



## Article

# A Hierarchical Fusion SAR Image Change-Detection Method Based on HF-CRF Model

Jianlong Zhang <sup>1</sup>, Yifan Liu <sup>1</sup>, Bin Wang <sup>1,\*</sup> and Chen Chen <sup>2</sup>

<sup>1</sup> School of Electronic Engineering, Xidian University, Xi'an 710071, China; jlzhang@mail.xidian.edu.cn (J.Z.); yifanl@stu.xidian.edu.cn (Y.L.)

<sup>2</sup> State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China; cc2000@mail.xidian.edu.cn

\* Correspondence: bwang@xidian.edu.cn

**Abstract:** The mainstream methods for change detection in synthetic-aperture radar (SAR) images use difference images to define the initial change regions. However, methods can suffer from semantic collapse, which makes it difficult to determine semantic information about the changes. In this paper, we proposed a hierarchical fusion SAR image change-detection model based on hierarchical fusion conditional random field (HF-CRF). This model introduces multimodal difference images and constructs the fusion energy potential function using dynamic convolutional neural networks and sliding window entropy information. By using an iterative convergence process, the proposed method was able to accurately detect the change-detection regions. We designed a dynamic region convolutional semantic segmentation network with a two-branch structure (D-DRUNet) to accomplish feature fusion and the segmentation of multimodal difference images. The proposed network adopts a dual encoder–single decoder structure where the baseline is the UNet network that utilizes dynamic convolution kernels. D-DRUNet extracts multimodal difference features and completes semantic-level fusion. The Sobel operator is introduced to strengthen the multimodal difference-image boundary information and construct the dynamic fusion pairwise potential function, based on local boundary entropy. Finally, the final change result is stabilized by iterative convergence of the CRF energy potential function. Experimental results demonstrate that the proposed method outperforms existing methods in terms of the overall number of detection errors, and reduces the occurrence of false positives.

**Keywords:** change detection; hierarchical fusion; CRF; D-DRUNet; local boundary entropy



**Citation:** Zhang, J.; Liu, Y.; Wang, B.; Chen, C. A Hierarchical Fusion SAR Image Change-Detection Method Based on HF-CRF Model. *Remote Sens.* **2023**, *15*, 2741. <https://doi.org/10.3390/rs15112741>

Academic Editor: Gerardo Di Martino

Received: 26 April 2023

Revised: 23 May 2023

Accepted: 23 May 2023

Published: 25 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Change detection using remote sensing technology is a valuable research technique in the field of Earth observation [1–3]. It quantitatively analyzes multi-temporal images of the same geographical area to determine surface change characteristics [4].

The needs and standards in the field of communication are increasing [5,6] in response to the expanding range of human trajectories. The proposed integrated space–air–ground network has the potential to greatly enhance the efficiency of acquiring various types of information data and improve the computational efficiency of such data [7,8]. Incorporating edge computing tasks can improve the overall availability and scalability of the system [9–11]. By combining edge computing technology with remote sensing and satellite communication networks, the quality of satellite communication can be improved and the processing capability of satellite tasks can be enhanced, all while ensuring efficient resource scheduling [12,13]. Overall, the proposed integrated space–air–ground network has the potential to greatly influence the efficiency of acquiring various types of information data as well as the computational efficiency of such data.

The rich remote sensing data acquired by airborne remote sensing satellites can be used to describe urban land use, cover types and their detailed changes over time [2,14,15].

The field of remote sensing image change detection mostly utilizes a single optical image data source as the research target [16]. However, the imaging quality of optical remote sensing images is highly susceptible to complex weather, as well as satellite performance. Synthetic-aperture radar (SAR) images contain richer pixel information and clearer detailed information, which can effectively improve the limitations of optical image-based methods in the field of remote sensing change detection. SAR is an active microwave remote sensing technique [17,18] that could operate in less restricted natural environments [19] and plays an important role in remote sensing. The remote sensing change-detection method based on SAR images demonstrates significant advantages in the integrated space–air–ground tasks, such as urban building change detection [20], forest fire location [21], and geological disaster monitoring [22–24].

Currently, the predominant techniques of image change detection are based on difference-image detection [25], which involves analyzing the difference images generated from simultaneous phase images to obtain the final binary change maps [26]. Recently, deep learning has been applied to remote sensing image change detection, with the aim of learning complex features by constructing multilayer network models and training on huge amounts of data [27–30]. How to handle deep-level change information [31] is one of the key challenges in applying deep learning to change detection in the field of remote sensing imagery.

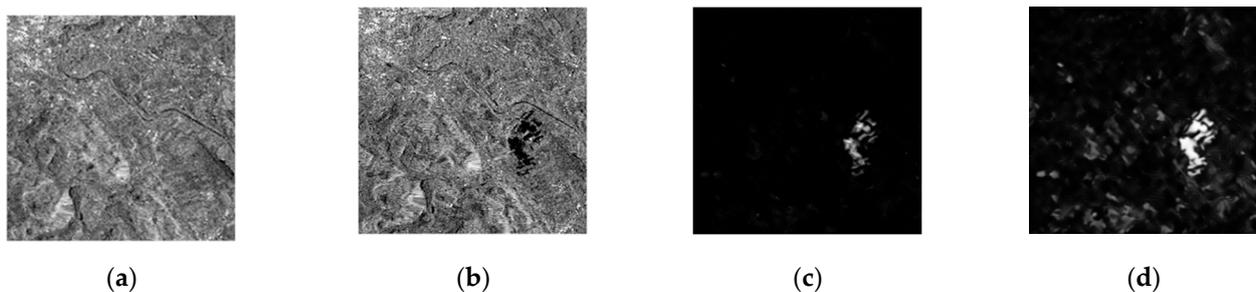
Taking advantage of the powerful feature-capture capability of the convolutional neural network (CNN) [32–34], Chen [28] proposed DSMS-CN and DSMSFCN methods for change detection of multi-temporal high-resolution remote sensing images based on multi-scale feature convolution units. Wu J et al. [29] proposed a deep supervised network (DSAHRNet) model. After the network extracts the decoding change information, the features are refined by parallel stacking of convolutional blocks, and more discriminative features can be obtained with the deep supervised module.

The attention mechanism is introduced in change-detection convolutional networks to focus on change information in complex information as well. Chen, J et al. [30] proposed DASNet, a change-detection model based on the dual attention full convolutional twin neural network, to obtain change-detection results by extracting rich features from dual temporal phase maps. Li et al. [35] introduced the pyramidal attention layer structure into the full convolutional network framework to further extract multi-scale variation information from the difference-feature maps processed by the original network encoder structure. Song et al. [36] proposed AGCDetNet by combining a fully convolutional network with an attention mechanism. The network takes into account the joint use of a spatial attention mechanism and channel attention mechanism. The paper verifies that AGCDetNet is able to enhance the discrimination of changing targets and backgrounds while improving the performance of feature representation for changing information. Lv et al. [37] proposed a hybrid attention semantic segmentation network (HAssNet), which incorporates a spatial attention mechanism and a channel attention mechanism based on a fully convolutional network [38]. This approach effectively utilizes multi-scale extracted features and global correlation to locate and segment targets in the image.

Nevertheless, most of the work only uses the deep features of CNN to build semantic feature descriptions, which ignores the fine-grained information contained in the shallow features [39]. Du et al. [40] designed a bilateral semantic fusion twin network (BSFNet) integrating shallow and deep semantic features in order to better map dual-temporal images to semantic feature domains for comparison, and obtained pixel-level change results with more complete structures. The UNet network [41] has outstanding performance in the field of semantic segmentation, and has been widely adopted into the field of remote sensing change detection. Zhi, Z et al. [42] proposed a UNet-based CLNet network with a cross-layer structure to improve the change-detection accuracy by improving the way of contextual information fusion. By incorporating the advantages of dense connections for multi-scale information mining within UNet++, Li et al. [43] introduced multiple sources of information to supplement the channel information of remote sensing images

in their framework. The resulting model exhibits excellent performance across various datasets. Chen et al. [44] combined the attention mechanism with UNet to design the Siamese\_AUNet twin neural network. The model performs well in solving the problems related to weak change detection and noise suppression. Furthermore, for change detection, the conditional random field (CRF) based on probabilistic graphical models (PGM) [45] and the Markov random field (MRF) [46] have been introduced. Zhang et al. [47] used the CRF model to improve the traditional change-detection method. A half-normal CRF (HNCRF) method is proposed to construct the interaction between pixel points in the spatial analysis of difference images, which is effective when the change region is small. Lv et al. [48] proposed a hybrid conditional random field (HCRF) model that combining traditional random field methods with object-based techniques. The improved model fully exploits the spectral spatial information, thereby enhancing the change-detection performance of high-spatial-resolution remote sensing images, improving the traditional change-detection method with the CRF model. A half-normal CRF (HNCRF) method is proposed to construct the interaction between pixel points in the process of the spatial analysis of difference images, which is effective when the change region is not significant. However, CRF ignores the image–global-distribution relationship. The localization accuracy of the fully connected conditional random field (FCCRF) [49,50], coupled with the recognition ability of the deep convolutional neural network shows better boundary localization in the change-detection results. Y. Shang et al. [51] introduced a novel approach to mitigate the issue of excessive feature smoothing in the fully connected conditional random field (FCCRF) model by incorporating region boundary constraints. This method involves obtaining a complete set of pixels in a multi-temporal image, and calculating the average pixel probability, to enable the refinement and classification of boundary information through the regional potential function. Gong et al. [52] proposed the patch matching method for fully connected CRF optimization, which combined with the results of semantic segmentation network to detect architectural changes in dual-temporal images. However, the post-processing method based on the front-end output still causes the loss of change-detection information. To address this limitation, Zheng et al. [53] proposed a new end-to-end deep twin CRFs network (PPNet) for high-resolution remote sensing images. The detection results obtained by PPNet are able to refine the edges of change regions and effectively eliminate noise.

