



Article

Self-Incremental Learning for Rapid Identification of Collapsed Buildings Triggered by Natural Disasters

Jiayi Ge ^{1,2}, Hong Tang ^{1,2,*} and Chao Ji ^{1,2}

¹ State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; 202021051203@mail.bnu.edu.cn (J.G.); jichao@mail.bnu.edu.cn (C.J.)

² Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

* Correspondence: hongtang@bnu.edu.cn

Abstract: The building damage caused by natural disasters seriously threatens human security. Applying deep learning algorithms to identify collapsed buildings from remote sensing images is crucial for rapid post-disaster emergency response. However, the diversity of buildings, limited training dataset size, and lack of ground-truth samples after sudden disasters can significantly reduce the generalization of a pre-trained model for building damage identification when applied directly to non-preset locations. To address this challenge, a self-incremental learning framework (i.e., SELF) is proposed in this paper, which can quickly improve the generalization ability of the pre-trained model in disaster areas by self-training an incremental model using automatically selected samples from post-disaster images. The effectiveness of the proposed method is verified on the 2010 Yushu earthquake, 2023 Turkey earthquake, and other disaster types. The experimental results demonstrate that our approach outperforms state-of-the-art methods in terms of collapsed building identification, with an average increase of more than 6.4% in the Kappa coefficient. Furthermore, the entire process of the self-incremental learning method, including sample selection, incremental learning, and collapsed building identification, can be completed within 6 h after obtaining the post-disaster images. Therefore, the proposed method is effective for emergency response to natural disasters, which can quickly improve the application effect of the deep learning model to provide more accurate building damage results.



Citation: Ge, J.; Tang, H.; Ji, C.

Self-Incremental Learning for Rapid Identification of Collapsed Buildings Triggered by Natural Disasters.

Remote Sens. **2023**, *15*, 3909. <https://doi.org/10.3390/rs15153909>

Academic Editors: Raffaele Albano, Ivanka Pelivan and Reza Arghandeh

Received: 5 July 2023

Revised: 27 July 2023

Accepted: 4 August 2023

Published: 7 August 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: building damage; remote sensing; self-incremental learning; sample selection; disaster emergency response

1. Introduction

The frequent occurrence of extreme natural disasters seriously threatens the safety of human life. Timely access to the distribution information of collapsed buildings is crucial to emergency response and post-disaster rescue efforts [1]. Currently, remote sensing technology provides an efficient solution for the accurate and rapid extraction of building damage. As a result, post-disaster remote sensing images with high spatial resolution have become indispensable basic data for identifying disaster damage in numerous studies [2,3]. Among these, optical imagery stands out as a common and accessible source of remote sensing data [3], with a wide variety of sensors facilitating easy data acquisition. Some studies have also utilized radar equipment mounted on drones to scan post-disaster buildings [4], which remain unaffected by post-disaster weather conditions and can be combined with optical images for comprehensive analysis [5]. Moreover, LiDAR data proves useful in detecting height changes in buildings, enabling precise extraction of collapsed parts [2].

The vast diversity of buildings in different regions presents a significant challenge in accurately identifying buildings and assessing their damage using a pre-trained model [6]. Currently, deep learning technology, particularly convolutional neural networks, has

achieved state-of-the-art results in the task of building damage extraction [3]. Most of the research in this area focuses on proposing or improving a model for change detection to extract building damage information from paired bitemporal images of pre- and post-disaster [3,7–9]. However, in the context of emergency response scenarios, relying on pre-disaster imagery can significantly impact both the effectiveness and the efficiency of damage assessment. Therefore, an alternative approach worth considering is that the distribution maps of buildings extracted from pre-disaster images should be prepared before any disaster occurs.

The building distribution maps can filter out complex background categories and provide key information, including building location and shape, which is obviously helpful for building damage identification. Currently, there are only a few studies that exclusively utilize pre-disaster building distribution maps in combination with post-disaster imagery, despite the availability of building footprint or rooftop data that covers the vast majority of the world [10]. Notable examples include Open Street Map (<http://www.openstreetmap.org/>, accessed on 13 October 2022) and Bing maps of Microsoft (<https://github.com/microsoft/GlobalMLBuildingFootprints>, accessed on 21 February 2023) open-access data. Admittedly, there is a real problem that these building distribution data cannot guarantee a high update frequency at present, resulting in long time intervals between the availability of pre-disaster building distribution maps and post-disaster images, potentially spanning several years.

In addition, it is still difficult to accurately identify buildings from post-disaster images because the training data may come from different sensors or from different geographical regions [11]. Therefore, simply applying a pre-trained model to post-disaster scenarios can lead to a considerable drop in generalization performance and poor recognition results [9].

Transfer learning is a common solution to adapt the original model for better performance on the target domain. Data-based transfer learning has been shown to improve the model's application by utilizing target domain data [12,13]. Hu et al. [14] demonstrated that using post-disaster samples effectively enhances the identification accuracy of damaged buildings. On the other hand, incremental learning is a model-based transfer learning approach that improves generalization in specific scenarios by adding new base learners [15,16]. Ge et al. [6] confirmed that incremental learning significantly saves transfer time during emergency response, as it focuses on training only on new data containing post-disaster information. Therefore, learning sufficient post-disaster samples incrementally can effectively and rapidly improve the model's performance. In this process, the key technology lies in selecting high-quality post-disaster samples with the assistance of building distribution maps.

To enhance both the accuracy and efficiency of building damage extraction during disaster emergency response, we propose the Self-incremental Learning Framework (SELF). This framework utilizes post-disaster samples selected from optical remote sensing images to rapidly improve the identification accuracy of collapsed buildings. As illustrated in Figure 1, the preparation involves building distribution maps and a building recognition model before any disaster occurs. Subsequently, after the disaster event, essential samples of disaster imagery are automatically selected based on the knowledge of building distribution and predicted probability maps generated by the pre-trained model. The model's generalization ability is then swiftly improved through self-supervised training using these selected samples in an incremental learning manner. This process enables us to obtain reliable building damage results efficiently and effectively.

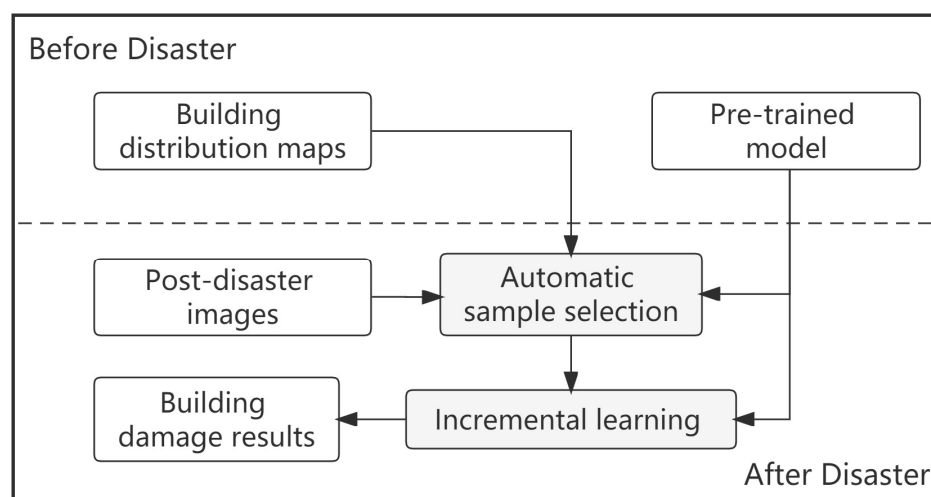


Figure 1. The SELF framework. The dotted line defines before and after a disaster.

This paper is organized as follows. The literature review related to this study is summarized in Section 2. The data and methods are introduced in Sections 3 and 4, respectively. The experimental results are shown in Section 5, and a discussion is conducted in Section 6. Some conclusions are drawn in Section 7.

2. Related Work

2.1. Building Damage Identification Methods

Most studies have focused on using the paired images, i.e., pre- and post-disaster images, to identify building damage [8,17]. Durnov [18] proposed a change detection method utilizing a Siamese structure that achieved top-ranking results in a competition focused on building damage identification. Subsequently, several similar change detection models were introduced [3,7], with a primary focus on optimizing the model structure. However, these methods necessitate the use of bitemporal images for damage detection, thus limiting the efficiency of disaster emergency response due to the reliance on pre-disaster images. Additionally, methods that combine multi-source images with various auxiliary data have been employed to extract high-precision disaster results [5]. For instance, Wang et al. [2] employed multiple types of data, including LiDAR and optical images, to extract the collapsed areas through the changes of the information of building height and corner points. However, the reality is that many types of specific data may be difficult to obtain in the short time after a disaster.

Methods that solely rely on post-disaster images aim to efficiently identify damaged buildings [19,20]. Based on the morphological and spectral characteristics of post-earthquake buildings, Ma et al. [21] proposed a method for depicting collapsed buildings using only post-disaster high-resolution images. Munsif et al. [22] achieved a lightweight CNN model, occupying just 3 MB, which can be deployed on Unmanned Aerial Vehicles (UAVs) with limited hardware resources, by utilizing several data augmentation techniques to enhance the efficiency and accuracy of multi-hazard damage identification. Nia et al. [23] introduced a deep model based on ground-level post-disaster images and demonstrated that using semantic segmentation results as the foreground positively impacted building damage assessment. Miura et al. [20] developed a collapsed building identification method using a CNN model and post-disaster aerial images, and achieved a damage distribution that was basically consistent with the inventories in earthquakes. However, existing studies have shown that the separability between collapsed buildings and the background is relatively low [24]. Solely using post-disaster information often falls short in accurately locating the damaged areas.

It is not easy to meet both the accuracy and efficiency requirements under emergency conditions by relying on bitemporal images or only post-disaster images. Therefore, com-

binning key pre-disaster knowledge (such as pre-disaster building distribution maps, a pre-trained model for buildings identification, and so on) with post-disaster images to identify damaged buildings quickly and accurately is a solution that is being developed in some studies [25]. For example, Galanis et al. [26] introduced the DamageMap model for wildfire disasters, which leverages pre-disaster building segmentation results and post-disaster aerial or satellite imagery for a classification task to determine whether buildings are damaged. At present, there are few studies that make full use of pre-disaster building distribution maps. Even though the building distribution data may not strictly correspond to each building in the post-disaster images, it can still provide much effective information about the location and shape of the buildings. Therefore, it is a promising way to devise methods to better apply the pre-disaster information in the future disaster response tasks.

2.2. Transfer Learning Methods

When a pre-trained model is directly applied to a target domain with significantly different features from the training data, there can be a considerable drop in accuracy. Transfer learning is used to address this practical problem. Current transfer learning methods can be categorized into the following three categories: (1) Data-based transfer learning [27] usually uses some samples of the target domain to enhance the model's performance in target applications. An example is the self-training method [28], which improves the model's generalization ability by automatically generating pseudo-labels. (2) Feature-based transfer learning [29,30] transforms the data of two domains into the same feature space, reducing the distance between the features of the source domain and the target domain, such as domain adversarial networks [31]. (3) Model-based transfer learning [32] usually adds new layers or integrates new base learners to optimize the original model, such as incremental learning [16].

Transfer learning in building damage extraction tasks aims to improve the performance of models in post-disaster scenes. However, these methods encounter challenges in practical applications, such as the scarcity of post-disaster samples, variations in image styles, and the unique features of buildings themselves. Hu et al. [14] conducted a comparison of three transfer learning methods for post-disaster building recognition and discovered that utilizing samples from disaster areas can significantly boost the recognition accuracy for various types of disasters. On the other hand, Lin et al. [33] proposed a novel method to filter historical data relevant to the target task from earthquake cases, aiming to improve the reliability of classification results.

In addition to transfer learning, data augmentation is often used to improve the generalization performance of the models [16,24], including applying various transformations to existing images, such as rotations, flips, or zooms, so that the model becomes more robust. Data synthesis is another valuable strategy that can address data scarcity by combining real data with computer simulations or generative models [34]. In fact, these methods can be combined with transfer learning to provide more precise and timely disaster information in emergency missions. Ge et al. [6] employed the generative network to transfer the style of remote sensing images under an incremental learning framework, and used data augmentation strategy to train the models, which improved the accuracy of building damage recognition.

2.3. Contributions of This Research

However, there is insufficient exploration on how to obtain and utilize important samples from post-disaster images efficiently and effectively. The aim of this paper is to fill this gap. The main contributions of this paper are twofold. (1) A knowledge-guided sample selection method is present, which uses a pre-trained model and pre-disaster building distribution maps to assist in sample selection from post-disaster images. (2) A self-incremental learning method is proposed by assembling self-training and incremental learning, which uses selected samples to realize the growth of the original model to quickly improve the accuracy of building damage extraction.

3. Data

3.1. Training Data: DREAM-B+

The DREAM-B+ dataset [6,16] is a large-scale building dataset comprising sampled remote sensing images and corresponding labels from over 100 cities worldwide. The dataset consists of 18,876 image tiles, each captured with RGB bands and having a high spatial resolution of either 0.5 m or 0.3 m. Each image tile has a size of 1024×1024 pixels. There are two categories in the ground-truth: building and background. The location of the images in this dataset is shown in Figure 2, and some examples are showcased in Figure 3.

