SAR Image Change Detection Based on Data Optimization and Self-Supervised Learning

WENHUI MENG1, LIEJUN WANG1,2, ANYU DU1, AND YONGMING LI1

1College of Information Science and Engineering, Xinjiang University, Ürümqi 830046, China
2Key Laboratory of Signal Detection and Processing, Xinjiang Uygur Autonomous Region, Xinjiang University, Ürümqi 830046, China

Corresponding author: Liejun Wang (wljxju@xju.edu.cn)

ABSTRACT In the SAR change detection algorithm based on self-supervised learning, speckle noise reduces the difference image (DI) quality. Therefore, the contrast of the DI is low, and its change area is not significant. Moreover, the preclassification algorithm with the poor robustness makes the classification results of the low-quality DI inaccurate. First, to improve the quality of the initial DI, we design an adaptive gamma correction algorithm that adjusts the contrast according to the mean value of the initial DI and the variation coefficient $\beta$. The contrast of the new DI generated by this algorithm is higher. Furthermore, to suppress the noise, we adopt a new algorithm based on popular ranking to obtain the saliency map of the new DI. Combining the initial DI with this saliency map, a high-quality DI with a low noise level is obtained. After that, we introduce the structure tensor into the fuzzy local information c-means clustering algorithm (FLICM) to classify the DI more accurately. The new clustering algorithm improves the accuracy of preclassification, especially its hierarchical version. Besides, we use the structure tensor to generate the structure maps of the original images. Finally, according to the prior information obtained from the preclassification, we use a convolution wavelet neural network (CWNN) to supervise and train the structure maps of the original images. Experimental results show that the DI generated by us is closer to the ground-truth than other methods. Our preclassification algorithm performs better. Our algorithm shows higher detection accuracy for SAR images with strong noise than some advanced change detection algorithms.

INDEX TERMS Adaptive gamma correction, convolution wavelet neural network, fuzzy local information c-means clustering, saliency map, structure tensor.

I. INTRODUCTION

In recent years, with the increasing number of topics, such as resource exploration and environmental monitoring, remote sensing image processing technology has attracted more and more attention (e.g., target detection [1], [2], classification [3], [4], segmentation [5], etc.). In these research subjects, remote sensing image change detection, as a technology that can analyze surface change from remote sensing data at different times, has a high application value in disaster assessment, land use change detection, and urbanization process analysis. For example, the Arctic Ocean sea ice cover change affects human shipping activities [6]. As a result, the transportation routes of human beings across the Arctic Ocean have changed. In the past few decades, internal variability has intensified the impact of greenhouse gases on sea ice cover change [7]. Therefore, the study of sea ice cover change is of great significance to predicting safe routes. Besides, the spatial distribution of vegetation has an important impact on geomorphic evolution [8]. The analysis of geomorphic change data helps establish a reasonable vegetation distribution, improving the ecological environment.

Remote sensing image data is not derived from a single model; each model (e.g., hyperspectral, multispectral, RGB, and SAR) has its own unique characteristics. Hyperspectral images contain image information and spectral information,
which can reflect the internal and external structure information of objects. It has attracted extensive attention in image classification and other fields. For example, the hyperspectral classification task is concerned in [4]. In [4], a new graph convolutional network (miniGCN) is combined with CNN to generate the FuNet by using different fusion strategies. From the perspective of feature extraction, CNN extracts information from HS images in a short-range region, while graph convolutional network (GCN) models the middle- and long-range spatial relationship between samples. From the perspective of network training, a minibatch strategy is adopted in GCN. From the perspective of input, the piecewise sample is sent to CNN for training, while the pixel-level sample is sent to the network with the help of the adjacency matrix in GCN. As far as the computational complexity, large-scale matrix multiplication makes GCN more complex than CNN in one layer. The miniGCN train in a minibatch fashion similar to CNN, and it greatly saves the training time and tends to achieve a better local optimal solution. The fused FuNet is better than the single miniGCN model or CNN model, and the FuNet shows excellent classification performance.

However, hyperspectral images are susceptible to the effects of material mixing. Although hyperspectral decommisioning technology [9] reduces this effect, it undoubtedly increases the time complexity. Multispectral images can be considered a case of hyperspectral images. Compared with hyperspectral images, the imaging bands are smaller, and the spectral resolution is lower, but the spatial resolution of multispectral images is higher. The application of a generative adversarial network (GAN) to multispectral image segmentation has been proven to achieve excellent results [5]. RGB images only contain the information of red, green, and blue bands, and other lost spectral information cannot be recovered. It should be noted that hyperspectral images, multispectral images, and RGB images belong to optical remote sensing images.

In the task of change detection, the optical remote sensing image can be affected by the shooting time (such as daytime and night). In addition, the cloud layer in the atmosphere causes the information loss of an optical remote sensing image. However, synthetic aperture radar (SAR) can work all day long and can get high-quality images under extreme weather conditions (such as penetrating clouds and dust). Therefore, we think that an SAR image has more advantages than other optical images in a change detection task. This paper mainly studies how to use the existing multi temporal SAR images to achieve change detection.

In general, change detection algorithm consists of three steps: 1) image preprocessing; 2) difference image (DI) formation of a pair of multi-temporal images; 3) DI image analysis to achieve segmentation of the changed regions [10]–[13].

In the SAR image change detection process, the noise of DI cannot be ignored. Therefore, it is important to generate the DI with a low noise level. Currently, the methods for generating DI mainly include the difference method [14], the ratio method [15], the image transformation method [16], and the log-ratio method [17]. While these methods are simple and fast, they exhibit some deficiencies like low contrast or high noise in the generated DI. Some scholars have improved these methods to obtain the ideal DI. Gong et al. [18] proposed a method that involved the combination of the ratio map and the mean ratio map. Although this algorithm can suppress the noise to a certain extent, it enlarges the gray value on the edges of the DI, which causes errors in subsequent classification. Despite the ability of the Gaussian logarithmic ratio method [19] to reduce the impact of noise, unfortunately, it reduces the local contrast of the image.

The salient image detection method can effectively reduce the noise of the DI, and retain the information on the changed area to a certain extent. Gao et al. [20] used the frequency domain saliency detection method [21], but the quality of the DI was unsatisfactory, and some small changed areas were lost. Context awareness saliency detection [22] evaluates the saliency value of each pixel on multiple scales, and can extract attractive and compact salient regions. On this basis, it is simplified in [23] to extract potential changed regions. However, this method is still unable to detect some relatively isolated changed areas.

To obtain the final change map, there is a need for the DI to be classified. There are two methods of classification: the supervised method and the unsupervised method. In the supervised classification algorithm, information on the real changes in the surface can be obtained in advance, but it is difficult to be achieved in actual real-life production. In contrast, this is not a problem for unsupervised classification.