Overall, deep learning has shown promising results in change detection of the remote sensing image, and the proposed models and techniques have significantly improved the accuracy and efficiency of the process. Although the change-detection methods for remote sensing imagery have made significant progress, there are still challenges in detecting the direction of areas of change by analyzing difference images, which can be summarized as follows. Firstly, difference images may cause a semantic collapse phenomenon. The original temporal image, as shown in Figure 1a,b, contains obvious feature classification information, i.e., semantic information. However, the difference operation, as shown in Figure 1c,d, can quickly locate the change region, but it also leads to a typical semantic collapse phenomenon where the semantic classification information disappears.



**Figure 1.** Plot of the difference of Berne data processed by different methods. (a) Multitemporal image 1. (b) Multitemporal image 2. (c) Log-ratio difference image. (d) Mean-ratio difference image.

Secondly, the multimodal difference images provide complementary information. As demonstrated in the two modal difference images of the Berne data in Figure 1c,d, the log-ratio difference image has less interference, but the change information is weak, which results in serious missed alarms. On the other hand, the mean-ratio difference image has prominent change information, but the strong interference leads to high false alarms. Thus, the interference and change performance of the two modal difference images are quite different, and improving the semantic perception of change detection through the complementary information of modal difference images is key to improving the overall performance of change detection.

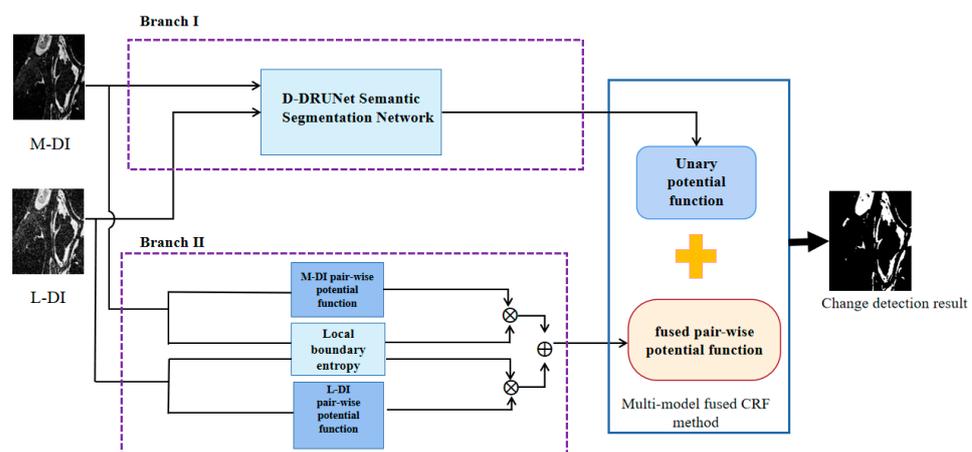
To address the challenges mentioned above, this paper proposes a hierarchical fusion SAR image change-detection method based on hierarchical fusion conditional random field (HF-CRF). The main contributions are as follows.

- Designing a dynamic region convolutional semantic segmentation module with a dual encoder–single decoder structure (D-DRUNet). It involves constructing a unary potential function by fusing multimodal difference-image features using neural networks, and enhancing the semantic perception capability of the CRF model.
- Introducing a boundary prior to constructing a pair-wise potential function based on multimodal dynamic fusion, and enhancing the boundary perception capability of the CRF model.

This paper consists of the following three parts: the Method section provides a detailed description of the principle and implementation steps of the proposed method; the Experiment section provides the experimental results and analysis; and the Conclusion section summarizes the article.

## 2. Materials and Methods

The principle framework of the proposed HF-CRF hierarchical fusion SAR image change-detection method is shown in Figure 2. The HF-CRF method employs the CRF as the fusion framework. To establish the unary potential function and the pair-wise potential function, the mean-difference image [54] and the logarithmic-difference image [55] are used through neural networks and local sliding windows, respectively. In Branch I, the D-DRUNet neural network is adopted to fuse multimodal difference-image features. The network structure is a dynamic convolutional UNet with a dual encoder and a single decoder, and the encoder completes the semantic-feature fusion at the bottom layer, and the dual-jump connection structure obtains the fused segmented image during the decoding process. In Branch II, a CRF pair-wise local boundary entropy potential function is constructed by using a local sliding window to extract the boundary a priori information of the multi-modal difference image. Finally, the CRF model iteratively reasons the fused energy potential function to obtain the optimized change-detection results.

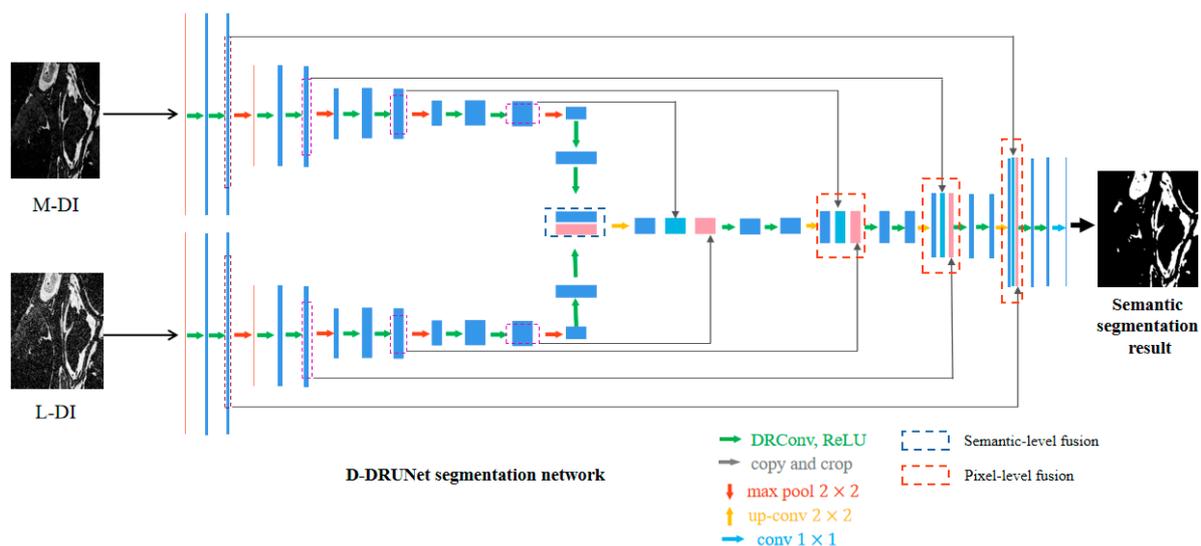


**Figure 2.** The schematic diagram of HF-CRF hierarchical fusion structure.

### 2.1. D-DRUNet Fusion Semantic Segmentation Network

UNet networks based on an encoder–decoder structure are commonly employed in medical imaging and change-detection segmentation tasks, due to their ability to learn from small datasets [56]. However, U-Net does not have multimodal feature fusion capability, due to the limitation of a single encoder structure. Additionally, the fixed convolutional kernel model limits its ability to generalize feature extraction, which results in difficulty in detecting change details and the increasing of missed alarms in detection.

In order to solve the above two problems, we designed a novel D-DRUNet segmentation network model, which mainly includes three features: adopting a dual encoder–single decoder structure to solve the network-level multimodal fusion design problem; introducing a dynamic region convolution kernel (DRConv) and designing a multiscale guide mask module to improve the feature extraction capability of the network; a hierarchical fusion mechanism to realize the bottom and upsampling stages in turn of the multi-level network feature fusion. The specific network structure is shown in Figure 3, and we elaborate it in the following sections.



**Figure 3.** D-DRUNet Fusion Semantic Segmentation Network: Rational Block Diagram.

#### 2.1.1. Dual-Encoder and Single-Decoder Structure

The proposed method utilizes the log-ratio difference map and mean-ratio difference map as dual encoder inputs, enabling the network to extract change features simultaneously from different modal difference images. The method employs different fusion strategies at different stages of image encoding and decoding. The encoder consists of a convolution operation and a downsampling operation, where the convolution process characterizes the image information and the downsampling process obtains the contextual information of the image. The shallow features obtained at the two-way encoding stage retain rich detailed information, while the deep features ensure semantic structure integrity. The single-way decoder side performs multiple upsampling operations for the fused features to recover the compressed feature maps in the spatial dimension in the original input size, layer by layer. The dual encoder–single decoder structure effectively improves the information perception capability of the network by designing a multimodal feature extraction and fusion approach.

