Article

SUNet: Change Detection for Heterogeneous Remote Sensing Images from Satellite and UAV Using a Dual-Channel Fully Convolution Network

Ruizhe Shao, Chun Du, Hao Chen * and Jun Li

Department of Cognitive Communication, College of Electronic Science and Technology, National University of Defense Technology, Changsha 410000, China; shaoruizhe@nudt.edu.cn (R.S.); duchun@nudt.edu.cn (C.D.); junli@nudt.edu.cn (J.L.)

* Correspondence: hchen@nudt.edu.cn

Abstract: Change Detection in heterogeneous remote sensing images plays an increasingly essential role in many real-world applications, e.g., urban growth tracking, land use monitoring, disaster evaluation and damage assessment. The objective of change detection is to identify changes of geo-graphical entities or phenomena through two or more bitemporal images. Researchers have invested a lot in the homologous change detection and yielded fruitful results. However, change detection between heterogenous remote sensing images is still a great challenge, especially for change detection of heterogenous remote sensing images obtained from satellites and Unmanned Aerial Vehicles (UAV). The main challenges in satellite-UAV change detection tasks lie in the intensive difference of color for the same ground objects, various resolutions, the parallax effect and image distortion caused by different shooting angles and platform altitudes. To address these issues, we propose a novel method based on dual-channel fully convolution network. First, in order to alleviate the influence of differences between heterogeneous images, we employ two different channels to map heterogeneous remote sensing images from satellite and UAV, respectively, to a mutual high dimension latent space for the downstream change detection task. Second, we adopt Hough method to extract the edge of ground objects as auxiliary information to help the change detection model to pay more attention to shapes and contours, instead of colors. Then, IoU-WCE loss is designed to deal with the problem of imbalanced samples in change detection task. Finally, we conduct extensive experiments to verify the proposed method using a new Satellite-UAV heterogeneous image data set, named HTCD, which is annotated by us and has been open to public. The experimental results show that our method significantly outperforms the state-of-the-art change detection methods.

Keywords: change detection; remote sensing; heterogeneous images; deep learning; fully convolution network

1. Introduction

Remote sensing image change detection refers to the process of analyzing two or more remote sensing images at different times to identify the changed geo-graphical entities or phenomena [1]. Heterogeneous images refer to images taken from different types of remote sensing platforms or sensors [2,3]. This paper aims at a special input situation of image change detection, which is heterogeneous images taken by satellites and Unmanned Aerial Vehicles (UAVs).

Remote sensing image change detection analyses multitemporal and multispace information simultaneously, and the result is very meaningful for a vast set of applications, including urban growth tracking [4,5], land use monitoring [6,7], disaster evaluation and damage assessment [8,9]. Therefore, change detection has attracted increasing attention from researchers all over the world.

The condition of conventional homologous image change detection is too strict for many real-world applications. The remote sensing images of a target area before (pre-

image) and after (post-image) the event to be analyzed should be taken by exactly the same sensor under very similar situations, and then the conventional homologous change detection methods can play its role. Hence, most of change detection research use satellite images as input, for the reason that they are taken from the same orbit and under similar situations. However, in many actual applications, this requirement is difficult to satisfy for the limitations of characteristics of satellite imaging, such as weather, solar altitude, restriction of orbital period and imaging width of payloads.

With the advancement of UAV, remote sensing images from UAVs are much easier to be obtained. Change detection between heterogeneous images from satellite and UAV is very import under a couple of applications. For example, in the application of emergency disaster evaluation and rescue, pre-disaster data is often satellite images for its wide coverage and rich historical accumulation, while the post-disaster data can be timeliness images from UAVs, since UAV optical images are often the fastest way to be obtained. In this emergency, satellite-UAV image change detection is the fastest way to assess the disaster damage situation and support the first rescue. Therefore, heterogeneous change detection between images from satellite and UAV has important application prospects, especially in various emergency situations.

At present, the research of heterogeneous change detection mainly focuses on the change detection between different sensors, such as optical images and SAR images. However, there are many differences between Satellite-UAV change detection and optical-SAR change detection. In application, SAR images and UAV images have their own advantages. The advantage of SAR is that it can provide full-time and full-weather observation for targets. While compared to SAR images, UAV optical images have the following two advantages. Firstly, compared with optical imaging, the SAR imaging system is more complex. Therefore, it is more difficult to install on a portable UAV platform to obtain the target area information at the first time in applications for emergency rescue. Secondly, the resolution of SAR image is much lower than UAV optical image, and it cannot provide color information such as UAV optical image [10]. This makes it inferior to UAV images in reflecting the detailed information of the target area. Therefore, in the application of emergency disaster evaluation and rescue, the portable UAV optical remote sensing system and high-resolution UAV images are still the most effective source information of the target area. Technically, optical-SAR heterogeneous change detection and Satellite-UAV change detection face different difficulties. Optical-SAR change detection needs to deal with more intensive image differences. While due to the high resolution of UAV image, Satellite-UAV change detection requires the network to fully consider detailed features to obtain pixel-level accurate change detection results.

To the best of our knowledge, there is no research work on satellite-UAV heterogeneous image change detection. The main challenges of it can be summarized as follows:

1. The intensive spectral difference between satellite images and UAV images. UAV remote sensing generally uses visible light cameras to obtain RGB three-spectrum images, while satellite remote sensing images generally can contain three or more spectra. Furthermore, due to the intensive difference in shooting angle, atmospheric propagation conditions, imaging sensors, etc., the two images might not be obtained with the same spectrum, or the imaging colors of the same ground object will be different. This poses challenges for change detection. The change detection model should be able to discriminate the real ground object changes from the interference caused by shooting condition differences.

2. Various ground resolution of satellite images and UAV images. Since the shooting altitude and the sensors on satellite and UAV are different, UAV images generally have higher resolution than satellite images. The ground resolution of VHR satellites is about 0.5–1 m [11], while the ground resolution of UAV images is generally less than 0.1 m [12]. Shrinking the image by downsampling and interpolation is a straightforward way to solve the problem, but that will lose a lot of detailed information and introduce some accumulated error.

3. The parallax and image distortion of the two kinds of remote sensing images are more serious than homologous images. Since the satellite images and UAV images might shoot from difference angles, the parallax and image distortion are inevitable. There may exist differences between the two images of the same object and the positional deviations are difficult to completely register.

4. The number of unchanged pixels (negative samples) are much more than that of changed pixels (positive samples). Satellite-UAV heterogeneous remote sensing image change detection can be regarded as a typical imbalanced learning problem.

To address these difficulties in heterogeneous change detection, we propose a novel heterogeneous change detection approach and conduct comprehensive experiments on a new dataset we collect and annotate. We name our approach Satellite-UAV heterogeneous remote sensing image change detection neural Network (SUNet).

The main contributions of this paper are as follows:

1. We propose a novel end-to-end deep neural network-based model to map heterogeneous images from satellite and UAV to a mutual high dimension latent space for change detection, in order to mitigate the influence of differences (color, resolution, parallax and image distortion) between heterogeneous images. A novel loss function is designed to address the problem of imbalanced learning issue.

2. We automatically extract the edges of ground objects and feed it into the proposed neural network as auxiliary information to enhance the performance of heterogeneous change detection.

3. To validate the proposed methods, we conducted comprehensive experiments on a new satellite-UAV heterogeneous image change detection dataset, named HTCD, which is collected and annotated by ourselves. The experiment results show that the proposed model significantly outweighs the state-of-the-art change detection methods. The HTCD dataset and our code is open access on https://github.com/ShaoRuizhe/SUNet-change_detection (accessed on 9 September 2021).

2. Related Work

2.1. Image Change Detection

Automatic change detection is a significant technology in remote sensing. Various change detection application scenarios bring various challenges. Researchers have proposed various imaginative change detection approaches to meet these challenges. Traditional approaches, such as change vector analysis (CVA) [13], Multivariate alteration detection (MAD) [14], slow feature analysis (SFA) [15], are simple but effective. They generally use an elaborate handcraft method to extract features from the multitemporal images and then compare the features to acquire the change map. These methods are simple, but the effect is not ideal for some complex situations such as high resolution or heterogeneous image input.