K-means clustering [24] and fuzzy c-means clustering (FCM) [25] are two common unsupervised clustering algorithms. However, these two methods do not consider the local information of the image. Fuzzy local information c-means clustering algorithm (FLICM) [26] considers the neighborhood information of an image, which introduces a new factor into the objective function of FCM, and this method is easily disturbed by isolated noise points. Gong et al. [27] proposed a noise suppressing improvement of the FLICM by replacing the spatial distance of the pixels with a local variation coefficient. Nevertheless, this procedure does not effectively use the intensity of the local pixels, leading to the loss of information. This is detrimental to learning-based change detection algorithms that rely on the accuracy of the initial clustering results.

The learning-based change detection algorithms can be divided into two types: one does not need to generate DI, such as [28], [29]. The other, however, needs to generate DI, such as [30], [31]. They are all trained in the original image. It is important to note that regardless of the type of network, fundamentally speaking, there is a need for it to learn the features of the original image. If we choose a suitable feature as the initial input, will the detection accuracy be improved? We found that the CWWNN is sensitive to structural information.
like the edge of the image. This inspired us to generate a new feature map to replace the original image as the input.

The purpose of our study is to find a more accurate change detection method. In the past, some algorithms are not effective because of the following aspects:

1) There is often a large amount of noise or insufficient contrast in the difference images (e.g., [14]–[19]), because they do not adopt effective noise suppression methods and contrast enhancement algorithms. To generate a higher quality DI, we use a new method to adaptively adjust the pixels with different gray values. The changed area in the adjusted DI is more significant. Then, we used an improved significant map generation method to obtain a significant map of the DI after adjusting the contrast. This significant map has less noise than the original and adjusted contrast graphs, and it has a good contrast. To avoid losing some isolated change regions, we combine the significant map with the initial DI. The final DI is close to the ground truth.

2) The performance of the preclassification algorithm is not good enough, so some training samples get wrong marks. This affects the accuracy of the final change map because the preclassification algorithm (e.g., FCM) adopted by these change detection algorithms does not fully consider the neighborhood information of the DI. The neighborhood information can improve the accuracy of classification. To make the preclassification more accurate, we improved the objective function of FLICM, which considers the spatial, intensity, and structural information of the image neighborhood. To further improve the classification performance, we adopt the idea of multilevel classification. The experimental results show that a better change map can be obtained by using our hierarchical clustering algorithm.

3) The previous algorithms do not mine the information of the training samples. It is well known that a CNN [29] extracts image features through a convolutional layer and completes image processing tasks with image depth features. Different from a CNN, a principal component analysis network (PCANet) uses a PCA filter to replace convolution operation in a GaborPCANet [32]. The PCA filter has completed the extraction of principal component characteristics. Furthermore, these features are used for training networks. Similarly, in a CWNN [31], twin-tree wavelet transforms replace convolution operations. The frequency domain features of the input image are used to train the network. Both the PCANet and CWNN mine the feature information of the original input image samples. At the same time, we notice that the CWNN is sensitive to the structural information of the image. This inspired us to generate samples from the structure map of the original image and send them to the CWNN for training. Many previous change detection algorithms only take samples generated from original images as the input of the network (e.g., [29]–[32]). Although these algorithms also achieve good results, our method shows better performance.

Our paper is divided into the following parts: first, we used a new method to construct a DI with less noise and high contrast, which reduces the difficulty of preclassification. Then, to further improve the accuracy of clustering, we adopted a new clustering algorithm. Finally, we verified the performance of the new sample generation method in experiments. The main contributions of this paper are as follows:

1) An adaptive gamma correction method is proposed; our gamma value varies with the noise level and gray value of the initial DI. This makes the difference in the different regions more obvious, with higher gray values in the changed areas. At the same time, much weak noise is eliminated. So, the contrast of the adjusted DI is higher, conducive to obtaining its saliency map.

2) To further suppress the isolated noise, we proposed a new saliency map generation method. It defines the probability of the saliency seeds and the calculation method of the saliency map with the high gray value of salient clusters, which is different from [33].

3) A fuzzy clustering algorithm based on the structural information, spatial location, and intensity in the DI neighborhood is proposed, which makes the classification of the DI more accurate.

4) According to the characteristics of structure tensor, we use it to generate a new feature map. The experimental results show that the samples generated by the structure map make the detection results more accurate.

II. RELATED WORK

A. STRUCTURE TENSOR

The structure tensor is a matrix derived from the gradient of an image, which can distinguish the flat area, edge area, and corner area of the image [34]. For an image $I$, its structure tensor (ST) can be expressed as:

$$ST = G(x, y) \cdot \begin{bmatrix} I_{xx}^2 & I_{xy}^2 \\ I_{xy}^2 & I_{yy}^2 \end{bmatrix} = \begin{bmatrix} j_{11} & j_{12} \\ j_{12} & j_{22} \end{bmatrix}$$  

where $I_{xx}$ and $I_{yy}$ represent the horizontal and vertical gradients of the original image $I$, respectively. $I_{xy}$ represents the gradient of the original image in the horizontal direction and then in the vertical direction. $G(x, y)$ is a non-negative weighting function, usually determined by a Gaussian function centered on a point $(x, y)$. Given the $ST$ of the original image $I$, two eigenvalues $\theta_1$ and $\theta_2$ can be obtained:

$$\theta_1 = \frac{1}{2} (j_{11} + j_{22} + \sqrt{(j_{11} - j_{12})^2 + 4j_{12}^2})$$

$$\theta_2 = \frac{1}{2} (j_{11} + j_{22} - \sqrt{(j_{11} - j_{12})^2 + 4j_{12}^2})$$  

where $\theta_1$ and $\theta_2$ are the eigenvalues of normal direction and tangent direction, respectively. In the corner region, $\theta_1$ and $\theta_2$ have larger positive values; In the smooth region, both $\theta_1$ and
$\theta_2$ are approximately equal to 0; In the edge region, $\theta_1$ has a larger positive value, while $\theta_2 \approx 0$.

B. SALIENT REGION DETECTION WITH DIFFUSION PROCESS

Given an image $P$, we define a graph $G = (S, E)$ with $N$ nodes $S = \{s_i|1 \leq i \leq N\}$ and edges $E = \{e_{ij}|1 \leq i, j \leq N\}$. $S$ is divided into $K$ clusters $C = \{c_1, c_2, \ldots, c_K\}$. $e_{ij}$ connects $s_i$ and $s_j$. $\omega_{ij}$ is the weight corresponding to $e_{ij}$. $W = [\omega_{ij}]_{N \times N}$ is defined as the affinity matrix of graph $G$. $D$ is the degree matrix which be defined as:

$$D = [d_{11}, d_{22}, \ldots, d_{NN}]$$

where $d_{ii} = \sum_j \omega_{ij}$.