#### 2.1.2. Layered Fusion Mechanism Design

We proposed a layered feature fusion mechanism for multimodal disparity images based on the encoder–decoder structure. This mechanism enhances the feature expression capability and regional change localization capability of the disparity maps. The proposed

method employs different fusion strategies at different stages of image encoding and decoding. The layered feature fusion is realized by two stages. In the first stage, the depth features extracted from the two difference images are fused in the bottom layer in a splicing way to achieve semantic fusion before up-sampling and decoding. In the second stage, the multimodal information of the corresponding resolution is supplemented by the two-way jump connection during the decoding up-sampling to achieve pixel fusion. In the decoding pixel-level fusion structure marked in the decoding side of Figure 3 network, the splicing features at the same level contain three parts: the logarithmic modal coding features and the mean modal coding features at the encoding side, and the up-sampling fusion features at the decoding side. In order to improve the fusion efficiency and reduce the computational effort, both semantic-level fusion and pixel-level fusion of features are used in the concatenation method.

### 2.1.3. Dynamic Convolution Kernel with Multi-Scale Guide Mask

Traditional UNet networks use the CNN structure, and the convolution kernel size needs to be determined in advance, which is difficult to adapt to changes in dynamic remote sensing image content. The D-DRUNet network introduces a dynamic region sensing method [57] and adopts the feature pyramid network (FPN) structure [58] to improve the dynamic convolutional guide mask generation method to dynamically divide the spatial dimensional distribution, according to the input features.

Dynamic region-aware convolution (DRConvs) consists of a learnable guide mask module and a filter generation module that automatically generates region-sharing patterns of filters, based on each input image’s features. The guide mask module divides the features with similar semantic information into the same region, which determines the distribution of filters in the spatial dimension; the filter module generates the corresponding filters that would be assigned to different regions, and different filters extract the information at different abstraction levels.

In particular, the D-DRUNet network is designed with the FPN structure to improve the guide mask generation of DRConvs by fusing features among three scales to increase the content localization capability of the guide mask, and the principle of the method is shown in Figure 4. Figure 5 shows the visualization result graph of the improved guide mask region segmentation on the Berne dataset. Figure 5a shows the Berne log and mean-difference maps, and Figure 5b shows the results of the guide mask delineation on the corresponding difference maps with higher false alarms; Figure 5c shows the region delineation results of the FPN structure guide mask, and the accuracy improvement is obviously significant.

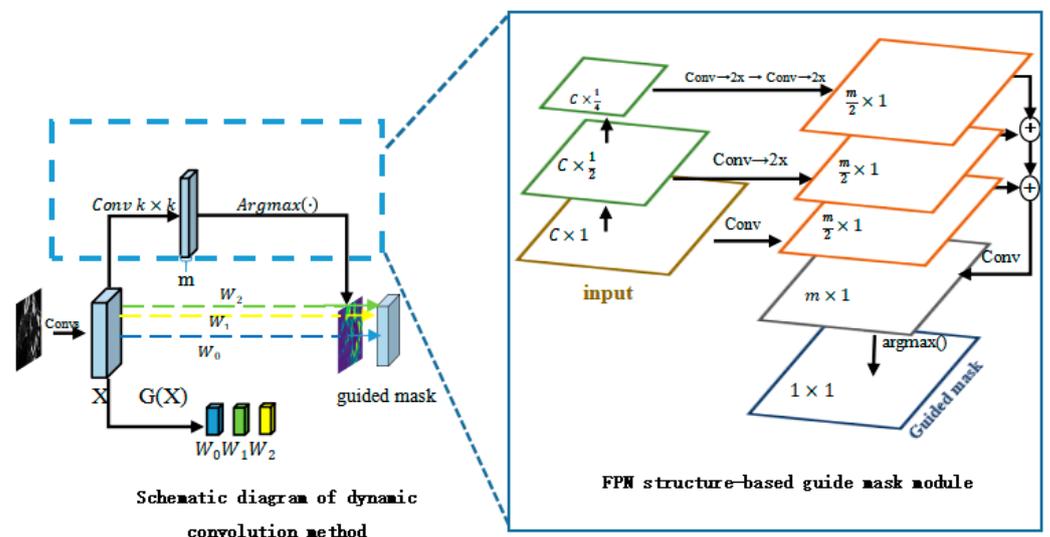
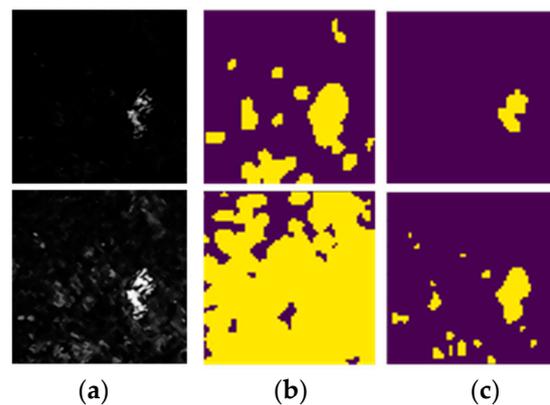
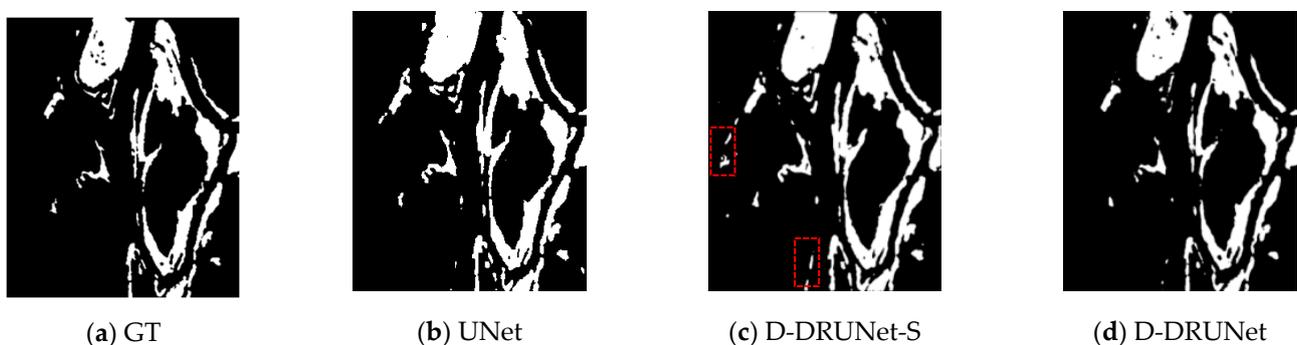


Figure 4. Improved dynamic convolution structure diagram.



**Figure 5.** Visualization of guide mask region division (the Berne data).

To verify the effect of the D-DRUNet network, we conducted local validation experiments on the Ottawa dataset, and the results are shown in Figure 6 and Table 1. The D-DRUNet-S method is a semantic segmentation model with a decoder structure that includes a single skip connection. It shows significant improvement in detection performance, compared to the UNet network. However, the detection results exhibit a high false alarm rate, as shown in the dashed box in Figure 6c. This is likely caused by severe reconstruction distortion due to the lack of encoding information from another modality, resulting in an imbalanced performance between false alarms and missed alarms, and poor Kappa performance. The D-DRUNet achieves a favorable trade-off between false positives and false negatives by adopting a double-skip-connection fusion method, resulting in a significant improvement in performance.



**Figure 6.** Experimental test result graph of Ottawa data section. The red dashed boxes mark the false alarms detected by the D-DRUNet-S method.

**Table 1.** Performance metrics of different methods on Ottawa dataset.

Method	FP	FN	OE	PCC	Kappa
U-Net	5076	3451	8527	91.60%	0.6969
D-DRUNet-S	3908	274	4182	95.88%	0.8583
D-DRUNet	1206	894	2100	97.93%	0.9229

## 2.2. Boundary Entropy Dynamic Fusion CRF Model

CRF [59] is a well-known discriminative model, widely utilized in various domains such as image segmentation [60]. The CRF model for change detection [61] comprises two components. Modeled as a probability distribution map, the unary potential function represents the intrinsic energy of each pixel, which is generated by processing the clustering algorithm with either the temporal phase map [62] or the semantic segmentation network [53]. The pair-wise potential function models the second-order neighborhood potential energy,

incorporating both the positional and color information as feature functions, which can be expressed using Equation (1) [59], where  $Z(X|Y)$  is the regularization constant,  $E(X|Y)$  is the energy function, and  $\psi_u(x_i)$ ,  $\psi_p(x_i, x_j)$  are the expressions for the one-dimensional potential function and the two-dimensional potential function, respectively.

$$\begin{aligned} P(X|Y) &= \frac{1}{Z(Y)} e^{-E(X|Y)} \\ E(X|Y) &= \sum_i \psi_u(x_i) + \sum_{i<j} \psi_p(x_i, x_j) \end{aligned} \quad (1)$$

The fully connected CRF model calculates the pair-wise potential function by considering all pixels in the image as neighborhood relations. It employs the mean-field inference model [45] to achieve the model solution through downsampling convolution. Herein, we introduced a hybrid pair-wise potential function using the fully connected CRF model. The segmentation result obtained from D-DRUNet serves as a unary potential function, while the multimodal difference map is incorporated as Equations (2)–(6).  $\alpha_i$  and  $\beta_i$  are the corresponding weights of the first pixel of the two modal difference images;  $P(x_i = l_k)$  is the probability that the predicted label  $x_i$  of the  $i$ -th pixel point in the network segmentation result is  $l_k$ ;  $u(x_i, x_j)$  is the class-label consistency function, which limits the energy which can be conducted from one to the other under the condition of consistent labels;  $\omega^m$  is the weight parameter,  $K_G^m(f_i, f_j)$  is the feature function, and the vector  $f_i$  and  $f_j$  is the feature representation of the pixel  $i$  and  $j$  under the same feature space.