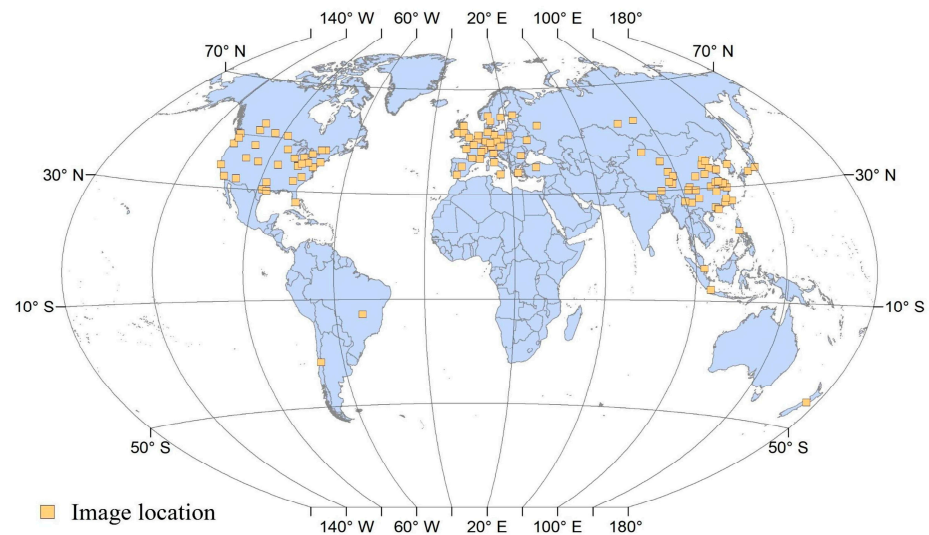


Figure 2. Geographical location of images in DREAM-B+ dataset, where each rectangle contains several sampled remote sensing images in this area.



Figure 3. Example images from the DREAM-B+ dataset.

The DREAM-B+ dataset is split into two sets for training and validation purposes. Specifically, 90% of the dataset is allocated for training a building recognition model, which serves as the prepared model in stage 1 before any disaster occurs. The remaining 10% of the dataset is used as a validation set to assess and validate the training process.

3.2. Test Data

The Yushu earthquake (Mw 6.9) occurred on April 14, 2010, with the epicenter very close to the urban area. This earthquake eventually resulted in about 14,700 deaths and many densely distributed houses were destroyed [35]. As shown in Figure 4a, the hardest-hit urban region of Yushu is used as an emergent disaster event to test both the effectiveness and efficiency of the proposed method. We obtained the post-disaster aerial images of this area with a resolution of 0.5 m. Due to the lack of available satellite images before the event, the building distribution map was visually interpreted from the pre-disaster images captured in 2004, and cross-validated by multiple domain experts in order to minimize the uncertainty of the map.

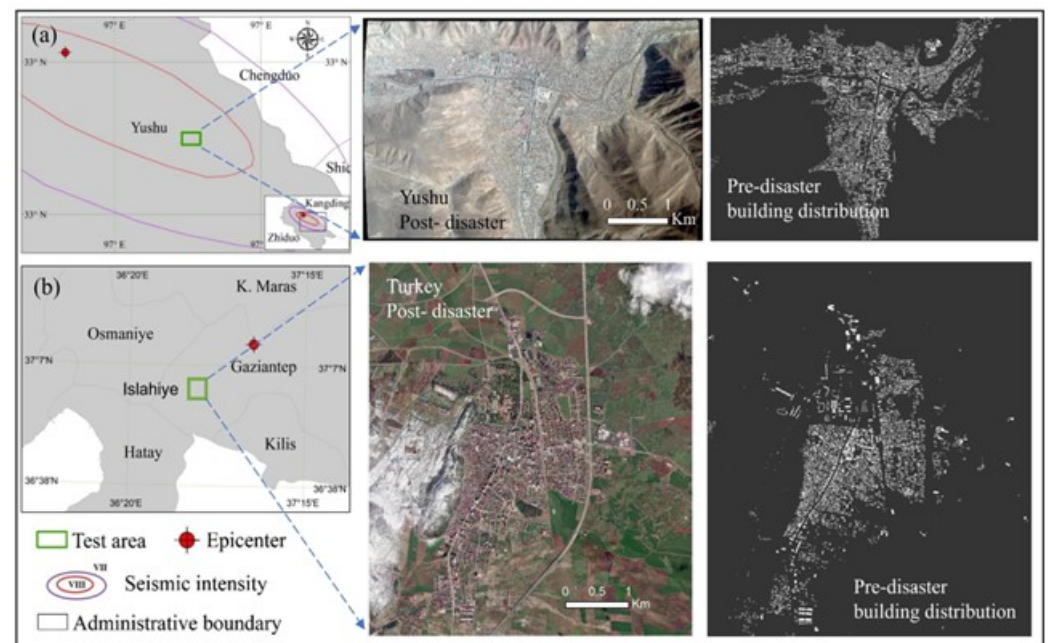


Figure 4. Location map and main data of the test areas. Yushu test area (a), and Turkey test area (b).

The Turkey earthquake (Mw 7.8) occurred on 6 February 2023, with the epicenter at 37.15°N, 36.95°E. This earthquake killed more than 40,000 people in Turkey and Syria. We obtained the post-disaster remote sensing images captured by Worldview-3, which have a spatial resolution of 0.3 m. As shown in Figure 4b, the Islahiye town serves as the second test area, which is close to the epicenter and has been severely affected. The pre-disaster building distribution map are from Microsoft's products [36], and the ground-truth of collapsed buildings are obtained by visual interpretation. The data details of Yushu and Turkey test areas are shown in Table 1.

Table 1. Details of the test data.

Cases	Data	Source	Bands	Acquisition Time	Resolution
Yushu	Post-disaster image	Aerial platform	RGB	April 2010	0.5 m
	Pre-disaster image	Quickbird		6 November 2004	0.6 m
	Pre-disaster building distribution map	Visual interpretation	/		0.5 m
Turkey	Post-disaster image	Worldview-3	RGB	February 2023	0.3 m
	Pre-disaster building distribution map	Microsoft	/	2023	

4. Methodology

4.1. Overview

The purpose of the proposed self-incremental learning framework (SELF) aims to rapidly enhance the identification accuracy of a pre-trained model by selecting and utilizing new samples from post-disaster images. The specific application process of the framework is shown in Figure 5. First, we need to prepare both the building distribution map and a pre-trained model (i.e., stage 1 model) for building identification. When the post-disaster images are available, the stage 1 model is then used to produce the probability maps of buildings on the post-disaster images. To improve the accuracy of identifying post-disaster buildings, the framework employs the knowledge-guided sample selection (K-SS) method to select new samples from the post-disaster images. The new samples then incrementally learned a new model, i.e., the stage 2 model, through an end-to-end gradient boosting algorithm (i.e., EGB-A). The stage 2 model is specifically designed to identify buildings from post-disaster images. Finally, pixel-level collapsed buildings are identified by comparing both pre- and post-disaster building maps.

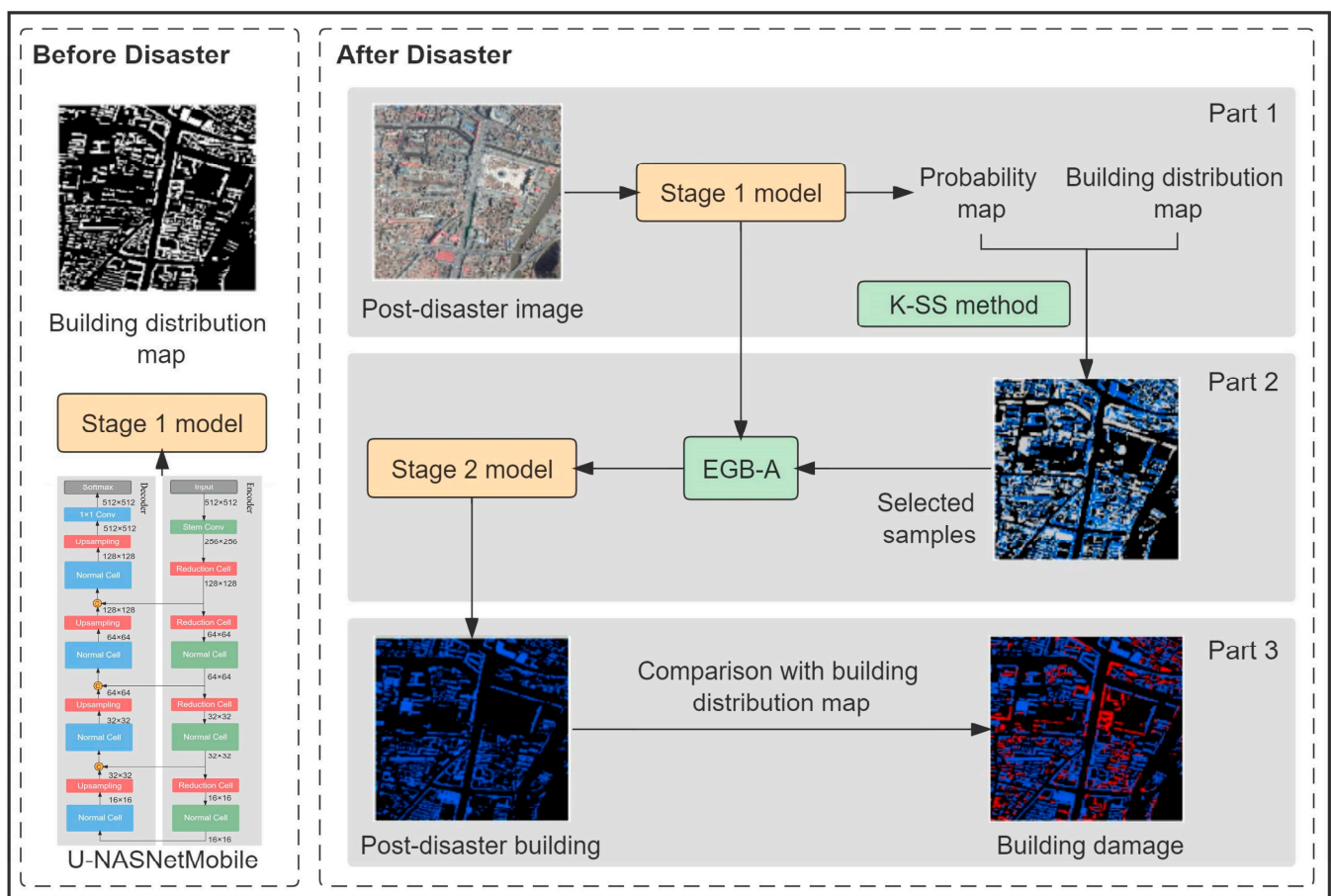


Figure 5. Collapsed building identification under the SELF framework.

4.2. Knowledge-Guided Sample Selection Method

As presented in Table 2, the pixels in both the pre-disaster building distribution maps and the post-disaster images are classified into two categories: building (positive class) and background (negative class). The post-disaster category “building” consists of buildings that have not collapsed, other buildings that consist of new buildings that only appear in post-disaster images, and some buildings that were missed in the building distribution maps. The post-disaster category “background” refers to pixels of both collapsed buildings and the pre-disaster background.

Table 2. Categories of pre- and post-disaster data.

Data	Category	Class Label	Detailed Category
Pre-disaster	Building Background	Positive Negative	/
Post-disaster	Building	Positive	Not collapsed building Other building
	Background	Negative	Collapsed building Original background

The location and shape information of each building provided by the pre-disaster building distribution maps should be fully utilized in the process of sample selection. The first core idea of the K-SS sample selection method is to analyze each building object individually. In existing studies, the entropy-based sample selection methods often screen an entire image or region, such as selecting the top 10% of the image with the highest probability value as positive samples. We believe that conducting a detailed analysis for each individual building can better consider the capabilities of the model for buildings with various features. In addition, buildings and their nearby background pixels are relatively critical samples, because the pixels near the classification boundary are often easily confused by the model, such as the edge of buildings and their junction with the background. Therefore, another idea of the K-SS method is to use the contrast between the probability values of the building and its surrounding area as the basis for selecting samples. Specifically, the probability values within a certain range, including buildings, are counted, and threshold segmentation is performed to maximize the variance between classes.

The complete K-SS sample selection method is shown in Algorithm 1. Please note that Figure 6 might be helpful for understanding the algorithm in a more intuitive way. One of the important steps is to double the minimum enclosing rectangle of each building object and use the Otsu algorithm [37] to perform threshold segmentation on the probability map in the enlarged area. The Otsu algorithm has the advantages of fast calculation speed and is not affected by image contrast. The principle of Otsu is to maximize the variance between classes and automatically generate the best segmentation threshold:

$$T = Otsu(\{P\}) \quad (1)$$

where $\{P\}$ is the set of image pixel values in the region to be segmented. The *Otsu* algorithm binarizes $\{P\}$ and returns a threshold T . Pixels whose value are greater than T are classified as foreground (i.e., not collapsed buildings and other buildings); otherwise, they are background (i.e., collapsed buildings and original background). In addition, three specific modules are designed for sample selection of categories: not collapsed buildings, collapsed buildings, and other buildings, respectively.

