

Article

# Multispectral Remote Sensing Image Change Detection Based on Twin Neural Networks

Wenhao Mo <sup>1,\*</sup>, Yuanpeng Tan <sup>1</sup>, Yu Zhou <sup>2</sup>, Yanli Zhi <sup>2</sup>, Yuchang Cai <sup>1</sup>  and Wanjie Ma <sup>2,\*</sup>

<sup>1</sup> China Electric Power Research Institute, Beijing 100192, China; tanyuanpeng@epri.sgcc.com.cn (Y.T.); caichangyu@epri.sgcc.com.cn (Y.C.)

<sup>2</sup> State Grid Jiangxi Electric Power Company, Nanchang 330096, China; zhouyu@jx.sgcc.com.cn (Y.Z.); zhiyl@jx.sgcc.com.cn (Y.Z.)

\* Correspondence: mowenhao@epri.sgcc.com.cn (W.M.); mawj@edu.cn (W.M.)

**Abstract:** Remote sensing image change detection can effectively show the change information of land surface features such as roads and buildings at different times, which plays an indispensable role in application fields such as updating building information and analyzing urban evolution. At present, multispectral remote sensing images contain more and more information, which brings new development opportunities to remote sensing image change detection. However, this information is difficult to use effectively in change detection. Therefore, a change-detection method of multispectral remote sensing images based on a Siamese neural network is proposed. The features of dual-temporal remote sensing images were extracted based on the ResNet-18 network. In order to capture the semantic information of different scales and improve the information perception and expression ability of the algorithm for the input image features, an attention module network structure is designed to further enhance the extracted feature maps. Facing the problem of false alarms in change detection, an adaptive threshold comparison loss function is designed to make the threshold more sensitive to the remote sensing images in the data set and improve the robustness of the algorithm model. Moreover, the threshold segmentation method of the measurement module is used to determine the change area to obtain a better change-detection map domain. Finally, our experimental tests show that the proposed method achieves excellent performance on the multispectral OSCD detection data sets.



**Citation:** Mo, W.; Tan, Y.; Zhou, Y.; Zhi, Y.; Cai, Y.; Ma, W. Multispectral Remote Sensing Image Change Detection Based on Twin Neural Networks. *Electronics* **2023**, *12*, 3766. <https://doi.org/10.3390/electronics12183766>

Academic Editor: Chiman Kwan

Received: 17 July 2023

Revised: 25 August 2023

Accepted: 30 August 2023

Published: 6 September 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/>).

**Keywords:** remote sensing image; change detection; void space pyramid pool; self-attention; adaptive threshold

## 1. Introduction

Multi-spectral remote sensing images contain information of multiple spectral bands, extending from visible light to thermal infrared bands, containing rich image information and spectral information. At the same time, the contour information of land objects in the image is relatively clear, and there is a certain correlation between each band, which can provide a stronger land surface information resolution effect and provide rich data support for the change detection of land objects.

Remote sensing image change-detection technology is designed to determine whether the surface objects have changed in this period through quantitative analysis from two remote sensing image pairs at different times and in the same region. It shows the spectral characteristics of the unchanged region, the changed region and the surface objects in the remote sensing image pairs of two different periods. Remote sensing image change detection can effectively update the change information of surface vegetation, buildings and other surface features, and plays an indispensable role in assessing the level of natural disasters, predicting the development trend of natural disasters, and monitoring land cover information [1–3].

Image change detection is an important research direction in computer vision and remote sensing. With the continuous progress of remote sensing technology and image-processing algorithms, it becomes more and more important to detect changes in images. Firstly, ref. [4] systematically investigated and summarized image change detection algorithms. It provides an overview of different methods, techniques, and evaluation methods, which provides an important reference for the research in the field of image change detection. Secondly, ref. [5] covers related topics such as acquisition, preprocessing, feature extraction, classification and change detection of multispectral satellite images, and provides specialized techniques and methods for multispectral image change detection. Ref. [6] is another book published by Springer in 2012 that focuses on 2D image change detection methods. The book presents various techniques and algorithms for analyzing images, including pixel-based and object-based change detection methods. Finally, ref. [7] discusses the approaches and challenges of using multispectral and hyperspectral images for practical change detection applications. It explores the potential of these types of images to detect changes in different environments and provides insights into limitations and future research directions in the field. In summary, these papers provide valuable information and in-depth research in the field of image change detection, covering all aspects of the field and various techniques and methods used. These research results are of great significance for further promoting the development of image change detection algorithms.

In the field of traditional algorithms, Fung et al. [8] based on the image difference algorithm, the principal component analysis method, the post-classification comparison algorithm and other major change detection algorithms, conducted in-depth research on threshold division. Mas [9] in Landsat MultiSpectral Scanner (MSS) tested image difference algorithm and selective principal component analysis (pca) [10,11], plant index difference algorithm, multiple data unsupervised classification algorithm and the changes after after classification comparison difference algorithm and image enhancement algorithm is a variety of combination, through calculating the Kappa coefficient evaluation results of the calculation of each algorithm accuracy, Pointed out that the merits of the various algorithms in different environment. Chen [12] proposed an object-based change detection algorithm. Object-based change detection can improve the influence of background information on the change of objects and further improve the effect of change detection by utilizing the characteristics of remote sensing images with high spatial resolution.

