric distortion, imaging system interference, speckle noise, and so on. In the field of single polarization SAR image change detection, the most difficult problem is how to overcome the speckle noise. The stochastic backscattering of the basic unit of SAR imaging system [7], [8] makes the phase angle lost continuity, which shows a granular signal strength related distortion in images. The intensity distortion associated with this granular signal is speckle noise. Speckle noise is appended to the original image in the form of multiplication, which seriously affects the interpretation of SAR images. Thus, how to effectively suppress speckle noise is still a challenging problem in the field of SAR image change detection.

In recent years, the change detection of SAR images is usually based on three steps [9]: 1) image preprocessing; 2) generating a difference image (DI); and 3) analyzing the DI. The preprocessing stage is designed to allow two SAR images to have consistent comparability in both the spatial and spectral domains [9], which is the basis for the latter algorithm. The focus of SAR image change detection is generating and analyzing the DI. The main purpose of these two steps is to better suppress the influence of speckle noise. The commonly used algorithms for generating DI are difference operator and ratio operator. The difference operator is the direct difference calculation for two SAR images. However, the imaging mechanism of SAR images is very different from that of optical images. The difference operator is effective for optical images, but the effect on SAR images is not significant. Since the model of speckle noise inherent in SAR images is multiplicative random noise, the difference operator cannot effectively suppress the speckle noise. Compared with the difference operator, the ratio operator is the ratio operation for two SAR images. It can better overcome the disadvantage of sensitivity to multiplicative noise. Nevertheless, the ratio method does not consider further information about the local, edge, and class conditional distributions of the image, so there is still room for improvement. Based on the ratio operator,

the logarithmic operator and mean operator are widely used. The nature of the logarithmic operator [10] itself can reduce the greater difference caused by the ratio calculation. It can further reduce the effect of the independent points of the background part in the unchanged class, but its ability to preserve the edge information of the two types of regions is weak. The mean operator [11] takes into account the spatial neighborhood information and has a very good inhibitory effect on the independent points.

The analysis of the DI can be seen as a process of image segmentation [12]. The most common methods are the threshold method and the clustering method. The threshold method is used to find the optimal threshold through a threshold selection method, and the pixels of DI are divided into changed class and unchanged class through the threshold. Unsupervised threshold methods include the Kittler & Illingworth (KI) algorithm [13] and the expectation maximization algorithm [14]. Both the two methods need to first establish a model to fit the class condition distribution of the unchanged class and the changed class. Finally, the minimum error rate criterion of Bayes is used to make the posterior probabilities of the two classes of distributions equal to the selected optimal threshold. On the basis of KI algorithm, a generalized KI (GKI) threshold selection algorithm [15] is introduced, which uses more model information than KIs index function. Inspired by the GKI algorithm, a histogram optimization method [16] has been developed to solve the difficulty of GKI processing when the histogram is unimodal. Su *et al.* [17] proposed a local fitting and semiexpectation maximization [locally fitting and semi-electromagnetic (EM)] algorithm to solve the tedious problem of initialization and iterative computation of EM algorithm. Although the threshold method is fast, convenient, and concise, it is not accurate enough. Minor differences in thresholds can lead to greater errors in the final results. Another method for image segmentation is clustering, which does not need to be modeled. It is more flexible than the threshold method. The fuzzy clustering algorithm can retain more information than the hard clustering algorithm and has a wider range of applications. Krinidis and Chatzis [18] proposed the local neighborhood information fuzzy C-means (FCM) clustering method (FLICM). In this method, neighborhood information is added to the cluster for reference. To improve the method, Gong *et al.* [19] proposed the improvement of FLICM reformulated fuzzy local-information C-means algorithm (RFLICM), which can be better used to solve the change detection of SAR images. Clustering method does not need to build a model, which is more flexible than the threshold method, while it is sensitive to noise because of insufficient consideration of spatial information. With the data acquisition channel and application range increasing, the above-mentioned methods cannot meet the demand well.

In recent years, the deep neural networks have gained great attention in the image understanding field due to their powerful ability to automatically learn complicated and relevant contextual features, which is mainly to construct the deep structure to learn the multilevel expression [20], [21]. The classic deep learning models include deep belief network (DBN) [22],

deep autoencoder [23], [24], convolutional neural network (CNN) [25], and the generative adversarial network [26]. DBN and deep autoencoder are unsupervised learning algorithms that can be trained by unlabeled training sets and then tuned with a small amount of tagged data [24]. CNNs have unique advantages in speech recognition and image processing [27] because of their special structure shared by local weights. Some scholars have applied deep learning to the problem of SAR image change detection and achieved good results. The method of combining threshold method with DBN is adopted in [28]. The threshold method is first used to convert the DI into the binary image. The training sample is then pulled into a column vector for training in the DBN. The method based on a threshold and fuzzy clustering is used to obtain the binary image in [29]. Similarly, the training sample is pulled into a column vector for training in the DBN. The above-mentioned deep learning methods are trained by pulling the training samples into column vectors, so that the spatial information of the samples is neglected, and the spatial features of the samples cannot be learned.

In this paper, we propose a novel method for SAR image change detection method based on CNN. The proposed method includes the following steps: 1) generating false labels by spatial fuzzy clustering; 2) generating training samples and select suitable training samples; 3) training the CNN to get a good classifier; and 4) using the trained CNN for the classification of changed and unchanged pixels. The task of change detection is often applied to disaster assessment and other related fields, where the prior knowledge is difficult to obtain, so it is clear that there is an urgent need to design unsupervised methods for this task. We use false labels to train the CNN, ingeniously convert a supervised training process into an unsupervised learning process, and apply it well to change detection problems. For the traditional methods, the raw images need to be filtered and other preprocessing operations. Although our method eliminates the tedious preprocessing process and does not require any filtering operations on the SAR images. The final maps can be obtained directly by the trained CNN. In addition, our method skips the stage of generating a DI, which can avoid the influence of generating DI on the results. Making full use of the spatial feature and neighborhood information of pixels, the unique structure of the CNN can learn the hierarchical features of the images and implement an end-to-end framework.

The rest of this paper is organized as follows. Section II shows the problem description and the proposed algorithm framework. In Section III, we provide a detailed introduction to the proposed algorithm. Section IV deals with the experiments on simulated SAR images and real multitemporal SAR images to verify the feasibility of our algorithm. Finally, the conclusions are presented in Section V.

II. PROBLEM DESCRIPTION AND ALGORITHM FRAMEWORK

Consider two coregistered intensity SAR images

$$I_1 = \{I_1(i, j), 1 \leq i \leq A, 1 \leq j \leq B\} \quad (1)$$

$$I_2 = \{I_2(i, j), 1 \leq i \leq A, 1 \leq j \leq B\}. \quad (2)$$

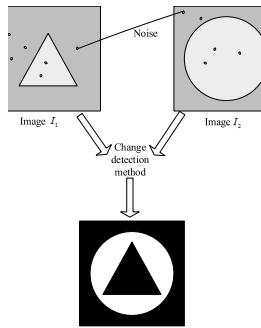


Fig. 1. Change detection problem.

The two SAR images have the same size $A \times B$. They are obtained from the same geographical location at different moments t_1 and t_2 . The two SAR images are contaminated by coherent noise, and the basic SAR image change detection problem can be described as shown in Fig. 1. We need to design efficient change detection algorithms to find areas of change in two SAR images.

Change detection is usually done in three steps: 1)image preprocessing; 2)generating a DI; and 3) analyzing the DI. The preprocessing stage is designed to allow two SAR images to have consistent comparability in both the spatial and spectral domains [9], which is the basis for the latter algorithm. In the first step, traditional methods still need to consider some noise reduction methods, such as median filtering [30], mean filtering [31], Gauss filtering [32], and so on. For most methods, filtering is essential and the original SAR image needs to be filtered to reduce the speckle noise [33]. In our approach, we do not need to do filtering operations and even do not need to generate the DI. The final maps can be obtained directly from the two original SAR images. The CNNs can learn complex relationships from two original SAR images to enable us to achieve the purpose of classification.