Building selection module is utilized to select samples of the category of not collapsed buildings and their surrounding background. The rationale behind designing this module is that the probability value of the building predicted by the stage 1 model is generally higher than the surrounding background. The Otsu algorithm can roughly distinguish the foreground and the background. Combining the threshold segmentation results and the pre-disaster building information, pixels with high confidence are selected as positive samples—that is, take the intersection of the post-disaster threshold segmentation results and pre-disaster building distribution maps in the same category. To minimize the inclusion of erroneous samples, other pixels are ignored because it is difficult to determine their actual classes.

Collapsed building selection module is designed to select samples of collapsed buildings and their surrounding background. In general, the features of building ruins are close to those of the background, so the probability value of the collapsed buildings predicted by the model is close to that of their surrounding areas. If there is no obvious contrast between the foreground and background probability values, the ratio of the two categories after

threshold segmentation is likely to be unbalanced. For not collapsed buildings, the areas of the foreground and background after segmentation should be relatively similar, because we doubled the minimum enclosing rectangle of each building. The very unbalanced area ratio of the two classes gives us greater confidence that the building has collapsed. Here, we assume that if the area ratio of two categories is more than four times, it is very unbalanced. Similarly, some samples of collapsed buildings are selected in combination with the pre-disaster building distribution maps, and other pixels are ignored.

There may be missing buildings in the pre-disaster building distribution maps or newly built buildings. Background screening module serves the purpose of filtering out possible buildings in the background area of the building distribution maps. First, find the average value P_b of the post-disaster probability corresponding to the pixels of the building category before the disaster is calculated. Then, ignore the pixels whose probability value is greater than P_b in the background category before the disaster, and the remaining areas have a great confidence that they belong to the background category in all the pre- and post-disaster data.

As depicted in Figure 6, after the screening by using the three modules, the final effective samples are labeled as positive or negative samples, and other pixels are ignored or invalid.

Algorithm 1 The K-SS method for post-disaster sample selection.

Definition:

The minimum enclosing rectangle of N

building objects : $r_1, r_2, \dots, r_N$.

Expand the area of r_n to get : $R_1, R_2, \dots, R_N$. ($Area_{R_n} = 2Area_{r_n}$).

In probability map : the value of pixels $\{P_i\}$, and the average value of the pixels corresponding to the building category pre-disaster : P_b

Sample selection:

for $n = 1$ to N do:

$T = Otsu(P_i \in R_n)$

As for $P_i \in R_n$:

Building selection:

(1) if $P_i > T$ and corresponds to the pre-disaster building category:
Not collapsed.

(2) if $P_i < T$ and corresponds to the pre-disaster background category:
Background.

(3) other regions: Ignored.

Collapsed building selection:

when $\frac{\text{Number of } P_i > T}{\text{Number of } P_i < T} < \frac{1}{4}$ or $\frac{\text{Number of } P_i > T}{\text{Number of } P_i < T} > 4$:

(1) if $P_i < T$: Collapsed.

(2) other regions: Ignored.

end for.

Background screening:

As for the regions except $R_1 \cup R_2 \cup \dots \cup R_N$:

(1) if $P_i < P_b$: Background.

(2) other regions: Ignored (may be buildings).

Output: Positive samples: Not collapsed.

Negative samples: Collapsed and background.

Invalid samples: Ignored.

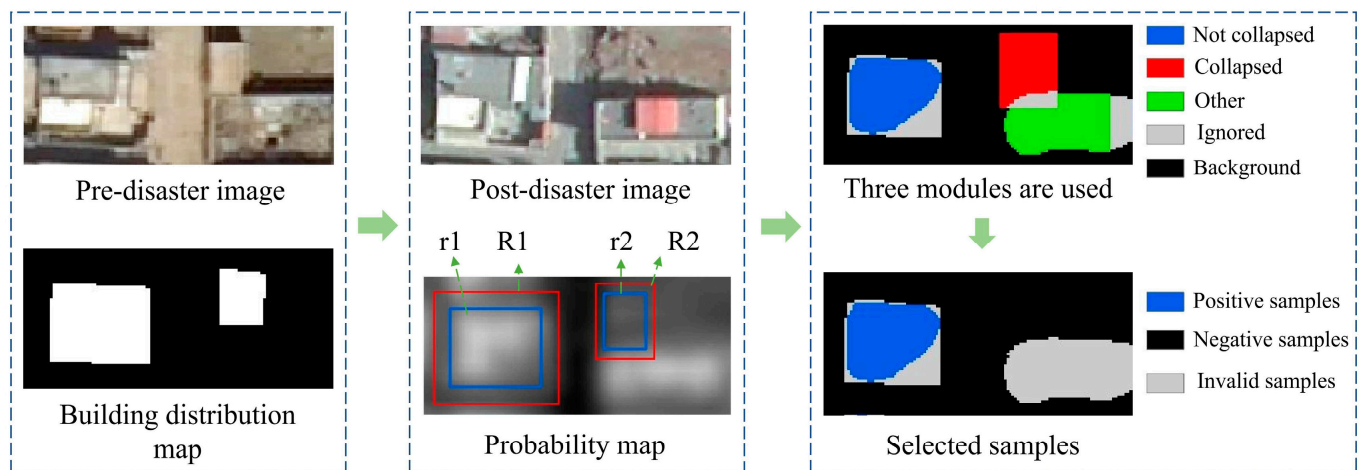


Figure 6. Schematic of the K-SS method. Buildings that existed in the pre-disaster images but not in the post-disaster images were considered as collapsed buildings, and vice versa were considered other buildings. The blue and red boxes are the minimum enclosing rectangles and enlarged rectangular areas of pre-disaster building objects, respectively.

4.3. Incremental Learning Using the EGB-A

The end-to-end gradient boosting (EGB) algorithm achieves incremental learning by integrating multiple base learners together [16]. The new base learner is trained on the newly collected data based on all existing base learners. This method has certain advantages in the disaster emergency process because it can utilize post-disaster data to urgently train a base learner and incorporate it into the original model in order to achieve rapid transfer learning for specific applications. The EGB-A method [6] is an improved version of the EGB for the building damage classification task, which alleviates the knowledge forgetting problem and optimizes the ability of adaptive learning. The training algorithm of the EGB-A method is shown in Algorithm 2. For additional application details of the method, we recommend referring to the papers of Ge et al. [6] and Yang and Tang [16].

Algorithm 2 Training algorithm of EGB-A [4].

Input: Training data, $X = \{x_0, \dots, x_M\}$, and labels, $Y = \{y_0, \dots, y_M\}$; base learner, $f(x; \theta)$; learning rate of base learner, v ; and softmax function, σ .

1: $F_0(x_0) = \sigma(f_0(x_0; \theta_0))$

2: $f_0(x_0) = \argmin_{\theta_0} L(y_0, F_0(x_0))$

3: for $m = 1$ to M do

4: $F_m(x_m) = \sigma(f_m(x_m) + \sum_{i=0}^{m-1} v_i \cdot f_i(x_m))$

5: $f_m(x_m; \theta_m) = \argmin_{\theta_m, v_0, \dots, v_{m-1}} L(y_m, F_m(x_m))$

6: end for

Output: $F_M(x) = \sigma(\sum_{i=0}^M v_i \cdot f_i(x; \theta))$, $v_M = 1$

In the SELF framework, the EGB-A method is used to incrementally train a new base learner based on the existing stage 1 model, utilizing the selected post-disaster samples. This process results in an ensemble model with two base learners in stage 2. One of the significant advantages of this approach is that the training process does not need to reuse pre-disaster datasets (e.g., DREAM-B+), which can save valuable time during emergency response.

The network architecture of each base learner in the SELF framework is based on U-NASNetMobile [16], as illustrated in Figure 7. It combines the neural architecture search structure in NASNet [38] and the upsampling module in the classic U-Net model [39] to perform semantic segmentation tasks. The U-NASNetMobile is suitable for ensemble models and disaster scenarios due to its small number of parameters and fast training speed [16].

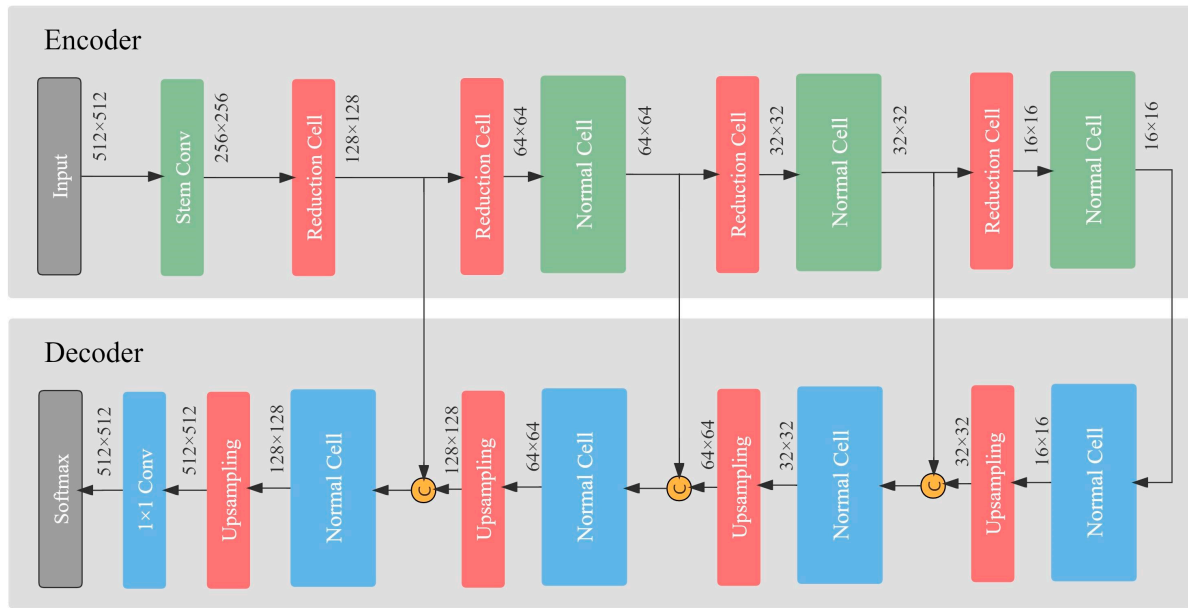


Figure 7. The structure of U-NASNetMobile [16].

4.4. Experimental Settings and Evaluation Metrics

The Adam optimizer [40] and the cosine learning rate annealing schedule [41] are employed to update the weights of the model. The default batch size is 4, and the maximum learning rate is 3×10^{-4} . During the training process, the parameters of the latter base learner are initialized with the parameters of the previous base learner to speed up the convergence. In addition, traditional data augmentation methods are also applied to prevent overfitting, including brightness variation, flipping, and random rotation. The experiments were run on the hardware device of NVIDIA Tesla K80 GPU.

For the identification of post-disaster buildings, the IoU metric [42] of the building category is employed to evaluate the accuracy. The F1 score, recall, precision, and OA (overall accuracy) are employed as reference evaluation metrics:

$$\text{IoU} = \frac{\text{Prediction} \cap \text{GroundTruth}}{\text{Prediction} \cup \text{GroundTruth}} \quad (2)$$

$$\text{F}_1 \text{ score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{OA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (6)$$

where TP, FP, TN, and FN are the pixel numbers of true positive, false positive, true negative, and false negative, respectively.

For the building damage extraction result, the Kappa metric [43] is employed to represent the evaluation accuracy. In addition, the OA, PA (producer accuracy of collapsed buildings), and UA (user accuracy of collapsed buildings) are provided for reference:

$$\text{Kappa} = \frac{p_0 - p_e}{1 - p_e} \quad (7)$$

$$PA = \frac{a_1 \cap b_1}{a_1} \quad (8)$$

$$UA = \frac{a_1 \cap b_1}{b_1} \quad (9)$$

$$p_e = \frac{a_1 \times b_1 + a_2 \times b_2 + a_3 \times b_3 + a_4 \times b_4}{n^2} \quad (10)$$

where p_0 is equal to the value of OA, that is, the number of pixels correctly classified divided by the total number of pixels; a_1 , a_2 , a_3 and a_4 are the real pixels of collapsed buildings, not collapsed buildings, other buildings, and the background, respectively; b_1 , b_2 , b_3 and b_4 are the predicted pixels of collapsed buildings, not collapsed buildings, other buildings, and background, respectively; and n is the total number of pixels.

5. Experimental Results

On the one hand, we quantitatively evaluate the results of our method in Section 5.1. On the other hand, the difference in building damage extraction is qualitatively highlighted in Section 5.2. Furthermore, failure examples are analyzed in Section 5.3.

5.1. Quantitative Evaluation

5.1.1. Post-Disaster Building Recognition

Table 3 presents the post-disaster building recognition accuracy on the test data using the stage 1 and stage 2 model, respectively. The results show that utilizing the selected post-disaster samples leads to a significant improvement in the IoU value after incremental learning. In the Yushu case, the IoU increased by 14%, and in the Turkey case, it increased by 7.23%. This improvement can be primarily attributed to the substantial increase in the recall metric of building recognition, although the precision might be slightly sacrificed. The optimized model can identify more buildings that have not collapsed after the disaster, which is the premise to ensure a better effect of building damage extraction.