$$E = \sum_i \psi_u(x_i) + \sum_{i<j} \alpha_i \psi_L(x_i, x_j) + \sum_{i<j} \beta_j \psi_M(x_i, x_j) \quad (2)$$

$$\Psi_u(x_i) = -\ln(P(x_i = l_k)) \quad (3)$$

$$\Psi_L(x_i, x_j) = u(x_i, x_j) \sum \omega^m K_G^m(f_i, f_j) \quad (4)$$

$$\Psi_M(x_i, x_j) = u(x_i, x_j) \sum \omega^m K_G^m(f_i, f_j) \quad (5)$$

$$K_G^m(f_i, f_j) = \exp\left(-\frac{1}{2}(f_i, f_j)^T \Lambda^{(m)}(f_i, f_j)\right) \quad (6)$$

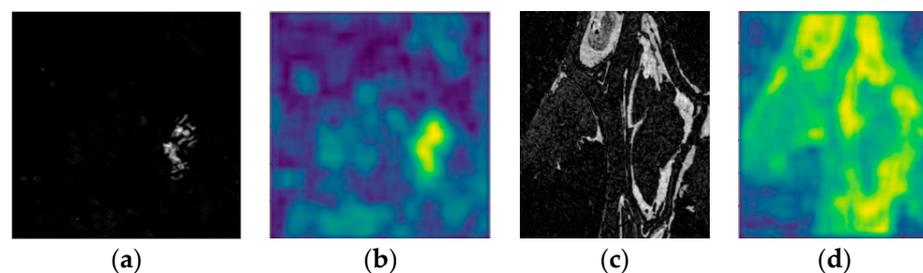
The pair-wise potential function in Equation (2) is a hybrid potential function that significantly influences the final iterative results. The weight coefficients  $\alpha_i$  and  $\beta_i$  play a key role in this function. Two factors must be considered while determining these weights. Firstly, the dynamic changes in the semantic content of the multimodal difference map should be reflected, and fixed weights may not adequately account for the dynamic content of the image. Secondly, the boundary information in the multimodal difference image should be strengthened, as it can accurately reflect the change region, and is crucial for the convergence of the CRF model.

To address these issues, we proposed a dynamic weight construction method based on local boundary entropy. Image information entropy is introduced to measure the dynamic change of image content in our method, which is defined as in Equation (7) [63], where  $p(n)$  is the proportion of the pixel points with gray value  $n$  in the image to the total pixel points in the image.

$$Q = -\sum_t p(n) \times \ln p(n) \quad (7)$$

Figure 7 displays the log-ratio modal difference maps and corresponding local information entropy feature maps for both the Berne and Ottawa datasets. It is observed that entropy values could reflect the semantic content changes in the difference images. To be specific, we first partition the images into sub-blocks, and calculate the corresponding image information entropy values for each sub-block. These entropy ratios serve as the foundation for determining dynamic weights. Additionally, an image-sharpening tech-

nique is utilized to enhance the boundary information in the difference maps, whereby its significance in the dynamic weight calculation is increased. We choose three different sharpening operators for performance comparison. Both the Sobel operator and the Prewitt operator use a pixel window to compute the first-order derivatives of the image in the x and y directions, respectively, with the edge operator for the convolution summation operation, thus processing each pixel point of the image to extract the edges of the image. The Laplacian operators are based on the second-order derivatives, specifically, the Laplacian operator with a four-neighborhood template is used in this paper. The parameters of operators are shown in the Table 2. Figure 8a,d show the sharpening results of the sharpening operators [64,65] on the Berne dataset, and (e) to (h) show the figures of the sharpening results on the Ottawa dataset. The comparison shows that the Sobel operator [66] has the optimal boundary strengthening effect on both datasets, as shown in the blue box. Therefore, we use the Sobel operator as the final sharpening operator.



**Figure 7.** Difference images of two data sets and the entropy value of respective local information. (a) Berne difference image. (b) Information entropy of Berne data. (c) Ottawa difference image. (d) Information entropy of Ottawa data.

**Table 2.** Parameters of the three image-sharpening operators.

	$S_x$	$S_y$
The Laplacian operator	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$
The Prewitt operator	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
The Sobel operator	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

In summary, the proposed local boundary entropy pair-wise potential function is constructed as follows.

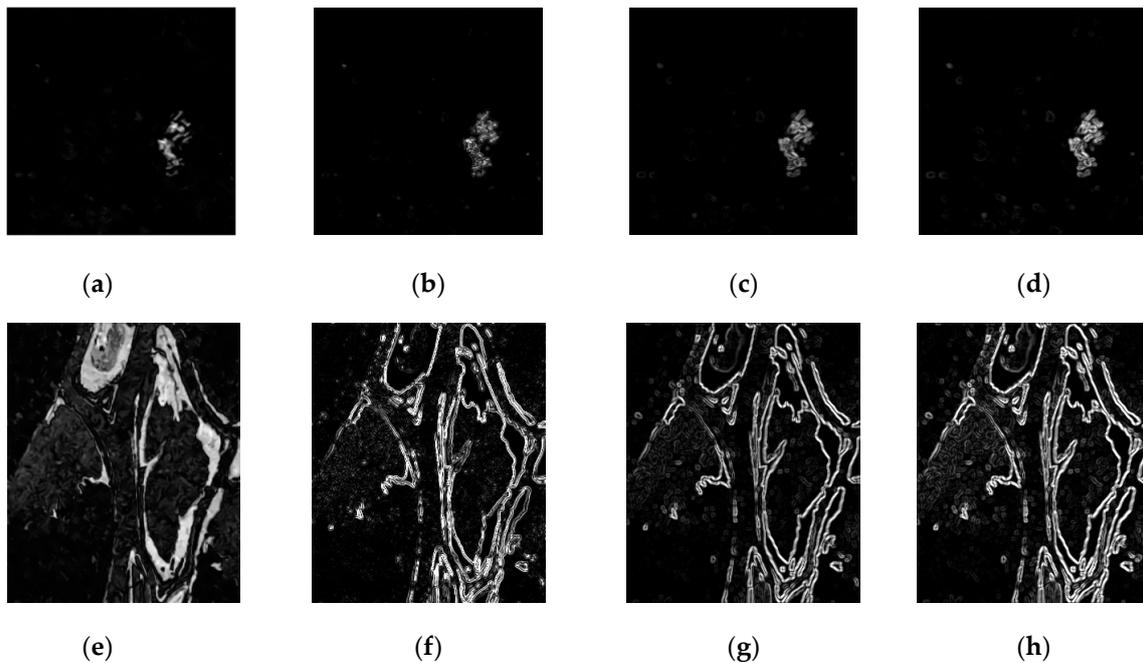
- First, sharpening image by using the Sobel operator to obtain sharpened log-difference images  $X_L$  and sharpened mean-difference images  $X_M$ , respectively.
- Second, sliding window with fixed block size on two image blocks  $X_l^{(i)}$  and  $X_m^{(i)}$  with the same position on the sharpened images  $X_L$  and  $X_M$ , respectively, and calculating them based on local information entropy, as shown in Equations (8) and (9), where sub-blocks  $Q_l$  and  $Q_m$  represent the image boundary entropy values on the  $i$ -th image block  $X_l^{(i)}$  and  $X_m^{(i)}$  corresponding to the modalities.

$$Q_l = -\sum_i p(n) \times \ln p(n) \quad (8)$$

$$Q_m = -\sum_j p(n) \times \ln p(n) \quad (9)$$

- Third, the entropy ratio method is used to calculate the log-difference modal counterpart weights  $\alpha$  and the mean-difference modal weights  $\beta$  in the mixed pair-wise potential function.

$$\alpha = \frac{Q_m}{Q_l + Q_m}, \beta = \frac{Q_m}{Q_l + Q_m}, \text{s.t.} \begin{cases} \alpha \in [0, 1] \\ \beta \in [0, 1] \\ \alpha + \beta = 1 \end{cases} \quad (10)$$



**Figure 8.** Two dataset showing sharpening-processing difference boundary maps. (a) Original difference image of Berne data. (b) Sharpening results with the Laplacian operator. (c) Sharpening results with the Prewitt operator. (d) Sharpening results with the Sobel operator. (e) Original difference image of Ottawa data. (f) Sharpening results with the Laplacian operator. (g) Sharpening results with the Prewitt operator. (h) Sharpening results with the Sobel operator.

Figure 9 shows the result plots of the D-DRUNet semantic segmentation network and the HF-CRF model for detection on the Ottawa dataset. By analyzing the detection results, it can be concluded that the dynamic fusion iterative structure CRF optimization significantly reduces the false detection phenomenon, with finer change boundaries. The improved CRF model uses a multimodal fusion method to achieve information complementarity between modes, and a more accurate change-detection region is obtained.



**Figure 9.** Results of different methods on Ottawa data.

### 3. Experiment and Analysis

#### 3.1. Datasets

In this study, we utilized eight publicly available SAR image sets to generate mean difference maps and log-ratio difference maps. To construct the training set, we employed data enhancement techniques. The effectiveness of our algorithm was then evaluated using four SAR datasets, with the composition of the training set being readjusted for each validation process. All datasets consist of two temporal phase maps and the corresponding GTs. The first dataset is a remotely sensed image of the Berne area with a pixel resolution of  $301 \times 301$ , as shown in Figure 10; the second dataset is a remotely sensed image of the Ottawa area with a pixel resolution of  $290 \times 310$ , as shown in Figure 11; the third dataset is a remote sensing image of the Mexico region with a pixel resolution of  $256 \times 256$ , as shown in Figure 12; the fourth dataset is a remote sensing image of the San Francisco region with a pixel resolution of  $512 \times 512$ , as shown in Figure 13.