The similarity between node $s_i$ and cluster center $c_j$ is defined as $a_{ij}$. Similarity matrix $H$ can describe the similarity between $S$ and $C$:

$$H = (D - \rho W)^{-1}A$$

where $A$ is the matrix form of $a_{ij}$. $(D - \rho W)^{-1}$ is the diffusion matrix. The mapping process from $A$ to $H$ is called diffusion process.

By calculating $H$, we can get the saliency value of each node and finally get the salient map.

C. CONVOLUTION WAVELET NEURAL NETWORK

CWNN is a kind of neural network that combines the convolutional neural network and wavelet transform. The pooling operation of CWNN is completed by a wavelet transform.

The advantage of this method is that the convolution wavelet neural network is sensitive to the texture features of the image, so it can effectively process the image under strong noise conditions. Because of this advantage, CWNN has been applied to SAR image segmentation and transform detection tasks, such as [31], [35].

In [31], a convolutional wavelet neural network with two convolution layers is used. The dual-tree wavelet transform is used in the pooling layer to replace the mean pooling and maximum pooling in the traditional convolution neural network. CWNN can effectively process SAR images under complex noise conditions, which brings convenience to change detection.

III. MATERIALS AND METHODS

Given SAR images $X_1$ and $X_2$ captured at the same place but different times, their sizes are all $H \times W$. To get the change map with noise suppression and prominent change area, first, we adjust the contrast of the initial difference image ($DI_{ini}$) to obtain a new difference image ($DI_{g}$). Then we calculate the saliency map of $DI_{g}$ and modify it to get a high-quality final difference image ($FDI$). Finally, we use a new fuzzy clustering algorithm to preclassify the $FDI$. The flowchart of proposed algorithm is shown in Figure 1. The known category include changed category (CC) and unchanged category (UC). Pixels belonging to the intermediate category (IC) need to be further classified.
A. ACQUISITION OF INITIAL DI

First of all, a neighborhood ratio-based method [28] has been proved to obtain a DI with prominent edges and obvious change regions, and this method has been widely used, such as [18], [36].

\[ DI_{ini}(x, y) = \Omega \cdot \frac{\min (X1(x, y), X2(x, y))}{\max (X1(x, y), X2(x, y))} + (1 - \Omega) \]

\[ \sum_{m,n\in\Omega} \frac{\min(X1(m, n), X2(m, n))}{\max(X1(m, n), X2(m, n))} \]

where \( \Omega \) is the ratio of mean and variance in neighborhood \( \Omega \).

B. ACQUISITION OF A HIGH-QUALITY DI

In SAR images, strong speckle noise often brings many difficulties on image processing. Although the complex noise and background are often mixed with the changed area, we found that the brightness of the changed area is high and the brightness of the invariant area is low. Although the brightness of some noises is higher than background area, it is lower than changed area. This inspires us adjusting the contrast of the initial DI. Gamma correction is a common method to adjust image contrast. It is usually used in the field of image enhancement. Given an image \( P \), the gamma correction can obtain the following results:

\[ P_{\gamma} = P'^{\gamma} \]

When \( \gamma > 1 \), the intensity of \( P \) is suppressed and when \( \gamma < 1 \), the brightness of \( P \) is enhanced. Figure 2 shows the results of gamma correction with a different \( \gamma \) on the initial DI.

Figure 2 shows that when \( \gamma < 1 \), the brightness of the image increases with the decrease in \( \gamma \). However, gamma that are too small will make the gray value of some invariant regions and noises very large, as shown in Figure 2 (b), which affects the accuracy of detection. When \( \gamma > 1 \), with the increase of \( \gamma \), some points with small gray value decrease. Although the points with high gray value are suppressed, the brightness of these areas, which contain changed areas, are still high. However, too large gamma may cause some pixels belonging to the changed region to be suppressed, as shown in Figure 2 (g).

According to the above analysis, in order to obtain the DI with prominent changed areas, we propose the use of adaptive gamma correction to adjust the contrast of the initial DI. On this basis, we calculate the saliency map of \( DI'_{init} \). After combining the saliency map of \( DI'_{init} \) and the initial DI, the FDI is obtained.

1) ADAPTIVE GAMMA CORRECTION

Although a fixed \( \gamma \) makes the calculation simple, it also brings serious distortion. Adaptive gamma correction can adaptively adjust the contrast of the image according to the brightness, space, and other information. To highlight the changed area and suppress the weak noises, we carry out adaptive gamma correction on \( DI_{ini} \) according to the following formula:

\[ DI_{g}(x, y) = DI'_{ini}(x, y), \gamma(x, y) = e^{30(\tau - DI(x, y))} \]

When \( DI(x, y) > \tau \), we think that the pixel is likely to be the CC, so it is enhanced; On the contrary, the pixel is more likely to be the background and noise, so it is suppressed. Obviously, the value of \( \tau \) is very important. In this paper, we set \( \tau \) according to the following rules:

\[ \frac{\overline{DI}_{ini}}{\beta} \leq \beta, \quad \text{then } \tau = \left( \frac{\sigma^2 + 1}{1 - \sigma} \right) \frac{\beta \overline{DI}_{ini}}{\epsilon} \]

\[ \frac{\overline{DI}_{ini}}{\beta} > \beta, \quad \text{then } \tau = (\sigma + 1) \cdot \epsilon \cdot e^{\sigma(\overline{DI}_{ini} - \mu)} \]

where \( \beta \) is the variation coefficient of the image, which can be used to represent the roughness of the image. Generally, the greater the noise of an image, the greater the \( \beta \). \( \epsilon \) is a small number to prevent the denominator from being 0. \( \sigma \) and \( \mu \) are manually set parameters. The detailed parameter analysis is given in Sec. IV (B).

Eqns. 8 and 9 are explained as follows: The larger the \( \overline{DI}_{ini} \), the larger the proportion of changed area and noise because of their high gray values.

In Eq. 8, when the \( \overline{DI}_{ini} \) is less than \( \beta \), we consider the noise level of \( DI_{ini} \) to be small. With the increase of \( \overline{DI}_{ini} \), the background areas of \( DI_{ini} \) get smaller. There are two situations in the process of \( \overline{DI}_{ini} \) increasing:

- \( 1 - \sigma \overline{DI}_{ini} < 0 \): \( \tau < 0 \) and \( \gamma(x, y) < 1 \), the DI is enhanced in any region. It should be noted that the brightness of the changed area is still high, as shown in Fig. 2 (b).
- \( 1 - \sigma \overline{DI}_{ini} > 0 \): \( \tau \) can be appropriately increased. If \( \tau - D(x, y) < 0 \) and \( \gamma(x, y) < 1 \), \( D(x, y) \) is more likely to belong to the region of changed, where the \( D(x, y) \) is enhanced. If \( \tau - D(x, y) > 0 \), \( \gamma(x, y) > 1 \), \( D(x, y) \) is more likely to belong to the background or weak noise, where the \( D(x, y) \) is suppressed.
As for Eq. 9, when $\overline{DI}_{ini}$ is greater than $\beta$, we consider
that $DI_{ini}$ has a larger noise level. At this time, the value of
$\tau$ should be set to a larger value to avoid excessive noise
enhancement. Finally, in order to prevent some changed areas
from being suppressed due to excessive $\tau$, it is necessary to
revise $\tau$ according to the following rules:

$$\text{if } \tau > \eta \text{ then } \tau = 2\overline{DI}_{ini}$$

(10)

where $\eta = 0.5$ which represents the midpoint of the normalized
image.