Table 3. Post-disaster building recognition accuracy in test cases predicted by the pre-trained model (stage 1 model) and the incrementally learned model (stage 2 model) using selected samples.

Cases	Stages	IoU	F1 Score	Recall	Precision	OA
Yushu	Stage 1	0.4286	0.6001	0.4953	0.7609	0.9603
	Stage 2	0.5686	0.7249	0.7108	0.7397	0.9676
Turkey	Stage 1	0.4998	0.6665	0.5971	0.7542	0.9788
	Stage 2	0.5721	0.7278	0.6705	0.7957	0.9822

5.1.2. Building Damage Extraction

Table 4 presents the building damage extraction accuracy on the test cases of stage 1 and stage 2. The Kappa coefficient and OA represent the comprehensive situation of the four categories of collapsed buildings, not collapsed buildings, other buildings, and background, and PA and UA specifically measure the accuracy of collapsed buildings. Due to the significant improvement in the accuracy of post-disaster buildings in stage 2, the damage results of Yushu and Turkey cases become more reliable, with Kappa values reaching 0.8267 and 0.7688, respectively. The UA metric shows that the accuracy of collapsed building identification has increased significantly in the Yushu case. The UA value of the Turkey case is relatively low because the not collapsed buildings identified by the model have incomplete edges, resulting in some extra collapsed building pixels. In addition, the number of collapsed buildings in the Turkey case was less compared to the Yushu case. Therefore, the influence of these misclassified pixels on the UA value will be more obvious.

Table 4. Building damage extraction accuracy in test cases predicted by the pre-trained model (stage 1 model) and the incrementally learned model (stage 2 model) using selected samples.

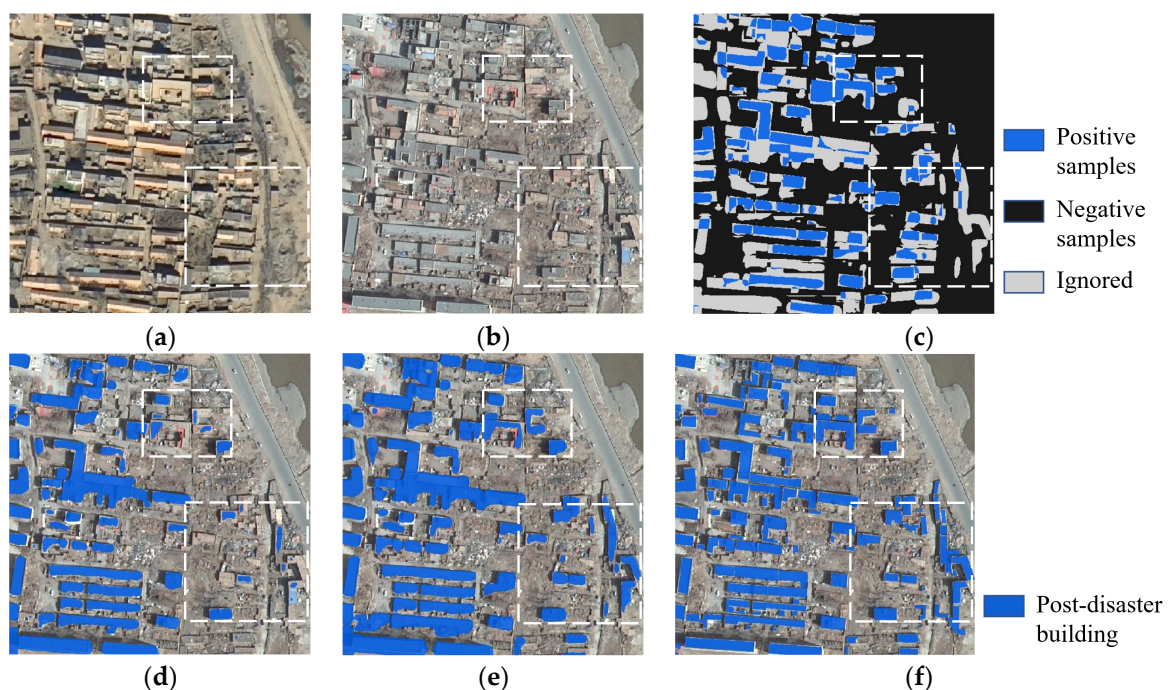
Cases	Stages	Kappa	OA	PA	UA
Yushu	stage 1	0.7379	0.9303	0.9126	0.6931
	stage 2	0.8267	0.9676	0.9021	0.7998
Turkey	stage 1	0.7281	0.9788	0.8656	0.2741
	stage 2	0.7688	0.9880	0.8387	0.2852

Overall, the proposed building damage extraction method is feasible. The K-SS method provides key post-disaster samples, which can effectively improve the performance of the model in disaster areas to obtain high-precision results.

5.2. Qualitative Analysis

As shown in Figure 8, we have visualized the results of post-disaster building and damage recognition in partial areas of Yushu to intuitively analyze the significance of incremental learning using post-disaster samples. It is evident that the pre-trained model was directly applied to post-disaster images, that some buildings were missed, and that the effect of damage extraction needs to be improved furthermore. The proposed sample selection method effectively identifies buildings that were not recognized in stage 1. The model learned the features of these samples in an incremental manner of EGB-A, which can keep the buildings that have been correctly identified in stage 1 as much as possible, and continue to optimize the recognition results. In the white boxes, we can see that the results in stage 2 have been significantly improved, and most of the missed buildings were identified.

Figure 9 displays the results of post-disaster building and damage identification at different stages in some areas of the Turkey case. Comparing the results, it is evident that the improved stage 2 model using post-disaster samples can more completely identify building edges and small buildings with white roofs. As a result, the model, after self-incremental learning, can predict a more accurate distribution of post-disaster buildings. Referring to the ground-truth, the stage 2 model using the selected samples has achieved more reliable building damage results.

**Figure 8.** Cont.

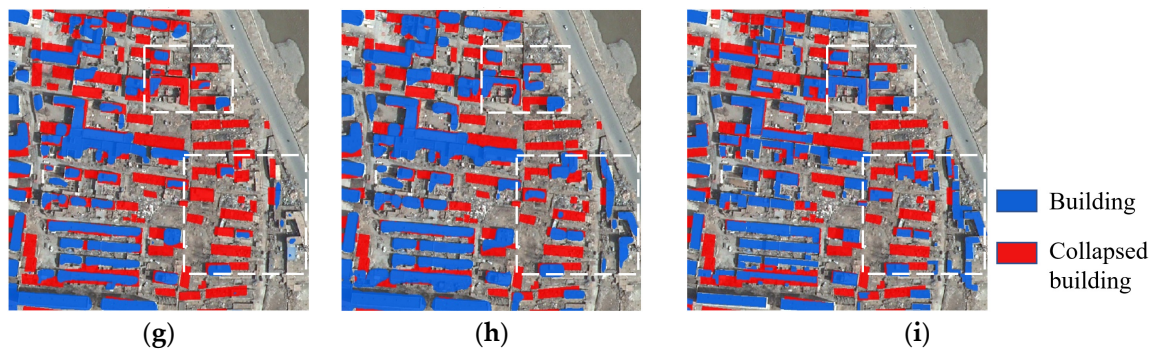


Figure 8. Comparison of post-disaster building identification (second row) and damage results (third row) at different stages in the Yushu test area. Results before optimization (stage 1), results after optimization (stage 2). The white boxes indicate noteworthy details. (a) Pre-disaster image; (b) post-disaster image; (c) selected samples; (d) stage 1; (e) stage 2; (f) ground-truth; (g) stage 1; (h) stage 2; (i) ground-truth.

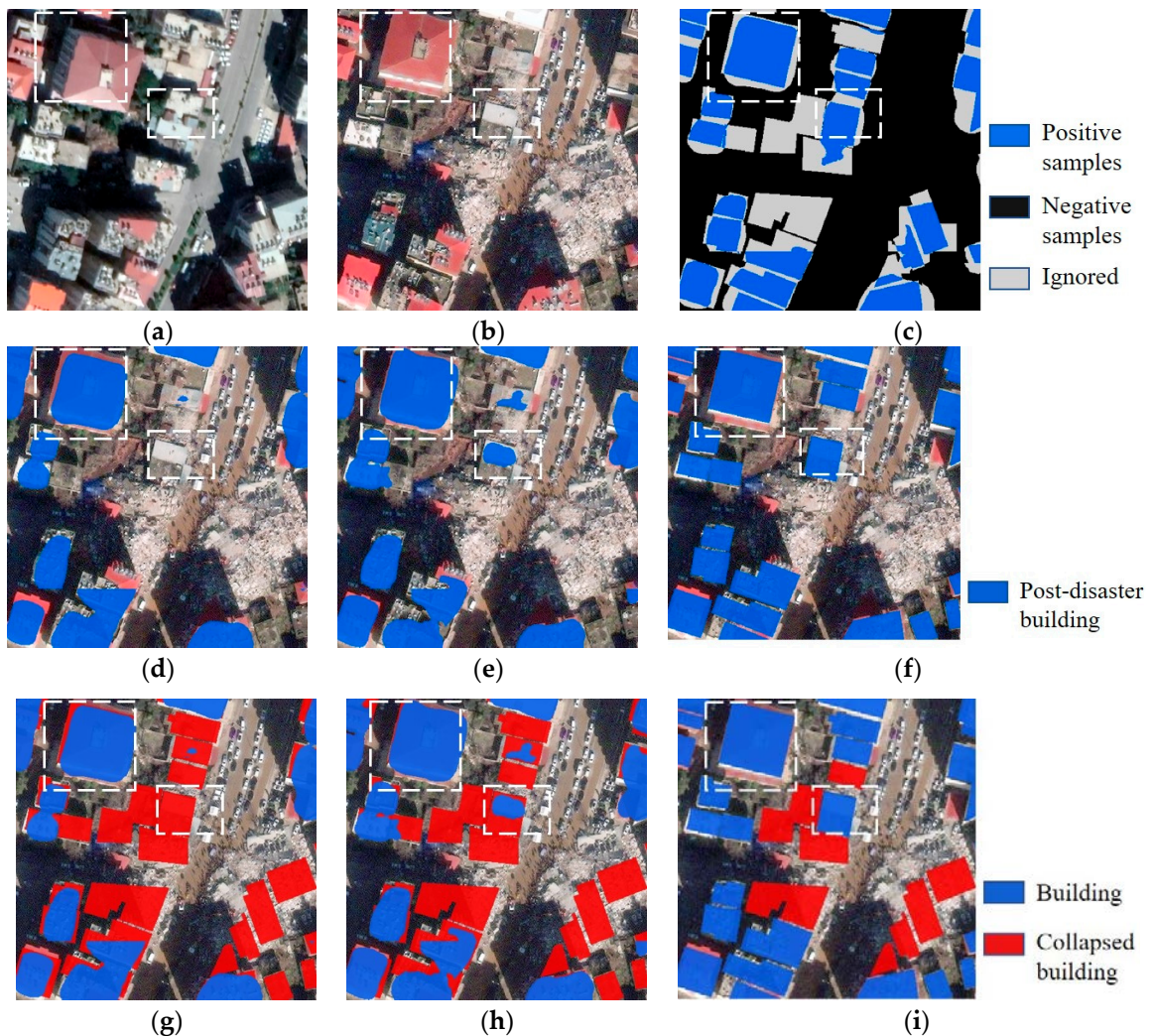


Figure 9. Comparison of post-disaster building identification (second row) and damage results (third row) at different stages in the Turkey test area. Results before optimization (stage 1), results after optimization (stage 2). The white boxes indicate noteworthy details. (a) Pre-disaster image; (b) post-disaster image; (c) selected samples; (d) stage 1; (e) stage 2; (f) ground-truth; (g) stage 1; (h) stage 2; (i) ground-truth.

The final damage extraction results (stage 2) of the entire Yushu test area are shown in Figure 10. We can see that the distribution of collapsed buildings identified by the model is similar to the ground-truth. From the map, it is evident that the buildings in the southwest of the urban area, specifically subfigure (1) in Figure 10, have sustained severe damage, with extensive areas of ruins. In contrast, the buildings near the center exhibit less concentrated collapse and seem to have experienced relatively less damage. At a subtle level, the results of building damage extraction have more red parts, indicating that the model has identified some undamaged buildings as collapsed buildings. Overall, the proposed method obtains results that are basically consistent with the real situation at the macro level.