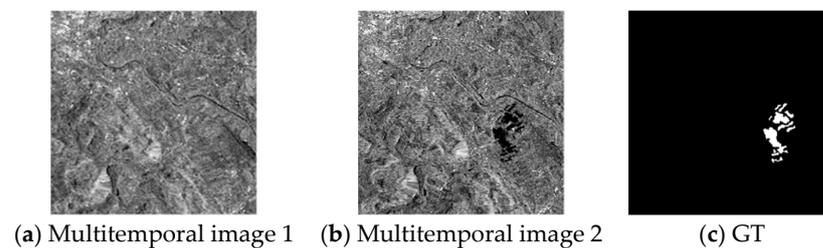


Figure 10. The Berne dataset.

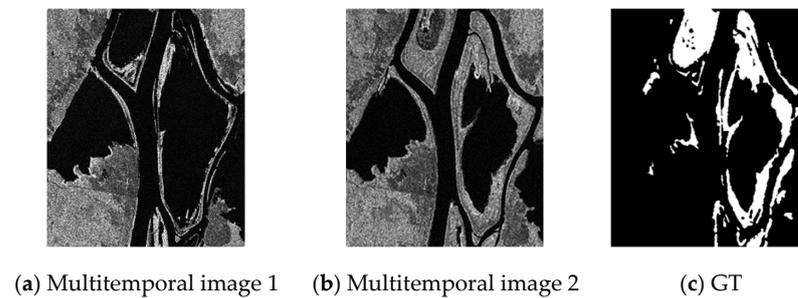


Figure 11. The Ottawa dataset.

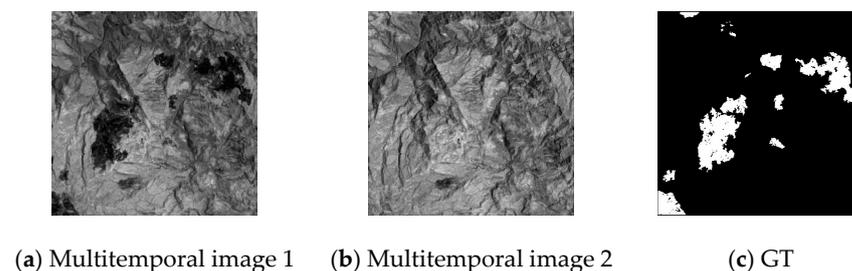


Figure 12. The Mexico dataset.

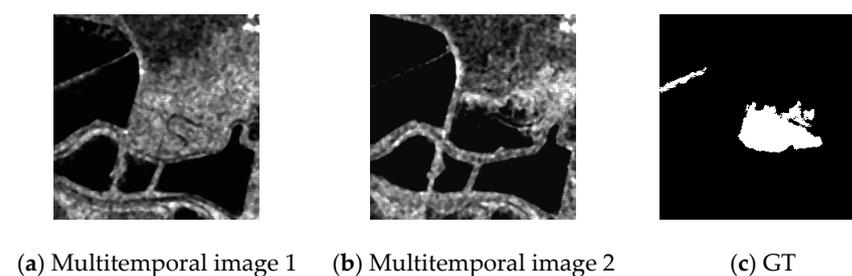


Figure 13. The San Francisco dataset.

### 3.2. Parameter Setting and Evaluation Indexes

D-DRUNet parameter settings: the number of guide mask regions  $m = 2$ ; the learning rate is set to 0.001; the batch size is set to 8. The image block size is  $256 \times 256$ ; the overlap step is 4; and the number of training iterations is 200.

Boundary entropy dynamic fusion CRF parameter setting:  $\omega^1 = 20$ ;  $\omega^2 = 20$ ;  $\theta_\alpha = 12$ ,  $\theta_\beta = 12$ ,  $\theta_\gamma = 2$ ; the number of iterations is 3; the boundary entropy sliding window size is  $25 \times 25$  and the step size is 1.

Simulation of experimental environment: Core (TM) i7-7820X CPU @ 3.60 GHz, Nvidia RTX 2080Ti  $\times 2$ , Ubuntu18.04, python3.7, pytorch1.4, cuda10.0.

The number of false alarms, FP; the number of missed detection, FN; the correct rate, PCC; and the Kappa coefficient, which measures the classification accuracy, are used as the evaluation metrics for the experiments in this paper, which could be formulated as follows.

$$PCC = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%,$$

$$Kappa = \frac{PR_0 - PR_C}{1 - PR_C}.$$

### 3.3. Ablation Experiments

We conducted ablation experiments on the Berne dataset to verify the effectiveness of the D-DRUNet fusion semantic segmentation network, and the experimental results are shown in Figure 14 and Table 3. Methods 1 and 2 are the change-detection results with unimodal difference images as input, while methods 3 and 4 are the change-detection results with bimodal difference images as input. Moreover, method 1 utilizes the conventional CNN convolutional kernel, and methods 2, 3, and 4 use the dynamic convolutional kernel. Table 4 shows the performance metrics of the change-detection results of the four methods. The comparison results in Figure 14a with the ground truth revealed that the segmentation results of method 1 had serious false alarms and blurred boundaries of the change region. However, method 2, which employed a single encoder, detected a more comprehensive change region, with significantly reduced false alarms. Method 3, with an asymmetric encoder and jump connection structure, destroyed the change-region structure and resulted in obvious leakage detection in the detection results. In contrast, method 4, which utilized a dual encoder and a dual jump structure, effectively reduced the leakage alarm, with more refined change boundaries. It can be seen that the D-DRUNet network incorporates multimodal difference information, and the detection effect is significantly improved.

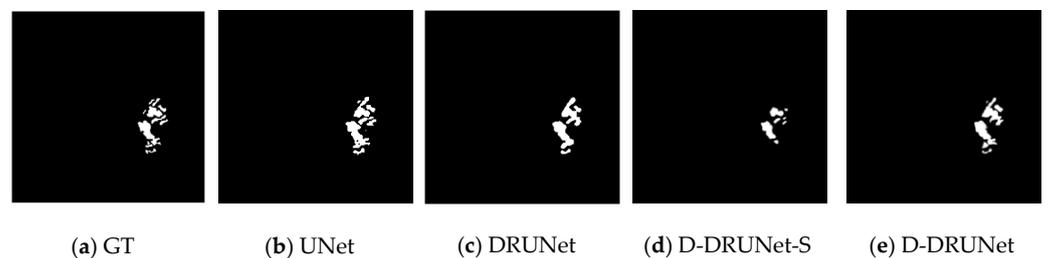


Figure 14. Ablation experiment results of D-DRUNet.

Table 3. Results of ablation experiments of the D-DRUNet.

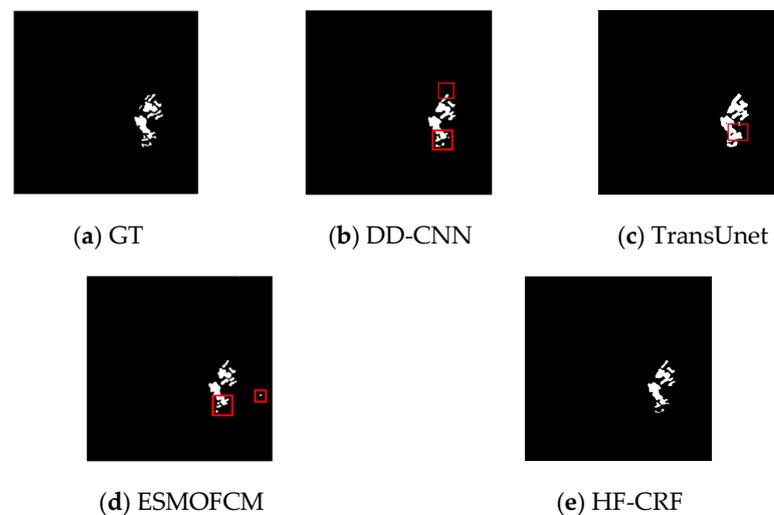
Method	DRConv	Single Encoder	Dual Encoder	Single-Skip Connection	Dual-Skip Connection	FP	FN	OE	PCC	Kappa
1		✓		✓		1384	25	1409	98.44%	0.6091
2	✓	✓		✓		311	137	448	99.51%	0.8172
3	✓		✓	✓		11	532	543	99.48%	0.7075
4	✓		✓		✓	230	152	382	99.58%	0.8379