Nowadays, deep learning techniques are demonstrating its promising prospective, especially in image processing field. Since neural networks are able to automatically and continuously tune their parameters to obtain better models, many change detection approaches based on deep learning have been demonstrated better performance than traditional approaches [16–18].

2.2. Deep Learning for Change Detection

According to difference change detection ideas, various change detection processes and model structures are proposed. Transformer [19] based approach uses encoders and decoders to construct the change detection network. Graph Convolutional Network (GCN) [20] based approach performs image segmentation first, then constructs image blocks into a graph and uses GCN to determine which blocks are changed. Gong [21] applies Generative Adversarial Networks (GAN) in change detection.

Siamese networks [18,22–24] use two subnets to extract the high-level features from the two input images, respectively. For homologous change detection tasks, the two subnets

can share their weights to reduce the parameters of the network and make sure the features are extracted in the same way. Chen [23] uses a convolutional network without pooling as subnet, and a RNN to detect changes. Chen [24] uses a dual attention mechanism to locate the changed areas after feature extraction.

Fully convolution network (FCN) [25–28] is widely used in both image classification and change detection. It uses deconvolution to obtain the change map from high-dimensional features, which makes FCN complete change detection task in the form of end-to-end. Liu [25] uses depth-wise separable convolution to make FCN lighter and outperform the original FCN. Li [28] appends an unsupervised noise modeling module after FCN to realize unsupervised change detection.

FCN is relatively mature and versatile among these methods and is therefore widely used. This paper is also based on FCN change detection model.

Loss function is one of the most crucial ingredients in deep learning methods. Early researchers employ the same loss functions of image segmentation tasks [29] to train change detection networks. However, due to the problem of imbalanced samples (unchanged pixels are generally much more than changed pixels) in change detection tasks, these loss functions tend to acquire poor training results. Therefore, various specialized loss functions for change detection have been proposed.

According to the optimization objective of change detection loss functions, they can be divided into three categories. Firstly, distribution-based losses are generally the extensions of Cross Entropy (CE) loss. For example, Weighted Cross Entropy (WCE) [30] adds weight to different classes to deal with the problem of imbalanced samples. Secondly, region-based losses aim to minimize the mismatch regions and maximize the overlap regions between the ground truth region and the predicted region. The mainstream region-based loss functions are Dice loss [31] and IoU loss [32], which are good at dealing with imbalanced sample problem [32]. Finally, combo loss [33–35] refers to the fusion of multiple loss functions through weighted addition and other methods to comprehensively utilize the advantages of multiple loss functions. In this way, the change detection network can be precisely optimized toward the right direction that constrained by the loss.

In our experiments, we found that the region-based loss functions, such IoU loss, cannot drive the model training stably, which often leads to non-convergence of training. However, CE loss is less effective in change detection, which is an imbalanced sample task. Therefore, this paper tries to study the combo loss to fit our problem best.

2.3. Heterogeneous Remote Sensing Change Detection

The existing research works on heterogeneous change detection mainly focus on the change detection between optical images and SAR images. There are two main ideas for heterogeneous change detection: homogeneous transformation and extracting the feature of contour and structures as auxiliary information.

Homogeneous transformation refers to the idea that transform the heterogeneous images into a homogenous domain, then use homogenous change detection methods to complete the change detection task. Based on this idea, Mercier proposes conditional copula [36] and homogeneous pixel transformation [37], Jiang [38] proposes Deep Homogeneous Feature Fusion for heterogeneous optical and SAR images.

Contour and structure feature refers to the idea that uses the feature of contours and structures in the image rather than the feature of colors and pixels in the image to detect changed areas, since the color of heterogeneous images is less reliable than contour and structure. Contour and structure features can play an essential role when pixels and color are incompetent for heterogeneous change detection. Sun [39] performs heterogeneous change detection by constructing and mapping a nonlocal patch-based graph (NLPG). On the basis of histogram of direction of lines (HODOL), the building changes are detected in [40]. In [20], images are segmented into patches of different shapes, and a graph is constructed on these patches and their structure, and the changed areas are determined based on this graph.

Heterogeneous image change detection is a field that gives full play to the creativity of researchers. In addition to the above two main ideas, there are also many ingenious methods. Touati, R. et al. [41] use a series of elaborate constraints and an energy-based model to yield a similarity map. Ayhan, B et al. [42] calculate the correlation between each pixel and other pixels of the two images based on the Pixel Pair, and further obtain the change map. These approaches can perform quite well in some heterogeneous change detection scenarios. However, the existing methods cannot be employed in satellite-UAV change detection tasks directly, since the challenges of optical-SAR change detection or other types of heterogenous are very different from those of satellite-UAV change detection.

3. Methodology

Given two heterogeneous remote sensing images for the same target area A: satellite image $I_{Sat} \in R^{W_{Sat} \times H_{Sat}}$ and UAV image $I_{UAV} \in R^{W_{UAV} \times H_{UAV}}$ captured on different dates t_1 and t_2, respectively, the purpose of change detection is to generate a binary change map D to identify the changed and unchanged geographical entities or phenomena in area A. To achieve this, we propose a novel and effective method, named SUNet.

In this section, first the overall pipeline of the proposed SUNet is given. Then, we elaborate on the detail design of the proposed change detection approach.

3.1. Overview

The framework of the proposed SUNet is shown in Figure 1. Our approach consists of three components: a dual-channel FCN (Section 3.2), an edge auxiliary information extraction method (Section 3.3) and the IoU-WCE loss (Section 3.4).
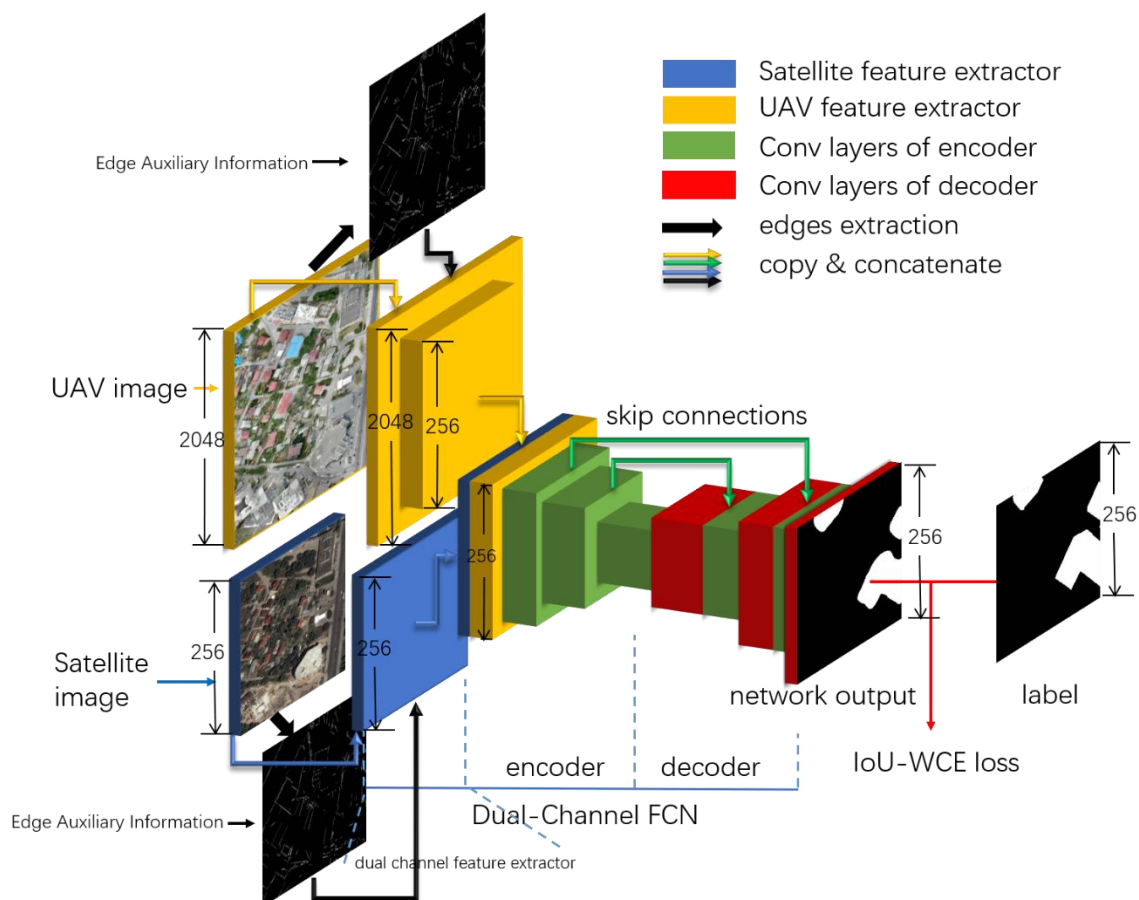
Figure 1. Overview of the proposed SUNet change detection approach.