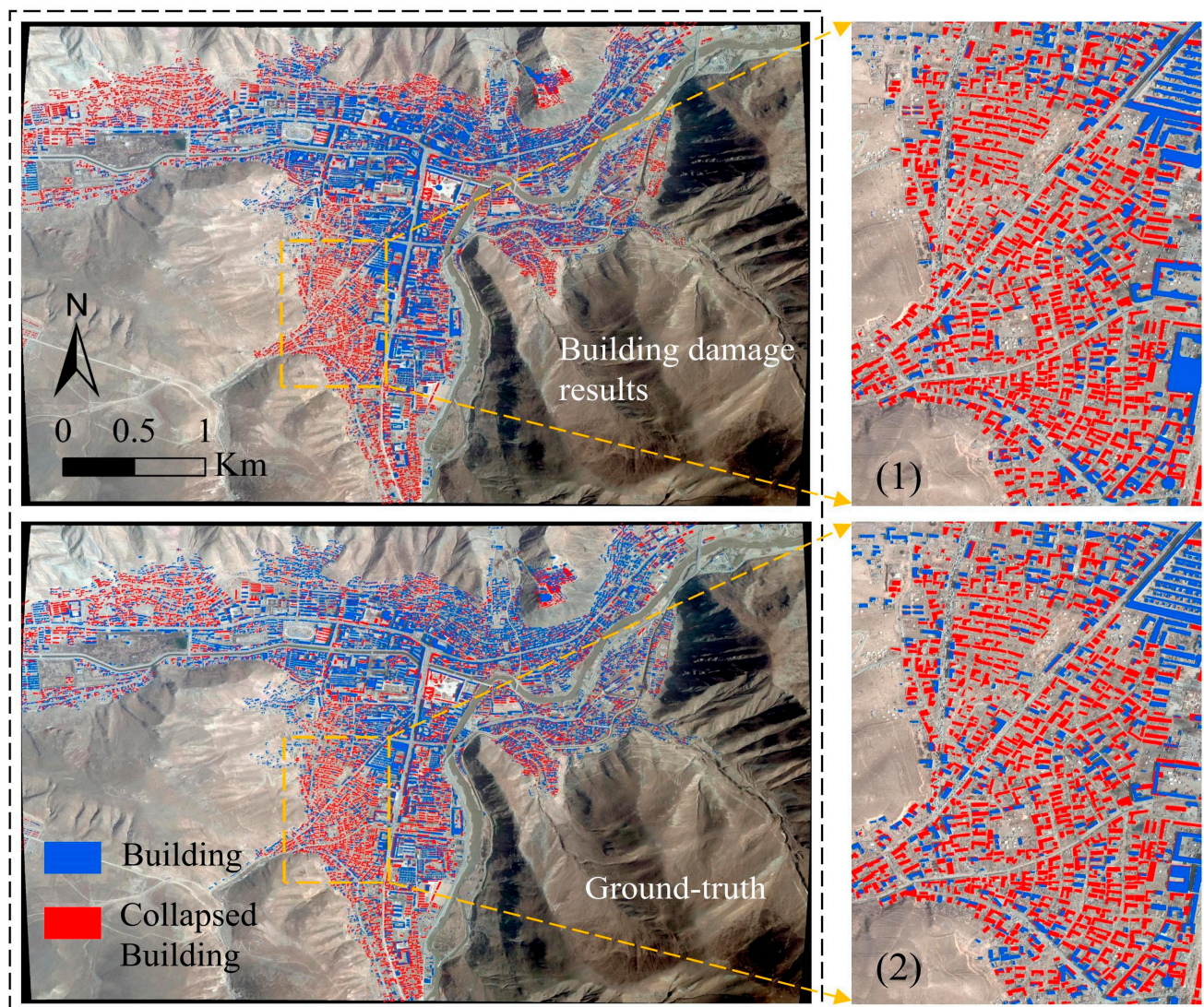


Figure 10. Building damage results of the Yushu case extracted by the SELF method (**top**) and the corresponding ground-truth (**bottom**). To highlight collapsed areas, we combined the categories of not collapsed buildings and other buildings in blue. The orange boxes show a severely damaged area. (1) The region enlarged from the building damage results; (2) The region enlarged from the ground-truth.

The final damage extraction results (stage 2) of the entire Turkey test area are shown in Figure 11. On the whole, there are not many collapsed buildings, and the building damage results show more collapsed pixels than the ground-truth. There is an area of concentrated damage in the middle of the town, which is the enlarged subfigure (1). It can be seen that

some edge pixels of intact buildings are misclassified as collapsed because there is still space for improvement in the recall value of post-disaster building identification. The collapsed buildings can basically be completely extracted.

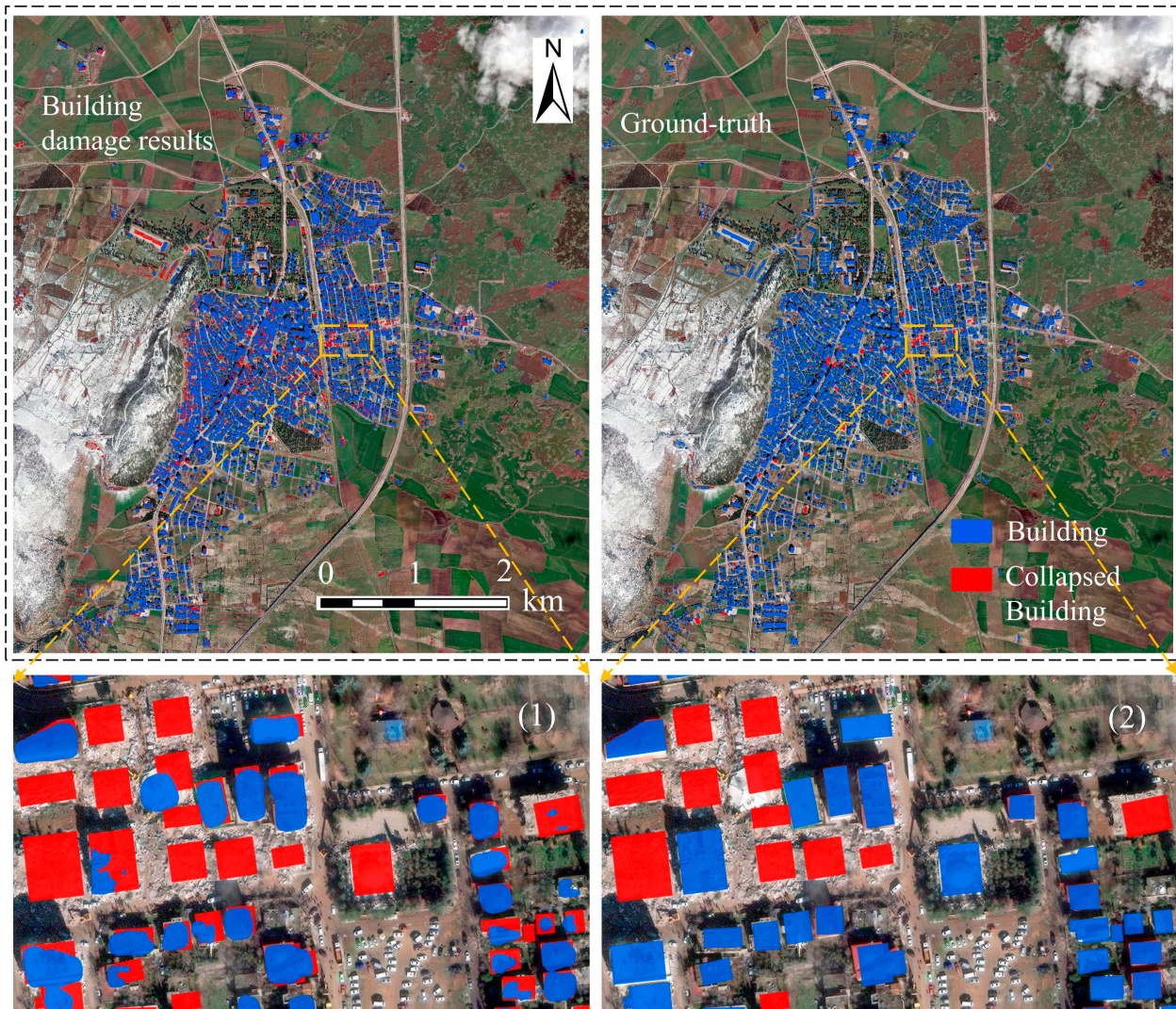


Figure 11. Building damage results of the Turkey case extracted by the SELF method (**left**) and the corresponding ground-truth (**right**). The orange boxes show a severely damaged area. (1) The region enlarged from the building damage results; (2) The region enlarged from the ground-truth.

5.3. Failure Example Analysis

Despite the improvements achieved by the stage 2 model using the SELF method, there are still some recognition errors. As shown in the first row of Figure 12, although the post-disaster samples are correctly selected, some buildings are still missed in the recognition results. This is related to the lack of buildings with similar features in the training set, and smaller buildings are generally harder to identify. The second row shows a case where the K-SS sample selection method fails. The reason is that although the building is damaged, there are still some building roof features that lead to a higher activation value in this area. The recognition results in this example are not affected by the wrong samples, indicating that impure sample does not necessarily reduce the recognition effect, and we need to avoid overfitting when using these samples for training. In addition, it is usually difficult for the model to identify roofs covered by shadows of high-rise buildings (the third line in Figure 12). To address this problem, it may be feasible to design data enhancement strategies or use generative networks to remove shadows.

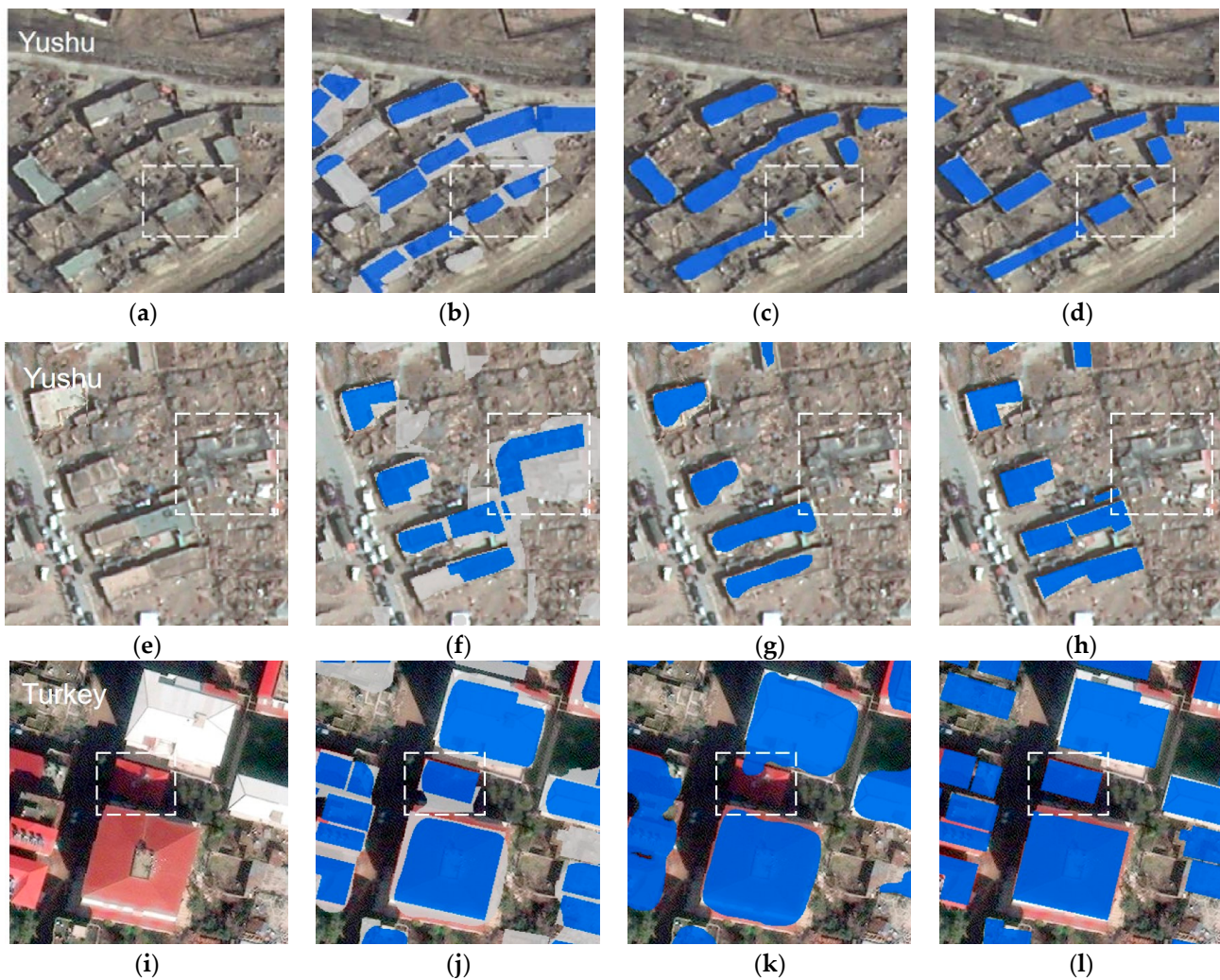


Figure 12. Failure examples in post-disaster building identification (stage 2) or sample selection step. The first line shows that the sample selection was correct but the building was not recognized. The second line shows the failure cases of sample selection. The third line shows the missed detection of buildings due to shadows. Blue: buildings or building samples. Gray: ignored samples. The white boxes indicate noteworthy areas. (a) Post-disaster image; (b) selected samples; (c) recognition results; (d) ground-truth; (e) post-disaster image; (f) selected samples; (g) recognition results; (h) ground-truth; (i) post-disaster image; (j) selected samples; (k) recognition results; (l) ground-truth.

6. Discussion

This section begins by comparing the accuracy of the proposed method with several state-of-the-art building damage extraction methods. Subsequently, using the Yushu case as an example, the impact of different sample selection methods is discussed, and the timeliness of the proposed approach is analyzed. Finally, the performance of the proposed method in multiple disaster types other than earthquakes is evaluated and analyzed.