With the rise of technologies like artificial intelligence and deep learning, in order to meet the challenges of high-dimensional data sets (higher spatial resolution and more spectral features), complex remote sensing image data structures (nonlinear and overlapping data) and high computational complexity brought by nonlinear optimization, supervised learning requires a large number of training samples, and the robustness of artificial intelligence-based remote sensing image detection models is difficult to guarantee. Zhang et al. [13] proposed a change detection method based on a deep twinned semantic network framework. This supervised network uses triplet loss function [14] for training, which can not only directly extract surface features from remote sensing images with multi-scale information, but also extract surface features from remote sensing images. Moreover, the features of inter-class separability and intra-class separability can be obtained from the learning of semantic relations to make the network more robust. A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection (STAnet) proposed by Chen et al. [15] designed two kinds of self-attention modules, the basic space-time attention module (BAM) and the pyramidal space-time attention module (PAM). BAM learns to capture the spatio-temporal dependence between any two locations (noting the weights) and calculates the response of each location by weighting the sum of features at all locations in space-time. PAM embedded BAM into a pyramid structure to generate multi-scale representations of attention. However, PAM pyramid attention can combine the shallow information of different areas of the image to generate multi-scale attention feature expression, so that each pixel participates in the self-attention mechanism of sub-regions of different scales. This attention mechanism can effectively extract more

fine-grained spatial information, but it is found in the experiment that this attention mechanism can improve the mutual correlation between pixels. As a result, the algorithm model pays more attention to the relationship between multiple pixels in a certain region in the generation of the difference feature map, rather than the single pixel judgment, which leads to the false alarm problem. This feature map, which combines global information and local cross-correlation information, improves the detection rate of changing pixels well, but the high cross-correlation between pixels also leads to the expansion of the contour and the reduction of the detail of the changing region. Therefore, the accuracy will be reduced.

To solve these problems, this paper proposes a change detection algorithm of multi-spectral remote sensing images based on twin neural networks. The feature extraction network based on ResNet-18, which like feature extraction network designed by STA-Net is used to extract the global information of different sizes of two-phase images. On this basis, an ASPP+BAM pyramid attention module is designed in order to retain the fine feature information of large- and small-scale objects, reduce the attention of the algorithm model to the inter-correlation between neighboring pixels, and make the model better understand the features of the image with a relatively small number of pixels in the changing region. At the same time, we propose the concept of adaptive threshold contrast loss function and obtain an adaptive threshold through an adaptive threshold generation network. Because an adaptive threshold is obtained through the convolution of feature graphs, the problem that the given threshold in the basic algorithm is not refined enough in the region of change can be avoided, and the problem of false alarms can be solved.

In summary, the main contributions of this paper include:

1. We designed a multi-level pyramid attention mechanism (ASPP+BAM) to reduce the attention of the algorithm model to the inter-correlation between neighboring pixels, so that the model can better understand the characteristics of the image with a relatively small proportion of pixels in the changing region, so as to better solve the false alarm problem.
2. In order to further alleviate the false alarm problem, we designed an adaptive threshold contrast loss function, which can better consider the texture characteristics of the changing region and the unchanged region when the adaptive threshold is used, so that the generated threshold can better judge the boundary of the changing region, sharpen the edge of the changing region, and reduce the expansion region.

## 2. Related Work

After many years of study and accumulation in the field of remote sensing image change detection, there are enough original remote sensing image data with high spatial resolution and multi-band. The statistical analysis method based on traditional mathematical theory has poor performance in processing a large number of remote sensing image data with high complexity and cannot provide rapid and accurate change detection results in many application scenarios.

Deep neural networks can effectively solve the problem of multi-spectral image resolution matching. Xu et al. [16] proposed to use an autoencoder to learn the corresponding feature relationship between multi-temporal high-resolution images, determine the threshold between changing pixels and non-changing pixels according to Otsu's thresholding method (OTSU) [17], discard the isolated points, and mark the part higher than the threshold as the changing region. Zhang et al. [18] designed a sparse autoencoder algorithm framework based on spatial resolution changes according to the characteristics of remote sensing images. The autoencoder is stacked to learn the feature representation from the local neighborhood of a given pixel in an unsupervised manner. Gong et al. [19] proposed a deep confidence network to generate change detection maps directly from two different simultaneous Synthetic Aperture Radar (SAR) images. Two years later, Gong [20] proposed a new framework for high-resolution remote sensing image change detection, which combined change feature extraction based on superpixels and hierarchical difference representation by neural networks. Liu et al. [21] proposed a dual-channel convolutional neural

network model for SAR image change detection, which simultaneously extracted image deep features from two different simultaneous SAR images. Zhan et al. [22] proposed a change detection method based on deep twin convolutional networks, in which the twin networks share node weights, so as to directly extract feature information from image pairs. Liu et al. [23] proposed an unsupervised deep convolutional coupling network for two different types of remote sensing images obtained by optical sensors and SAR images, focusing on the complementary feature points of optical images and SAR images.

However, using only the neural network based on deep learning to extract the change information, it cannot deal with the deeper multi-scale and multi-level change characteristics. In order to adapt objects of various sizes to generate better representations, STAnet uses neural networks to extract global information while adding a self-attention module to capture spatiotemporal dependencies at different scales. Ding et al. [24] proposed a new two-branch end-to-end network. It innovatively introduces cross-layer plus a skip connection module guided by a spatial attention mechanism to aggregate multi-layer context information, which improves network performance. In order to dig deeper into multi-scale and multi-level features and improve detection accuracy, Li et al. [25] designed a pyramid attention layer using the spatial attention benefit mechanism. Wang et al. [26] proposed a high-resolution feature difference attention network for change detection. In this network, a multi-resolution parallel structure is introduced to make comprehensive use of image information of different resolutions to reduce the loss of spatial information, and a differential attention module is proposed to improve the sensitivity of differential information to maintain the change information of a building.