Change detection problems are often used in disaster assessment and environmental monitoring. In most cases, it is difficult to obtain priori knowledge or accurate annotation for the raw images in these practical applications. It is clear that there is an urgent need to design unsupervised methods for this task. Although training the CNN is a supervised learning fashion, we obtain the training samples by the unsupervised fuzzy clustering. Therefore, the process of the whole algorithm can be regarded as a completely unsupervised process. The flowchart of our method is shown in Fig. 2, we first generate false labels through the spatial fuzzy clustering algorithm. Then we select the appropriate samples to train the CNN. After the network training is completed, all the samples are fed into CNN to obtain the classification results (the changed class and the unchanged class).

III. METHODOLOGY

A. Preclassification

In order to preclassify two SAR images and generate false labels, we first need to compare the similarity between the two SAR images. We compute the similarity between the two

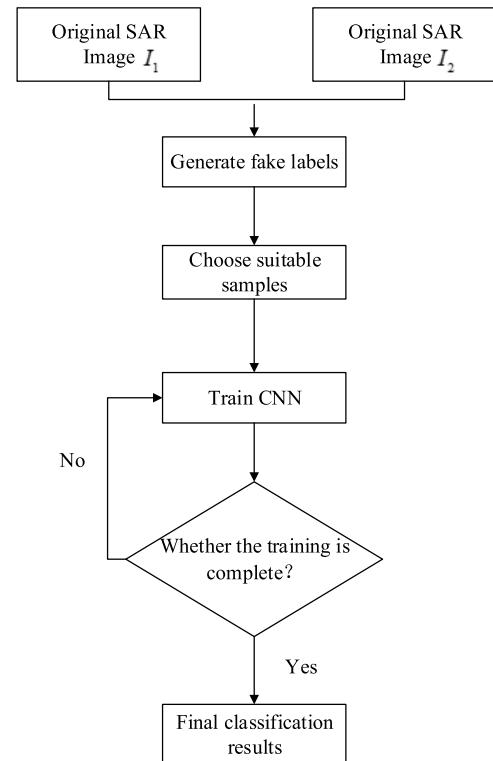


Fig. 2. Flowchart of the proposed method.

images by the similarity matrix

$$S_{ij} = \frac{|I_{ij}^1 - I_{ij}^2|}{I_{ij}^1 + I_{ij}^2} \quad (3)$$

where $S_{ij} \in [0, 1]$, I_{ij}^t represents the gray value of the position (i, j) in the images on the t moment. This is similar to generating a DI. However, the similarity matrix is used only in the preclassification stage. In the final classification process, the similarity matrix of two images is not used, but the original two SAR images are used. After the similarity matrix is generated, we use spatial fuzzy clustering algorithm to preclassify the similarity matrix so that the false labels could be obtained. The FCM clustering algorithm was first introduced by Dunn [34] and later extended by Bezdek [35], which uses the degree of membership to determine the extent of each data belonging to the clustering center. This algorithm is an iterative clustering method that produces an optimal c partition by minimizing the weighted within group sum of squared error objective function [35]

$$J_{\text{FCM}} = \sum_{j=1}^N \sum_{i=1}^c u_{ij}^m \|x_j - v_i\|^2. \quad (4)$$

The fuzzy clustering divides the N vectors into c fuzzy sets and finds the clustering centers of each set, which makes the cost function of the nonsimilarity indexes reaches the minimum. $X = (x_1, x_2, \dots, x_N)$ represents an image with N pixels, and the N pixels are divided into c cluster centers, where x_i represents the intensity values of the i pixels. c is the number of clusters with $2 \leq c < N$. Since the task of the

change detection eventually falls into two categories (changed classes and unchanged classes), so c is set 2. v_i stands for the i clustering center and u_{ij} represents the fuzzy membership relation between the pixel x_j and the i cluster center. $\|\cdot\|$ stands for the Euclidean distance, and m is the weighting exponent on each fuzzy membership. In this paper, we make $m = 2$. FCM randomly initializes the clustering center and updates u_{ij} and v_i to minimize the cost function.

FCM is flexible for initial condition when the original data have a good separation of clusters [36]. However, the fuzzy clustering is sensitive to speckle noise because it does not take into account spatial information and cannot effectively suppress the speckle noise. In terms of the pixels in images, the adjacent elements, which have similar intensities and characteristic value, are highly likely to be assigned into the same cluster center [37]. Inspired by this principle, we introduce the spatial information into fuzzy clustering to effectively suppress the speckle noise in SAR images. The spatial function is defined as follows:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik} \quad (5)$$

where $NB(x_j)$ stands for the adjacent elements of x_j , and h_{ij} stands for the possibility that x_j belongs to the i cluster center. By adding h_{ij} to the fuzzy membership, the formula is obtained as

$$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{kj}^p h_{kj}^q} \quad (6)$$

where p and q , respectively, determine the initial fuzzy membership degree and the relative weight of the spatial function. The specific methods of the spatial fuzzy clustering algorithm are as Algorithm 1.

After the spatial fuzzy clustering operation is applied to the similarity matrix, the corresponding false label matrix will be obtained. The false label matrix is a binary matrix. The two values in the matrix represent the changed class and the unchanged class, respectively. However, not all the false labels are accurate, and we need to select some suitable samples for training the CNNs. Before that, we need to generate training samples.

B. Generate Training Samples

For change detection problem of SAR images, the original samples are two SAR images, which are typically too large to be processed in only one pass through a CNN model. We need to convert the two SAR images into many smaller images of the same size, and the conversion method is shown in Figs. 3 and 4.

As shown in Fig. 3, there are two simulated SAR images and each square in the picture represents each pixel point in the image. Before generating the sample, we first determine a neighborhood size. For example, we assume that the neighborhood size is 3×3 . Then we start with the first pixel of the first SAR image which is point I_{11}^1 as shown in Fig. 3(a). We crop the pixel I_{11}^1 and its neighborhood pixels in the first

Algorithm 1 SFCM

1. Provide the initial values for the centroids v_i where $i = 1, \dots, c$.
2. Compute the membership function as follows:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_j - v_k\|}{\|x_j - v_i\|} \right)^{2/(m-1)}}, i = 1, \dots, c \text{ and } j = 1, \dots, N$$

3. Calculate the spatial function as:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}, i = 1, \dots, c \text{ and } j = 1, \dots, N$$

4. Compute the new membership function which incorporates the spatial function as:

$$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{kj}^p h_{kj}^q}, i = 1, \dots, c \text{ and } j = 1, \dots, N$$

5. Set $u_{ij} = u'_{ij}$, for all $j = 1, \dots, N$ and $i = 1, \dots, c$

6. Calculate the new centroids as follows:

$$v_i = \sum_{j=1}^N u_{ij}^m x_j / \sum_{j=1}^N u_{ij}^m \text{ for } i = 1, \dots, c$$

7. If $|u_{ij}(\text{new}) - u_{ij}(\text{old})| < \epsilon$, then stop, otherwise go to step 2.
-

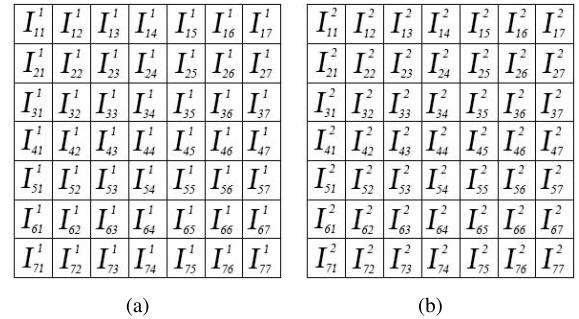


Fig. 3. Simulated SAR images. (a) First SAR image. (b) Second SAR image.

SAR image to form a smaller image whose size is 3×3 . For pixels located at the edge, we supplement 0 to their neighborhood. For example, there are no pixels on the top and left of the pixel I_{11}^1 shown in Fig. 3; therefore, we add 0 to the left and top of it. Similarly, for the second SAR image, the corresponding small size picture is obtained in the same way, and then the two extracted pictures are stacked together. As shown in Fig. 4, a training sample is obtained. The label for this newly generated sample is replaced by the label for I_{11}^1 . Next, the pixels I_{12}^1 and I_{12}^2 and their neighborhood pixels are taken in the same way, then we stack these two small images and obtain the second training sample. The operation is repeated until all the pixels are traversed. Fig. 4 shows the process of generating samples at several different locations. According to this conversion method, if the number of pixels in the SAR image is 500, we will get 500 samples.