Algorithm 1 shows the process of adaptive gamma cor-
rection. Figure 3 (b) shows the $DI_g$. It can be seen that
the backgrounds are suppressed, and the most changed regions
are enhanced. However, at the same time, some isolated
noises are also enhanced.

**Algorithm 1** Adaptive Gamma Correction for the Initial DI

**Input:** Initial difference image $DI_{ini}$, parameters $\sigma$, $\mu$ and $\eta$.

**Do:**
1. Calculate the coefficient of variation $\beta$ and parameter $\alpha$;
2. Obtain $\tau$ by Eq. 9 if $\overline{DI}_{ini} \geq \beta$;
3. Obtain $\tau$ by Eq. 8 if $\overline{DI}_{ini} < \beta$
4. Retire $\tau$ by Eq. 10 if $\tau > \eta$
5. Obtain the corrected difference image $DI_g$ via Eq. 7.

**Output:** The corrected difference image $DI_g$.

2) ACQUISITION OF A FINE DI

Although the variable region of the DI after adaptive gamma
correction is more obvious, some noises are also enhanced.
Generally, the variation region is salient, and the salient level
of isolated noise is low, so we can get the salient map of
gamma-corrected DI to eliminate the isolated noise.

A saliency map generation method based on popular
ranking is adopted in [33]. After $N$ superpixels $S = [s_1, s_2, \ldots, s_K]$ are obtained by clustering the original images, [33] divides these superpixels into $K$ clusters
$
C = [c_1, c_2, \ldots, c_K].$

Then, the salient seeds of foreground and background are
calculated respectively by using the compactness of these
clusters. Finally, these seeds are used for diffusion prop-
gagation to obtain a salient map of foreground and back-
ground. Using this method to calculate the salient map is fast.

Meanwhile, the result is very accurate when detecting the
compact change areas. However, it cannot effectively detect
the isolated change areas, which affects the detection accu-
acy. We improve it based on [33].

The author of [33] thinks that salient objects are sur-
rounded by background. The tighter the cluster, the more
prominent it is. The probability of salient seeds is calculated
as:

$$p = 1 - \text{Norm}(sv)$$

(11)

$$sv(j) = \frac{\sum_{i=1}^{N} h_{ij} n_i \left\| b_i - \varphi_j \right\|}{\sum_{i=1}^{N} h_{ij} n_i} \text{ for } j = 1, 2 \ldots K$$

(12)

where $sv(j)$ stands for spatial variance of $j$th cluster. $h_{ij}$
represents the similarity of nodes $i$ and $j$ after diffusion. $n_i$
represents the number of pixels belonging to superpixel $s_j$.
$\varphi_j$ represents the spatial average of $j$th cluster. Norm $(^*)$
stands for the normalization operation. $b_i$ is the centroid of
each superpixel.

The author of [33] thinks that the spatial structure of larger
clusters is more compact. However, for some isolated small
areas, the compact density is obviously lower. As a result,
some relatively isolated changed areas cannot be detected.
Fortunately, we found that: given a normalized DI, the higher
the gray value, the more salient the pixel. According to this
character, we proposed a new method to calculate the proba-
bility of salient seeds by the following formula:

$$p = 1 - \text{Norm}(sv \cdot \frac{1-M_j}{2})$$

(13)

where $M_j$ represents the mean value of the $j$th cluster.

Obviously, the closer the $M_j$ is to 1, the greater the proba-
bility that it is a salient seed.

After obtaining foreground seeds and background seeds,
the foreground salient map $Sa_{fg民办}$ and background salient
map $Sa_{bg民办}$ are calculated based on a diffusion process.
To make the salient target evenly covered, the fine salient map
is obtained by the following operations:

$$Sa_{fg民办} (i) = \sum_{j=1}^{N} q(i,j) Sa_{fg民办} - mr(j)$$

(14)

$$q(i,j) = m(i,j) \sum_{k=1}^{N} m(i,k)$$

(15)

$$m(i,j) = w_{ij} d_{ij}$$

(16)
where \( m(i, j) \) is the weight of \( S_{fg} - m(j) \). \( S_{fg}(i) \) is the salient value of superpixel \( s_i \). \( q(i, j) \) is the normalized form of \( m(i, j) \). \( w_{ij} \) is the weight of the edge connecting the two superpixels \( s_i \) and \( s_j \). \( d_{ij} \) is the degree of superpixel \( s_j \).

In [33], when \( S_{fg}(i) \) is calculated, the weight of the superpixels closer to \( s_j \) is greater. This makes the distribution of a salient map more uniform. This operation makes the gray values of salient pixels closer to each other. However, if a background pixel is wrongly estimated as a salient pixel, its values of salient pixels closer to each other. To correct the pixel whose saliency is misjudged, we calculate the weight \( w_{ij} \) by using the proximity between the average value of each superpixel and 1:

\[
 w_{ij} = e^{-d_{ij}} \quad (17)
\]

The method for calculating the background map is the same as that in [33]. Finally, the final salient map is obtained by integrating the foreground map and background map:

\[
 Sa = \text{Norm}(S_{fg} + S_{bg}) \quad (18)
\]

After obtaining the salient DI \( Sa \), we modified it to get the final DI:

\[
 DI_f = Sa \ast DI_{ini} \quad (19)
\]

The comparison of salient map and DI obtained by using [33] and the improved method in this paper is shown in Figure 3. Although the method of [33] can retain a large range of salient regions, it has poor detection effect for some isolated salient regions. However, the saliency map obtained by the improved method can not only preserve a large range of salient regions but also avoid the loss of some isolated salient regions.

3) ACQUISITION OF CHANGE MAP

The comparison of salient map and DI obtained by using [33] and the improved method in this paper is shown in Figure 3. Although the method of [33] can retain a large range of salient regions, it has poor detection effect for some isolated salient regions. However, the saliency map obtained by the improved method can not only preserve a large range of salient regions but also avoid the loss of some isolated salient regions.

