



Revolutionizing building damage detection: A novel weakly supervised approach using high-resolution remote sensing images

Wenfan Qiao, Li Shen, Qi Wen, Quan Wen, Shiyang Tang & Zhilin Li

To cite this article: Wenfan Qiao, Li Shen, Qi Wen, Quan Wen, Shiyang Tang & Zhilin Li (2024) Revolutionizing building damage detection: A novel weakly supervised approach using high-resolution remote sensing images, International Journal of Digital Earth, 17:1, 2298245, DOI: [10.1080/17538947.2023.2298245](https://doi.org/10.1080/17538947.2023.2298245)

To link to this article: <https://doi.org/10.1080/17538947.2023.2298245>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 28 Dec 2023.



Submit your article to this journal



Article views: 1855



View related articles



View Crossmark data



Citing articles: 8 View citing articles

Revolutionizing building damage detection: A novel weakly supervised approach using high-resolution remote sensing images

Wenfan Qiao  ^a, Li Shen  ^{a*}, Qi Wen  ^{b*}, Quan Wen^c, Shiyang Tang^d and Zhilin Li^a

^aFaculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, People's Republic of China; ^bTechnology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing, People's Republic of China; ^cTencent Technology (Beijing) Co., Ltd., Beijing, People's Republic of China; ^dState Grid Smart Grid Research Institute Co., Ltd., Beijing, People's Republic of China

ABSTRACT

Rapidly estimating post-disaster building damage via high-resolution remote sensing (HRRS) imagery is essential for initial disaster relief. However, the complex appearance of building damage poses challenges for existing methods. Specifically, relying solely on post-disaster images lacks building boundary guidance, while change detection methods using dual-temporal imageries are prone to introducing false changes. To address these issues, this paper presents a novel weakly supervised approach that leverages pre- and post-disaster HRRS images for building damage detection. The contributions of this paper are twofold. Firstly, a unique framework is proposed to utilize dual-temporal images. Precisely, the proposed method initially extracts fine-grained sub-building-level individuals from pre-disaster images by combining a fully convolutional neural network (FCN)-based method with superpixel segmentation. Then, these details serve as cues to effectively guide the detection of damaged building areas on post-disaster images, thereby enhancing accuracy. Secondly, we propose a weakly supervised method that solely relies on labeling building damage based on image patches but can ultimately yield pixel-level building damage results. Experiments conducted using HRRS images captured during the 2010 Haiti earthquake demonstrate that the proposed method outperforms existing methodologies. This effort of this paper will contribute to the sustainable development of cities and human settlements.

ARTICLE HISTORY

Received 5 September 2023
Accepted 1 December 2023

KEYWORDS

Convolutional neural network; superpixel segmentation; weakly supervised semantic segmentation; high-resolution remote sensing image; building damage detection

1. Introduction

The capacity for disaster prevention and mitigation is a crucial indicator of sustainable development. Sustainable Development Goal 11 (SDG 11) within the 2030 Agenda for Sustainable Development emphasizes the imperative of 'making cities and human settlements inclusive, safe, resilient, and sustainable', and places particular focus on the development of sustainable and resilient buildings. The built-up area serves as the primary hub for human activities, and it is

CONTACT Li Shen  rsshenli@outlook.com  No.999 Xi'an Road, Pidu District, Chengdu 611756, People's Republic of China;
Qi Wen  whistlewen@aliyun.com  No. 9 Deng Zhuang South Road, Haidian District, Beijing, 100094, People's Republic of China

*Li Shen and Qi Wen are both corresponding authors

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

also the region that experiences the most severe casualties and property losses in the aftermath of natural disasters, such as earthquakes (Amirebrahimi et al. 2016). Therefore, after an earthquake, rapidly and accurately obtaining information about building damage can provide technical support and a decision-making basis for emergency rescue and post-disaster reconstruction (Fiedrich, Gehbauer, and Rickers 2000; Janalipour and Mohammadzadeh 2018). Typically, such damage information is gathered through field surveys. Nevertheless, these conventional investigative approaches are not only time-consuming and labor-intensive but also entail risks due to the harsh conditions in post-disaster regions.

The increasing availability of high-resolution remote sensing (HRRS) images offers the potential for meticulous detection of building damage across expansive territories, reducing the need for direct on-site observations (Ajmar et al. 2013). In HRRS images, the damaged regions within buildings often exhibit distinct characteristics that distinguish them from typical geographical elements. These characteristics include: (1) irregular and chaotic patterns without well-defined boundaries, and (2) notable inter-class resemblance combined with intra-class diversity. For these reasons, the automated extraction of building damage from HRRS images remains a challenging task (Ehrlich et al. 2009; Fan et al. 2017), and great attention is devoted to this work. Based on the variation in the utilized data, existing approaches can be broadly categorized into two groups: those employing solely post-disaster data and those utilizing both pre- and post-disaster data (Ge, Gokon, and Meguro 2020).