6.1. Comparison of Building Damage Extraction Methods

The proposed method is compared with the Incre-Trans method [6] since both approaches aim to enhance building damage recognition through transfer learning within an incremental framework. The Incre-Trans method transfers the style of historical disaster images to the current post-disaster style and applies incremental learning using the transferred data on the pre-trained stage 1 model to obtain an improved stage 2 model. In contrast, our SELF method utilizes pre-disaster knowledge and post-disaster samples instead of image styles to improve the accuracy and efficiency of emergency response.

In addition to the Incre-Trans method, our proposed approach was compared with two existing deep learning change detection methods, BDANet [8] and ChangeOS [3], in the Yushu case. Both BDANet and ChangeOS require paired pre- and post-disaster images for their operation. However, in the Turkey case, we did not have access to pre-disaster sub-meter images during the emergency situation. This also reflects the limitations of damage detection methods that rely on pre-disaster imagery.

As shown in Table 5, the Kappa coefficient of our method in the Yushu case reaches 0.8267, which is higher than that of the Incre-Trans method, indicating that using post-disaster samples for self-training can improve the accuracy of emergency recognition more effectively than transferring the style of images. Most metrics of the SELF method are significantly higher than that of the two other change detection methods because the generalization performance of the existing model is poor when it is directly applied to the non-preset location, so it is necessary to improve the results according to the characteristics of disaster areas. In the Turkey case, our proposed SELF method achieves a slightly higher Kappa coefficient compared to the Incre-Trans method. Overall, the proposed SELF method can indeed deliver more reliable building damage results.

Table 5. Building damage extraction accuracy in test areas predicted by the pre-trained model (stage 1 model) and the incrementally learned model (stage 2 model) using selected samples.

Cases	Methods	Kappa	OA	PA	UA
Yushu	SELF	0.8267	0.9676	0.9021	0.7998
	Incre-Trans	0.7521	0.9508	0.8468	0.7617
	BDANet	0.5819	0.9365	0.4752	0.4350
	ChangeOS	0.4672	0.8926	0.3487	0.4785
Turkey	SELF	0.7688	0.9880	0.8387	0.2852
	Incre-Trans	0.7582	0.9814	0.8694	0.2829

Different types of methods have their advantages and limitations. As shown in Table 6, both the SELF and Incre-Trans are based on the post-classification comparison framework, which allows them to use single-temporal building datasets to train models. In contrast, BDANet and ChangeOS require paired pre- and post-disaster images, which may limit their accuracy due to dataset availability. As shown in Figure 13, the direct application of the BDANet and ChangeOS models does not perform well in the Yushu scene. Specifically, BDANet misidentifies some intact buildings as the collapsed category. The results of ChangeOS show a large area of adhesion, mainly because the object-level change detection approach used in ChangeOS tends to group multiple densely distributed buildings as a single object. As a result, the individual collapsed areas within the group cannot be effectively detected. The advantage of the two change detection models lies in their optimization of the network structure, which can enhance the accuracy of building damage identification. However, they lack specific strategies to quickly improve the recognition effect in emergency response scenarios, such as utilizing incremental learning. The SELF method has the potential to serve as an alternative to Incre-Trans, as the selected samples already contain post-disaster style information. Hence, there is no need to transfer the style of historical disaster images.

Table 6. Characteristics of different methods.

Methods	Type	Incremental Learning	Required Images	Result Level
SELF	Post-classification comparison	Yes	Post-disaster	Pixel level
Incre-Trans				
BDANet	Change detection	No	Pre- and post-disaster	Object level
ChangeOS				

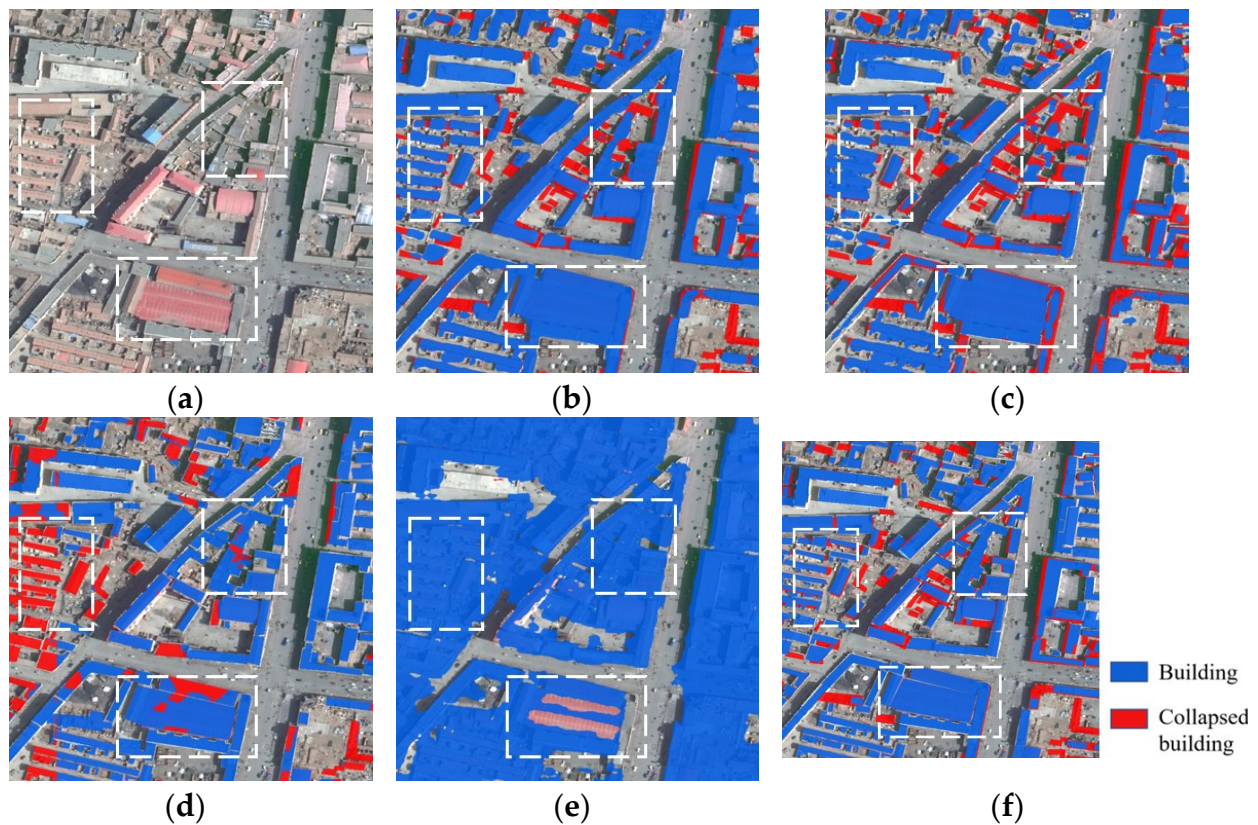


Figure 13. Comparison of building damage results of the Yushu case extracted by different methods. The white boxes indicate noteworthy details: (a) post-disaster image; (b) SELF; (c) Incre-Trans; (d) BDANet; (e) ChangeOS; (f) ground-truth.

6.2. Other Sample Selection Methods

In the entropy-based sample selection methods, the top $n\%$ pixels with the highest confidence are usually selected as samples for self-supervised training. Furthermore, Hu et al. [14] employed the top 10% of the most certain pixels in probability maps predicted by a building identification model as the post-disaster samples for transfer learning. Our method utilizes the building distribution maps to guide the selection of samples. In order to make an objective comparison, we also added the building distribution maps on the basis of the method of Hu et al. [14], and realized the following sample selection approach as a benchmark: for the probability maps predicted by the stage 1 model, the top $n\%$ of the post-disaster pixels corresponding to the building category pre-disaster are selected as positive samples, the top $n\%$ of the post-disaster pixels corresponding to the background category pre-disaster are selected as negative samples, and the remaining pixels are ignored.

The proportion of selected samples has an important impact on the effect of transfer learning [44]. Therefore, we conducted experiments when n was equal to 50, 70, 90 and 99, respectively. For example, if $n = 50$, the first 50% of the pixels with the highest confidence are effective samples, and the remaining pixels are invalid. If $n = 99$, most of the pixels are indiscriminately employed as effective samples, that is, the pixels of the building category pre-disaster are almost all selected as positive samples of post-disaster, and the same is true for the background category. We denote the above sample selection approaches as Top 50%, Top 70%, Top 90%, and Top 99%, respectively, and compare them with the proposed K-SS method in two aspects: (1) qualitative comparison and analysis of the selected samples, and (2) under the same parameter settings, we compare the application effect of the stage 2 models optimized after incremental learning using the selected samples, that is, the accuracy of post-disaster building recognition and building damage extraction.

6.2.1. Different Numbers of Selected Samples

In order to obtain as many post-disaster samples as possible, we should reserve pixels with high confidence. At the same time, the pixels with low confidence should not be labeled as wrong categories. Figure 14 shows the samples obtained from different proportions and the methods presented in this paper. It can be seen that if a small proportion of high-confidence samples (Top 50% and Top 70%) are selected, many pixels in the Yushu urban area are invalid, and many buildings and their nearby background are missed. As the proportion of effective samples increases, more background pixels are selected. At the same time, the cases of collapsed buildings being wrongly selected as positive samples are also increasing. If the pre-disaster building distribution data are applied to post-disaster images almost without filtering (Top 99%), many collapsed buildings are selected as positive samples and other buildings are negative samples, which leads to a lot of incorrect information.

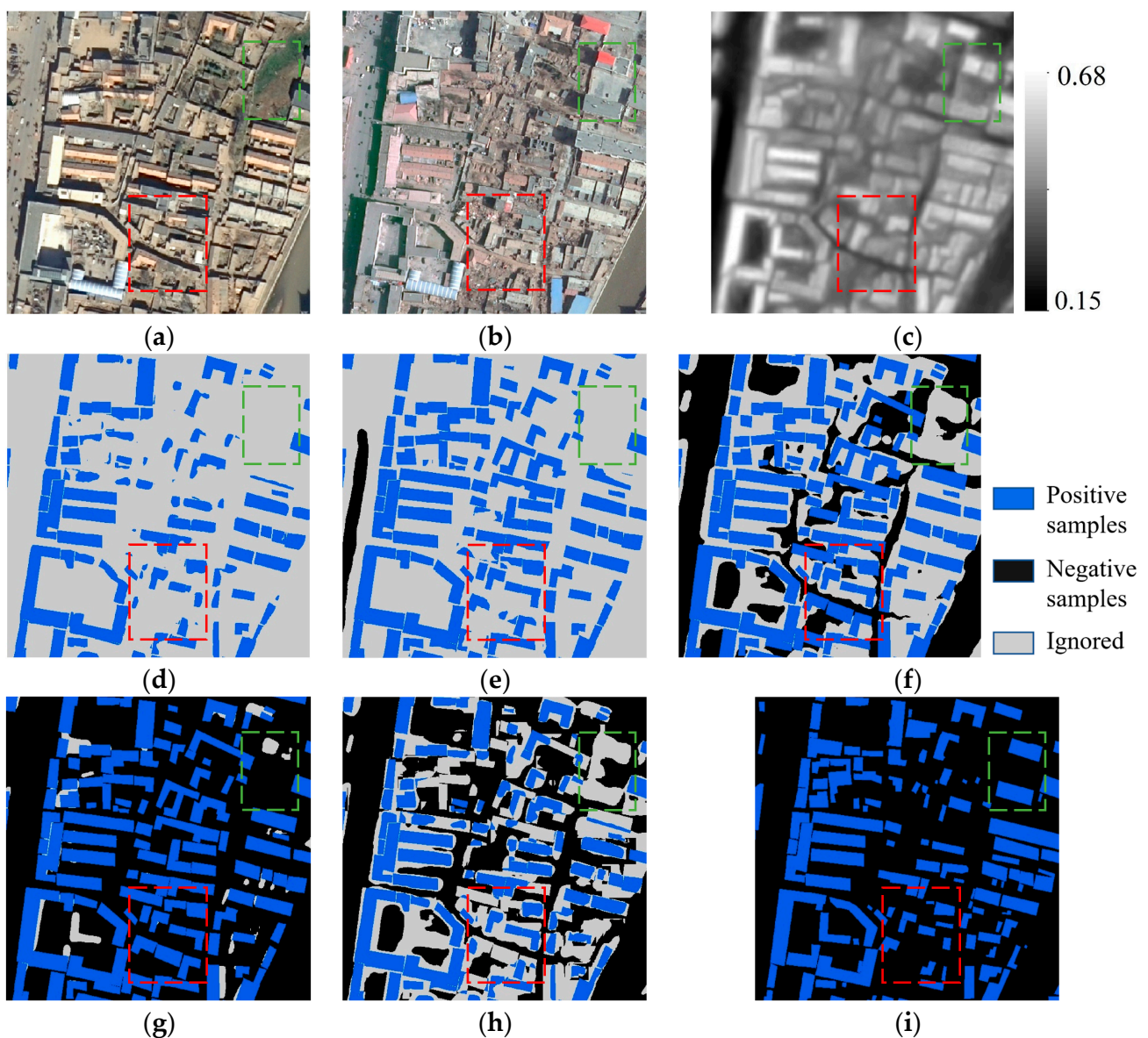


Figure 14. Comparison of the selected post-disaster samples in the Yushu case. (a) Pre-disaster image; (b) post-disaster image; (c) probability map of post-disaster buildings predicted in stage 1; (d) Top 50%; (e) Top 70%; (f) Top 90%; (g) Top 99%; (h) ours; (i) ground-truth of post-disaster buildings. The red box indicates collapsed buildings in the area, and the green box indicates other buildings in the area.