4) FUZZY LOCAL INFORMATION C-MEANS BASED ON MULTIPLE FEATURES (MFFLCM)

To obtain the change map, the DI needs to be further classified. The K-means algorithm and FCM algorithm are widely used in image classification. However, these two methods do not consider the local information of the image. Although FLICM [26] considers the neighborhood information of the image, a new factor is introduced into the objective function of FCM. However, this method is easily disturbed by isolated noise points. Inspired by [34], we decided to introduce image structure tensor, spatial, and intensity information to improve the accuracy of classification. The objective function of FLICM is as follows:

\[
 J_m = \sum_{i=1}^{N} \sum_{k=1}^{c} \left[ u_{ki}^m \| x_i - v_k \|^2 + G_{ki} \right] \quad (20)
\]

\[
 G_{ki} = \sum_{j \in \Omega_i \cap j \neq i} \frac{1}{d_{ij} + 1} (1 - u_{ki})^m \| x_j - v_k \|^2 \quad (21)
\]

where \( N \) represents the number of pixels in the neighborhood and \( c \) represents the number of clusters. In addition, \( v_k \) represents the clustering center of the \( k \)th cluster and \( u_{ki} \) represents the membership degree of the \( i \)th pixel to the \( k \)th cluster. \( G_{ki} \) is the fuzzy factor.

The new fuzzy factor proposed in this paper is defined as:

\[
 G_{ki} = \sum_{j \in \Omega_i \cap j \neq i} \frac{1}{1 + w_1 \cdot w_2} (1 - u_{ki})^m \| x_j - v_k \|^2 \quad (22)
\]

where \( w_1 \) combines the pixel position and intensity information in the neighborhood \( \Omega_i \), and \( w_2 \) is obtained from the relationship between the structure tensor of the local neighborhood pixel and the central pixel. They are defined as:

\[
 w_1 = e^{-\| x_j - x_i \|^2} \quad (23)
\]

\[
 w_2 = e^{-\| \theta_{ij} - \theta_{i1} \|^2} \quad (24)
\]

where \( d_{ij} \) represents the Euclidean distance between the pixel \( j \) and the center pixel \( i \). \( \theta_{ij} \) and \( \theta_{i1} \) are respectively the eigenvalues of the normal direction of pixel \( j \) and center pixel \( i \) in \( \Omega_i \). \( w_1 \) shows that the weight far away from the center point is small, and the weight closer to the center point is large. \( w_2 \) is inversely proportional to \( | \theta_{ij} - \theta_{i1} | \), which represents the weight of the point with a large difference from the center pixel, which is small. If there is a large difference between \( j \) and \( i \), \( w_1 \) and \( w_2 \) will reduce the probability of \( j \) and \( i \) being judged as the same class. This fuzzy factor, which combines position, gray value information, and structure tensor information, makes the final classification result more reliable. Substituting Eq. 21 into Eq. 19, we can obtain:

\[
 u_{ki} = \frac{1}{\sum_{j=1}^{c} \left( \frac{\| x_i - v_k \|^2 + G_{ji}}{w_1 \cdot w_2} \right) (m-1)} \quad (25)
\]

\[
 v_k = \frac{\sum_{i=1}^{N} u_{ki}^m x_i}{\sum_{i=1}^{N} u_{ki}^m} \quad (26)
\]

The result of clustering determines the selection of samples, which directly affects the accuracy of detection. To make the clustering result more reliable, we used the same method as [36] to classify DI, and the specific process is shown in Algorithm 2.

5) RECLASSIFICATION

After clustering, the final DI is divided into three categories: \( CC \), \( UC \), and \( IC \). Then we use the prior information of \( CC \) and \( UC \) to train the CWN. Finally, the trained CWN is used to
classify IC. Because CWNN is more sensitive to the texture features of images, which inspires us to mine the structure information of images further. So, we use the structure map of the original image as the sample. It is defined as:

\[
X_{st} = \frac{\theta_X}{2}
\]

where \(\theta_X\) represents the structural tensor in the tangent direction of image \(X\).

It can be seen from Fig. 4 that, compared with the original image, the structure map can obviously highlight the edges and corners of the original image. It is helpful for the neural network to further mine the structural information in training.

In flat areas, \(\theta_X\) is small. So, \(I_1\) and \(I_2\) are generally dark. In order to facilitate visual analysis, we use gamma correction to process \(I_1\) and \(I_2\). The results are shown in Figure 4. It can be seen from Figure 4 that, compared with the original image, the structure map highlights the edges and corners of the original image. It is helpful for the neural network to further mine the structural information in training.

In addition, to make up for the shortage of samples, we use the same method as [31] to generate some virtual samples. The flow chart is shown in Figure 5. For simplicity, we replace \(I_1\) and \(I_2\) with \(GST1\) and \(GST2\), respectively.

The structure of CWNN is shown in Figure 6. It consists of two convolution layers, two wavelet pooling layers, one input layer, one output layer, and one fully connected layer.

The difference between CWNN and CNN is the operation of the pooling layer. In the pooling layer of CWNN, the output of the upper layer is decomposed into two low-frequency sub-bands and six high-frequency sub-bands by dual-tree complex wavelet transform (DT-CWT). To eliminate noise interference, the output of the pooling layer is the mean value of the two low-frequency sub-bands. The input training sample is a series of image patches with a size of 28*28, stitched together from the structure pictures of two original images, and the image patches with a size of 28*14. The first convolution layer contains six convolution kernels with a size of 5 * 3. The second convolution layer contains 12 convolution kernels with a size of 5 * 3. The last convolution layer contains 96 units, and the final output layer contains two units, which represent two different classes.

The IC is divided into \(CC\) and \(UC\) after inputting CWNN. We combine the results of the pre-classification to get the final change map.

IV. EXPERIMENT

In this section, we evaluate the performance of our algorithm through experiments. To maintain the fairness of the comparison, all the experiments are conducted on a PC running Windows 10 with 4G RAM and a 2.6 GHz CPU.

The experiment is divided into the following parts: in the first part, we introduce our datasets. In the second part, we analyze the setting of parameters. To prove the advantages of the DI generation method proposed in this paper, we use different methods to generate DI and analyze their performance. In the fourth part, to prove the effectiveness of the proposed clustering algorithm, we compare it with several commonly used clustering algorithms. Finally, we compare with several advanced algorithms to prove the effectiveness of the proposed algorithm.

The comparison algorithms used in this paper include: Li et al. [29], Gao et al. [31], GaborPCANet [32], NRELM [36], NR-CR [37], and ESMOFCM [38]. Our experiments are carried out on four public datasets: Yellow River I, Yellow River II, Mexico and Ottawa. Besides, five evaluation indexes (TN, TP, OE, PCC (%), and Kappa (%)) are used for quantitative analysis. Especially, Kappa and PCC are the two most convincing data in the change detection task.