### 3. Approach

In this paper, a change detection method of multispectral remote sensing images based on twin neural networks is proposed. The pre-trained Deep residual network-18 (ResNet-18) network, which removes the global pooling layer and the fully connected layer, is used for feature extraction of two-phase remote sensing images. In order to capture semantic information of different scales and improve the ability of the algorithm to perceive and express the information of the input image objects, we designed the void space pyramid pool and self-attention module to further enhance the feature map extracted from the feature. In order to make the threshold more sensitive to remote sensing images in the data set, the robustness of the algorithm model is improved. We use an adaptive threshold to generate the network and apply the global judgment threshold generated in real time from the feature map to the contrast loss function. Finally, the threshold segmentation method of the measurement module is used to determine the change region.

#### 3.1. Feature Extraction Module

As shown in Figure 1, the feature extraction module will obtain the corresponding feature map from the output feature map of each Block through a  $1 \times 1$  convolution. Then, the feature map of these four blocks will be up-sampled to change the size of the feature map of Block 1, and finally, the four feature maps will be spliced together. The feature map of remote sensing images is obtained by a  $3 \times 3$  convolution and  $1 \times 1$  convolution.

#### 3.2. Attention Module Network Structure

In order to reduce the attention of the algorithm model to the inter-correlation between neighboring pixels, the model can better understand the features of the image with a relatively small proportion of pixels in the changing region. We are fully utilized the input characteristics under the premise of all pixel information, improve receptive field by ASPP cross-regional pixel related information, including global attention since the branch and the smallest scale of regional branch, the attention to maximize degree for two kinds of large scale and small scale fine characteristic information of the object, convolution ASPP branch of a combination of three empty, Maximum use of the input feature information, the characteristics of large scale and small scale object refinement of retain information,

and less adjacent area of the cross correlation of pixels. The ASPP+BAM pyramid attention module uses two self-attention modes of different scales to reduce the feature expression of medium-scale objects and reduce the cross-correlation of adjacent pixels. Combined with ASPP's cross-region pixel correlation information processing, these five forms of branch output feature maps are stacked, channel-fused, and residually connected with input features. The multi-scale and fine-grained ground object information recognition ability of remote sensing images is obtained, which weakens the cross-correlation between adjacent pixels, effectively alleviates the problem that the recall rate is much higher than the accuracy rate and false alarm, and then improves the comprehensive performance of change detection. The ASPP+BAM pyramid attention module network diagram is shown in Figure 2.

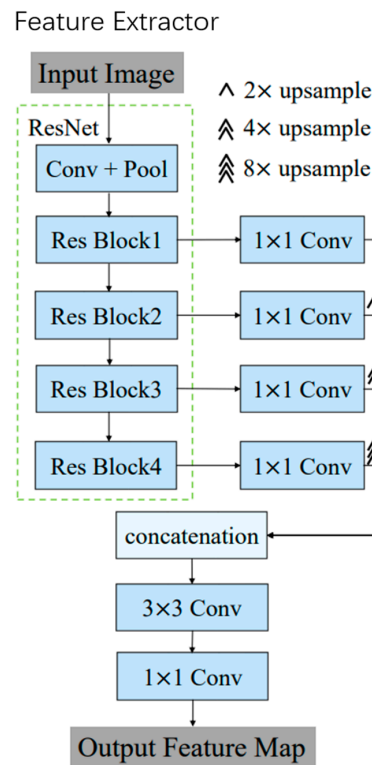


Figure 1. Network structure of feature extraction.

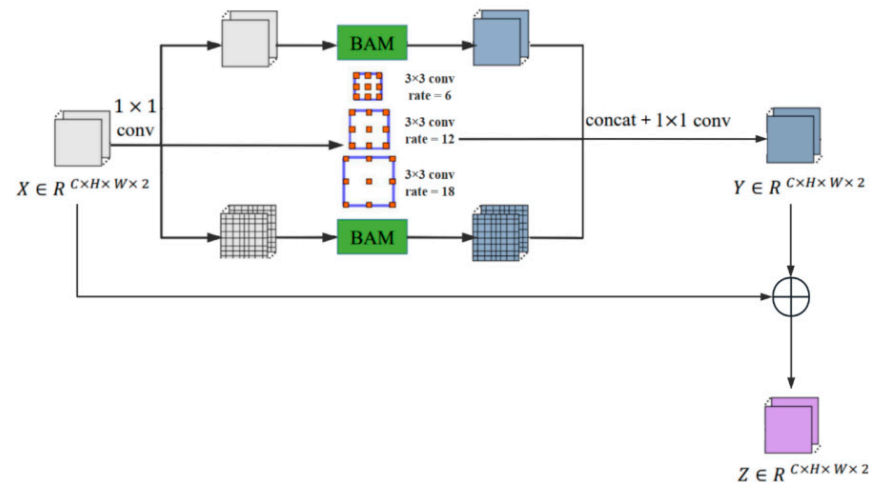


Figure 2. Structure of ASPP+BAM pyramid attention module network.

In order to solve the problem of serious false alarms caused by the comparison between the prediction change detection graph and the ground truth, we proposed the adaptive threshold contrast loss function. The adaptive threshold was obtained through the adaptive threshold generation network, as shown in Figure 3. Therefore, the problem that the given threshold value in the basic algorithm is not finely divided in the changing region can be avoided. The adaptive threshold value is applied to the loss function, and the optimized loss function is obtained, as shown in Formula (1). The adaptive threshold will be updated with loss in the network, which is equivalent to putting higher requirements on the distance graph, making the difference between the changing pixel pairs larger, and thus solving the false alarm problem.

$$L(D^*, M^*, (Adaptive)margin) = \frac{1}{2} \frac{1}{n_u} \sum_{b,i,j} (1 - M_{b,i,j}^*) D_{b,i,j}^* + \frac{1}{2} \frac{1}{n_c} \sum_{b,i,j} M_{b,i,j}^* \max(0, (Adaptive)margin - D_{b,i,j}^*) \tag{1}$$

$$n_u = \sum_{b,i,j} (1 - M_{b,i,j}^*) D_{b,i,j}^* \tag{2}$$

$$n_c = \frac{1}{n_c} \sum_{b,i,j} M_{b,i,j}^* \tag{3}$$

where b, i, j represents the first batch of b, the row first j column of pixels.  $n_u$  and  $n_c$  represent not change and change to the number of pixels, the Numbers can be calculated by the following formula.