SUNet employs the proposed Dual-Channel FCN to detect changes between two images of different types. The satellite images and UAV images are directly input into Dual-Channel FCN, without interpolation and scaling. In the Dual-Channel FCN, the feature maps of the two input images are firstly generated through two different feature extraction channels separately, and then the feature maps are concatenated before fed into the encoder of FCN to obtain the high-level feature map. Afterwards, the high-level feature map will pass through the decoder to acquire the change map.

In order to use the edge and shape information to help the model detect heterogeneous changes, the edge auxiliary information is extracted from the image through a Hough-based edge extraction method. Then, it is fed into the model along with the feature maps which are the output of the two channels.

In the training phase, we use the proposed IoU-WCE loss to minimize the difference between the output change map and the labeled change map, because IoU-WCE loss can excellently solve the difficulty of imbalanced samples that plague the general CE loss1.

3.2. Construction of Dual-Channel FCN

As mentioned in Section 1, the ground resolution of satellite and UAV images are always different. Therefore, for the same target area, the sizes of the two input images are different, which cannot be processed by general homogeneous change detection models.

One straightforward way is to use image interpolation methods to reduce the size of UAV images, such as bilinear interpolation and nearest interpolation. However, the simple image downsampling and interpolation methods have some drawbacks with respect to the heterogeneous change detection problem.

1. Image interpolation uses only a few surrounding pixels and loses some information in the UAV image. The edges of the ground objects in the image may be blurred, texture and detail information may be lost, which may lead to a degeneration in the accuracy of subsequent change detection.
2. Image interpolation methods cannot overcome other challenges of the heterogeneous input images, such as the spectral differences and parallax mentioned in Section 1.

To address the above problems, we propose a Dual-Channel FCN to map the two types of images into a mutual high dimension space and then do change detection with FCN. FCN-based change detection methods [43] need two input images of the same size, which means that the heterogenous images of different sizes have to be scaled to the same size with interpolation. While the proposed Dual-Channel FCN can directly input images of different sizes, and extract feature maps from them through two different feature extraction channels, which are composed of two different multilayer convolutional networks, respectively. The UAV feature extraction channel converts the high-resolution UAV image into a high-dimensional feature map that carries both context information and detailed information. Afterwards, the two feature maps are concatenated over their channels and fed into the following FCN. Compared with image interpolation, convolutional network has larger receptive field and allows the network to propagate the context information of high-resolution layers into high-dimensional feature layers, thereby the contextual information can be used to deal with spectral differences and parallax. Given two remote sensing images $I_{UAV} \in R^{W_{UAV} \times H_{UAV}}$ and $I_{Sat} \in R^{W_{Sat} \times H_{Sat}}$, and the edge auxiliary information $E_{UAV} \in \{0,1\}^{W_{UAV} \times H_{UAV}}$ and $E_{Sat} \in \{0,1\}^{W_{Sat} \times H_{Sat}}$ which will be detailed in Section 3.3, the formulas for calculating the feature maps $F_{UAV} \in R^{W_{Sat} \times H_{Sat}}$ and $F_{Sat} \in R^{W_{Sat} \times H_{Sat}}$ through two channels are as follows:

$$F_{UAV} = C_{UAV}(cat(I_{UAV}, E_{UAV})) \tag{1}$$

$$F_{Sat} = C_{Sat}(cat(I_{Sat}, E_{Sat})) \tag{2}$$

where $cat\ (\cdot)$ represents the operation of concatenating two inputs over their channels, $C_{UAV}\ (\cdot)$ and $C_{Sat}\ (\cdot)$ are feature channels for UAV and satellite, respectively.

Inspired by U-net [30] and FCNs [26,43,44], we use "skip connection" concept in our Dual-Channel FCN model. Compared with FCN, Dual-Channel FCN adds the same size feature layer of the encoder to the input of the decoder convolutional layer, as illustrated in Figure 1. The motivation for "skip connections" is to complement large-scale and global information of the encoded features with the detail and localized information that are present in the earlier feature layers of the encoder. Through the fusion of local and global information, the model can not only correctly judge the change area and position, but also find the precise and detailed boundary. To be specific, the encoder and decoder of Dual-Channel FCN are formulated as follows:

$$e_i = \begin{cases} \mathcal{H}_{ei}\{cat(\boldsymbol{F_{UAV}}, \boldsymbol{F_{Sat}})\} & i = 0 \\ \mathcal{H}_{ei}\{AvePool(\boldsymbol{e_{i-1}})\} & 4 > i > 0 \end{cases} \quad (3)$$

$$d_j = \begin{cases} \mathcal{H}_{dj}\{cat(TConv(AvePool(\boldsymbol{e_{3-j}})), \boldsymbol{e_{3-j}})\} & j = 0 \\ \mathcal{H}_{dj}\{cat(TConv(\boldsymbol{d_{j-1}}), \boldsymbol{e_{3-i}})\} & 4 > j > 0 \end{cases} \quad (4)$$

$$\boldsymbol{D} = argmax\left(\boldsymbol{d_3}^{2 \times W_{out} \times H_{out}}\right) \quad (5)$$

where $\mathcal{H}(\cdot)$ denotes the operation of the convolution block, $\boldsymbol{e_i}$ and $\boldsymbol{d_j}$ are feature layers in encoder and decoder, respectively, $AvePool(\cdot)$ is average pooling, $TConv(\cdot)$ is the operation of transposed convolution. $\boldsymbol{D} \in \{0,1\}^{1 \times W_{out} \times H_{out}}$ is the output binary change map.

3.3. Extraction of Edge Auxiliary Information

Auxiliary information including edges and shapes is very helpful to the change detection between UAV images and satellite images. Firstly, although the Dual-Channel structure enlarges the receptive field and preserves more context information compared to downsampling and interpolation, the feature still only contains local information of a small region. If this local and detail information can be complemented with global information, the obtained feature map will represent the original image more completely. The edge auxiliary information is extracted based on global features—the shape and contour of the ground objects. Secondly, as stated in Section 1, since there is a variety of intensive spectral difference between satellite images and UAV images, the color information is less reliable than the shape and contour. Therefore, for the task of Satellite-UAV heterogenous change detection, if the change detection algorithm can pay more attention to the shape and contour, it will reduce the interference of the spectral difference and find the real changes. Finally, for application scenarios of urban area administration and human rescue supporting, artificial entity changes are more attention-worthy than natural feature changes. Since the latter might be caused by seasons and climate, while the former can directly reflect the factors we are concerned about, such as the extent of house damage caused by natural disasters or urban development. Generally speaking, artificial entities (such as houses and roads) will have neat straight edges, while natural features (such as trees, etc.) will have irregular edges. Based on this, the proposed model can pay more attention to artificial features with the help of auxiliary edge information, so as to better meet the needs of these applications.

Hough algorithm [45] is a classic and effective method for straight line extraction. However, it can satisfy the effect of the auxiliary information mentioned above on the Dual-Channel structure. First, in the process of mapping pixels in the entire image to the parameter space and finding extreme values, global information can be obtained to complement the local information provided by image convolution. Secondly, the straight line extracted by the Hough transform conforms to the straight edge feature of artificial entities.

In this paper, the edges of the two images are extracted using Canny algorithm [45], and then Hough algorithm is applied to extract straight edges as auxiliary information. Hough algorithm is delicately tuned to acquire two suitable groups of parameters for satellite and UAV images, respectively. The extracted edge auxiliary information is input

into the model by concatenating with the input images over the channels. An example of edge auxiliary extraction is displayed in Figure 2.