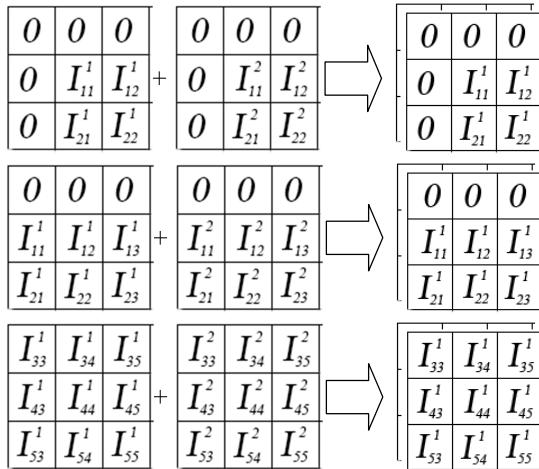


Fig. 4. Example to illustrate how to generate samples.

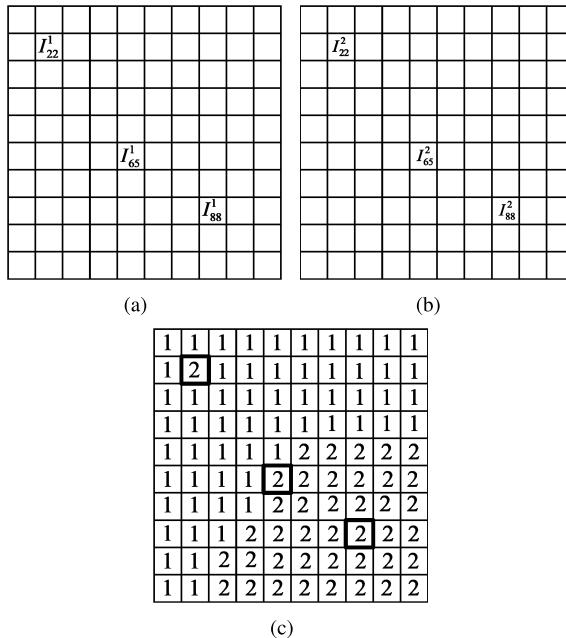


Fig. 5. Example to illustrate how to choose suitable samples. (a) First simulated SAR image. (b) Second simulated SAR image. (c) False label matrix.

C. Sample Selection

The results of preclassification are not entirely correct. We need to select some suitable samples, where the pixels have higher probability of being correctly classified to train the CNN. The method of sample selection is presented as follows.

Suppose that I_{ij}^1 represents the pixel point on the position of (i, j) on the first SAR image and that I_{ij}^2 represents the pixel point on the position of (i, j) on the second SAR image. BI represents the false label matrix generated in step A. The false labels corresponding to the two-pixel points I_{ij}^1 and I_{ij}^2 are $BI(i, j)$ in BI . N_{ij}^1 represents the pixels that consist of the pixel I_{ij}^1 and its neighbor pixels. The neighborhood size is $n \times n$. As shown in Fig. 5, the pixel I_{65}^1 represents the intensity of the pixel in the sixth row and fifth column in

the first SAR image. The pixel I_{65}^2 represents the intensity of the pixel in the sixth row and fifth column in the second SAR image. N_{65}^1 and N_{65}^2 are stacked to form a training sample, and the corresponding label is the value in row 6 and column 5 of the BI matrix. The numbers 1 and 2 in the false label matrix represent two types of labels (changed areas and unchanged areas) obtained by the spatial FCM (SFCM) algorithm. Then we need to select some appropriate samples as the training samples. For instance, the label of N_{65}^1 and N_{65}^2 is 2, which is about half the same as the adjacent label; therefore, the stacked image can be sent as a sample into the CNN for training. For the points N_{88}^1 and N_{88}^2 , the label is 2 and the adjacent labels are basically the same, thus the image stacked by N_{88}^1 and N_{88}^2 can also be selected as a sample for training. In terms of the points N_{22}^1 and N_{22}^2 whose label is 2, while the label is basically different from the adjacent labels, this kind of point should be eliminated. In short, we use the following formula to determine whether the sample points are selected:

$$\frac{Q(BI_{\varepsilon\eta} = BI(i, j))}{n \times n} > \alpha \quad (7)$$

where $BI_{\varepsilon\eta}$ stands for the corresponding label value of N_{ij}^m ($m = 1, 2$). $Q(BI_{\varepsilon\eta} = BI_{ij})$ represents the number of pixels in N_{ij}^m ($m = 1, 2$) that have the same labels as I_{ij}^m ($m = 1, 2$). The coefficient α determines whether the N_{ij}^m ($m = 1, 2$) is selected as a sample. Through this formula, we can obtain a batch of training samples with false labels to train the CNN.

D. Convolutional Neural Network Establishment

Training a CNN is the core step of the whole algorithm. CNN is an important deep learning architecture for image understanding, which is inspired by the natural visual perception mechanism of biology. Considered from the aspect of feature extraction, CNN is a hierarchical network structure which can obtain an effective representation of the original image with very little preprocessing [27], [38]. Generally, the basic structure of CNN consists of two types of layers [39]. One is the feature extraction layer, the input of each neuron is connected with the local accepted domain of the previous layer and the feature of the part is extracted. Once the local feature is extracted, the position relation between it and other features is also determined. The second is the feature mapping layer, and each computing layer of the network consists of multiple feature maps. Sigmoid function is used as the activation function in the feature mapping structure, which makes it easier to express the entire network, otherwise, the stacking of several linear operating layers can only function as a linear mapping and cannot form a complex function. Each convolution layer is usually followed by a pooling layer, which is intended to simplify the information from the convolution layer and reduce the size of the feature map. CNN is primarily applied to identify displacement, scaling, and other forms of distortion-invariant 2-D graphics [25]. Since the feature detection layer of CNN is trained by training samples, it avoids explicit feature extraction while using CNN, and implicitly learns from the training data. Moreover, since the weights of neurons on the same feature mapping surface are the same,

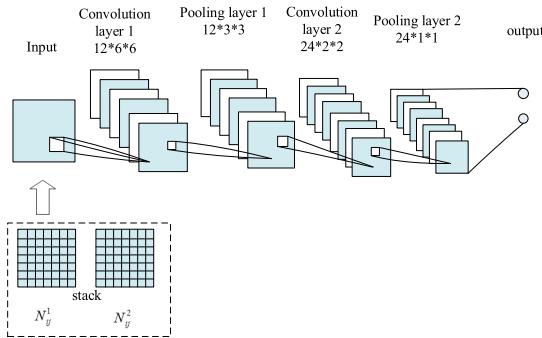


Fig. 6. Structure of the CNN.

the network can learn in parallel, which is also a major advantage of CNNs relative to other structures. The special structure of CNN with its local weight sharing has a unique advantage in speech recognition and image processing. Its layout is closer to the actual biological neural network. Weight sharing reduces the complexity of the network, especially the image of the multidimensional input vector can be input directly to the network, which avoids the complexity of data reconstruction in the process of feature extraction and classification.

The structure of CNN for change detection is shown in Fig. 6. The first layers are the alternation of the convolution layer and the pooling layer. Finally, the softmax layer classifies the results into two classes (the changed class and the unchanged class). For example, the previously selected samples N_{ij}^1 and N_{ij}^2 are stacked to form an image with two channels and fed into the CNN. After several iterations of training, the trained network is obtained. Then, all samples (selected and unselected) are fed into the network for classification (changed class and unchanged class) and the final detection results can be obtained.

IV. EXPERIMENTAL STUDY

In this section, we first conduct an experiment on a simulated SAR image data set to demonstrate the performance of the proposed method for suppressing the noise. Then two real data sets (the Ottawa data set and the Farmland C data set) are used to prove the effectiveness of the samples obtained by preclassification, also for comparison with a supervised deep learning algorithm. Furthermore, we test our method on six real data sets and conduct six group comparison experiments with other methods, including a clustering algorithm SFCM, a threshold algorithm GKI [15], and two deep learning algorithms [28], [29] based on DBN.