The adaptive threshold generation network is a convolutional neural network that takes two feature maps output by the ASPP+BAM pyramid attention module as input and output values. At the beginning, the feature maps obtained after feature extraction and pyramid attention module processing are stacked on two two-phase remote sensing images. Then, the channel fusion is carried out through a convolutional layer, and then three convolutional pooling activation modules are passed through; each module contains a convolutional layer with a convolution kernel size and step size of 2, and a maximum pooling layer has an activation function. The reason for using the maximum pooled layer is that this paper hopes to better consider the texture features of the changing region and the unchanged region when obtaining the adaptive threshold through the feature graph, so that the generated threshold can better judge the boundary of the changing region, thus alleviating the false alarm problem. The reason for using the Relu activation function is that the distance in the distance graph is non-negative. Therefore, the non-negative distance is considered in the threshold value.

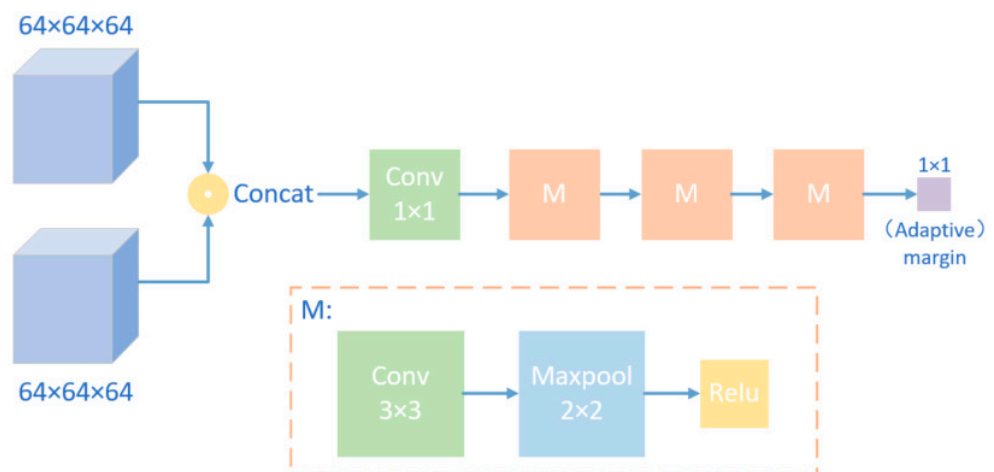
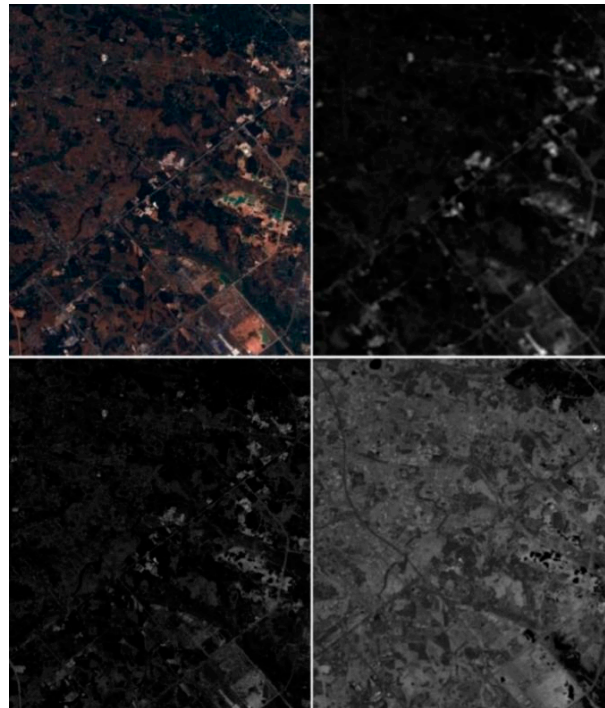


Figure 3. Adaptive threshold generation network.



#### 4. Data Sets

In the experiment of Change Detection, we used the publicly available change detection data set—the Onera Satellite Change Detection data set (OSCD) [27]. The OSCD data set was built using images from the Sentinel-2 satellites. The satellite captures images of various resolutions between 10 m and 60 m in 13 bands between ultraviolet and short-wavelength infrared. Twenty-four regions of approximately  $600 \times 600$  pixels at 10 m resolution with various levels of urbanization where urban changes were visible were chosen worldwide. The images of all bands were cropped according to the chosen geographical coordinates, resulting in 26 images for each region, i.e., 13 bands for each of the images in the image pair. Figure 4 shows the visible light image of a certain area and its corresponding B1, B4, and B11, three single-band remote sensing images of OSCD.



**Figure 4.** Remote sensing image of a certain area (**upper left** is visible light, **upper right** is B1 band, **lower left** is B4 band, **lower right** is B11 band).