Figure 2. The visualized result of edge auxiliary information extraction. Left: the input (UAV) image. Right: the extracted information.

3.4. IoU-WCE Loss

As mentioned in Section 2.1, compared with general image segmentation, an important difference of change detection is the imbalanced ratio of the two classes—"changed" class and "unchanged" class. This is likely to cause the misdirection of the model optimization and training. For this problem, we propose two solutions based on loss function improvement.

3.4.1. Introduction of WCE

Based on the idea of minimizing KL divergence, cross entropy is defined by:

$$L_{ce}(i) = -\sum_{c=1}^{C} g(i,c) log p(i,c) \tag{6}$$

where C is the number of classes, $g(i,c)$ represents the ground truth binary map, indicating if pixel i is in class c, and $p(i,c)$ is the predicted probability of pixel i is in class c.

To address the problem of imbalanced samples, WCE [30] increases the weight of sub-categories to prevent them from being ignored:

$$L_{wce}(i) = -\sum_{c=1}^{C} w_c g(i,c) log p(i,c) \tag{7}$$

where w_c is the weight of class c.

In this paper, to balance the changed and unchanged class, we use the ratio of the pixels as the weight. The proposed WCE loss for change detection can be formulated as follows:

$$L_{wce}(i) = -[P_u g(i,u) log p(i,u) + P_c g(i,c) log p(i,c)] \tag{8}$$

where P_u and P_c are the frequencies of the changed and unchanged pixels, respectively. u and c in g and p represents the pixel i is unchanged or changed.

3.4.2. IoU-WCE Combo Loss

In the field of image segmentation, region-based loss functions are proposed to handle the imbalanced sample problem [29]. At present, the mainstream region-based loss func-

tions are Dice loss and IoU loss [29]. Through experiments (Section 4.5), we found that in our heterogenous change detection task, using IoU-loss can acquire a better training result. The idea of IoU loss is to minimize the IoU of the network output. IoU loss is defined by the following equation.

$$L_{iou} = 1 - \frac{\sum_{i \in I} g_i p_i}{\sum_{i \in I}(g_i + p_i - g_i p_i)} \quad (9)$$

IoU loss can handle imbalanced sample problems such as change detection well [32]. However, in IoU loss (Equation (9)), the model output p appears in the denominator and the unstable term $\frac{1}{p^2}$ will turn up in the derivative of p during the backpropagation process. This term lays a hidden trouble to the training process. Therefore, we propose IoU-WCE combo loss to smoothly optimize our model while overcoming the problem of unbalanced sampling:

$$L_{total} = \alpha L_{iou} + (1 - \alpha) L_{wce} \quad (10)$$

where α is a constant weight to balance the two weights. It will be further discussed in the experiment (Section 4.5.3).

4. Experiment

4.1. Dataset

The objective of our HTCD dataset is to provide a standardized way to compare the efficacy of various heterogenous change detection models and algorithms proposed by creative researchers. The HTCD dataset covers Chisinau and its surrounding area, which is an area of approximately 36 square kilometers. The dataset contains two image scenes that are shot by satellite in 2008 and UAV in 2020, respectively. The dataset is labeled focused on urban areas. Urban changes, including buildings, roads and other urban man-made features, are carefully labeled, while natural changes are ignored. Some examples of the dataset are shown in Figure 3. Since the dataset contains pixel-wise ground truth change labels, accurate to every changing object in the area, the dataset can be applied to evaluate or train various supervised learning models for heterogenous change detection.
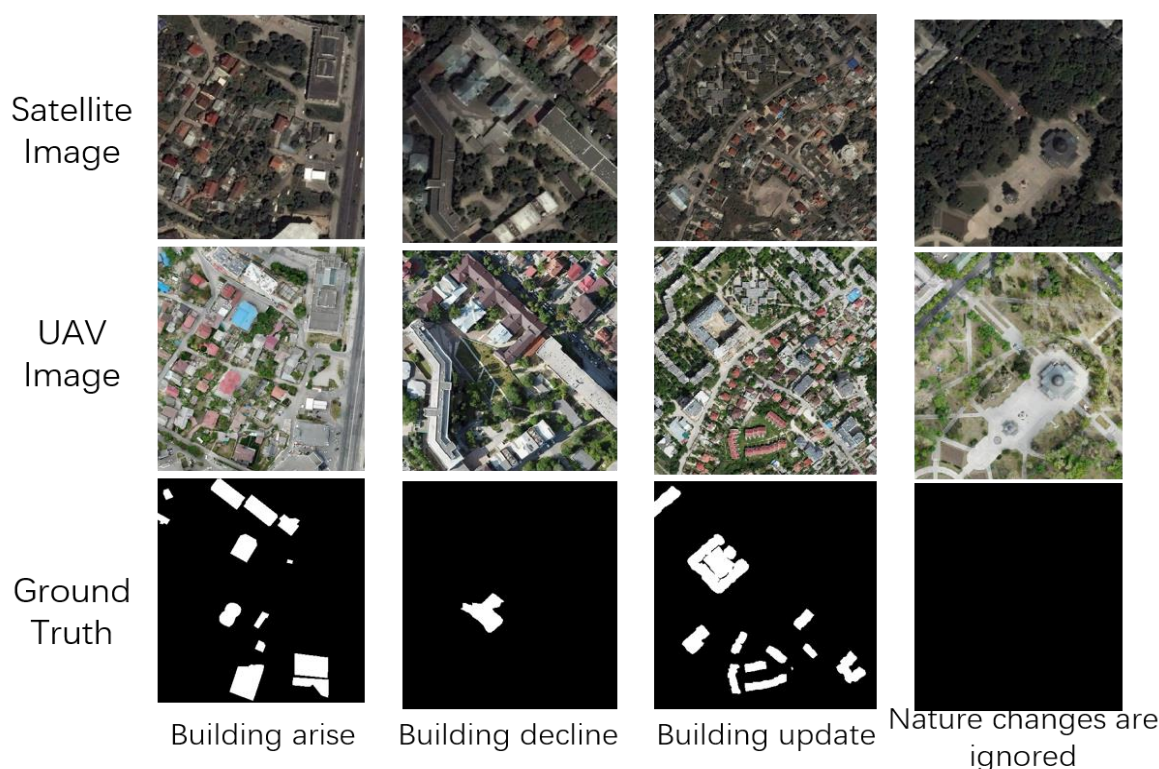
Figure 3. Example Image Pairs of HTCD Dataset. White-changed. Black-unchanged.

The HTCD dataset was built using the satellite images from Google Earth (https://www.google.com/earth/ (accessed on 10 April 2021)) and UAV images from Open Aerial Map (https://map.openaerialmap.org/ (accessed on 10 April 2021)). The size of the satellite image is 11 K × 15 K pixels. While the UAV image is consisted of 15 image blocks, in total 1.38 M × 1.04 M pixels. The ground resolutions of them are 0.5971 m and 7.465 cm, respectively. Images and labels are all stored in GeoTiff format with location information, for the convenience of further analysis and research.

The image is registered by manually selecting control points and using the polynomial method. After registration, the pixel-wise ground truth labels were manually generated by comparing the two images. The operations including image registration, comparison and annotation are completed using QGIS (https://www.qgis.org/en/site/ (accessed on 9 September 2021)).

4.2. Setup

4.2.1. Metrics

We calculate 5 metrics to evaluate the performance of the proposed method: precision (*Pre*), recall (*Rec*), F1 score (*F1*), overall accuracy (*OA*) and Intersection-over-Union (*IoU*). *P* and *R* represent lower false detection and omission, respectively. While the overall evaluation metrics of the prediction results are given by *F1* and *OA*. The larger their values, the better the prediction results. The above-mentioned metrics are mainly introduced from prediction problems. In addition to these metrics, we also adopt *IoU*, a pivotal metric in image segmentation problem, which is a good assessment of the consistency between the detected changes and ground truth. They are expressed as follows:

$$Pre = \frac{TP}{TP + FP} \tag{11}$$

$$Rec = \frac{TP}{TP + FN} \tag{12}$$

$$F1 = \frac{2P \cdot R}{P + R} \tag{13}$$

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$IoU = \frac{\{PreChange\} \cap \{GTChange\}}{\{PreChange\} \cup \{GTChange\}} \tag{15}$$

where *TP* is the number of true positives, *FP* is the number of false positives, *TN* is the number of true negatives and *FN* is the number of false negatives. {*PreChange*} and {*GTChange*} are the set of detected change pixels and ground-truth change pixels, respectively.