![Algorithm 2 Hierarchical MFLLICM (HMFFLICM)](image)

**Input:** final difference image \(FDI\), clustering number \(N\), \(\sigma_1 = 1.25\), \(\sigma_2 = 1.1\).

**Do:**
1. Divide \(DFI\) into two clusters by MFLLICM: \(CC\) and \(UC\). The number of pixels belonging to \(CC\) and \(UC\) are \(T_{CC}\) and \(T_{UC}\) respectively.
2. Divide \(DFI\) into \(N\) classes by MFLLICM.
3. Sort \(N\) classes from large to small based on their average: \(N_{C1}, N_{C2}, \ldots, N_{CN}\). The corresponding numbers are \(n_1, n_2, \ldots, n_N\).
4. Set the threshold of \(CC\) and \(UC\), respectively:
   \[T_{CC} = \sigma_1 \cdot T_{CC} \quad \text{and} \quad T_{UC} = \sigma_2 \cdot T_{UC}.\]
5. Parameters set
   \[i = i + 1, \quad cft = cft + n_i, \quad clt = cft + n_i, \quad i = 1, 2, \ldots, N.\]
6. Select the first \(i\) clusters from \(N\) clusters, if \(cft < T_{CC}\), \(N_{Cl}\) belongs to \(CC\). Select the last \(i\) clusters from \(N\) clusters, if \(cft < T_{MC}\), \(N_{Cl}\) belongs to \(UC\).
7. Other clusters of the \(M\) clusters are set as the intermediate class \(IC\).

**Output:** The pre-classification results \{\(CC\), \(IC\), \(UC\)\}.
A. DESCRIPTION OF DATASETS
The Yellow River I data was taken by RADARSAT-2 in The Yellow River Estuary, China, in 2008 and 2009. The original size of the initial captured images is 7666 * 7692. It contains Farmland C data and Farmland D data. Their sizes are 257 * 289 and 306 * 291, respectively. The two sets of images depict the changes of two farmland from 2008 to 2009, respectively. Figure 7 shows their corresponding original images and ground-truth.

The Yellow River II data was obtained in the same way as the Yellow River I data. It includes Inland water data and Coastline data. Their sizes are 291 * 444 and 450 * 280, respectively. Figure 8 shows their corresponding original images and ground-truth.

The Mexico data was taken by LANDSAT 7 in 2000 and 2005. The fire caused changes in the surface of the earth. Figure 9 shows the original images and ground-truth. Their resolution is 512 * 512. As can be seen from Figure 9, in addition to some large areas of change, the change map also contains some scattered change areas. This detection task is also very challenging.

B. PARAMETERS SETTING
In order to ensure the balance between the number of virtual samples generated by the CC and UC, the number of training samples of each kind is set to 1000. In the process of testing, the size of image block \( \omega \) is an important parameter. Table 1 shows the detection accuracy with different \( \omega \).

Table 1. Change detection results by using different patch.

| Size | \( \omega=3 \) | \( \omega=5 \) | \( \omega=7 \) | \( \omega=9 \) | \( \omega=11 \) |
|------|-------------|-------------|-------------|-------------|-------------|
| PCC (%) | 98.268      | 98.268      | 98.291      | 98.268      | 98.267      |
| Kappa (%) | 83.600      | 83.601      | 84.398      | 83.600      | 83.599      |

Ottawa data is obtained by the RADARSAT SAR sensor in Ottawa, Canada. Its original images were taken in July and August 1997, respectively. The flood led to changes in the topography of the area. Figure 10 shows the original images and ground-truth. Their resolution is 290 * 350.
C. ANALYSIS OF DI

To prove the advantages of the DI proposed in our paper, we compare it with several different methods. These methods include logarithmic ratio method, neighborhood logarithmic ratio method, and neighborhood ratio method. We calculate the structure similarity (SSIM) of each DI and ground truth to evaluate its performance. The experiment is carried out on four datasets, and the average value of SSIM obtained by each method is shown in Table 2.

As seen from Table 2, the DI generation method proposed in this paper is far better than other methods on Yellow River I and Yellow River II datasets. At the same time, the best results are obtained on the Ottawa dataset. For the Mexico dataset, the difference is tiny. Figure 12 shows a comparison of the different DIs. It is obviously that our DI is closely approaching the ground-truth.

To further prove the performance of the DI in this paper, we use different methods to detect DI. These methods include MFFLICM, MFFLICM-ELM and MFFLICM-CWNN. The results are shown in Figure 14. It can be seen that the results we proposed are far better than others.

D. PERFORMANCE OF CLUSTERING ALGORITHM

To verify the effectiveness of MFFLICM, we compare the performance of MFFLICM with FCM and FLICM.

First, we use these three algorithms to divide the DI into two categories and get the mean value of the detection results on four datasets (Yellow River I, Yellow River II, Mexico, and Ottawa). The results are shown in Table 3. Although FLICM achieves better results, there is not much difference between our method and FLICM. Then, we use these three algorithms in hierarchical clustering and then divide the DI into three categories. The IC is sent to the CWNN for training.

Table 4 shows the mean value of the detection results on four datasets. After using the hierarchical clustering method, our method achieves better results, especially for Kappa. For example, Figure 13 shows the detection results in Coastline with different hierarchical clustering algorithms.
E. QUANTITATIVE ANALYSIS

Before training, the structure maps of the original images are used to generate samples, and the results obtained by this method are more accurate. Table 5 shows the average results with different samples on four datasets.

Although the improvement brought by this method is not much, it indicates the feasibility of using the image structure feature training method.

After that, we compare our algorithm with several advanced algorithms. Table 6 shows that our algorithm is much better than other algorithms in Farmland C, because our algorithm is better than other algorithms in terms of TP, OE, PCC, and Kappa. As far as Farmland D, our algorithm is only inferior to Gao et al. [31]. In general, our algorithm achieves excellent results in Yellow River I dataset. Our method suppresses the noise and background of the initial DI, as shown in Figure 15 and Figure 16. Furthermore, our salient map retains the changed area and eliminates the noise to a great extent. The changed areas are highlighted. This makes our DI closer to the ground truth and provides convenience for our detection.

The detection results on Coastline are shown in Table 7. Our method is improved by 39.96%, 66.07%, 63%, 52.63%, 71.2% and 21.8% over Li et al. [29], Gao et al. [31], Gabor-PCANet, NRELM, NR-CR and ESMOFCM in terms of the value of Kappa, respectively. Besides, the data of our algorithm on TN, OE and PCC is much better than others.

Figure 17 shows that other algorithms perform extremely poorly. Because the speckle noise in their DI is not suppressed, and their pre-classification algorithms are not robust. Although ESMOFCM looks close to the ground-truth, it loses some details. The DI generated by us can suppress the noise well. Our clustering algorithm fully combines all kinds of effective information in the image neighborhood, making our pre-classification results more stable. So, our result is far better than other algorithms.

In Inland water data, our detection results are close to Gao et al. [31]. However, as can be seen from Figure 18, our method has fewer false detection areas.