A. Introduction to Data Sets

The first data set contains SAR images with a different equal number of looks (ENL). In detail, the data set consists of five pairs of simulated SAR images with ENL = 1, 2, 3, 4, and 5, respectively. ENL indicates the level of noise influence, and a higher value means a smaller degree of influence. As shown in Fig. 7, we select two pairs of simulated images for comparison experiments.

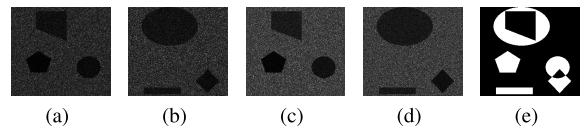


Fig. 7. Simulated data set with two pairs of SAR images with different ENLs. (a) and (b) Pairs of SAR images with ENL = 2. (c) and (d) Pairs of SAR images with ENL = 3. (e) Reference image.

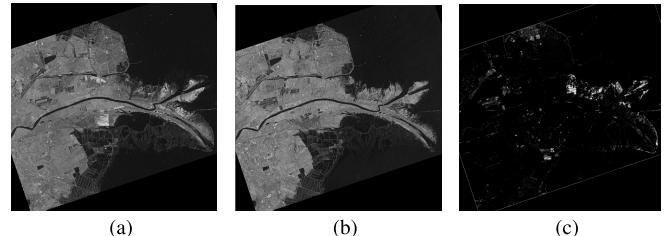


Fig. 8. Multitemporal images relating to Yellow River Estuary. (a) Image acquired in June 2008. (b) Image acquired in June 2009. (c) Ground truth.

We select six real data sets to verify the effectiveness of our algorithm. The first three data sets are taken from the Yellow River data set, as shown in Fig. 8. The Yellow River data set was taken by the Radarsat-2 satellite in June 2008 and June 2009, and the Yellow River data set is 7666×7692 in size, reflecting changes in the region of the Yellow River estuary from 2008 to 2009. Since the size of the Yellow River data set is large, we choose three representative regions from the Yellow River data set, which are the Inland water data set, the FarmLand C data set, and the FarmLand D data set. The corresponding sizes are 291×444 , 306×291 , and 257×289 . The fourth data set is the Ottawa data set, which was taken by RADARSAT-SAT satellite in May 1997 and August 1997, with a resolution of 12 m and an image size of 290×350 . The fifth data set is Taiwan Shimen reservoir upstream data set, the size is 400×400 . This data set was taken in August 2004 and September 2004 by the FORMOSAT satellite. It is a response to the changing situation before and after Aere typhoon. The major change is ground collapse and a large number of bare lands. The last comparison experiment is conducted on a data set consists of an SAR image and an red-green-blue (RGB) optical image, which is our attempt at heterogeneous change detection problems. The image is a farm in Shuguang village, Dongying City, China. The size of the data set is 500×500 . It was obtained in June 2008 and September 2012, respectively. These raw images of these six data sets and their ground truth are shown in Figs. 9–14.

B. Evaluation Criteria and Experimental Setting

The evaluation criteria for change detection generally include false negatives (FN), which represents the changed pixels that undetected, and false positives (FP), which represents the unchanged pixels wrongly detected as changed. The total number of errors is represented by overall error (OE), and the formula is as follows [40]:

$$OE = FN + FP. \quad (8)$$

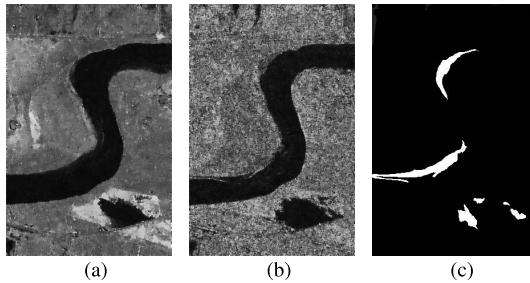


Fig. 9. Multitemporal images relating to Inland water of Yellow River Estuary. (a) Image acquired in June 2008. (b) Image acquired in June 2009. (c) Ground truth

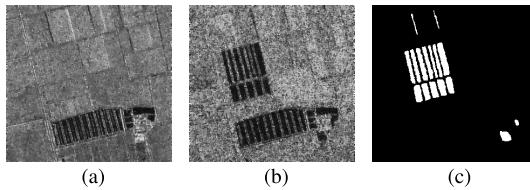


Fig. 10. Multitemporal images relating to Farmland C of Yellow River Estuary. (a) Image acquired in June 2008. (b) Image acquired in June 2009. (c) Ground truth.

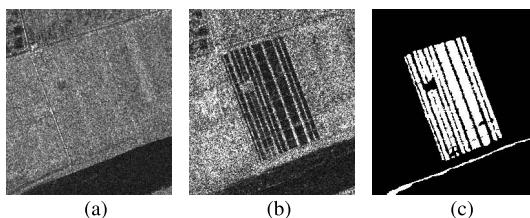


Fig. 11. Multitemporal images relating to Farmland D of Yellow River Estuary. (a) Image acquired in June 2008. (b) Image acquired in June 2009. (c) Ground truth.

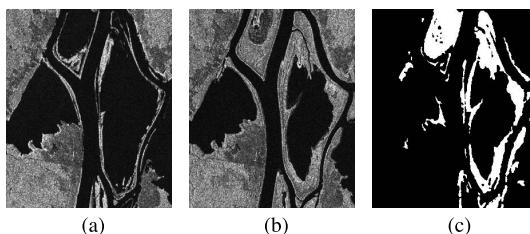


Fig. 12. Multitemporal images relating to Ottawa. (a) Image acquired in July 1997. (b) Image acquired in August 1997. (c) Ground truth.

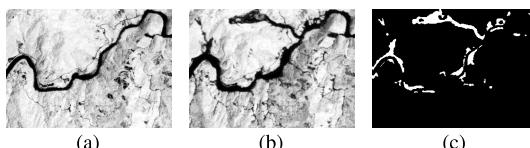


Fig. 13. Multitemporal images relating to Taiwan Shimen. (a) Image acquired in August 2004. (b) Image acquired in September 2004. (c) Ground truth.

The correct rate of the classification is represented by percentage correct classification (PCC), and the PCC formula is as follows:

$$PCC = (TP + TN) / (TP + FP + TN + FN) \quad (9)$$

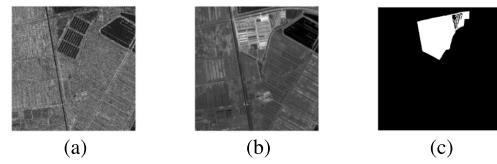


Fig. 14. Multitemporal images relating to Shuguang village. (a) SAR image acquired in 2008. (b) Optical image acquired in 2012. (c) Ground truth.

where TP is the true positives, which represents the number of pixels that are correctly detected as the changed area, and TN is the true negatives, which represents the number of pixels that are correctly detected as the unchanged area. For accuracy assessment, Kappa statistic is a measure of accuracy or agreement based on the difference between the error matrix and the change agreement [41]. Kappa coefficients are calculated as follows:

$$\text{Kappa} = \frac{PCC - PRE}{1 - PRE} \quad (10)$$

where

$$PRE = \frac{(TP + FP) \times (TP + FN) + (FN + TN) \times (FP + TN)}{(TP + TN + FP + FN)^2}. \quad (11)$$

In the experiment, we use the structure of the CNN as shown in Fig. 6. The parameters of the CNN model are as follows. There are two convolution layers and two pooling layers. The size of convolution kernel is 2×2 and the stride size is 1. Zero padding of 1 is used in the input layer, and then the sample size of the input layer is 7×7 . Both the two pooling layers are mean-pooling layers with a kernel size of 2×2 . The size of feature maps output from the first convolution layer is 6×6 , and the size of feature maps output from the first pooling layer is 3×3 . The size of feature maps output from the second convolution layer is 2×2 , and the size of feature maps output from the second pooling layer is 1×1 . The parameter of α controls the selection of the sample, where we set α to 0.6. n stands for the size of the neighborhood in the sample selection, where we set it to 5. Then expand 0 at the outermost level of the sample and get the final input sample size of 7×7 . The number of iterations of CNN is 5.