4.2.2. Training Details

During training, HTCD dataset was clipped into 3772 image tile pairs consisting of $W_{Sat} \times H_{Sat} = 256 \times 256$ satellite images and $W_{UAV} \times H_{UAV} = 2048 \times 2048$ UAV images. Furthermore, 60%, 20% and 20% of them are randomly selected as the training set, validation set and testing set, respectively. The training set was input into the model with 5 images as a batch. Adam optimizer with a learning rate of 0.001 was adopted to adjust the weights. The weight of IoU loss α (Equation (10)) is set to 0.66.

Our change detection approach is implemented via OpenCV (https://opencv.org/ (accessed on 15 September 2021)) and Pytorch (http://pytorch.org/ (accessed on 9 September 2021)) on a deep learning server powered by 2× GeForce GTX 2080Ti GPU, 256 GB RAM, 2 × 10 GB VRAM.

4.3. Baselines

To the best of our knowledge, there are no existing research results on Satellite-UAV heterogeneous change detection. Therefore, we selected the latest change detection method including both homologous and heterogeneous models as the baseline.

- Symmetric Convolutional Coupling Network (SCCN) [46]. SCCN is a method for heterogenous change detection tasks. It extracts feature maps from SAR and optical images from satellites. Next, the high-quality different map can be obtained by comparing and fusing these feature maps. SCCN is a wide-used heterogenous change detection model.
- Spatial–Temporal Attention neural Network (STANet) [47]. STANet is a homologous change detection method for very high resolution (VHR) satellite images. It first uses a Siamese FCN to extract the bitemporal image feature maps, then uses the proposed self-attention module to update the feature. Finally, the features are compared and the change map is generated.
- Fully Convolutional Early Fusion (FC-EF) [43]. FC-EF is a method for homologous change detection tasks. It uses the idea of FCN image segmentation for change detection. It stacks the two images as the input to the model, and uses the skip connection concept introduced in U-Net [30] to fuse the features of different scales and obtain the final change map.
- BiDateNet [48]. BiDateNet is a model for homologous change detection between satellite images. It integrated the Recurrent Neural Network (RNN) into the encoder of U-Net to help the network learn the temporal change pattern.
- The combination of Siamese network and Nested U-Net for Change Detection (SNUNet-CD) [49]. SNUNet-CD is a method for homologous change detection between VHR satellite images. It uses nested information transmission between the encoder and the decoder to alleviate the loss of localization information in the deep layers of the neural network. In addition, ECAM is proposed in this work to suppress semantic gaps and localization errors.

To make the above methods be able to process the HTCD satellite-UAV heterogeneous data, we used downsampling and linear interpolation to reduce the UAV image to the same size as the satellite image before training and testing these baseline methods. When testing SCCN, in addition to the above operation. We also converted three-channel RGB images into grayscale images. During the training process, we used hyperparameters as the experiments of the method proposers.

4.4. Performance

4.4.1. Experiment Results

To evaluate the performance of the proposed method, we conducted comparative experiments with the baselines introduced in Section 4.3. The results are shown in Table 1 and Figure 4.

Table 1. Result of SUNet and Baselines on HTCD Dataset. The metrics are Recall (*Rec*), Precision (*Pre*), F1-score (*F1*), Overall-Accuracy (*OA*), and Intersection over Union (*IoU*).

Method	*Rec*	*Pre*	*F1*	*OA*	*IoU*
SCCN	0.053	0.324	0.091	0.828	0.048
STANet	0.817	0.656	0.728	0.985	0.572
FC-EF	0.774	0.907	0.835	0.991	0.717
BiDateNet	0.813	0.930	0.863	0.993	0.767
SNUNet-CD	0.334	0.227	0.249	0.947	0.156
SUNet	0.864	0.973	0.910	0.996	0.842
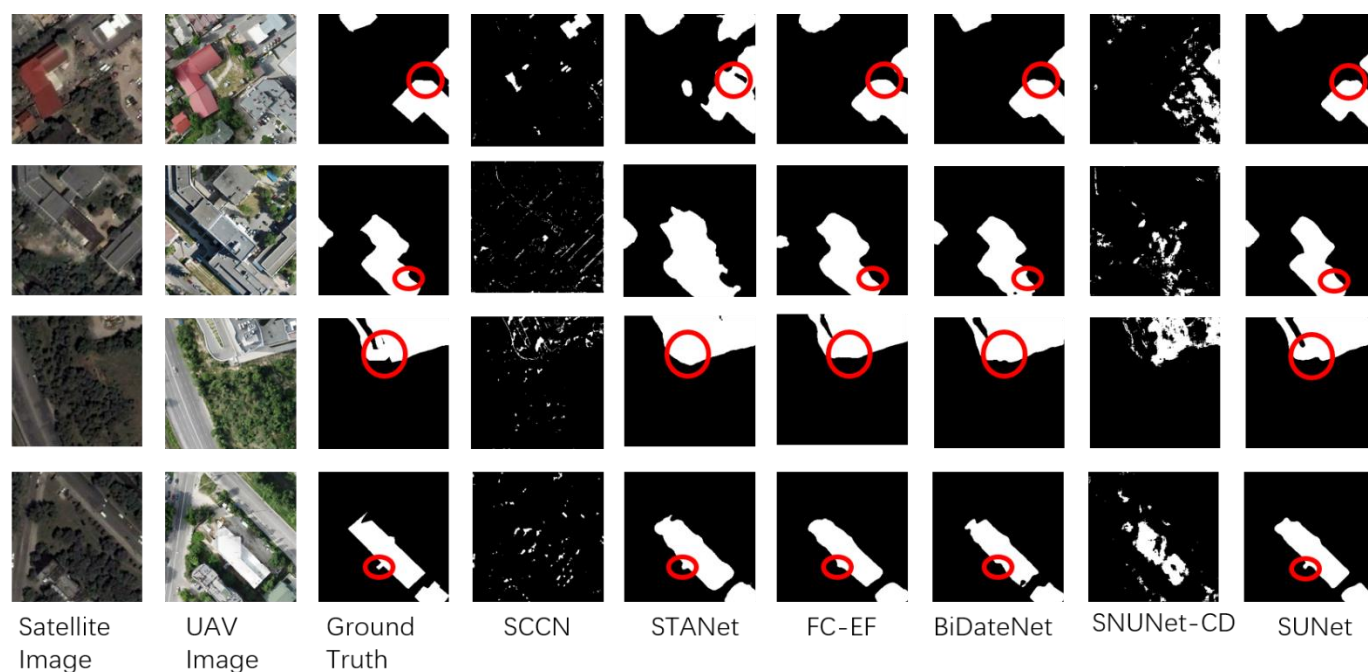
Figure 4. Visualized result of baselines and SUNet on HTCD dataset.

Table 1 shows that our method has a significant improvement over other methods on the heterogeneous dataset, which proves that our method is effective for change detection task between satellite and UAV remote sensing images.

In order to intuitively explain why the proposed SUNet is superior to other advanced methods in the evaluation metrics in Table 1. We visualized some of the outputs in Figure 4.

4.4.2. Discussions on the Performance of SUNet

In Section 4.4.1, the outstanding performance of the proposed method is proved. As you can see from Table 2 and Figure 4, compared to the homologous methods, including STANet, FC-EF, BiDateNet and SNUNet-CD, SUNet, using the dual-channel structure and edge auxiliary information, obtained more accurate detection results in the details of the changed areas. Among them, BiDateNet, STANet and SNUNet-CD use a Siamese network with shared weights, which can reduce the number of parameters and improve the training efficiency in homologous change detection, but this structure seems less effective for heterogeneous change detection task. The SCCN method for optical-SAR image change detection performs poorly in the satellite-UAV task. One of the reasons is that SCCN is an unsupervised method, therefore, its results are worse than those supervised methods. However, the performance of SCCN in satellite-UAV change detection task is also worse than that of the optical-SAR task [46]. This fact reveals the difference between satellite-UAV change detection and optical-SAR change detection, and focused research on satellite-UAV change detection is necessary to handle its unique challenges.