Table 8 shows our method is respectively improved by 4.89%, 3.92%, 5.11%, 6.25%, 2.39% and 1.79% over Li et al. [29], Gao et al. [31], GaborPCANet, NRELM, NR-CR and ESMOFCM in terms of the value of Kappa on the Mexico data. Compared with Li et al. [29], Gao et al. [31], GaborPCANet, NRELM, NR-CR and ESMOFCM, our algorithm improves the PCC value by 6.63%, 0.54%, 0.57%, 0.68%, 0.3% and 10.50%, respectively.

Our algorithm also achieves the best results in OE and TP. As shown in Figure 19, our method does not lose isolated change regions, such as the region marked by the red circle. It owes to our salient map detection method which takes advantage of the high pixel value of the changed region.
Our algorithm also achieves the best results on the Ottawa data, as shown in Table 9. Compared with Li et al. [29], Gao et al. [31], GaborPCANet, NRELM, NR-CR and ESMOFCM, our algorithm improves the
TABLE 6. Change detection results on Yellow River I dataset.

| Dataset   | Method   | TN     | TP     | OE     | PCC (%) | Kappa (%) |
|-----------|----------|--------|--------|--------|---------|-----------|
| Farmland C | Li et al. [29] | 83369  | 4497   | 1180   | 98.67   | 87.70     |
|           | Gao et al. [31] | 83452  | 4536   | 1058   | 98.81   | 88.93     |
|           | GaborPCANet | 83759  | 3798   | 1489   | 98.32   | 82.75     |
|           | NRELMT    | 83688  | 3486   | 1872   | 97.90   | 77.77     |
|           | NR-CR     | 83718  | 2997   | 2331   | 97.38   | 70.72     |
|           | ESOMOFCM  | 83115  | 4144   | 1787   | 97.99   | 81.20     |
|           | Proposed  | 83579  | 4612   | 855    | 99.04   | 91.01     |
| Farmland D | Li et al. [29] | 59277  | 11602  | 3394   | 95.43   | 84.46     |
|           | Gao et al. [31] | 60003  | 11855  | 2415   | 96.75   | 88.79     |
|           | GaborPCANet | 58998  | 11892  | 3383   | 95.44   | 84.76     |
|           | NRELMT    | 60237  | 9567   | 4469   | 93.98   | 77.57     |
|           | NR-CR     | 58576  | 11118  | 4579   | 93.83   | 79.16     |
|           | ESOMOFCM  | 58057  | 9286   | 6930   | 90.67   | 67.21     |
|           | Proposed  | 59789  | 11298  | 3186   | 95.71   | 85.05     |

TABLE 7. Change detection results on Yellow River II dataset.

| Dataset   | Method   | TN     | TP     | OE     | PCC (%) | Kappa (%) |
|-----------|----------|--------|--------|--------|---------|-----------|
| Coastline | Li et al. [29] | 121269 | 1249   | 3482   | 97.24   | 40.79     |
|           | Gao et al. [31] | 111429 | 1297   | 13274  | 89.46   | 14.68     |
|           | GaborPCANet | 113478 | 1341   | 11181  | 91.12   | 17.75     |
|           | NRELMT    | 118337 | 1322   | 6341   | 94.97   | 28.12     |
|           | NR-CR     | 103831 | 1337   | 20832  | 83.46   | 9.55      |
|           | ESOMOFCM  | 124638 | 573    | 789    | 99.37   | 58.95     |
|           | Proposed  | 124649 | 918    | 433    | 99.66   | 80.75     |
| Inland water | Li et al. [29] | 123729 | 3181   | 2294   | 98.22   | 72.59     |
|           | Gao et al. [31] | 122860 | 3772   | 2572   | 98.46   | 73.64     |
|           | GaborPCANet | 125688 | 3095   | 2421   | 98.57   | 70.93     |
|           | NRELMT    | 123672 | 3065   | 2467   | 98.54   | 70.33     |
|           | NR-CR     | 123892 | 2864   | 2448   | 98.55   | 69.11     |
|           | ESOMOFCM  | 118559 | 3962   | 6683   | 94.83   | 51.84     |
|           | Proposed  | 122957 | 3690   | 2557   | 98.47   | 73.24     |

FIGURE 15. Change detection results of Farmland C. (a) Li et al. [29]. (b) Gao et al. [31]. (c) GaborPCANet. (d) NRELMT. (e) NR-CR. (f) ESOMOFCM. (g) Proposed. (h) Ground-truth image of Farmland C.

FIGURE 16. Change detection results of Farmland D. (a) Li et al. [29]. (b) Gao et al. [31]. (c) GaborPCANet. (d) NRELMT. (e) NR-CR. (f) ESOMOFCM. (g) Proposed. (h) Ground-truth image of Farmland D.

PCC value by 0.26%, 0.05%, 0.31%, 3.31%, 0.93% and 0.74%, respectively. In addition, our method is improved by 1.16%, 0.17%, 1.33%, 15.13%, 3.32% and 3.07% over Li et al. [29], Gao et al. [31], GaborPCANet, NRELMT, NR-CR and ESOMOFCM in terms of the value of Kappa, respectively. The performance of each algorithm on the
Ottawa data has little difference, as shown in Figure 20. However, from the objective index point of view, our algorithm has the best performance.

V. DISCUSSION

A. RUNNING TIME

In this part, the running time of each algorithm is discussed.

As reported in Table 10, NRELM needs minimal running time, and its running time is far less than the others. Thanks to its simple network, NRELM does not need to learn too many parameters. Unfortunately, it also makes the detection accuracy not high. NR-CR is the second fastest algorithm, but its results contain much noise, as shown in the change map of Figure 15 to 20. Li et al. [29] is faster than the rest of the algorithms. However, it does not show a greater advantage in the detection accuracy. GaborPCANet, Gao et al. [31] and proposed method run longer because their networks are more complex and need longer training time. ESMOFCM spends the longest time because it takes much time to achieve multi-objective clustering. Compared with Li et al. [29], NRELM, and NR-CR, our algorithm takes more time. Because NRELM and NR-CR are relatively simple in generating DI and training network. In Li et al. [29], DI is not necessary, and the clustering algorithm is simpler than ours. Although CWNN is used in Gao et al. [31] and our algorithm, our algorithm spends less time. Because our high-quality DI needs fewer iterations when clustering.
TABLE 8. Change detection results on Mexico dataset.

| Method    | TN   | TP   | OE   | PCC (%) | Kappa (%) |
|-----------|------|------|------|---------|-----------|
| Li et al. [29] | 230111 | 21103 | 10930 | 91.80   | 77.20     |
| Gao et al. [31] | 229957 | 21609 | 10578 | 97.89   | 78.17     |
| GaborPCANet | 230107 | 21015 | 1122  | 97.86   | 76.98     |
| NRELM      | 230176 | 20507 | 11461 | 97.75   | 75.84     |
| NR-CR      | 229768 | 22391 | 9985  | 98.13   | 79.70     |
| ESMOFCM    | 227539 | 24298 | 10307 | 87.93   | 80.30     |
| Proposed   | 229720 | 23461 | 8963  | 98.43   | 82.09     |