In terms of comparison experiment, a clustering algorithm SFCM, a threshold algorithm GKI [15], and two deep learning algorithms [28], [29] based on DBN are presented as comparative methods, respectively.

C. Experiment Results and Analysis

1) Results on the Simulated Data Set: The results of the simulated data set are shown in Figs. 15 and 16. We compare the proposed method with GKI and SFCM. Fig. 15(a) and (b) shows the detection results of GKI algorithm with ENL = 2 and 3, respectively, and Fig. 15(c) and (d) shows the results of SFCM algorithm with ENL = 2 and 3, respectively. From the detection results of these two traditional algorithms, whether the ENL value is 2 or 3, there are too many noise points in the image, and the noise distribution is extensive, which indicates that the traditional algorithm cannot cope well with

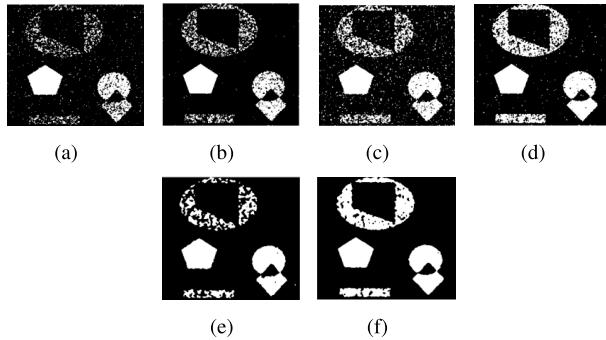


Fig. 15. Results on the simulated data set. (a) and (b) Results of GKI algorithm with ENL = 2 and 3, respectively. (c) and (d) Results of SFCM algorithm with ENL = 2 and 3, respectively. (e) and (f) Results of our method with ENL = 2 and 3, respectively.

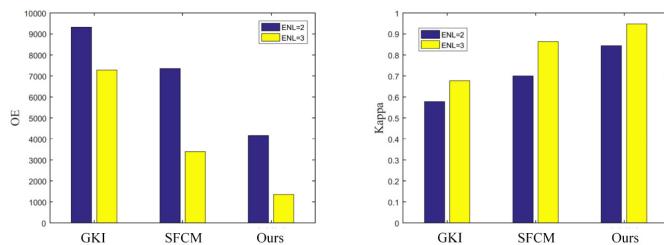


Fig. 16. Relationship between change detection results and ENL. (a) Values of OE. (b) Values of Kappa.

the speckle noise. The results of our algorithm are shown in Fig. 15(e) and (f), and the noise spots are obviously reduced, especially for the simulation image with ENL = 3. Except for a small number of noise points existing in the elliptical region, there are basically no noise points in other regions. As shown in Fig. 16, we exhibit the values of OE and Kappa at different ENL values in the form of a histogram. There are three groups of results, GKI, SFCM, and our method, and in each group, blue column shows the evaluation of result with ENL=2 and yellow column shows the evaluation of result with ENL=3. Whether the ENL value is 2 or 3, we find that the proposed method performs much better than GKI and SFCM, especially for the OE. This experiment demonstrates that the proposed method has a better performance than GKI (threshold method) and SFCM (clustering method) in restraining the impact of noise. On the other hand, the results also show that our method avoids the tedious procedure of preprocessing and does not require filtering operations on the raw images.

2) Results Under Supervised Learning and Unsupervised Learning: In these two comparison methods, training and testing are implemented on two deep neural networks with the same topology. Note that, in the proposed algorithm, the samples used for training are the pixels that satisfy the rules for selecting samples, rather than all pixels in the image. In the supervised network, we use the same training set to ensure fairness. However, the training samples have different labels in different methods. In the supervised network, the labels are given according to the ground truth, while the labels in the proposed algorithm are given by the procedure of preclassification.

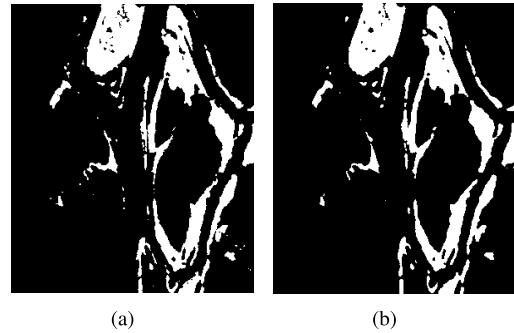


Fig. 17. Change detection results of Ottawa data set achieved by (a) supervised learning and (b) unsupervised learning.

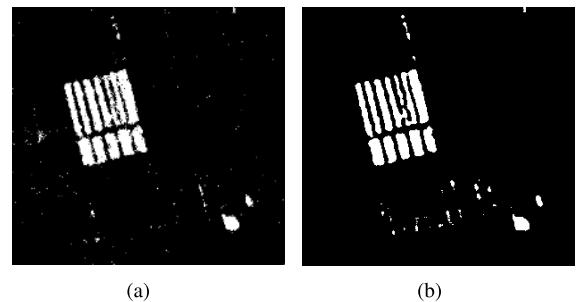


Fig. 18. Change detection results of Farmland C data set achieved by (a) supervised learning and (b) unsupervised method.

Figs. 17 and 18 show the detection results on the Ottawa data set and the Farmland C data set. From the visual results, we can find that the proposed algorithm achieves good performance, even close to the results of the supervised method, with only a few pixels being detected incorrectly. In addition, the numerical results of the two comparison methods are shown in Table I. From the perspective of various evaluation criteria, the results of the proposed algorithm have reached a satisfactory level of performance. Compared with the results of the supervised method, the FN of both data sets even achieved better results. Through this experiment, we can conclude that the proposed algorithm can achieve good performance compared to the supervised method. For the task of change detection which always lacks priori knowledge and labeled data, unsupervised learning is necessary and important.

3) Results on the Real Data Sets: The change detection results obtained by the proposed algorithm and the four comparative methods on the Inland water data set are shown in Fig. 19. The numerical results of evaluation criteria are listed in Table II. The SFCM method takes spatial information into account, while it still fails to obtain satisfactory results due to its sensitivity to noise. Although the operation of the GKI algorithm is simple, the result map obtained by GKI is rough. It can be seen from Fig. 19(b) that still exist a lot of white noise spots, because an optimal threshold should be concerned in the process of modeling. Any suboptimal selection of threshold may result in noise points on the final map. For the results generated by the methods based on deep neural network, we can find that there is an obvious performance boost. The deep neural network can extract meaningful features from the

TABLE I

CHANGE DETECTION RESULTS OF OTTAWA DATA SET AND FARMLAND C DATA SET OBTAINED BY SUPERVISED AND UNSUPERVISED LEARNINGS

Dataset	Criteria	Method	
		Supervised learning	Unsupervised learning
Ottawa	<i>FP(%)</i>	0.48	0.56
	<i>FN(%)</i>	0.78	0.77
	<i>OE(%)</i>	1.12	1.33
	<i>PCC(%)</i>	98.88	98.67
	<i>Kappa(%)</i>	95.45	95.00
Farmland C	<i>FP(%)</i>	0.40	0.41
	<i>FN(%)</i>	1.04	1.00
	<i>OE(%)</i>	1.35	1.41
	<i>PCC(%)</i>	98.64	98.59
	<i>Kappa(%)</i>	87.65	87.09

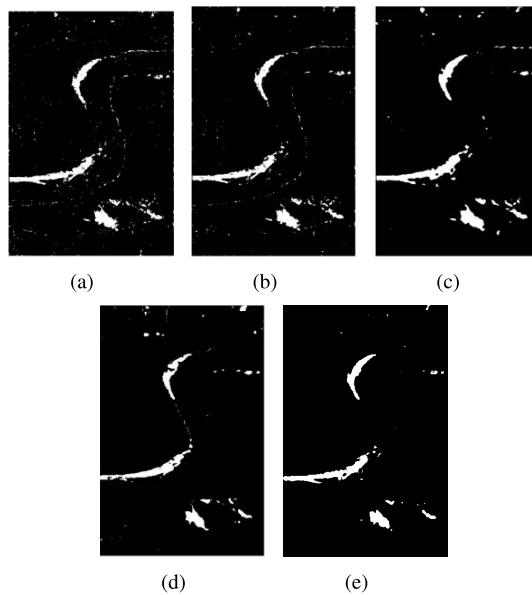


Fig. 19. Change detection results of Inland water data set achieved by (a) SFCM, (b) GKI, (c) Threshold+DBN, (d) JFCM+DBN, and (e) proposed method.

raw image and suppress the speckle noise. On the other hand, we can see from Table II that the proposed method based on convolutional network achieves better performance compared with the results obtained by the method based on DBN. We can infer that the spatial information is erased when the feature maps are flattened into a 1-D vector in the layers of DBN. Although the spatial and neighborhood information are preserved during the process of convolutional operation.