Our method ignores these hard-to-judge buildings and collapsed pixels, which is acceptable. By utilizing the contrast of probability values, the K-SS method can select negative samples around building objects. These samples close to the classification boundary are beneficial for the model to learn effective features and prevent overfitting. The characteristics of the proposed method is that it fully combines the pre-disaster knowledge and the recognition ability of the model to design specific methods for different situations, which can provide accurate, diverse, and critical post-disaster samples.

6.2.2. Overfitting

The number of iterations required for model training is related to the number of samples. We evaluated the IoU of post-disaster building recognition when the above methods were trained to several different epochs using the selected samples in the Yushu case, as shown in Figure 15. The training process was performed based on stage 1 through the EGB-A incremental framework. It can be seen that if a small number of samples are selected (Top 50% and Top 70%), it only takes a few epochs to stop iterations, and too many training epochs are prone to over-fitting and lead to a decrease in accuracy. Using more samples can indeed lead to higher accuracy after sufficient training. The IoU value of our method is always at a higher position.

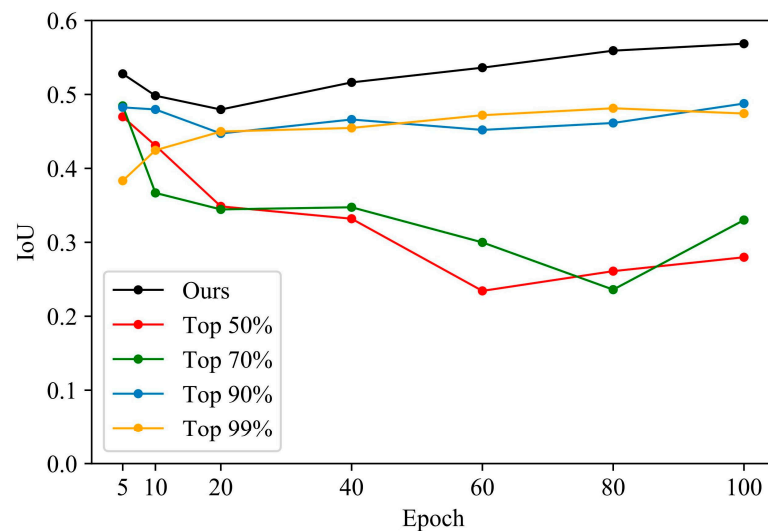


Figure 15. The post-disaster building recognition accuracy in the Yushu case when the different numbers of samples are trained to different epochs.

The detailed comparison of post-disaster building recognition accuracy among the methods with the highest IoU is presented in Table 7. Our method stands out by achieving the most reliable results, as indicated by the IoU and F1 metrics. When selecting 70% of the reliable samples, the method achieves the highest recall value, primarily due to the misidentification of some collapsed buildings as not collapsed, leading to a relatively lower precision. In contrast, using 99% of the reliable samples achieves the lowest recall, which may be due to the confusion of wrong samples that makes it more difficult to correctly identify post-disaster buildings.

Table 7. Comparison of post-disaster building recognition accuracy in the Yushu case.

Methods	Epoch	IoU	F1 Score	Recall	Precision	OA
Top 50%	5	0.4698	0.6393	0.7853	0.5391	0.9467
Top 70%	5	0.4845	0.6528	0.8104	0.5465	0.9482
Top 90%	100	0.4877	0.6556	0.7807	0.5651	0.9507
Top 99%	80	0.4813	0.6499	0.5339	0.8301	0.9654
Ours	100	0.5686	0.7249	0.7108	0.7397	0.9676

The damage is obtained by comparing with the pre-disaster building distribution maps, and the evaluation results are shown in Table 8. Similarly to the post-disaster recognition results, the lowest producer accuracy was achieved when selecting 70% reliable samples. This is because there are fewer negative samples around the buildings, making the post-disaster recognition results more prone to crossing the classification boundary and misclassifying background pixels around the buildings. Conversely, lower recall for post-disaster buildings leads to lower user precision for collapsed buildings. Overall, our method achieved more accurate damage extraction results and reached a Kappa coefficient of 0.8267, which depends on the sample selection methods we designed at the object level.

Table 8. Comparison of building damage accuracy in the Yushu case.

Methods	Epoch	Kappa	OA	PA	UA
Top 50%	5	0.7431	0.9467	0.7923	0.8299
Top 70%	5	0.7509	0.9482	0.7860	0.8420
Top 90%	100	0.7585	0.9507	0.8044	0.8236
Top 99%	80	0.8051	0.9654	0.9548	0.7152
Ours	100	0.8267	0.9676	0.9021	0.7998

6.3. Timeliness Analysis

The efficiency of building damage extraction is also crucial for emergency response. We evaluate the time required by the proposed SELF framework for the complete pipeline in the Yushu case, as shown in Figure 16. The timeliness estimation starts from the moment when the post-disaster images are obtained. First, the stage 1 model is used to predict the probability maps, and then the K-SS method is employed to select post-disaster samples, which takes about half an hour in total. Subsequently, the samples will be used for incremental learning to obtain an optimized stage 2 model. This process takes 5 h for 100 epochs of training. Finally, the stage 2 model is applied to complete the final building damage extraction. Overall, the proposed method can provide optimized damage results within 6 h, which is better than the timeliness of the method proposed by Ge et al. [6], which takes about 8 h to complete a similar emergency task.

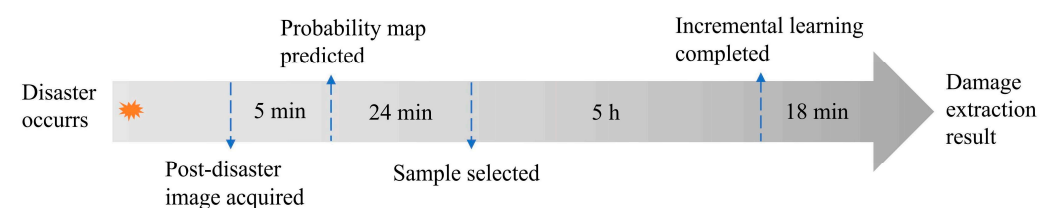


Figure 16. Timeliness estimation of the SELF framework for emergency response.

This efficiency evaluation result was calculated under the hardware conditions of the NVIDIA Tesla K80 GPU with 12G video memory and the software environment of the TensorFlow-GPU version 1.12.0. Additionally, when using the selected samples to train the stage 2 model, the parameters of the stage 1 model are used for initialization to speed up the convergence. The timeliness of this method can meet the needs of the emergency period (24 h after the earthquake) with the main goal of rescuing the buried people [45]. In actual rescue missions, better hardware conditions are expected to further reduce time consumption.

6.4. Performance in Other Natural Disasters

In order to evaluate the effect of the SELF method in a broader range of disaster scenarios besides earthquakes, this section selects the wildfire, tornado, and tsunami disasters from the xBD dataset [46] for verification. The post-disaster images of the three disaster cases have sub-meter spatial resolutions and RGB bands, and their details are shown in Table 9.

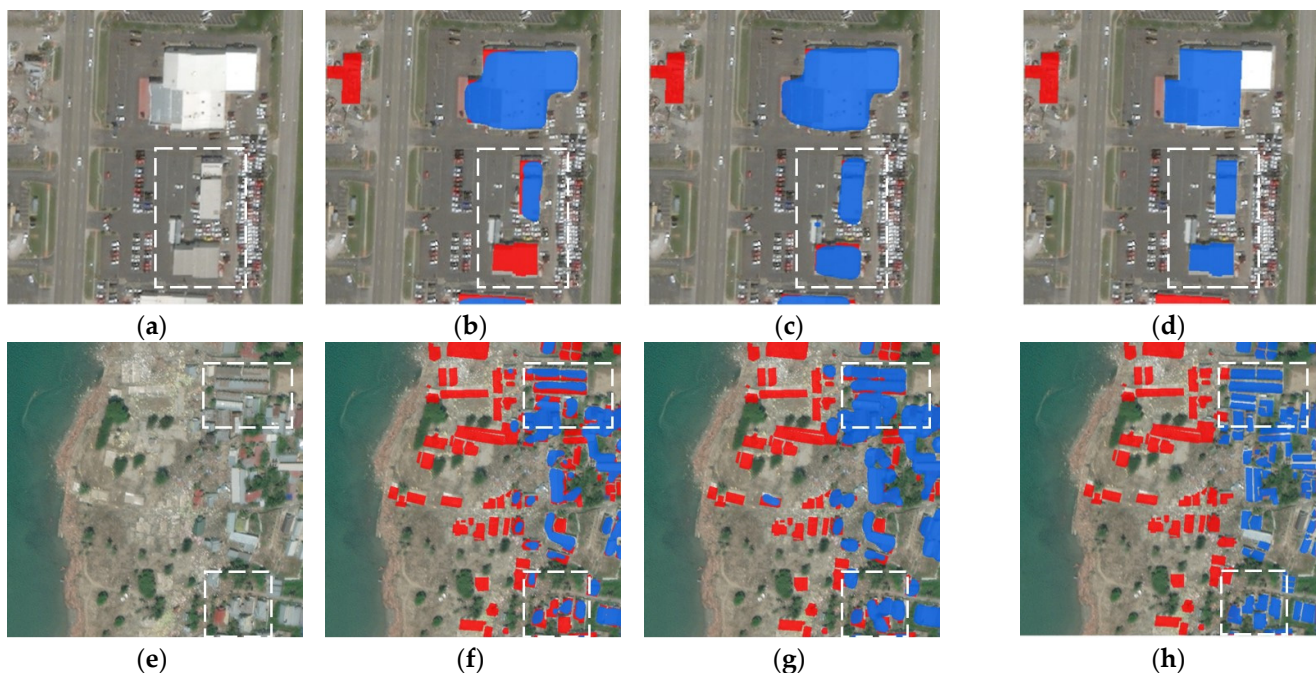
Table 9. Details of the disaster cases.

Cases	Event Date	Country
Joplin, MO Tornado	22 May 2011	America
Santa Rosa Wildfires	8–31 October 2017	America
Palu Tsunami	18 September 2018	Indonesia

The experimental results are shown in Table 10. It can be seen that after the self-incremental learning, the Kappa coefficient of building damage identification for different disaster types has been improved to a certain extent in stage 2. Among them, the Kappa value increased the most in the tornado disaster, reaching 4.78%, while it only increased by 1.98% in the wildfire disaster. It is worth noting that the common feature of these three cases is that the PA values in stage 2 have decreased to varying degrees, while other metrics have increased. Combined with the visualization results in Figure 17, this phenomenon can be explained by the self-incremental learning of the post-disaster samples significantly increasing the pixels of the intact building, which may misclassify some collapsed buildings. In the case of the Palu tsunami, due to the dense distribution of buildings, the recognition results of stage 2 are somewhat adhesions. Overall, the proposed SELF method can effectively improve the emergency identification results of building damage in multiple hazards.

Table 10. Building damage extraction accuracy of the disaster cases predicted by the pre-trained model (stage 1 model) and the incrementally learned model (stage 2 model) using selected samples.

Cases	Stages	Kappa	OA	PA	UA
Joplin, MO Tornado	Stage 1	0.7444	0.9559	0.8794	0.3915
	Stage 2	0.7922	0.9618	0.8290	0.5587
Santa Rosa Wildfires	Stage 1	0.7419	0.9706	0.8969	0.4008
	Stage 2	0.7617	0.9707	0.8910	0.5224
Palu Tsunami	Stage 1	0.6530	0.9151	0.8395	0.2325
	Stage 2	0.6936	0.9177	0.7969	0.3634

**Figure 17.** Cont.

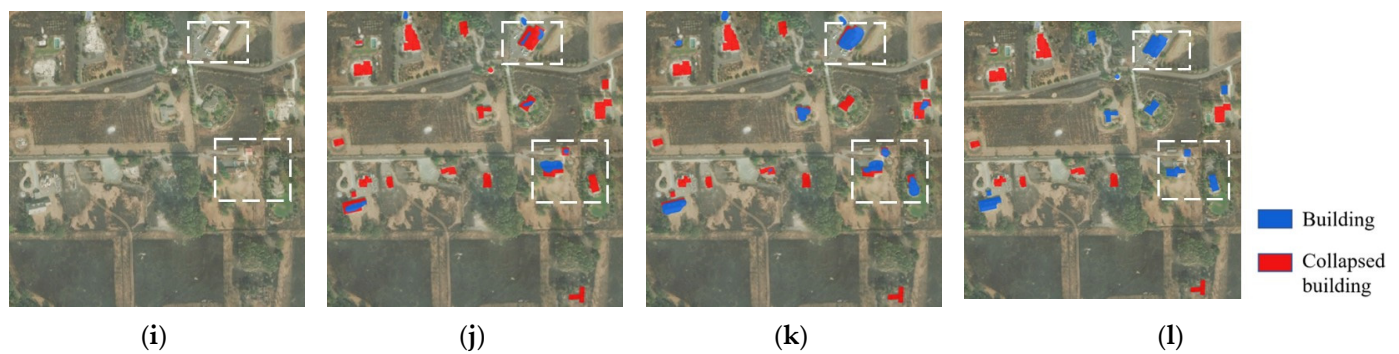


Figure 17. Comparison of building damage identification results before optimization (stage 1) and after optimization (stage 2). The first to third rows are the Joplin MO Tornado, Palu Tsunami, and Santa Rosa Wildfires, respectively. The white boxes indicate noteworthy details. (a) Post-disaster image; (b) stage 1; (c) stage 2; (d) ground-truth; (e) post-disaster image; (f) stage 1; (g) stage 2; (h) ground-truth; (i) post-disaster image; (j) stage 1; (k) stage 2; (l) ground-truth.