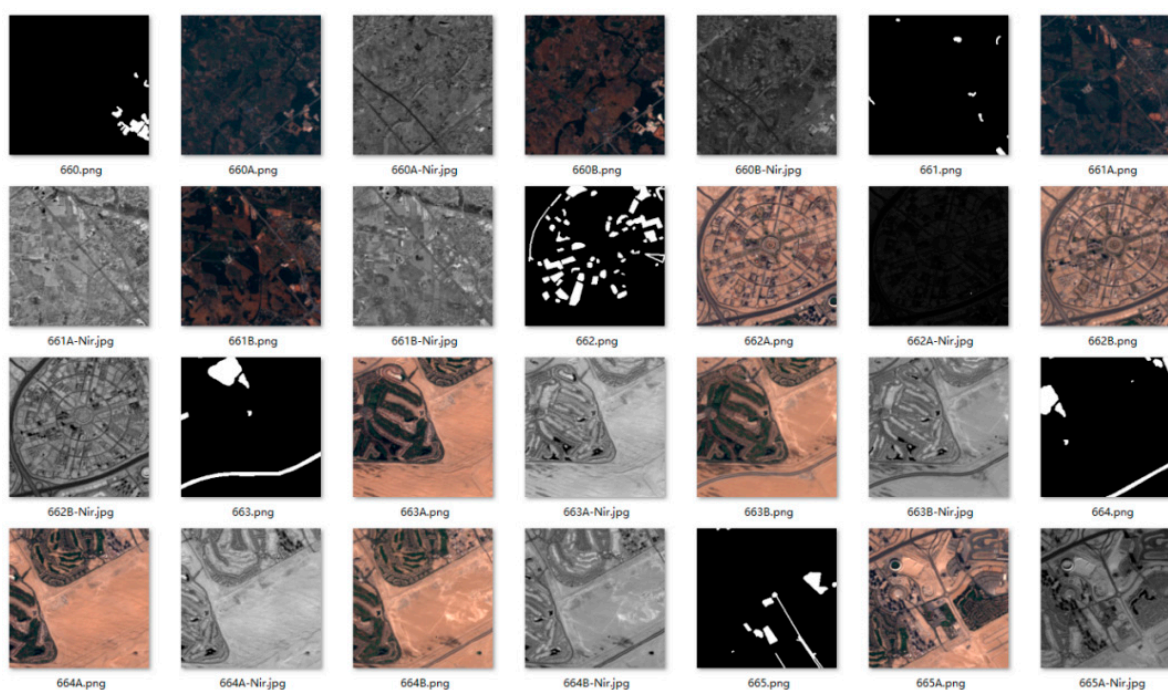
The remote sensing images of this data set are mainly taken in urban areas, and the main marked objects are urban roads and buildings, reflecting the evolution of the city, while ignoring natural changes (such as plant growth or tidal changes). Since the OSCD data set is pixel-level ground truth change-labeling for the change region, the data set supports various complex supervised learning algorithms to learn, understand, and solve many problems in the field of change detection. In the experiment, four bands of remote sensing images were used as the experimental objects, of which B02 represents the central band of 490 mm, the surface resolution of 10 m blue; B03 represents the central band of 560 mm, the surface resolution of 10 m green; B04 represents the central band of 665 mm, the surface resolution of 10 m red; and B08 represents the central band of 865 mm.

In the process of model training, the data set is divided into the training set and the test set according to the ratio of 7:3. Since the amount of data in the original data set is not enough to support the training requirements, it is necessary to strengthen the data set by pre-processing image enhancement.

Firstly, for the multi-spectral remote sensing images taken at each geographical location,  $256 \times 256$  patches are cropped out in the images, and then image enhancement is carried out. The image enhancement scheme includes: rotation, flip, noise addition, imager addition three kinds. In this paper, three rotation angles of  $90^\circ$ ,  $180^\circ$  and  $270^\circ$  are used

to rotate the image midpoint clockwise, so as to avoid the situation that other rotation angles lead to the change of image size and the need for numerical filling. After image enhancement of the training set, a total of 1680  $256 \times 256$  patches are obtained for network training, and the test set has 1200  $256 \times 256$  patches for testing the model training effect.

In the process of data pre-processing, it is necessary to downsample the patch pair of category (a) two-phase remote sensing image of the cropped negative sample (without changing the region) image and to remove some public samples from the sample image without changing the region. At the same time, it must be ensured that the ratio of (a) two-phase remote sensing image patch set that does not contain changing regions and (b) two-phase remote sensing image patch set that contains a small amount of changing regions is not greater than 3 after the sum of the number of patches and the (c) two-phase remote sensing image patch set that contains a large number of changing regions. In this way, the ratio of the overall pixel amount in the change region to the total pixel value of the remote sensing image can be ensured to be greater than 1.5%, which is close to the normal level of the change region map proportion in the OSCD training set. Meanwhile, the distribution of positive and negative samples can be more uniform during training, and the overfitting probability of the model can be reduced. Figure 5 shows the slice patch data set of remote sensing image change detection generated after the above consideration and operation.