For the Farmland C data set, Fig. 20 shows the final maps of the five methods and Table III lists the numerical results. As expected, GKI exhibits the worst performance. It can be seen from Fig. 20(b) that a large amount of misclassified white noise exists in the lower region. For the three methods based on deep learning, both joint classifier based on FCM (JFCM)+DBN and our method achieve satisfactory results,

TABLE II

CHANGE DETECTION RESULTS OF INLAND WATER DATA SET

Criteria \ Method	FP(%)	FN(%)	OE(%)	PCC(%)	Kappa(%)
SFCM	1.39	0.73	2.12	97.88	69.70
GKI	1.06	0.77	1.82	98.18	72.55
Threshold+DBN	1.27	0.52	1.79	98.21	74.74
JFCM+DBN	0.61	0.79	1.40	98.60	77.44
ours	0.49	0.84	1.33	98.67	78.11

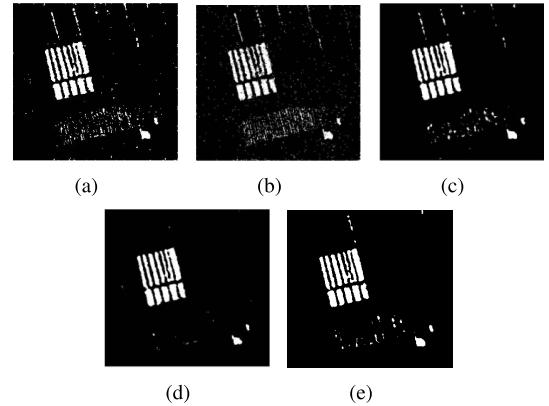


Fig. 20. Change detection results of Farmland C data set achieved by (a) SFCM, (b) GKI, (c) Threshold+DBN, (d) JFCM+DBN, and (e) proposed method.

TABLE III
CHANGE DETECTION RESULTS OF FARMLAND C DATA SET

Criteria \ Method	FP(%)	FN(%)	OE(%)	PCC(%)	Kappa(%)
SFCM	1.29	1.08	2.37	97.63	79.60
GKI	1.95	1.76	3.71	96.29	68.03
Threshold+DBN	0.65	1.03	1.67	98.33	84.95
JFCM+DBN	0.19	1.20	1.38	98.62	86.90
ours	0.41	1.00	1.41	98.59	87.09

while the DBN with threshold gains the lower accuracy due to the influence caused by the selection of threshold. Numerically speaking, the results obtained by JFCM+DBN and our method are very close. The results on Farmland D data set are shown in Fig. 21 and Table IV. The geographic information of this data set is relatively complex and contains more noise points. The SFCM method gains the worst performance under different evaluation metrics. Fig. 21(a) shows that the final map obtained by SFCM contains many missed regions. For this data set, GKI is less sensitive to noise. Although it can suppress a lot of noise, it results in the loss of details. The Kappa obtained by the proposed method equal to 84.06% is higher than that of 78.94% by JFCM+DBN, and achieves a big promotion.

The results of the Ottawa data set are shown in Fig. 22 and listed in Table V. The Ottawa data set is a relatively

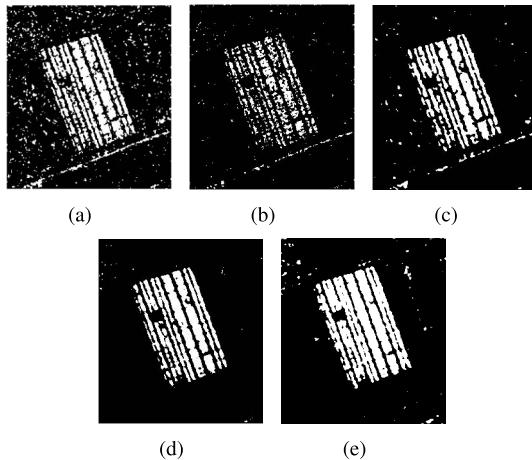


Fig. 21. Change detection results of Farmland D data set achieved by (a) SFCM, (b) GKI, (c) Threshold+DBN, (d) JFCM+DBN, and (e) proposed method.

TABLE IV
CHANGE DETECTION RESULTS OF FARMLAND D DATA SET

Method \ Criteria	FP(%)	FN(%)	OE(%)	PCC(%)	Kappa(%)
SFCM	4.41	4.64	9.05	90.94	69.30
GKI	0.42	10.64	11.06	88.94	52.14
Threshold+DBN	0.88	5.29	6.17	93.83	76.99
JFCM+DBN	0.27	5.29	5.56	94.44	78.94
ours	1.73	2.88	4.61	95.39	84.06

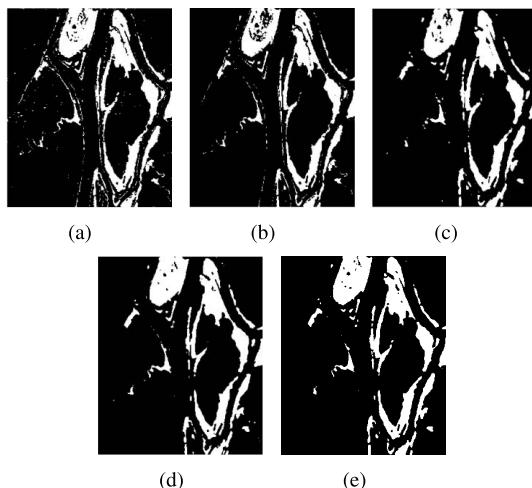


Fig. 22. Change detection results of Ottawa data set achieved by (a) SFCM, (b) GKI, (c) Threshold+DBN, (d) JFCM+DBN, and (e) proposed method.

simple data set, where the influence of noise is less great. Therefore, we can see that these all five methods perform well, especially the Kappa and PCC yielded by our method reached the best results 95.00% and 98.67%, respectively.

As for the Taiwan Shimen data set, the results are shown in Fig. 23 and listed in Table VI. From the visual results, we can see that the final map generated by our method

TABLE V
CHANGE DETECTION RESULTS OF OTTAWA DATA SET

Method \ Criteria	FP(%)	FN(%)	OE(%)	PCC(%)	Kappa(%)
SFCM	2.02	1.03	3.06	96.94	88.81
GKI	0.72	2.23	2.95	97.05	88.46
Threshold+DBN	0.79	1.30	2.09	97.91	92.03
JFCM+DBN	0.94	0.76	1.70	98.30	93.65
ours	0.56	0.77	1.33	98.67	95.00

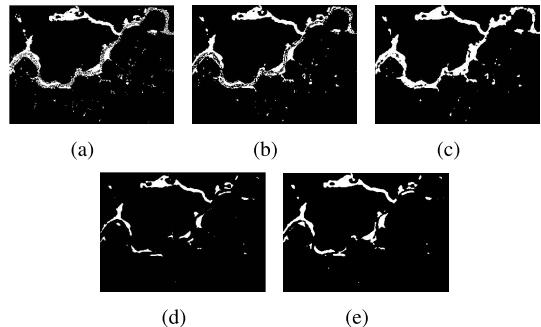


Fig. 23. Change detection results of Taiwan shimen data set achieved by (a) SFCM, (b) GKI, (c) Threshold+DBN, (d) JFCM+DBN, and (e) proposed method.