**Table 4.** Performance metrics of different methods on Ottawa dataset.

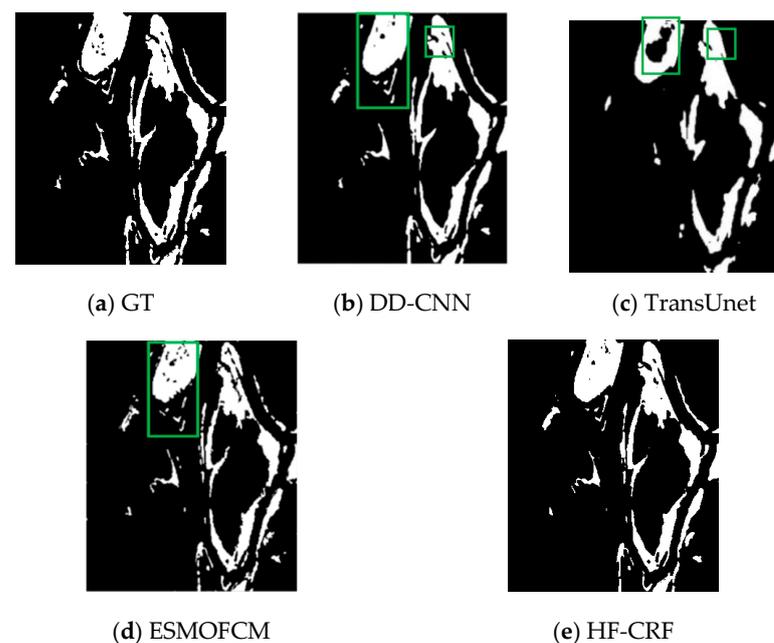
Method	FP	FN	OE	PCC	Kappa
D-DRUNet	1206	894	2100	97.93%	0.9229
HF-CRF	208	1163	1371	98.65%	0.9480

### 3.4. Comparison Experiments

In this paper, we select the mainstream DD-CNN [67], Trans Unet [68], and ES-MOFCM [69] in change detection to verify the performance of the HF-CRF model proposed in this paper on the Berne, the Ottawa, the Mexico, and the San Francisco datasets; the change-detection results are shown in Figures 15–18, respectively, and the performance index analysis results are shown in Tables 5–8.



**Figure 15.** Detection results of different methods on Berne data. The red boxes indicate the parts where the performance of comparison methods that have poor performance, including the blurred boundary and the false alarm phenomenon.



**Figure 16.** Detection results of different methods on Ottawa data. The green boxes indicate the parts of the comparison methods that have poor performance, including blurred boundaries and missed alarms.

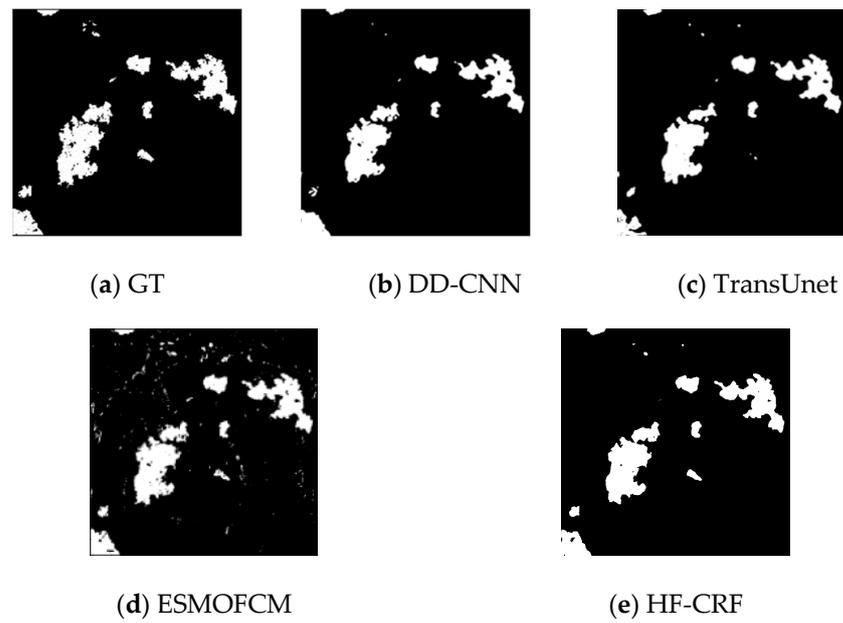


Figure 17. Detection results of different methods on Mexico data.

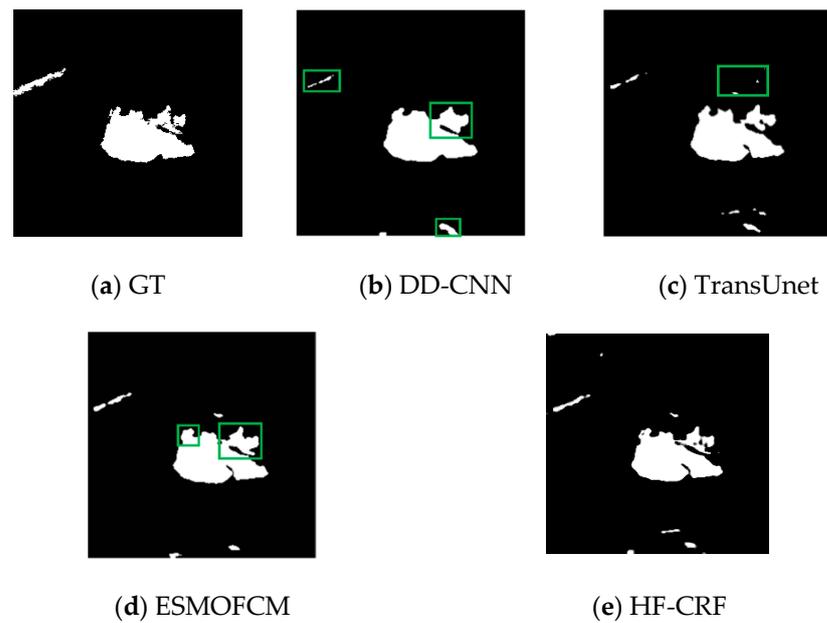


Figure 18. Detection results of different methods on San Francisco data. The green boxes indicate the parts where the performance of comparison methods that have poor performance, including the blurred boundary and the false alarm phenomenon.

Table 5. Performance metrics of different methods on the Berne dataset.

Method	FP	FN	OE	PCC	Kappa
DD-CNN	130	200	330	99.64%	0.8596
TransUnet	350	49	399	99.39%	0.7925
ESMOFCM	112	199	311	99.65%	0.8643
Our Method	105	187	292	99.68%	0.8673

**Table 6.** Performance metrics of different methods on the Ottawa dataset.

Method	FP	FN	OE	PCC	Kappa
DD-CNN	848	842	1690	98.33%	0.9375
TransUnet	881	2543	3424	95.78%	0.8474
ESMOFCM	540	1932	2472	97.56%	0.9051
Our Method	208	1163	1371	98.65%	0.9480

**Table 7.** Performance metrics of different methods on the Mexico dataset.

Method	FP	FN	OE	PCC	Kappa
DD-CNN	1076	3692	4768	98.18%	0.8918
TransUnet	1543	3674	5217	98.01%	0.8827
ESMOFCM	3637	6670	10,307	87.93%	0.8030
Our Method	1648	2174	3822	98.54%	0.9165

**Table 8.** Performance metrics of different methods on the San Francisco dataset.

Method	FP	FN	OE	PCC	Kappa
DD-CNN	428	393	821	98.75%	0.9060
TransUnet	321	509	830	98.73%	0.9058
ESMOFCM	295	437	732	98.88%	0.9170
Our Method	359	354	713	98.91%	0.9181

The detection results of the four methods on the Berne dataset are shown in Figure 15b–e. The comparison of the red areas shows that the TransUnet method has the blurred contours and higher false alarms, followed by DD-CNN, while ESMOFCM has obviously missed alarms, and the proposed method maintains clearer and more accurate segmentation boundaries, with more complete contours. The performance on the Ottawa dataset is shown in Figure 16b–e. As shown in the green box area 1 of Figure 16, the TransUnet model forms larger joint holes in the hole area with serious leakage, the ESMOFCM model produces more holes and leakage in the green area 2, the DD-CNN and HF-CRF models have more complete contours overall with less leakage, while the contours of HF-CRF model are clearer than the DD-CNN model. From the results in Table 6, the overall results of the ESMOFCM model and TransUnet are not satisfactory; the DD-CNN model has more balanced false and leaky alarms and higher Kappa coefficients, and the HF-CRF model has slightly higher false alarms and low leaky alarms but the lowest total errors, so it achieves the highest detection performance, which is due to the better segmentation baseline achieved by the D-DRUNet network. The dynamic second-order potential function further refines the contours and edges by iteration, but the balance of the HF-CRF false alarms is not enough, and the number of missing detections is higher. The next improvement step of the proposed method in this paper will also focus on reducing the missed alarms.

Figure 17 shows the detection results of different models on the Mexico dataset. The ESMOFCM model has more false alarms and a poor detection performance, due to background interference. In contrast, the DD-CNN, TransUnet, and HF-CRF models have more complete contours. However, the DD-CNN and TransUnet models have blurred boundary information and missed alarms in certain areas. On the other hand, the proposed method captures more detailed information and has a clear-edge detection effect. According to Table 7, the proposed method achieves optimal performance in terms of both PCC and Kappa evaluation indexes by reducing false detections and missed alarms.

The detection results of the methods in this paper on the San Francisco data are presented in Figure 18. All four methods obtain more complete change-region detection results. Both modal difference images of the San Francisco dataset are strongly disturbed by the background when modal complementarity is weak, so the performance advantage of HF-CRF for change-detection results on this dataset is diminished, while the FCM structure

based on traditional algorithms is suitable for such data. Our method has a significant improvement over the neural network-based comparison method, and achieves better results in terms of generalization ability for multiple datasets. The method proposed in this article demonstrates the best performance in terms of boundary detection accuracy and the missed alarms. As shown in Figure 18, the portion circled in green highlights the superior accuracy of this method in boundary segmentation compared to the other three methods, with more precise edge contours.