**Figure 5.** Slice patch set of remote sensing image change detection.

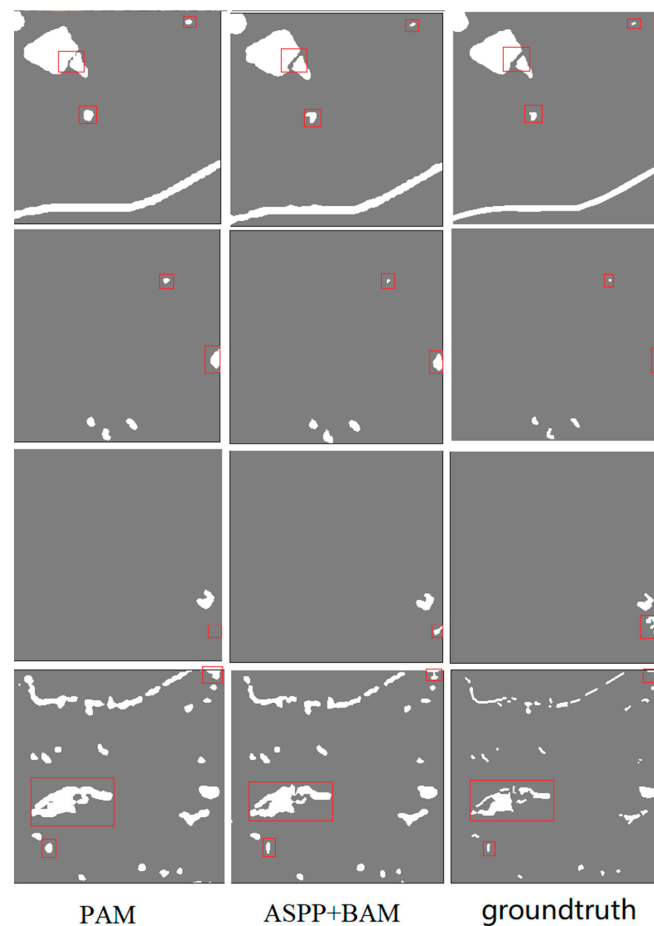
## 5. Experiments

In all experiments, supervised learning was used to train the remote sensing image change detection network model. During model training, all tests in this study were carried out on Intel(R) Core(TM) i7-7700K cpus @ 3.60 GHz and NVIDIA GeForce RTX 3090 (24 GB). The ResNet-18 network whose model is pre-trained by ImageNet [28] is adopted, and on this basis, subtle adjustments are made to remove the global pooling layer and the fully connected layer. The images are passed into the neural network with a fixed resolution of 56, and the adaptive moment estimation optimizer (Adam) is used to update and iterate the network node weights. The training batchsize is 8, iterating over 200 epochs with an initial learning rate of  $1 \times 10^{-3}$ , maintaining the same learning rate for the first 100 epochs and linearly decaying to 0 for the remaining 100 epochs.



### 5.1. Comparative Experiment of Feature Extraction

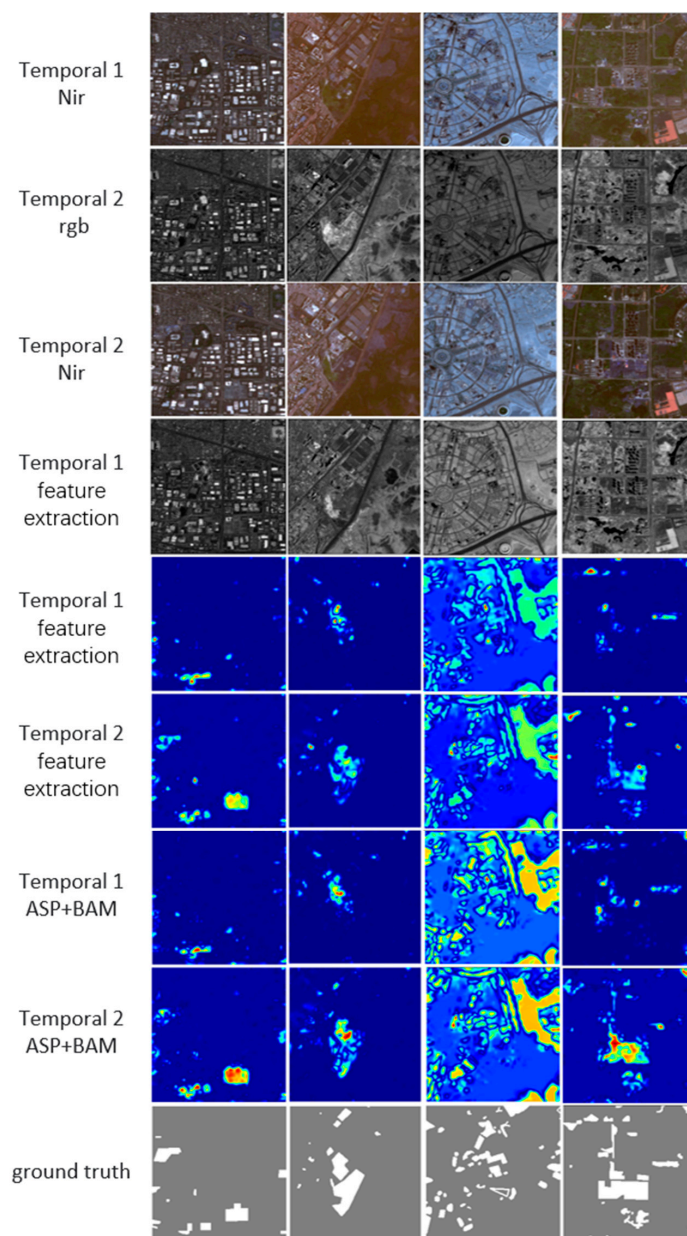
In order to more clearly show the effect of the ASPP+BAM pyramid attention mechanism change detection, this section also compares the change detection region prediction graph of STA-Net (with PAM) and STA-NET (with ASPP+BAM) pyramid attention modules with the ground truth. As shown in Figure 6, from the three columns of PAM, ASPP+BAM and ground truth, it can be found that the detail of the attention module of ASPP+BAM is better than that of the pyramid self-attention module of PAM, because the cross-correlation between adjacent pixels is reduced. It reduces the attention of the algorithm when judging the changes of the adjacent pixels of the changing pixels and focuses the attention of the algorithm model more on the difference between the feature map pixels of the same position of two remote sensing images with different phases at the same time. It can be proved from the above experiments that the combination of the ASPP and BAM self-attention mechanism can effectively combine the multi-scale features of the change region in the remote sensing image, so that the final output feature map has a better expression effect and is conducive to more refined annotation of the change region.