TABLE VI
CHANGE DETECTION RESULTS OF TAIWAN SHIMEN DATA SET

Method \ Criteria	FP(%)	FN(%)	OE(%)	PCC(%)	Kappa(%)
SFCM	2.88	0.70	3.58	96.42	70.93
GKI	2.42	0.29	2.71	97.29	77.95
Threshold+DBN	4.08	0.09	4.17	95.83	70.11
JFCM+DBN	0.05	1.73	1.78	98.21	79.90
ours	0.64	0.75	1.39	98.61	86.51

contains fewer noise points and most of the details are retained. In Table VI, our method exhibits the best PCC and Kappa.

This data set contains a pair of heterogeneous images (i.e., SAR and optical images). Due to the distinct feature representations of ground objects in different types of images, the change detection based on heterogeneous images is more challenging. The results of the Shuguang village data set are shown in Fig. 24 and Table VII lists the quantitative results. We can see that the final maps generated by these five methods contain many noise points. Although, in the numerical results, the proposed method achieved competitive results, especially on PCC and Kappa.

In summary, the proposed method has three compelling advantages. First, it avoids the tedious preprocessing operations such as some noise reduction methods. Second, our method skips the step of generating a DI, which can reduce the influence of DI on the final map. Third, it implements an unsupervised learning process from preclassification to obtaining the detection result, which is necessary and important

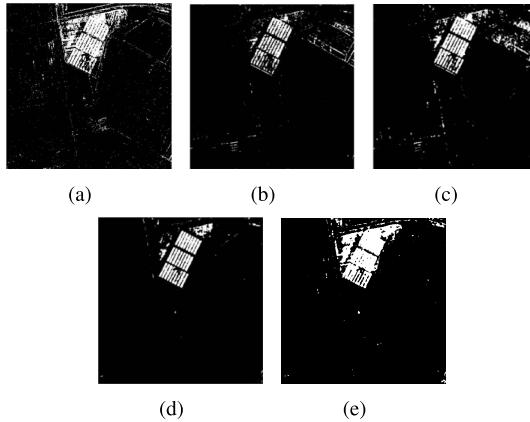


Fig. 24. Change detection results of Shuguang village data set achieved by (a) SFCM, (b) GKI, (c) Threshold+DBN, (d) JFCM+DBN, and (e) proposed method.

TABLE VII
CHANGE DETECTION RESULTS OF SHUGUANG VILLAGE DATA SET

Method	Criteria	FP(%)	FN(%)	OE(%)	PCC(%)	Kappa(%)
SFCM		1.93	3.28	5.21	94.79	65.51
GKI		0.87	6.20	7.07	92.92	40.16
Threshold+DBN		1.62	5.17	6.79	93.21	48.92
JFCM+DBN		0.39	5.38	5.76	94.24	52.35
ours		0.94	2.92	3.86	96.14	73.48

for the task of change detection. From the results of the comparative experiments on the above-mentioned six data sets, we can see that the proposed algorithm achieves state-of-the-art performance on the final map as well as on different evaluation criteria.

D. Analysis of Parameters

We have analyzed the following parameters: the number of convolutional kernels per layer, the sample selection parameter, and the relative weightage of the initial membership and the spatial function, respectively, in the SFCM algorithm. The selection of these parameters will affect the final classification results.

We first conduct a comparison experiment on the Farmland C data set with different numbers of convolutional kernels. Fig. 25 shows the influence of a different number of convolutional kernels on PCC and Kappa. The 6–12 in the abscissa represents the number of convolutional kernels in the first convolutional layer is 6 and in the second convolutional layer is 12. From Fig. 25, we can find that when the number of the convolutional kernels in the first layer is 12 and that in the second layer is 24, the results achieve the best performance on PCC and Kappa. On the other hand, we can see that when the number of convolutional kernels is too large, the performance will become worse. We theorize that a large number of convolutional kernels tend to lead to overfitting in the case of limited training data.

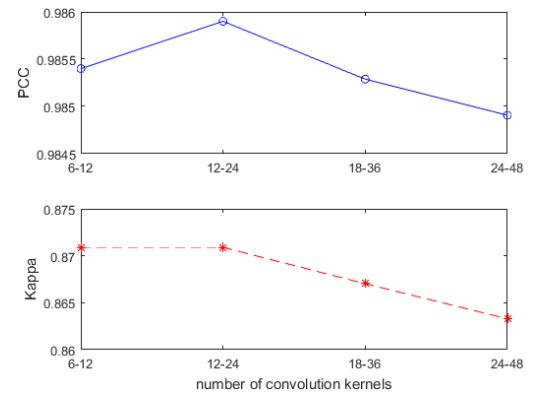


Fig. 25. Changes in PCC and Kappa with different numbers of convolution kernels on Farmland C data set.

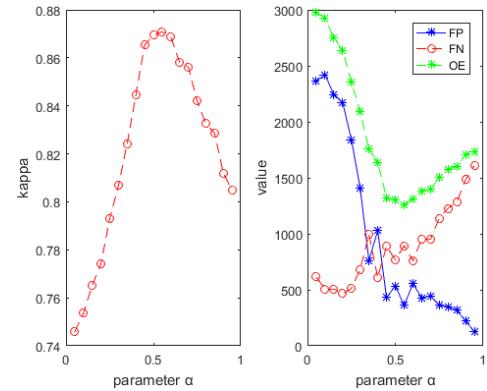


Fig. 26. Relationship between the parameter α and the criteria on Farmland C data set.

In the step of selecting suitable samples, α is an important parameter. We set α to 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, and 0.95 for testing it. The relationship between the parameter α and the criteria on the Farmland C data set is shown in Fig. 26. Where FN is short for false negatives, which represents the changed pixels that undetected. FP is short for false positives, which represents the unchanged pixels wrongly detected as changed. The total number of errors is represented by OE. The smaller the OE is, the fewer pixels are classified incorrectly. From Fig. 26, we can see that the result is better when α is between 0.4 and 0.6. When α is too large or too small, the classification results become worse. It can be seen from Fig. 26 that when α is small, the FP is large. Because when α is too small, the result is sensitive to noise, which yields a high FP. When α is large, the FN is large. Because when α is too large, the diversity of the sample is decreased, which yields a high FN.

In the SFCM algorithm, the parameters p and q determine the relative weight of the initial membership and the spatial function, respectively. We set the ratio of p and q to 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, and 4:1 for testing it. The relationship between the parameter and the criteria on the farmland C data set is shown in Fig. 27. From Fig. 27, when the proportion of p is large, the result is not very good. Due to the lack of spatial information, the result is easily affected by speckle noise. When the proportion of q is large, the result is not good either.

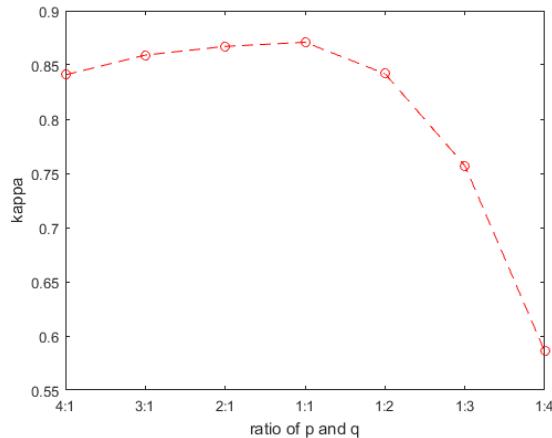


Fig. 27. Relationship between the ratio of p to q and the criteria on Farmland C data set.

Because too much spatial information is considered, the nature of the image is ignored. Thus, it cannot obtain satisfactory results either. The proper addition of spatial information will reduce the effects of the speckle noise. Therefore, when the ratio of p and q is close to 1:1, the classification result is better.

V. CONCLUSION

This paper has presented a novel change detection method based on spatial fuzzy clustering and CNN. Different from the traditional methods, our method does not need filter operation and it avoids the tedious preprocessing operations. In addition, our method does need to generate the DI, which reduces the influence of the DI on the final detection result. The well-designed structure of CNN can learn effective spatial characteristics from the raw images. From the results of the comparative experiments on six data sets, we conclude that the proposed algorithm achieves state-of-the-art performance on the final map as well as on different evaluation criteria.

The comparison experiments on six real data sets, which have different features, have demonstrated the effectiveness and stable of our method. In the future, we will do research on change detection based on heterogeneous images. It has good application prospects and it can reduce the cost of change detection.