The HF-CRF network fuses the semantic information of the modal difference maps at multiple levels to ensure the integrity of the changing semantics, and designs a CRF model with a fusion iteration structure based on local boundary entropy to realize the information complementation between the multimodal difference maps and enhance the optimization capability of the CRF model.

### 3.5. Efficiency Comparison Experiments

The neural-network-based comparison methods DD-CNN, as well as TransUNet in Section 3.4, are selected for the comparison of network computational efficiency. The network parameters (Params) and the number of floating-point operations (FLOPs) are used as effective indicators of the computational complexity of the network model. The specific calculation results are shown in Table 9.

**Table 9.** Comparison of module parameters and computation amount.

Method	Params	FLOPs
DD-CNN	11.2 M	73,279.01 M
TransUNet	93.19 M	64,357.53 M
Our method	63.27 M	85,928.02 M

We propose a new change-detection model that outperforms existing methods in terms of evaluation metrics. Unlike the DD-CNN approach, our model does not incorporate residual structures, resulting in a significantly larger number of parameters. However, compared to TransUNet, our model has fewer parameters while achieving better performance. This trade-off between model complexity and detection performance is a key contribution of our work.

Our proposed model utilizes a two-branch modal encoding strategy and a two-skip connection structure, which allows for effective fusion of multimodal differences. This approach introduces a new avenue for multimodal fusion in change detection.

## 4. Conclusions

To address the problem of semantic collapse caused by disparity computation, we proposed a hierarchical fusion SAR image change-detection model based on HF-CRF. This model adopts a hierarchical structure to compensate for the lost semantics, uses a D-DRUNet neural network to realize the fused semantic segmentation of multimodal disparity maps, constructs the first-order potential function of CRF, and uses local boundary entropy to realize the fused second-order potential function, which accurately reflects the dynamic semantic changes of multi-modal images. The CRF model is driven to converge with the change boundary accurately by minimizing the energy function. To verify the effectiveness of the method, we conduct experiments on the publicly available SAR dataset. The experimental results show that the HF-CRF model proposed in this paper achieves superior results on the test dataset compared with both traditional methods, as well as deep learning methods. In future work, we will combine self-supervised learning and the Siamese network structure to directly locate remote sensing change-detection regions in the spatiotemporal domain, to address the semantic loss caused by discrepant operations.

**Author Contributions:** Conceptualization, J.Z. and Y.L.; methodology, J.Z. and Y.L.; software, Y.L.; validation, J.Z., Y.L., B.W. and C.C.; formal analysis, J.Z.; investigation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, B.W.; supervision, C.C.; funding acquisition, J.Z. and C.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Key Research and Development Program of Shaanxi (No. 2023-ZDLGY-54, 2023-GHZD-44, 2021ZDLGY02-09, 2019ZDLGY13-07, 2019ZDLGY13-04), the National Natural Science Foundation of China (62072360, 61902292, 61971331, 62001357, 62072359, 62172438), the Natural Science Foundation of Guangdong Province of China (2022A1515010988), the Xi'an Science and Technology Plan (20RGZN0005) and the Key Project on Artificial Intelligence of Xi'an Science and Technology Plan (2022JH-RGZN-0003, 2022JH-RGZN-0103, 2022JH-CLCJ-0053).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We are grateful to those who participated in the data processing and manuscript revisions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yin, S.J.; Wu, C.Q.; Wang, Q.; Ma, W.D.; Zhu, L.; Yao, Y.J.; Wang, X.L.; Wu, D. A review of the research progress of multi temporal remote sensing image change detection methods. *Spectrosc. Spectr. Anal.* **2013**, *33*, 3339–3342.
2. Singh, A. Review Article Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sens.* **1989**, *10*, 989–1003. [[CrossRef](#)]
3. Liu, X.; Li, J.; Sahli, H.; Meng, Y.; Huang, Q. Improving unsupervised flood detection with spatio-temporal context on HJ-1B CCD data. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 4402–4405.
4. Lv, N.; Chen, C.; Qiu, T.; Sangaiah, A.K. Deep Learning and Superpixel Feature Extraction Based on Contractive Autoencoder for Change Detection in SAR Images. *IEEE Trans. Ind. Inform.* **2018**, *14*, 5530–5538. [[CrossRef](#)]
5. Wang, B.; Xu, K.; Zheng, S.; Zhou, H.; Liu, Y. A Deep Learning-Based Intelligent Receiver for Improving the Reliability of the MIMO Wireless Communication System. *IEEE Trans. Reliab.* **2022**, *71*, 1104–1115. [[CrossRef](#)]
6. Chen, C.; Xiao, T.; Qiu, T.; Lv, N.; Pei, Q. Smart-Contract-Based Economical Platooning in Blockchain-Enabled Urban Internet of Vehicles. *IEEE Trans. Ind. Inform.* **2019**, *16*, 4122–4133. [[CrossRef](#)]
7. Guo, T.; Yu, K.; Aloqaily, M.; Wan, S. Constructing a prior-dependent graph for data clustering and dimension reduction in the edge of AIoT. *Futur. Gener. Comput. Syst.* **2021**, *128*, 381–394. [[CrossRef](#)]
8. Liu, Y.; Li, D.; Wan, S.; Wang, F.; Dou, W.; Xu, X.; Li, S.; Ma, R.; Qi, L. A long short-term memory-based model for greenhouse climate prediction. *Int. J. Intell. Syst.* **2021**, *37*, 135–151. [[CrossRef](#)]
9. Ju, Y.; Chen, Y.; Cao, Z.; Liu, L.; Pei, Q.; Xiao, M.; Ota, K.; Dong, M.; Leung, V.C.M. Joint Secure Offloading and Resource Allocation for Vehicular Edge Computing Network: A Multi-Agent Deep Reinforcement Learning Approach. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 5555–5569. [[CrossRef](#)]
10. Chen, C.; Yao, G.; Liu, L.; Pei, Q.; Song, H.; Dustdar, S. A Cooperative Vehicle-Infrastructure System for Road Hazards Detection with Edge Intelligence. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 5186–5198. [[CrossRef](#)]
11. Chen, C.; Wang, C.; Liu, B.; He, C.; Cong, L.; Wan, S. Edge Intelligence Empowered Vehicle Detection and Image Segmentation for Autonomous Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2023**, 1–12. [[CrossRef](#)]
12. Zhang, Y.; Chen, C.; Liu, L.; Lan, D.; Jiang, H.; Wan, S. Aerial Edge Computing on Orbit: A Task Offloading and Allocation Scheme. *IEEE Trans. Netw. Sci. Eng.* **2022**, *10*, 275–285. [[CrossRef](#)]
13. Xiao, T.; Chen, C.; Wan, S. Mobile-Edge-Platooning Cloud: A Lightweight Cloud in Vehicular Networks. *IEEE Wirel. Commun.* **2022**, *29*, 87–94. [[CrossRef](#)]
14. Khelifi, L.; Mignotte, M. Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis. *IEEE Access* **2020**, 126385–126400. [[CrossRef](#)]
15. Tamiminia, H.; Salehi, B.; Mahdianpari, M.; Quackenbush, L.; Adeli, S.; Brisco, B. Google Earth engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm.* **2020**, *164*, 152–170. [[CrossRef](#)]
16. Zhang, C.; Feng, Y.; Hu, L.; Tapete, D.; Pan, L.; Liang, Z.; Cigna, F.; Yue, P. A domain adaptation neural network for change detection with heterogeneous optical and SAR remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *109*, 102769. [[CrossRef](#)]
17. Maître, H. *Processing of Synthetic Aperture Radar (SAR) Images*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
18. Nicolas, J.; Hégarat-Masclé, S.L. *Processing of Synthetic Aperture Radar Images*; John Wiley & Sons, Ltd.: New York, NY, USA, 2010.
19. Quan, S.; Xiong, B.; Xiang, D.; Zhao, L.; Zhang, S.; Kuang, G. Eigenvalue-Based Urban Area Extraction Using Polarimetric SAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 458–471. [[CrossRef](#)]
20. Brunner, D.; Lemoine, G.; Bruzzone, L. Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2403–2420. [[CrossRef](#)]
21. Liu, R.; Kuffer, M.; Persello, C. The Temporal Dynamics of Slums Employing a CNN-Based Change Detection Approach. *Remote Sens.* **2019**, *11*, 2844. [[CrossRef](#)]