**Figure 6.** Comparison of the STANet change area prediction graph with ground truth using PAM and ASPP+BAM.

In this paper, the two-phase remote sensing image is input into the twin neural network for visual analysis of the feature map output by the module of feature extraction and pyramid attention. The visual map is shown in Figure 7; Nir is near-infrared light, and RGB is visible light. The fifth and sixth lines are the feature map output by the remote sensing image of time phase 1 and time phase 2 after passing the feature extraction module. After the weight of each channel is superimposed, the feature weight visualization diagram with the resolution of the original Figures 5 and 6 is up-sampled. The 7th and 8th rows

are the feature maps of the remote sensing images of time phase 1 and time phase 2 with the output dimension after the feature map is extracted and processed by the ASPP+BAM pyramid attention module. The feature weights of each channel are superimposed and then up-sampled to 56 resolution. By comparing the visual feature maps of rows 5 and 7 or columns 6 and 8, it can be found that in the urban area of remote sensing images, the feature maps processed by the ASPP+BAM pyramid module will have more detailed features than the feature maps without attention mechanism processing. The pyramid attention module can help the algorithm to extract more refined image features.



**Figure 7.** Visual results of two-phase remote sensing image feature map.

For the remote sensing images of rgb+Nir-4ch, after the setting of the attention module and the loss function threshold was improved, a horizontal comparison was made again with Siam U-Net and CBAM and MAS-Net algorithms; the results are shown in Table 1 below. Since the improvements of the ASPP+BAM and the adaptive threshold are aimed at false alarms, a certain recall rate is sacrificed in exchange for the improvement of accuracy and the improvement of the F1 value of the algorithm's comprehensive effect. It can be

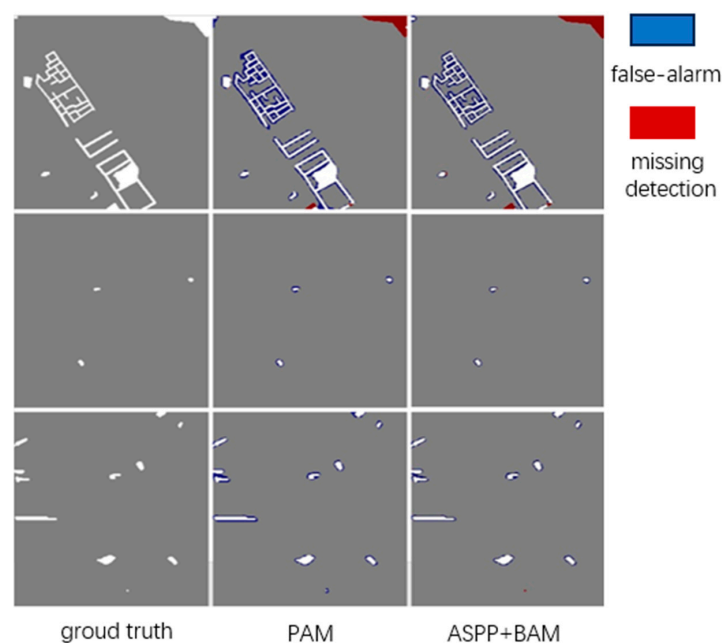
found in the table that, compared with the original algorithm, the improved algorithm has a 5.4% increase in accuracy when the recall rate is only 0.9% lower. The F1 value is improved by 2.5%, and the accuracy and F1 value are much higher than other algorithms compared with the same period.

**Table 1.** Comparison between improved STA-Net and other remote sensing image change detection algorithms (rgb+Nir-4ch).

Method	Precision	Recall	F1_Score
Siam U-Net and CBAM [27]	0.597	0.765	0.671
MAS-Net [29]	0.656	0.692	0.674
STA-Net and PAM [16]	0.667	0.753	0.707
Ours	0.721	0.744	0.732

### 5.2. Experimental Results of False Alarm Problem

In this section, the adaptive moment estimation optimizer is still used to update and iterate the weights of network nodes. The batchsize is trained to be 8, and 200 epochs are iterated. The initial learning rate is  $1 \times 10^{-3}$ . The same learning rate is maintained in the first 100 epochs, and the linear attenuation is reduced to 0 in the remaining 100 epochs. The experimental results are shown in Figure 8. As can be seen from the figure, the contrast loss function using the ASPP+BAM pyramid attention module and adaptive threshold solves a lot of pixel adhesion problems in the prediction graph of changing regions and has a better improvement effect on the false alarm problem. However, due to the addition of the adaptive threshold, the network with the ASPP+BAM pyramid self-attention mechanism and the contrast loss function solves the problem of false alarms. There will be a small increase in missed detection, which is reflected in the evaluation index of the algorithm. The detection recall rate of the changed pixels will decrease to a certain extent, the accuracy rate will be greatly improved, and the overall effect of the algorithm will be improved; that is, the F1 value will increase, which also echoes the previous experimental results. Therefore, this experiment can explain the previous questions about the false alarm problem. It can be concluded that the ASPP+BAM pyramid attention module and adaptive threshold contrast loss function both improve the change detection algorithm model.



**Figure 8.** Experimental comparison diagram of false alarm problem.

## 6. Conclusions

In order to obtain change information from asynchronous remote sensing image pairs more quickly, this paper designs a multi-spectral remote sensing image change detection system based on a twin network, conducts in-depth research on the false alarm problem of STA-Net, makes targeted improvements in two aspects of the attention mechanism and loss function, and integrates an ASPP and BAM self-attention mechanism. An ASPP+BAM pyramid attention module is proposed. Aiming at the threshold value in the contrast loss function, a convolutional network was designed to generate an adaptive threshold to alleviate the false alarm problem. The ablation experiment, feature visualization output and false alarm situation were compared for the optimization of the two aspects, which proved that the comprehensive performance of the improved algorithm framework was improved. The current remote sensing image change detection algorithm framework in this paper can identify the changed area and the unchanged area, and the subsequent specific change-type judgment of the changed area needs to be determined manually. In the future, we will consider studying the refined change-type judgment and determine the specific change type of the remote sensing image by adding post-processing.