Table 2. Ablation Study Result. Where Info denotes the Auxiliary information proposed in Section 3.1, and DC denotes the Dual-Channel structure proposed in Section 3.2. '√' sign means using this proposed component, while '×' sign means removing this component from the proposed change detection method. The metrics are Recall (*Rec*), Precision (*Pre*), F1-score (*F*1), Overall-Accuracy (*OA*) and Intersection over Union (*IoU*).

Method	Info	DC	Loss	*Rec*	*Pre*	*F*1	*OA*	*IOU*
NoInfo-NoDC-WCE	×	×	WCE	0.774	0.907	0.835	0.991	0.717
NoInfo-NoDC	×	×	WCE + IoU	0.911	0.855	0.883	0.994	0.790
HM-NoDC	×	×	WCE + IoU	0.935	0.835	0.877	0.994	0.789
NoDC	√	×	WCE + IoU	0.920	0.877	0.893	0.995	0.815
NoInfo	×	√	WCE + IoU	0.892	0.913	0.898	0.995	0.823
WCE	√	√	WCE	0.868	0.849	0.859	0.993	0.752
IoU	√	√	IoU			not converge		
Dice-WCE	√	√	WCE + Dice	0.825	0.937	0.872	0.994	0.781
SUNet	√	√	WCE + IoU	0.864	0.973	0.910	0.996	0.842

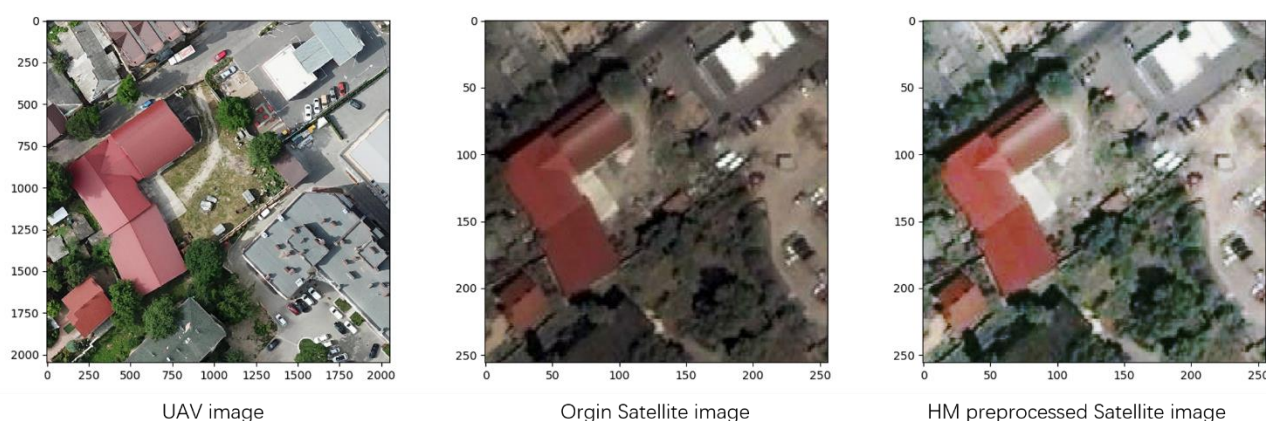*4.5. Ablation Study*

To better study the utility of each component in the method, we designed ablation experiments to evaluate the performance of the auxiliary information, Dual-Channel FCN structure and IoU-WCE combo loss.

4.5.1. Ablation Study Result for the Components in SUNet

We designed ablation experiments for every proposed component in SUNet. First, to evaluate the performance improvement brought by Edge Auxiliary Information, we remove the Edge Auxiliary Information and directly input the images into the model. We mark this modification as NoInfo. Furthermore, Histogram Matching (HM) [50] is a commonly used color correction method in the field of remote sensing images. In order to explore whether this method can be used as a preprocessing method to alleviate the impact of the spectral difference between satellite and UAV images, we also conducted an experiment to input the preprocessed images of histogram matching into the network. Figure 5 shows the effect of histogram matching preprocessing. This experiment is denoted as HM (Histogram Matching). Second, in order to verify the effect of the Dual-Channel structure, we use bilinear interpolation instead of Dual-Channel to reduce the UAV image to the same size as the satellite image. We mark this modification as NoDC. Third, we replace WCE-IoU loss with WCE, IoU or Dice-WCE loss, respectively, and mark this modification as WCE, IoU or Dice-IoU. Finally, in order to observe the correlation effects of these components, we also conducted experiments in which multiple components were modified simultaneously in accordance with the above-mentioned methods. While conducting these experiments, the weight of IoU loss in the combo loss (α in Equation (10)) is 0.66. The result is demonstrated in Table 2.
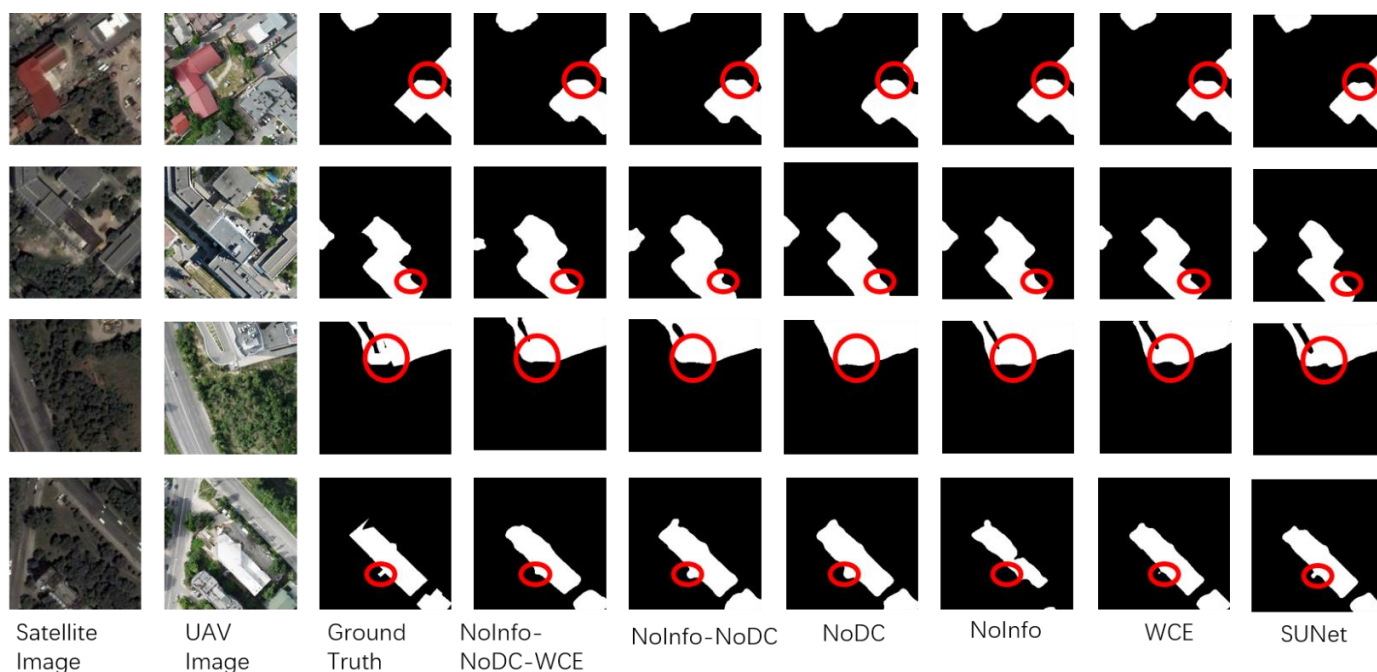
The result shows that our approach does improve the performance of the model, and each component contributes to the performance improvement. Experiment of row 2, 3, 4, 5 and 9 proves the effect of auxiliary information. For interpolation-based structure and dual-channel structure, the auxiliary information improves IoU by 3.2% and 2.3%, respectively. Experiment of row 2, 4, 5 and 9 also proves the effect of Dual-Channel FCN structure. In the presence and absence of auxiliary information, the dual-channel structure improves IoU by 4.2% and 3.3%, respectively. Experiments of row 1, 2, 6, 7 and 9 prove the effect of IoU-WCE loss. For the presence and absence of auxiliary information and Dual-Channel structure, compared with WCE loss, IoU-WCE loss improved by 12.0% and 10.2%. The last two rows of the table prove that in the training of SUNet, the effect of IoU-WCE loss is better than that of Dice-WCE loss.