22. Lv, Z.Y.; Shi, W.; Zhang, X.; Benediktsson, J.A. Landslide Inventory Mapping from Bitemporal High-Resolution Remote Sensing Images Using Change Detection and Multiscale Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1520–1532. [[CrossRef](#)]
23. Xiao, P.; Zhang, X.; Wang, D.; Yuan, M.; Feng, X.; Kelly, M. Change detection of built-up land: A framework of combining pixel-based detection and object-based recognition. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 402–414. [[CrossRef](#)]
24. Fujita, A.; Sakurada, K.; Imaizumi, T.; Ito, R.; Hikosaka, S.; Nakamura, R. Damage detection from aerial images via convolutional neural networks. In Proceedings of the the 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), Nagoya, Japan, 8–12 May 2017; pp. 5–8.
25. Zhang, M.; Shi, W. A Feature Difference Convolutional Neural Network-Based Change Detection Method. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 7232–7246. [[CrossRef](#)]
26. Haigang, S.; Wenqing, F.; Wenzhuo, L.; Kaiming, S.; Chuan, X. Review of Change Detection Methods for Multi-temporal Remote Sensing Imagery. *Geomat. Inf. Sci. Wuhan Univ.* **2018**, *43*, 1885–1898.
27. Zhang, L.; Zhang, L.; Du, B. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *Geosci. Remote Sens.* **2016**, *4*, 22–40. [[CrossRef](#)]
28. Chen, H.; Wu, C.; Du, B.; Zhang, L. Deep Siamese Multi-scale Convolutional Network for Change Detection in Multi-temporal VHR Images. In Proceedings of the 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Shanghai, China, 5–7 August 2019.
29. Wu, J.; Xie, C.; Zhang, Z.; Zhou, Y. A Deeply Supervised Attentive High-Resolution Network for Change Detection in Remote Sensing Images. *Remote Sens.* **2022**, *15*, 45. [[CrossRef](#)]
30. Chen, J.; Yuan, Z.; Peng, J.; Chen, L.; Huang, H.; Zhu, J.; Liu, Y.; Li, H. DASNet: Dual Attentive Fully Convolutional Siamese Networks for Change Detection in High-Resolution Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1194–1206. [[CrossRef](#)]
31. Di Lu, D.; Cheng, S.; Wang, L.; Song, S. Multi-scale feature progressive fusion network for remote sensing image change detection. *Sci. Rep.* **2022**, *12*, 11968. [[CrossRef](#)]
32. Cheng, G.; Han, J.; Lu, X. Remote Sensing Image Classification: Benchmark and State of the Art. *Proc. IEEE* **2017**, *105*, 1865–1883. [[CrossRef](#)]
33. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. UNet++: A nested U-Net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Cham, Switzerland, 2018; pp. 3–11.
34. Chen, H.; Shi, Z. A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. *Remote Sens.* **2020**, *12*, 1662. [[CrossRef](#)]
35. Li, S.; Huo, L. Remote Sensing Image Change Detection Based on Fully Convolutional Network with Pyramid Attention. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Brussels, Belgium, 11–16 July 2021.
36. Song, K.; Jiang, J. AGCDetNet: An Attention-Guided Network for Building Change Detection in High-Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 4816–4831. [[CrossRef](#)]
37. Lv, N.; Zhang, Z.; Li, C.; Deng, J.; Su, T.; Chen, C.; Zhou, Y. A hybrid-attention semantic segmentation network for remote sensing interpretation in land-use surveillance. *Int. J. Mach. Learn. Cybern.* **2023**, *14*, 395–406. [[CrossRef](#)]
38. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
39. Qu, J.; Su, C.; Zhang, Z.; Razi, A. Dilated Convolution and Feature Fusion SSD Network for Small Object Detection in Remote Sensing Images. *IEEE Access* **2020**, *8*, 82832–82843. [[CrossRef](#)]
40. Du, H.; Zhuang, Y.; Dong, S.; Li, C.; Chen, H.; Zhao, B.; Chen, L. Bilateral Semantic Fusion Siamese Network for Change Detection from Multitemporal Optical Remote Sensing Imagery. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 6003405. [[CrossRef](#)]
41. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted, Munich, Germany, 5–9 October 2015*; Springer International Publishing: Berlin, Germany, 2015.
42. Zhi, Z.; Yi, W.; Zhang, Y.; Xiang, S.; Peng, X.; Zhang, B. CLNet: Cross-layer convolutional neural network for change detection in optical remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 247–267.
43. Li, H.; Zhu, F.; Zheng, X.; Liu, M.; Chen, G. MSCDUNet: A Deep Learning Framework for Built-Up Area Change Detection Integrating Multispectral, SAR, and VHR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 5163–5176. [[CrossRef](#)]
44. Chen, T.; Lu, Z.Y.; Yang, Y.; Zhang, Y.; Du, B.; Plaza, A. A Siamese Network based U-Net for Change Detection in High Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 2357–2369. [[CrossRef](#)]
45. Krähenbühl, P.; Koltun, V. Efficient inference in fully connected crfs with gaussian edge potentials. *arXiv* **2011**, arXiv:1210.5644.
46. Gong, M.; Su, L.; Jia, M.; Chen, W. Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images. *IEEE Trans. Fuzzy Syst.* **2013**, *22*, 98–109. [[CrossRef](#)]
47. Zhang, K.; Lv, X.; Chai, H.; Yao, J. Unsupervised SAR Image Change Detection for Few Changed Area Based on Histogram Fitting Error Minimization. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–19. [[CrossRef](#)]
48. Lv, P.; Zhong, Y.; Zhao, J.; Zhang, L. Unsupervised change detection based on hybrid conditional random field model for high spatial resolution remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4002–4015. [[CrossRef](#)]
49. Campbell, N.D.F.; Subr, K.; Kautz, J. Fully-Connected CRFs with Non-Parametric Pairwise Potential. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition IEEE, Portland, OR, USA, 23–28 June 2013.

50. Cao, G.; Zhou, L.; Li, Y. A new change-detection method in high-resolution remote sensing images based on a conditional random field model. *Int. J. Remote Sens.* **2016**, *37*, 1173–1189. [[CrossRef](#)]
51. Shang, Y.; Cao, G.; Zhang, Y. Change Detection Based on Fully-Connected Conditional Random Field with Region Potential in Remote Sensing Images. In Proceedings of the IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 5107–5120.
52. Gong, J.; Hu, X.; Pang, S.; Li, P. Patch Matching and Dense CRF-Based Co-Refinement for Building Change Detection from Bi-Temporal Aerial Images. *Sensors* **2019**, *19*, 1557. [[CrossRef](#)]
53. Zheng, D.; Wei, Z.; Wu, Z.; Liu, J. Learning Pairwise Potential CRFs in Deep Siamese Network for Change Detection. *Remote Sens.* **2022**, *14*, 841. [[CrossRef](#)]
54. Kuruoglu, E.E.; Zerubia, J. Modeling SAR images with a generalization of the Rayleigh distribution. *IEEE Trans. Image Process.* **2004**, *13*, 527–533. [[CrossRef](#)]
55. Inglada, J.; Mercier, G. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1432–1445. [[CrossRef](#)]
56. Zhang, J.; Cui, M.; Wang, B. SAR Image Change Detection Method Based on Neural-CRF Structure. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS 2021, Brussels, Belgium, 11–16 July 2021; pp. 3797–3800. [[CrossRef](#)]
57. Chen, J.; Wang, X.; Guo, Z.; Zhang, X.; Sun, J. Dynamic region-aware convolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 8064–8073.
58. Zhang, J.; Liu, Y.; Wang, B.; Chen, C. A SAR Remote Sensing Image Change Detection Method Based on DR-UNet-CRF Model. In Proceedings of the 2022 IEEE International Conference on Smart Internet of Things (SmartIoT), Suzhou, China, 19–21 August 2022; pp. 180–184. [[CrossRef](#)]
59. Zhong, P.; Wang, R.S. Modeling and Classifying Hyperspectral Imagery by CRFs with Sparse Higher Order Potentials. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 688–705. [[CrossRef](#)]
60. Zheng, S.; Jayasumana, S.; Romera-Paredes, B.; Vineet, V.; Su, Z.; Du, D.; Huang, C.; Torr, P.H. Conditional random fields as recurrent neural networks. In Proceedings of the IEEE international conference on computer vision, Santiago, Chile, 7–13 December 2015; pp. 1529–1537.
61. Zhong, Y.; Zhao, J.; Zhang, L. A Hybrid Object-Oriented Conditional Random Field Classification Framework for High Spatial Resolution Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7023–7037. [[CrossRef](#)]
62. Zhou, L.; Cao, G.; Li, Y.; Shang, Y. Change Detection Based on Conditional Random Field with Region Connection Constraints in High-Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3478–3488. [[CrossRef](#)]
63. Kriminger, E.; Cobb, J.T.; Principe, J.C. Online active learning for automatic target recognition. *IEEE J. Ocean. Eng.* **2015**, *40*, 583–591. [[CrossRef](#)]
64. Zhao, E.; Sun, L.; Wang, C.; Xia, X. Image sharpening method based on anti-heat conduction equation and Sobel operator. In Proceedings of the Eighth International Symposium on Multispectral Image Processing and Pattern Recognition, Wuhan, China, 26–27 October 2013.
65. Wang, F.; Chen, W.; Qiu, L. Hausdorff derivative laplacian operator for image sharpening. *Fractals Interdiscip. J. Complex Geom. Nat.* **2019**, *27*. [[CrossRef](#)]
66. He, C.H.; Zhang, X.; Hu, Y. A study on the improved algorithm for Sobel on image edge detection. *Opt. Tech.* **2012**, *38*, 323–327.
67. Cao, X.; Ji, Y.; Wang, L.; Ji, B.; Jiao, L.; Han, J. SAR image change detection based on deep denoising and CNN. *IET Image Process.* **2019**, *13*, 1509–1515. [[CrossRef](#)]
68. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv* **2021**, arXiv:2102.04306.
69. Liu, R.; Wang, R.; Huang, J.; Li, J.; Jiao, L. Change detection in SAR images using multiobjective optimization and ensemble strategy. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1585–1589. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.