**Author Contributions:** Methodology, W.M. (Wenhao Mo); Software, Y.Z. (Yu Zhou) and Y.Z. (Yanli Zhi); Formal analysis, Y.T.; Investigation, Y.Z. (Yanli Zhi); Resources, Y.Z. (Yu Zhou); Writing—Original draft, W.M. (Wanjie Ma); Writing—Review & editing, Y.C.; Visualization, W.M. (Wanjie Ma); Supervision, Y.T.; Project administration, W.M. (Wenhao Mo). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study are available in this article.

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

## References

1. Xiao, J.; Guo, H.; Zhou, J.; Zhao, T.; Yu, Q.; Chen, Y.; Wang, Z. Tiny object detection with context enhancement and feature purification. *Expert Syst. Appl.* **2023**, *211*, 118665. [\[CrossRef\]](#)
2. Xiao, J.; Wu, Y.; Chen, Y.; Wang, S.; Wang, Z.; Ma, J. LSTFE-Net: Long Short-Term Feature Enhancement Network for Video Small Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023 (CVPR2023), Vancouver, CB, USA, 18–22 June 2023; pp. 14613–14622.
3. Xie, H.G.; Yang, M.; Yan, B.L.; Hou, K.Y.; Jiang, D. SRPAR: Anchor-free detector with aspect ratio priority for slender objects. *J. Electron. Image* **2022**, *31*, 043001. [\[CrossRef\]](#)
4. Radke, R.J.; Andra, S.; Al-Kofahi, O.; Roysam, B. Image change detection algorithms: A systematic survey. *IEEE Trans. Image Proc.* **2005**, *14*, 294–307. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Ünsalan, C. *Multispectral Satellite Image Understanding*; Springer: Berlin/Heidelberg, Germany, 2011.
6. İlsever, M.; Ünsalan, C. *Two-Dimensional Change Detection Methods*; Springer: Berlin/Heidelberg, Germany, 2012.
7. Kwan, C. Methods and Challenges Using Multispectral and Hyperspectral Images for Practical Change Detection Applications. *Information* **2019**, *10*, 353. [\[CrossRef\]](#)
8. Fung, T.; LeDrew, E. For change detection using various accuracy. *Photogramm. Eng. Remote Sens.* **1988**, *54*, 1449–1454.
9. Mas, J.F. Monitoring land-cover changes: A comparison of change detection techniques. *Int. J. Remote Sens.* **1999**, *20*, 139–152. [\[CrossRef\]](#)
10. Kwarteng, P.; Chavez, A. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogramm. Eng. Remote Sens.* **1989**, *55*, 339–348.
11. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510. [\[CrossRef\]](#)
12. Chen, G.; Hay, G.J.; Carvalho, L.M.; Wulder, M.A. Object-based change detection. *Int. J. Remote Sens.* **2012**, *33*, 4434–4457. [\[CrossRef\]](#)
13. Zhang, M.; Xu, G.; Chen, K.; Yan, M.; Sun, X. Triplet-based semantic relation learning for aerial remote sensing image change detection. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 266–270. [\[CrossRef\]](#)
14. Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015, Boston, MA, USA, 7–12 June 2015; pp. 815–823.
15. 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. [\[CrossRef\]](#)

16. Xu, Y.; Xiang, S.; Huo, C.; Pan, C. Change detection based on auto-encoder model for VHR images. In Proceedings of the MIPPR 2013: Pattern Recognition and Computer Vision 2013, Portland, OR, USA, 23–28 June 2013; p. 891902.
17. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
18. Zhang, P.; Gong, M.; Su, L.; Liu, J.; Li, Z. Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 24–41. [[CrossRef](#)]
19. Gong, M.; Zhao, J.; Liu, J.; Miao, Q.; Jiao, L. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 125–138. [[CrossRef](#)] [[PubMed](#)]
20. Gong, M.; Zhan, T.; Zhang, P.; Miao, Q. Superpixel-based difference representation learning for change detection in multi-spectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2658–2673. [[CrossRef](#)]
21. Liu, T.; Li, Y.; Cao, Y.; Shen, Q. Change detection in multitemporal synthetic aperture radar images using dual-channel convolutional neural network. *J. Appl. Remote Sens.* **2017**, *11*, 042615. [[CrossRef](#)]
22. Zhan, Y.; Fu, K.; Yan, M.; Sun, X.; Wang, H.; Qiu, X. Change detection based on deep siamese convolutional network for optical aerial images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1845–1849. [[CrossRef](#)]
23. Liu, J.; Gong, M.; Qin, K.; Zhang, P. A deep convolutional coupling network for change detection based on heterogeneous optical and radar images. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *29*, 545–559. [[CrossRef](#)]
24. Ding, Q.; Shao, Z.; Huang, X.; Altan, O. DSA-Net: A novel deeply supervised attention-guided network for building change detection in high-resolution remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102591. [[CrossRef](#)]
25. 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; pp. 4352–4355. [[CrossRef](#)]
26. Wang, X.; Du, J.; Tan, K.; Ding, J.; Liu, Z.; Pan, C.; Han, B. A high-resolution feature difference attention network for the application of building change detection. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102950. [[CrossRef](#)]
27. Daudt, R.C.; Le Saux, B.; Boulch, A.; Gousseau, Y. Urban change detection for multispectral earth observation using convolutional neural networks. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 2115–2118.
28. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
29. Li, J.; Zhu, S.; Gao, Y.; Zhang, G.; Xu, Y. Change Detection for High-Resolution Remote Sensing Images Based on a Multi-Scale Attention Siamese Network. *Remote Sens.* **2022**, *14*, 3464. [[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.