UAV image

Orgin Satellite image

HM preprocessed Satellite image

Figure 5. An Example of Histogram Matching Preprocess.

4.5.2. Visualization of Ablation Study Result

To further understand the effect of the components in SUNet more intuitively, we visually demonstrate some examples in Figure 6.


Satellite Image | UAV Image | Ground Truth | NoInfo-NoDC-WCE | NoInfo-NoDC | NoDC | NoInfo | WCE | SUNet

Figure 6. Visualization of Ablation Study Experiments. "NoInfo" below these columns indicates that the edge auxiliary information is removed from the proposed changed detection method. "NoDC" means downsampling and interpolation is employed to downscale the UAV images and Dual-Channel structure is not used. "WCE" indicates that the WCE loss function is used instead of IoU-WCE loss function.

As can be seen from Figure 6, all three ingredients in our method have contributed to the improvement of the detection result. The Dual-Channel structure and the edge auxiliary information make the boundary more refined, and IoU-WCE loss makes the changed area be detected more completely.

4.5.3. Experiments on α

As discussed in Section 3.4.2, the IoU-WCE loss use a constant parameter α to balance the two loss functions. Since the value of α will have a great influence on the training result, we conducted an experiment to analysis the relationship between the performance of SUNet and the value of α to find the best selection of α. Figure 7 shows the performance

of SUNet with different α. According to the result in Figure 7, 0.66 is the most suitable value of α for SUNet.
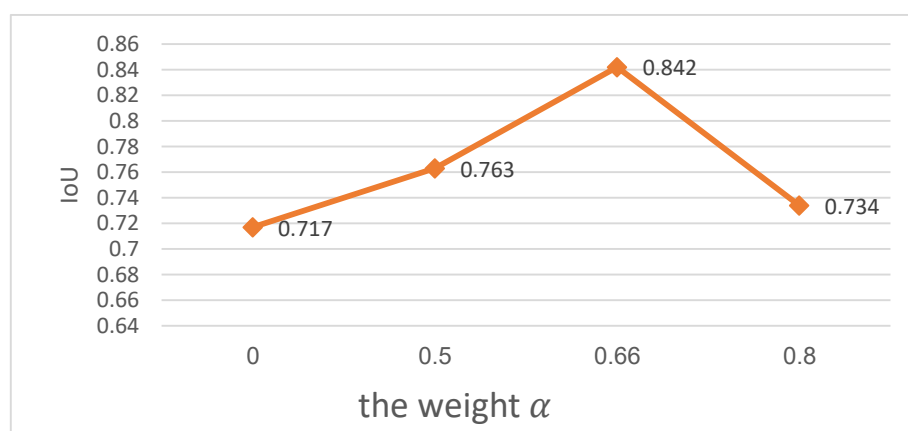
Figure 7. The influence of α on SUNet4.5.4 Discussions on the ingredients of SUNet.

In Sections 4.5.1 and 4.5.2, the ingredients of the proposed change detection method are confirmed.

Table 2 and Figure 5 confirm the discussion in Section 3.2. Compared with downsampling and interpolation methods, the multiple convolutional layers of the UAV channel in Dual-Channel structure retain more detailed information, so more detailed detection result is obtained. Figure 6 has a very intuitive proof of this. Table 2 also proves this point from the perspective of metric evaluation.

Comparing the 5th and 6th columns, the 7th and 9th columns of Figure 6, it can be found that when other factors are the same, the auxiliary information can help the model to better find the edges of the changed part, and make the detection result more accurate. This is consistent with what is stated in Section 3.2. Edge information can be used as global information to supplement the local information of the convolution model. In addition, HM's experiments proved the complexity of the color difference between UAV images and satellite images. As mentioned in Section 1, their complex spectral differences are caused by the differences in imaging conditions, such as different imaging spectra, atmospheric propagation conditions and imaging sensors. However, histogram matching is good at correcting the imaging color differences caused by the lighting conditions, so it is difficult to be used as the preprocessing method to greatly improve the network performance. Without the information about shape and structure, the complex spectral difference of the heterogeneous image will confuse the FCN network, and ultimately acquire poor results.

Table 2 shows that among the three methods we proposed, the loss function brings the most improvement, indicating that the loss function is indeed crucial to the training of the model, but it does not mean that other methods are not as important as the loss function. Since in the final model, the three methods work together to achieve the best performance. We tested serval different loss functions in our experiments. Firstly, for SUNet's change detection task, the combo losses (including IoU-WCE loss and Dice-WCE loss) have obtained better performance than WCE loss. As described in Section 3.4, the two proposed methods, adding weight to CE loss and adding the IoU loss, overcome the challenge of unbalanced samples in the change detection task, and correctly guide the model training. Secondly, since in IoU loss (demonstrated in Equation (9)), the model output p appears in the denominator, the unstable term $\frac{1}{p^2}$ will turn up in the derivative of p during the backpropagation process. This led to the result of non-convergent training. While by compounding with CE loss, IoU-WCE loss is more stable than IoU loss, and the problem of non-convergence is successfully overcome. Finally, The poor effect of Dice loss may be because it directly ignoring the background regions [51], which causes information loss during training.

5. Conclusions and Future Work

In this paper, we propose SUNet, which is a deep-learning-based approach for change detection in Satellite-UAV heterogenous images, which significantly improves the accuracy of Sat-UAV heterogeneous change detection. Through extensive experiments, we found that although satellite and UAV images are all optical images, there are many significant differences between them, such as intensive spectral difference, various ground resolution, the parallax and image distortion. The experiments also prove that the method proposed in this article can specifically address these differences and overcome the challenges they brought to significantly improve the change detection performance. We also provide a new dataset, HTCD for Satellite-UAV change detection tasks. We do hope that this dataset can help researchers explore more creative approaches in the field.

Our future work is to study semi-supervised or unsupervised satellite-UAV change detection methods, which have wider application scenarios.

Author Contributions: Conceptualization, R.S. and J.L.; methodology, R.S. and C.D.; software, R.S.; validation, R.S., C.D. and H.C.; data curation, H.C.; writing—original draft preparation, R.S.; writing—review and editing, C.D. and J.L.; supervision, H.C.; project administration, H.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/ShaoRuizhe/SUNet-change_detection (accessed on 9 September 2021).

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

References

1. Singh, A. Review article digital change detection techniques using remotely-sensed data. *Int. J. Remote Sens.* **1989**, *10*, 989–1003. [CrossRef]
2. Zhan, T.; Gong, M.; Jiang, X.; Li, S. Log-based transformation feature learning for change detection in heterogeneous images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1352–1356. [CrossRef]
3. Kwan, C.; Ayhan, B.; Larkin, J.; Kwan, L.; Bernabé, S.; Plaza, A. Performance of change detection algorithms using heterogeneous images and extended multi-attribute profiles (EMAPs). *Remote Sens.* **2019**, *11*, 2377. [CrossRef]
4. Ansari, R.A.; Buddhiraju, K.M.; Malhotra, R. Urban change detection analysis utilizing multiresolution texture features from polarimetric SAR images. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100418. [CrossRef]
5. Tian, S.; Ma, A.; Zheng, Z.; Zhong, Y. Hi-UCD: A large-scale dataset for urban semantic change detection in remote sensing imagery. *arXiv* **2020**, arXiv:2011.03247. preprint.
6. Lu, M.; Chen, J.; Tang, H.; Rao, Y.; Yang, P.; Wu, W. Land cover change detection by integrating object-based data blending model of Landsat and MODIS. *Remote Sens. Environ.* **2016**, *184*, 374–386. [CrossRef]
7. Thonfeld, F.; Steinbach, S.; Muro, J.; Kirimi, F. Long-term land use/land cover change assessment of the Kilombero catchment in Tanzania using random forest classification and robust change vector analysis. *Remote Sens.* **2020**, *12*, 1057. [CrossRef]
8. Weber, E.; Kané, H. Building disaster damage assessment in satellite imagery with multi-temporal fusion. *arXiv* **2020**, arXiv:2004.05525, preprint.
9. Yan, Z.; Huazhong, R.; Danyang, G. Research on the Detection Method of Building Seismic Damage Change. In Proceedings of the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 2906–2908.
10. Reigber, A.; Schreiber, E.; Trappschuh, K.; Pasch, S.; Müller, G.; Kirchner, D.; Geßwein, D.; Schewe, S.; Nottensteiner, A.; Limbach, M. The high-resolution digital-beamforming airborne SAR system DBFSAR. *Remote Sens.* **2020**, *12*, 1710. [CrossRef]
11. Wolniewicz, W. Assessment of geometric accuracy of VHR satellite images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 19–23.
12. Zongjian, L. UAV for mapping—Low altitude photogrammetric survey. *Int. Arch. Photogramm. Remote Sens. Beijing China* **2008**, *37*, 1183–1186.
13. Bruzzone, L.; Prieto, D.F. Automatic analysis of the difference image for unsupervised change detection. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 1171–1182. [CrossRef]