REFERENCES

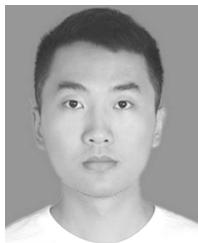
- [1] R. S. Lunetta and C. D. Elvidge, *Remote Sensing Change Detection*, vol. 310. London, U.K.: Taylor & Francis, 1999.
- [2] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 142–154, 2006.
- [3] R. Manonmani and S. G. M. Divya, "Remote sensing and GIS application in change detection study in urban zone using multi temporal satellite," *Int. J. Geomatics Geosci.*, vol. 1, no. 1, p. 60, 2010.
- [4] O. Yousif and Y. Ban, "Improving SAR-based urban change detection by combining MAP-MRF classifier and nonlocal means similarity weights," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 10, pp. 4288–4300, Oct. 2014.
- [5] J.-F. Zhang, L.-L. Xie, and X.-X. Tao, "Change detection of earthquake-damaged buildings on remote sensing image and its application in seismic disaster assessment," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, vol. 4, Jul. 2003, pp. 2436–2438.
- [6] T. Hame, I. Heiler, and J. S. Miguel-Ayanz, "An unsupervised change detection and recognition system for forestry," *Int. J. Remote Sens.*, vol. 19, no. 6, pp. 1079–1099, 1998.
- [7] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL, USA: CRC Press, 2009.
- [8] Y. Chen, L. Jiao, Y. Li, and J. Zhao, "Multilayer projective dictionary pair learning and sparse autoencoder for PolSAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 6683–6694, Dec. 2017.
- [9] L. Bruzzone and D. F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452–466, Apr. 2002.
- [10] 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.
- [11] J. Ingla and M. Grégoire, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.
- [12] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993.
- [13] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986.
- [14] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. B, Methodol.*, vol. 39, no. 1, pp. 1–22, 1977.
- [15] G. Moser and S. B. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2972–2982, Oct. 2006.
- [16] H. Hu and Y. Ban, "Unsupervised change detection in multitemporal SAR images over large urban areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 8, pp. 3248–3261, Aug. 2014.
- [17] L. Su, M. Gong, B. Sun, and L. Jiao, "Unsupervised change detection in SAR images based on locally fitting model and semi-EM algorithm," *Int. J. Remote Sens.*, vol. 35, no. 2, pp. 621–650, 2014.
- [18] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328–1337, May 2010.
- [19] 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.
- [20] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
- [21] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the Art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
- [22] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [23] S. Lange and M. Riedmiller, "Deep auto-encoder neural networks in reinforcement learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2010, pp. 1–8.
- [24] B. Du, W. Xiong, J. Wu, L. Zhang, L. Zhang, and D. Tao, "Stacked convolutional denoising auto-encoders for feature representation," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 1017–1027, Apr. 2016.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [26] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [27] F. Zhang, B. Du, and L. Zhang, "Scene classification via a gradient boosting random convolutional network framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1793–1802, Mar. 2016.
- [28] J. Zhao, M. Gong, J. Liu, and L. Jiao, "Deep learning to classify difference image for image change detection," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2014, pp. 411–417.
- [29] 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. 2015.
- [30] H. Hwang and R. Haddad, "Adaptive median filters: New algorithms and results," *IEEE Trans. Image Process.*, vol. 4, no. 4, pp. 499–502, Apr. 1995.
- [31] S. Rakshit, A. Ghosh, and B. U. Shankar, "Fast mean filtering technique (FMFT)," *Pattern Recognit.*, vol. 40, no. 3, pp. 890–897, 2007.

- [32] J.-M. Geusebroek, A. W. M. Smeulders, and J. van de Weijer, "Fast anisotropic Gauss filtering," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 938–943, Aug. 2003.
- [33] N. Saxena and N. Rathore, "A review on speckle noise filtering techniques for SAR images," *Int. J. Adv. Res. Comput. Sci. Electron. Eng.*, vol. 2, no. 2, p. pp–243, 2013.
- [34] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, vol. 3, no. 3, pp. 32–57, 1973.
- [35] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms," *Adv. Appl. Pattern Recognit.*, vol. 22, no. 1171, pp. 203–239, 1981.
- [36] D. L. Pham and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 737–752, Sep. 1999.
- [37] B. K. Tripathy, A. Basu, and S. Goyal, "Image segmentation using spatial intuitionistic fuzzy C means clustering," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res.*, Dec. 2015, pp. 1–5.
- [38] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2014, pp. 818–833.
- [39] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1646–1654.
- [40] P. L. Rosin and E. Ioannidis, "Evaluation of global image thresholding for change detection," *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2345–2356, 2003.
- [41] G. H. Rosenfield and K. Fitzpatrick-Lins, "A coefficient of agreement as a measure of thematic classification accuracy," *Photogram. Eng. Remote Sens.*, vol. 52, no. 2, pp. 223–227, 1986.



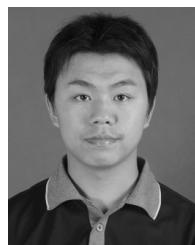
Yangyang Li (M'08–SM'18) received the B.S. and M.S. degrees in computer science and technology and the Ph.D. degree in pattern recognition and intelligent system from Xidian University, Xi'an, China, in 2001, 2004, and 2007, respectively.

She is currently a Professor with the School of Artificial Intelligence, Xidian University. Her research interests include quantum-inspired evolutionary computation, artificial immune systems, and deep learning.



Cheng Peng received the B.E. degree in intelligent science and technology from Xidian University, Xi'an, China, in 2016, where he is currently pursuing the Ph.D. degree in pattern recognition and intelligent system with the Key Laboratory of Intelligent Perception and Image Understanding, Ministry of Education of China, School of Artificial Intelligence.

His research interests include deep learning, remote sensing image processing, and evolutionary computation.



Yanqiao Chen (S'17) was born in Shijiazhuang, Hebei, China, in 1991. He received the B.S. degree from Hangzhou Dianzi University, Hangzhou, China, in 2014.

Since 2014, he has been taking successive post-graduate and doctoral programs with the Key Laboratory of Intelligent Perception and Image Understanding, Ministry of Education, International Research Center of Intelligent Perception and Computation, International Collaboration Joint Laboratory in Intelligent Perception and Computation, Xidian University, Xi'an, China. His research interests include deep learning, sparse representation, and polarimetric synthetic aperture radar image classification.



Licheng Jiao (SM'89–F'17) received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 1982, the M.S. and Ph.D. degrees in electronic engineering from Xian Jiaotong University, Xi'an, China, in 1984 and 1990, respectively.

From 1990 to 1991, he was a Post-Doctoral Fellow with the National Key Laboratory for Radar Signal Processing, Xidian University, Xi'an. Since 1992, he has been a Professor with the School of Electronic Engineering, Xidian University. He is currently the Director of the Key Laboratory of Intelligent Perception and Image Understanding, Ministry of Education of China, Xidian University. He is in charge of about 40 important scientific research projects. He has authored or co-authored more than 20 monographs and 100 papers in international journals and conferences. His research interests include image processing, natural computation, machine learning, and intelligent information processing.

Dr. Jiao is a member of the IEEE Xian Section Executive Committee, the Chairman of awards and recognition committee, the Vice Board Chairperson of the Chinese Association of Artificial Intelligence, the Councilor of the Chinese Institute of Electronics, the Committee Member of the Chinese Committee of Neural Networks, and an Expert of academic degrees committee of the state council.



Linhao Zhou received the B.S. degree in optical information science and technology from Yanshan University, Qinhuangdao, China, in 2015, and the M.S. degree in circuits and systems from Xidian University, Xi'an, China, in 2018.

He is currently a Researcher with the 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, China. His research interests include machine learning, deep learning, and remote sensing image processing.



Ronghua Shang (M'09) received the B.S. degree in information and computation science and the Ph.D. degree in pattern recognition and intelligent systems from Xidian University, Xi'an, China, in 2003 and 2008, respectively.

She is currently a Professor with Xidian University. Her research interests include evolutionary computation, image processing, and data mining.