TABLE 9. Change detection results on Ottawa dataset.

| Method    | TN   | TP   | OE   | PCC (%) | Kappa (%) |
|-----------|------|------|------|---------|-----------|
| Li et al. [29] | 84815 | 14842 | 1843  | 98.18   | 93.08     |
| Gao et al. [31] | 84259 | 15609 | 1632  | 98.39   | 94.07     |
| GaborPCANet | 84736 | 14872 | 1892  | 98.13   | 92.91     |
| NRELM      | 85448 | 11111 | 4941  | 95.13   | 79.11     |
| NR-CR      | 83660 | 15521 | 2519  | 97.51   | 90.92     |
| ESMOFCM    | 84660 | 14492 | 2339  | 97.70   | 91.17     |
| Proposed   | 84334 | 15586 | 1580  | 98.44   | 94.24     |

TABLE 10. Running time of each algorithm.

| Method    | Farmland C | Farmland D | Coastline | Inland water | Mexico | Ottawa | Average |
|-----------|------------|------------|-----------|--------------|--------|--------|---------|
| Li et al. [29] | 563.1s     | 561.0s     | 534.6s    | 563.4s       | 1065.0s| 609.8s | 649.5s  |
| Gao et al. [31] | 1096.2s    | 1088.6s    | 1097s     | 1100.0s      | 1103.8s| 1105.2s| 1098.6s |
| GaborPCANet | 579.2s     | 1161.9s    | 1693.2s   | 668.2s       | 1486.5s| 709.11s| 1049.7s |
| NRELM      | 62.2s      | 4.9s       | 8.3s      | 8.9s         | 16.7s  | 7.3s   | 8.7s    |
| NR-CR      | 22.9s      | 79.8s      | 228.8s    | 56.2s        | 463.4s | 67.7s  | 153.0s  |
| ESMOFCM    | 6218.3s    | 9116.5s    | 1838.9s   | 9309.3s      | 18940.5s| 7250.3s| 8785.6s |
| Proposed   | 1066.3s    | 1095.5s    | 1094.8s   | 1076.2s      | 1098.9s| 1083.0s| 1085.7s |

GaborPCANet spends less time in training, so it is slightly faster than our algorithm.

B. FUTURE DEVELOPMENTS

According to our research, an excellent difference image and high-performance pre-classification algorithm can effectively improve the detection accuracy. Therefore, we hope to find a better way to generate the DI and a better pre-classification method. Although CWNN is an outstanding network, we hope to find a better classification network. These are the focus of our future work.

VI. CONCLUSION

In this paper, we first divide the generation process of DI into two stages: contrast adjustment and refinement correction. Contrast adjustment makes the changed area more pronounced so that it can be detected easily. Refinement correction suppresses the isolated strong noise and makes the FDI closer to ground truth. Then, a new fuzzy clustering method is proposed for classification. In the experiment, we found that its multilevel operation has a better performance. Finally, we use a new method to generate the sample. These samples are taken from the structure map of the original images. The results on four public datasets show that our algorithm maintains good performance for SAR images with complex noise. Our algorithm proves that the DI optimization method is feasible in the change detection task. At the same time, selecting an appropriate image feature to train the network can also improve the detection results. However, it should be noted that the performance of our algorithm depends on the accuracy of preclassification. Finding a more reliable classification algorithm is the focus of our future work.

REFERENCES

[1] N. Yokoya and A. Iwasaki, “Object detection based on sparse representation and Hough voting for optical remote sensing imagery,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 8, no. 5, pp. 2053–2062, May 2015.
[2] W. Liu, L. Ma, J. Wang, and H. Chen, “Detection of multiclass objects in optical remote sensing images,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 5, pp. 791–795, May 2019.
[3] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chauussot, Q. Du, and B. Zhang, “More diverse means better: Multimodal deep learning meets remote-sensing imagery classification,” IEEE Trans. Geosci. Remote Sens., early access, Aug. 24, 2020, doi: 10.1109/TGRS.2020.3016820.
[4] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chauussot, “Graph convolutional networks for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., early access, Aug. 18, 2020, doi: 10.1109/TGRS.2020.3015157.
[35] F. Gao, X. Wang, J. Dong, and S. Wang, “Synthetic aperture radar image change detection based on frequency-domain analysis and random multigraphs,” *J. Appl. Remote Sens.*, vol. 12, no. 1, 2018, Art. no. 016010.
[36] X. Hou and L. Zhang, “Salient detection: A spectral residual approach,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[37] S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-aware saliency detection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915–1926, Oct. 2012.
[38] J. Geng, X. Ma, X. Zhou, and H. Wang, “Saliency-guided deep neural networks for SAR image change detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7365–7377, Oct. 2019.
[39] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Noida, India: Pearson Education, 2004.
[40] A. Ghosh, N. S. Mishra, and S. Ghosh, “Fuzzy clustering algorithms for unsupervised change detection in remote sensing images,” *Inf. Sci.*, vol. 181, no. 4, pp. 699–715, Feb. 2011.
[41] 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.
[42] 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.
[43] 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.
[44] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, “A deep learning method for change detection in synthetic aperture radar images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5751–5763, Aug. 2019.
[45] Y. Li, G. Liu, T. Li, L. Jiao, G. Lu, and N. Marturi, “Application of data driven optimization for change detection in synthetic aperture radar images,” *IEEE Access*, vol. 8, pp. 11426–11436, 2020.
[46] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, “Sea ice change detection in SAR images based on convolutional-wavelet neural networks,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1240–1244, Aug. 2019.
[47] 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.
[48] L. Zhou, Z. Yang, Z. Zhou, and D. Hu, “Salient region detection using fusion process on a two-layer sparse graph,” *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5882–5894, Dec. 2017.
[49] W. Zhang and C. Sun, “Corner detection using multi-directional structure tensor with multiple scales,” *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 438–459, 2020.
[50] Y. Duan, F. Liu, L. Jiao, P. Zhao, and L. Zhang, “SAR image segmentation based on convolutional-wavelet neural network and Markov random field,” *Pattern Recognit.*, vol. 64, pp. 255–267, Apr. 2017.
[51] 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, Dec. 2016, Art. no. 046019.
[52] Y. Gao, F. Gao, J. Dong, and S. Wang, “Sea ice change detection in SAR images based on collaborative representation,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 7320–7323.
[53] R. Liu, R. Wang, J. Huang, J. Li, and L. Jiao, “Change detection in SAR images using multiobjective optimization and ensemble strategy,” *IEEE Geosci. Remote Sens. Lett.*, early access, Jul. 7, 2020, doi: 10.1109/LGRS.2020.3005140.