14. Nielsen, A.A.; Conradsen, K.; Simpson, J.J. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sens. Environ.* **1998**, *64*, 1–19. [CrossRef]

15. Wu, C.; Du, B.; Zhang, L. Slow feature analysis for change detection in multispectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 2858–2874. [CrossRef]

16. Liu, T.; Abd-Elrahman, A.; Zare, A.; Dewitt, B.A.; Flory, L.; Smith, S.E. A fully learnable context-driven object-based model for mapping land cover using multi-view data from unmanned aircraft systems. *Remote Sens. Environ.* **2018**, *216*, 328–344. [CrossRef]

17. Radhika, K.; Varadarajan, S. A neural network based classification of satellite images for change detection applications. *Cogent Eng.* **2018**, *5*, 1484587. [CrossRef]

18. Wang, M.; Tan, K.; Jia, X.; Wang, X.; Chen, Y. A deep siamese network with hybrid convolutional feature extraction module for change detection based on multi-sensor remote sensing images. *Remote Sens.* **2020**, *12*, 205. [CrossRef]

19. Chen, H.; Qi, Z.; Shi, Z. Efficient Transformer based Method for Remote Sensing Image Change Detection. *arXiv* **2020**, arXiv:2103.00208, preprint.

20. Wu, J.; Li, B.; Qin, Y.; Ni, W.; Zhang, H.; Sun, Y. A Multiscale Graph Convolutional Network for Change Detection in Homogeneous and Heterogeneous Remote Sensing Images. *arXiv* **2020**, arXiv:2102.08041, preprint.

21. Gong, M.; Niu, X.; Zhang, P.; Li, Z. Generative adversarial networks for change detection in multispectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2310–2314. [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. Chen, H.; Wu, C.; Du, B.; Zhang, L.; Wang, L. Change detection in multisource VHR images via deep Siamese convolutional multiple-layers recurrent neural network. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2848–2864. [CrossRef]

24. Chen, J.; Yuan, Z.; Peng, J.; Chen, L.; Haozhe, H.; Zhu, J.; Liu, Y.; Li, H. DASNet: Dual attentive fully convolutional siamese networks for change detection of high resolution satellite images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *14*, 1194–1206. [CrossRef]

25. Liu, R.; Jiang, D.; Zhang, L.; Zhang, Z. Deep depthwise separable convolutional network for change detection in optical aerial images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 1109–1118. [CrossRef]

26. Sefrin, O.; Riese, F.M.; Keller, S. Deep Learning for Land Cover Change Detection. *Remote Sens.* **2021**, *13*, 78. [CrossRef]

27. Li, C.; Kong, X.; Wang, F.; Wang, Y.; Zhang, M. Research on change detection of high resolution remote sensing image based on U-type neural network. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Xiamen, China, 17–19 November 2020; p. 012166.

28. Li, X.; Yuan, Z.; Wang, Q. Unsupervised deep noise modeling for hyperspectral image change detection. *Remote Sens.* **2019**, *11*, 258. [CrossRef]

29. Ma, J. Segmentation loss odyssey. *arXiv* **2020**, arXiv:2005.13449, preprint.

30. 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 Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.

31. Milletari, F.; Navab, N.; Ahmadi, S. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.

32. Rahman, M.A.; Wang, Y. Optimizing intersection-over-union in deep neural networks for image segmentation. In Proceedings of the International Symposium on Visual Computing, Las Vegas, NV, USA, 12–14 December 2016; pp. 234–244.

33. Taghanaki, S.A.; Zheng, Y.; Zhou, S.K.; Georgescu, B.; Sharma, P.; Xu, D.; Comaniciu, D.; Hamarneh, G. Combo loss: Handling input and output imbalance in multi-organ segmentation. *Comput. Med Imaging Graph.* **2019**, *75*, 24–33. [CrossRef] [PubMed]

34. Wong, K.C.; Moradi, M.; Tang, H.; Syeda-Mahmood, T. 3D segmentation with exponential logarithmic loss for highly unbalanced object sizes. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018; pp. 612–619.

35. Caliva, F.; Iriondo, C.; Martinez, A.M.; Majumdar, S.; Pedoia, V. Distance map loss penalty term for semantic segmentation. *arXiv* **2020**, arXiv:1908.03679, preprint.

36. Mercier, G.; Moser, G.; Serpico, S.B. Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1428–1441. [CrossRef]

37. Liu, Z.; Li, G.; Mercier, G.; He, Y.; Pan, Q. Change detection in heterogenous remote sensing images via homogeneous pixel transformation. *IEEE Trans. Image Process.* **2017**, *27*, 1822–1834. [CrossRef]

38. Jiang, X.; Li, G.; Liu, Y.; Zhang, X.-P.; He, Y. Change detection in heterogeneous optical and SAR remote sensing images via deep homogeneous feature fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 1551–1566. [CrossRef]

39. Sun, Y.; Lei, L.; Li, X.; Tan, X.; Kuang, G. Structure Consistency-Based Graph for Unsupervised Change Detection With Homogeneous and Heterogeneous Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–21. [CrossRef]

40. Pang, S.; Hu, X.; Cai, Z.; Gong, J.; Zhang, M. Building change detection from bi-temporal dense-matching point clouds and aerial images. *Sensors* **2018**, *18*, 966. [CrossRef]

41. Touati, R.; Mignotte, M. An energy-based model encoding nonlocal pairwise pixel interactions for multisensor change detection. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 1046–1058. [CrossRef]

42. Ayhan, B.; Kwan, C. A new approach to change detection using heterogeneous images. In Proceedings of the IEEE Ubiquitous Computing, Electronics & Mobile Communication Conference, New York, NY, USA, 10–12 October 2019; pp. 10–12.

43. Daudt, R.C.; Le Saux, B.; Boulch, A. Fully convolutional siamese networks for change detection. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 4063–4067.

44. 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.

45. Sonka, M.; Hlavac, V.; Boyle, R. *Image Processing, Analysis, and Machine Vision*; Cengage Learning: Boston, MA, USA, 2014.

46. 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] [PubMed]

47. 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]

48. Papadomanolaki, M.; Verma, S.; Vakalopoulou, M.; Gupta, S.; Karantzalos, K. Detecting urban changes with recurrent neural networks from multitemporal Sentinel-2 data. In Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 214–217.

49. Chen, H.; Li, W.; Shi, Z. Adversarial Instance Augmentation for Building Change Detection in Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–16. [CrossRef]

50. Camps-Valls, G.; Tuia, D.; Bruzzone, L.; Benediktsson, J.A. Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Signal Process. Mag.* **2013**, *31*, 45–54. [CrossRef]

51. Wang, L.; Wang, C.; Sun, Z.; Chen, S. An improved dice loss for pneumothorax segmentation by mining the information of negative areas. *IEEE Access* **2020**, *8*, 167939–167949. [CrossRef]