7. Conclusions

This paper proposes a novel solution to address the challenges of limited recognition accuracy in building damage extraction due to the restricted generalization capability of pre-trained models and the difficulty in obtaining a large number of labeled disaster area samples in a short period after disasters. The main contributions of this paper are as follows, enabling rapid enhancement of collapsed building extraction for disaster emergency response:

- (1) The proposed SELF framework can rapidly enhance the building recognition ability of the pre-trained model through self-training by using automatically selected post-disaster samples. The experimental results on the Yushu earthquake and Turkey earthquake show that the Kappa accuracy of the building damage extracted by the optimized model is increased by 6.48% on average compared with the initial stage. In terms of efficiency, the framework can complete the entire process within 6 h and provide a more reliable building damage distribution map.
- (2) The K-SS sample selection method can automatically select high-quality post-disaster image samples with the assistance of pre-disaster building distribution map. The designed sample selection modules are based on the probability maps and the Otsu segmentation method, which realizes the targeted screening of collapsed buildings, not collapsed buildings, and other buildings. Compared with other similar sample selection methods, using the samples provided by K-SS can achieve a more significant improvement in accuracy.
- (3) The experimental results demonstrate that leveraging the difference in activation values between buildings and their surrounding backgrounds is an effective strategy for selecting key samples for self-training. The building location and shape information provided by the pre-disaster building distribution maps can realize more accurate judgment of the sample category from the object level.

The method presented in this paper does have certain limitations. Firstly, the effectiveness of the selected samples relies on the quality and availability of building distribution data, while, currently, high-quality building footprint or roof outline products are still missing in some regions. Secondly, the SELF framework has been verified on earthquakes and three other natural disasters, while the application to man-made hazards remains to be explored.

Author Contributions: Conceptualization, H.T.; funding acquisition, H.T.; investigation, H.T. and J.G.; methodology, J.G. and C.J.; software, J.G.; supervision, H.T. and C.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by National Natural Science Foundation of China Major Program (42192580, 42192584).

Data Availability Statement: Publicly available xBD dataset used in this study can be found here: <https://xview2.org/dataset>, (accessed on 5 July 2023).

Acknowledgments: This research is carried out under programming in Python and deep learning framework Tensorflow.

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

References

1. Motosaka, M.; Mitsuiji, K. Building damage during the 2011 off the Pacific coast of Tohoku Earthquake. *Soils Found.* **2012**, *52*, 929–944. [CrossRef]
2. Wang, X.; Li, P. Extraction of urban building damage using spectral, height and corner information from VHR satellite images and airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 322–336. [CrossRef]
3. Zheng, Z.; Zhong, Y.; Wang, J.; Ma, A.; Zhang, L. Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters. *Remote Sens. Environ.* **2021**, *265*, 112636. [CrossRef]
4. Dong, L.; Shan, J. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS J. Photogramm. Remote Sens.* **2013**, *84*, 85–99. [CrossRef]
5. Brunner, D.; Lemoine, G.; Bruzzone, L. Earthquake damage assessment of buildings using vhr optical and sar imagery. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2403–2420. [CrossRef]
6. Ge, J.; Tang, H.; Yang, N.; Hu, Y. Rapid identification of damaged buildings using incremental learning with transferred data from historical natural disaster cases. *ISPRS J. Photogramm. Remote Sens.* **2022**, *195*, 105–128. [CrossRef]
7. Bai, Y.; Hu, J.; Su, J.; Liu, X.; Liu, H.; He, X.; Meng, S.; Mas, E.; Koshimura, S. Pyramid pooling module-based semi-siamese network: A benchmark model for assessing building damage from xBD satellite imagery datasets. *Remote Sens.* **2020**, *12*, 4055. [CrossRef]
8. Shen, Y.; Zhu, S.; Yang, T.; Chen, C.; Pan, D.; Chen, J.; Xiao, L.; Du, Q.F. BDANet: Multiscale convolutional neural network with cross-directional attention for building damage assessment from satellite images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–14. [CrossRef]
9. Yang, W.; Zhang, X.; Luo, P. Transferability of convolutional neural network models for identifying damaged buildings due to earthquake. *Remote Sens.* **2021**, *13*, 504. [CrossRef]
10. Hu, Y.; Liu, C.; Li, Z.; Xu, J.; Han, Z.; Guo, J. Few-Shot Building Footprint Shape Classification with Relation Network. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 311. [CrossRef]
11. Wang, C. Investigation and analysis of building structure damage in Yushu Earthquake. *Build. Struct.* **2010**, *40*, 106–109. [CrossRef]
12. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Adv. Neural Inf. Process. Syst.* **2014**, *3*, 2672–2680. [CrossRef]
13. Park, T.; Efros, A.; Zhang, R.; Zhu, J. Contrastive learning for unpaired image-to-image translation. In Proceedings of the Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020. Proceedings, Part IX 16. [CrossRef]
14. Hu, Y.; Tang, H. On the generalization ability of a global model for rapid building mapping from heterogeneous satellite images of multiple natural disaster scenarios. *Remote Sens.* **2021**, *13*, 984. [CrossRef]
15. Ring, M.B. Continual Learning in Reinforcement Environments. Ph.D. Thesis, University of Texas at Austin, Austin, TX, USA, 1994. Available online: <https://www.researchgate.net/publication/2600799> (accessed on 10 April 2023).
16. Yang, N.; Tang, H. GeoBoost: An incremental deep learning approach toward global mapping of buildings from VHR remote sensing images. *Remote Sens.* **2020**, *12*, 1794. [CrossRef]
17. Weber, E.; Kan, H. Building disaster damage assessment in satellite imagery with multi-temporal fusion. *arXiv* **2020**, arXiv:2004.05525. [CrossRef]
18. Durnov, V. xview2 First Place Framework. 2020. Available online: https://github.com/DIUX-xView/xView2_first_place (accessed on 10 April 2023).
19. Li, X.; Yang, W.; Ao, T.; Li, H.; Chen, W. An improved approach of information extraction for earthquake-damaged buildings using high-resolution imagery. *J. Earthq. Tsunami* **2011**, *5*, 389–399. [CrossRef]
20. Miura, H.; Aridome, T.; Matsuoka, M. Deep learning-based identification of collapsed, non-collapsed and blue tarp-covered buildings from post-disaster aerial images. *Remote Sens.* **2020**, *12*, 1924. [CrossRef]
21. Ma, J.; Qin, S. Automatic depicting algorithm of earthquake collapsed buildings with airborne high resolution image. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; IEEE: Munich, Germany, 2012; pp. 939–942. [CrossRef]

22. Munsif, M.; Afridi, H.; Ullah, M.; Khan, S.D.; Cheikh, F.A.; Sajjad, M. A lightweight convolution neural network for automatic disasters recognition. In Proceedings of the 2022 10th European Workshop on Visual Information Processing (EUVIP), Lisbon, Portugal, 11–14 September 2022. [\[CrossRef\]](#)
23. Nia, K.R.; Mori, G. Building damage assessment using deep learning and ground-level image data. In Proceedings of the 2017 14th Conference on Computer and Robot Vision (CRV), Edmonton, AB, Canada, 16–19 May 2017. [\[CrossRef\]](#)
24. Qing, Y.; Ming, D.; Wen, Q.; Weng, Q.; Xu, L.; Chen, Y.; Zhang, Y.; Zeng, B. Operational earthquake-induced building damage assessment using CNN-based direct remote sensing change detection on superpixel level. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102899. [\[CrossRef\]](#)
25. Tilon, S.; Nex, F.; Kerle, N.; Vosselman, G. Post-disaster building damage detection from earth observation imagery using unsupervised and transferable anomaly detecting generative adversarial networks. *Remote Sens.* **2020**, *12*, 4193. [\[CrossRef\]](#)
26. Galanis, M.; Rao, K.; Yao, X.; Tsai, Y.; Ventura, J. Damagemap: A post-wildfire damaged buildings classifier. *Int. J. Disaster Risk Reduct.* **2021**, *65*, 102540. [\[CrossRef\]](#)
27. Liu, X.; Liu, Z.; Wang, G.; Zhang, H. Ensemble transfer learning algorithm. *IEEE Access* **2017**, *6*, 2389–2396. [\[CrossRef\]](#)
28. Gu, X.; Zhang, C.; Shen, Q.; Han, J.; Plamen, P.A.; Peter, M.A. A self-training hierarchical prototype-based ensemble framework for remote sensing scene classification. *Inf. Fusion* **2022**, *80*, 179–204. [\[CrossRef\]](#)
29. Hoffman, J.; Tzeng, E.; Park, T.; Zhu, J. CyCADA: Cycle-consistent adversarial domain adaptation. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018. [\[CrossRef\]](#)
30. Tasar, O.; Giros, A.; Tarabalka, Y.; Alliez, P.; Clerc, S. Dagnet: Unsupervised, multisource, multitarget, and life-long domain adaptation for semantic segmentation of satellite images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1067–1081. [\[CrossRef\]](#)
31. Na, J.; Jung, H.; Chang, H.; Hwang, W. Fixbi: Bridging domain spaces for unsupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021. [\[CrossRef\]](#)
32. Zhao, Z.; Chen, Y.; Liu, J.; Shen, Z.; Liu, M. Cross-people mobile-phone based activity recognition. In Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, Barcelona, Spain, 16–22 July 2011. [\[CrossRef\]](#)
33. Lin, Q.; Ci, T.; Wang, L.; Mondal, S.; Yin, H.; Wang, Y. Transfer learning for improving seismic building damage assessment. *Remote Sens.* **2022**, *14*, 201. [\[CrossRef\]](#)
34. Antoniou, A.; Storkey, A.; Edwards, H. Data augmentation generative adversarial networks. *arXiv* **2017**, arXiv:1711.04340. [\[CrossRef\]](#)
35. Wang, J.; Xu, C.; Shen, W. The Coseismic Coulomb Stress Changes Induced by the 2010 Mw 6.9 Yushu Earthquake, China and Its Implication to Earthquake Hazards. *Geomat. Inf. Sci. Wuhan Univ.* **2012**, *37*, 1207–1211. (In Chinese) [\[CrossRef\]](#)
36. Robinson, C.; Gupta, R.; Fobi Nsutezo, S.; Pound, E.; Ortiz, A.; Rosa, M.; White, K.; Dodhia, R.; Zolli, A.; Birge, C.; et al. Turkey Building Damage Assessment. 2023. Available online: <https://www.microsoft.com/en-us/research/publication/turkey-earthquake-report/> (accessed on 10 April 2023).
37. Otsu, N. A thresholding selection method from gray level histogram. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [\[CrossRef\]](#)
38. Zoph, B.; Vasudevan, V.; Shlens, J.; Le, Q.V. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8697–8710. Available online: <https://arxiv.org/abs/1707.07012> (accessed on 10 April 2023).
39. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241. [\[CrossRef\]](#)
40. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980. Available online: <https://arxiv.org/abs/1412.6980> (accessed on 10 April 2023).
41. Loshchilov, I.; Hutter, F. SGDR: Stochastic gradient descent with warm restarts. *arXiv* **2016**, arXiv:1608.03983. Available online: <https://arxiv.org/abs/1608.03983> (accessed on 11 April 2023).
42. Everingham, M.; Eslami, S.; Gool, L.V.; Williams, C.; Winn, J.; Zisserman, A. The pascal visual object classes challenge: A retrospective. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [\[CrossRef\]](#)
43. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Guo, Y.; Ding, G.; Yue, G.; Wang, J. *Semi-Supervised Active Learning with Cross-Class Sample Transfer*; AAAI Press: Washington, DC, USA, 2016; pp. 1526–1532. Available online: <https://dl.acm.org/doi/abs/10.5555/3060832.3060834> (accessed on 11 April 2023).
45. Wang, H.; Sun, G.; Ouyang, C.; Liu, J. Phases of earthquake emergency response period. *J. Catastrophology* **2013**, *28*, 166–169. (In Chinese)
46. Gupta, R.; Hosfelt, R.; Sajeev, S.; Patel, N.; Goodman, B.; Doshi, J.; Heim, E.; Choset, H.; Gaston, M. xBD: A dataset for assessing building damage from satellite imagery. *arXiv* **2019**, arXiv:1911.09296. Available online: <https://arxiv.org/abs/1911.09296> (accessed on 11 April 2023).

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.