For the first kind of approaches, their essence lies in achieving the automatic classification and recognition of building damage targets in remote sensing images. Currently, traditional methods commonly involve the development of carefully designed visual features to achieve effective differentiation between building damage and other background geographical elements within the feature space (Dong and Shan 2013). However, the primary challenges associated with these methods pertain to their heavy dependence on domain-specific prior knowledge and the limitations arising from artificially engineered features. Deep learning-based methods, exemplified by convolutional neural networks (CNNs) (Krizhevsky, Sutskever, and Hinton 2012) enable the automation of feature learning, thereby significantly enhancing the model's generalization capability (Gu et al. 2018; Lin et al. 2021; Long, Shelhamer, and Darrell 2015). They have already demonstrated substantial potential in extracting post-disaster building damage (Cha, Choi, and Büyüköztürk 2017; Nex et al. 2019; Sun et al. 2017). Nevertheless, damaged regions within buildings often exhibit characteristics of weak targets with vague boundaries and significant intra-class variance, which can render them susceptible to blending into the complex background environment of post-disaster imagery. Utilizing CNN-based sliding window approaches, without the guidance of prior knowledge, struggles to efficiently and comprehensively define the search space for potential damaged building candidate areas. Therefore, Ji, Liu, and Buchroithner (2018) employed auxiliary pre-disaster building vector data as guidance to enhance the processing efficiency of the CNN. Although effective, this method encounters challenges in obtaining pre-disaster building vector data during the emergency response phase. Furthermore, since CNN models are trained on image patch samples, they are limited to achieving category labeling at the image patch level. Consequently, in order to acquire more detailed pixel-level information on building damage, researchers have also explored semantic segmentation models trained on finely annotated samples (Bai, Mas, and Koshimura 2018; Song et al. 2020) and object detection models (Ma et al. 2019). However, the meticulous annotation of building damage samples is a time-consuming, costly, and error-prone process, making it impractical for disaster mitigation and emergency response needs.

In situations where both pre- and post-disaster data are available, the utilization of remote sensing image change detection methods has become a widely adopted approach to extracting damaged building areas. The core concept revolves around detecting and extracting building damage information by analyzing spectral and morphological differences presented on remote sensing images captured at different time intervals (Dubois and Lepage 2014; Ghaffarian et al. 2019; Janalipour and Taleai 2017; Moya et al. 2020; Sublime and Kalinicheva 2019; Wang and Li 2020;

Zheng et al. 2021). Still, change detection can only be considered as an initial step in the process of extracting damage information. The identified areas of change still require further evaluation and validation to ascertain whether they genuinely correspond to building damage regions.

In summary, we believe that leveraging both pre- and post-disaster HRRS images for building damage extraction is justified. Pre-disaster images hold the potential to provide essential cues for achieving efficient localization of candidate areas of building damage on post-disaster images. Nevertheless, the key factor lies in the design of effective guiding strategies. Furthermore, a critical issue that demands immediate attention is how to transition from weakly supervised image patch annotations to pixel-level refined extraction results.

With the aforementioned considerations in mind, this paper presents a novel weakly supervised approach that leverages pre- and post-disaster HRRS images for building damage detection. The introduced methodology employs a fully convolutional neural network (FCN)-based method to extract architectural features from pre-disaster images and utilizes a CNN-based method for extracting building damage from post-disaster images. The insights derived from the former are harnessed to complement the latter process. Overall, the main contributions of this paper are manifested in the following two aspects:

- (1) Firstly, this paper introduces a novel framework for building damage detection, utilizing both pre- and post-disaster images. Differing from the utilization approach in change detection, our method extracts fine-grained sub-building-level individuals from pre-disaster images by combining an FCN-based method with superpixel segmentation firstly. These details then serve as cues to effectively guide the detection of damaged building areas on post-disaster images, thereby enhancing accuracy. Such an approach serves a dual purpose. On one hand, it addresses the limitations of change detection methods. On the other hand, by constructing building damage candidate clues at a sub-building level, it enhances the precision of pinpointing damaged areas. This is particularly beneficial since building damage tends to occur locally.
- (2) Secondly, we propose a weakly supervised method that solely relies on labeling building damage information based on image patches but can ultimately yield pixel-level building damage results. Dissimilar to weakly supervised methods that generate class activation maps (CAMs) based on only post-disaster images (Qiao et al. 2023), we achieve the transition from image patch-level labeling to pixel-level detection results by employing a voting mechanism. Specifically, we use the building damage candidate clues obtained from pre-disaster images to constrain the building damage identification results from post-disaster images output by the CNN model trained using image-level annotations firstly. The final pixel-wise building damage map are then obtained using the confidential vector through a voting strategy.

Furthermore, both the CNN and FCN models employed in our approach incorporate a wide multi-scale block to combine multi-scale deep features and extend the network's width. This strategy is intended to enhance the models' capability to effectively represent multi-scale object features.

Experimental evaluations conducted on Google Earth image data captured during the 2010 Haiti earthquake demonstrate that the proposed method yields superior performance compared to existing methodologies.

The subsequent sections of this paper are structured as follows. Section 2 offers an overview of the relevant literature. In Section 3, we introduce the processing framework and principal components of the proposed approach. Section 4 presents the results of our experiments and comparisons. Section 5 encompasses a comprehensive discussion, while Section 6 concludes the paper with final remarks.

2. Related works

Numerous approaches have been proposed for automated detection of building damage after a catastrophic event using remotely sensed data. Additionally, many approaches have been proposed to

address the issue of structure health monitoring and structure damage identification in buildings (Minh et al. 2023a, 2023b; Sang-To et al. 2023; Tiachacht et al. 2021; YiFei et al. 2023). A comprehensive review of current literature is beyond the range of this paper. Here, we focus on methods based on optical HRRS images. According to the difference of the data utilized, existing approaches can be generally divided into two categories, i.e. employing only post-disaster data and applying both pre- and post-disaster data (Ge, Gokon, and Meguro 2020).

2.1. Using only post-disaster data

The rich detailed information on HRRS images makes it possible to identify building damage using only post-disaster data. Due to the complexity of building damage, constructing a suitable feature representation is critical for improving the detection accuracy.

Li, Zhang, and Wu (2012) fused both morphological texture and spectral information through a nonlinear function to characterize the debris of collapsed building structures. Moya et al. (2019) created a set of features by constructing a 3D gray level co-occurrence matrix, which was then fed to a SVM classifier for detecting collapsed buildings. Several scholars explored the relationship between buildings and their cast shadows in remote sensing images, which might be used as cues to infer building damage (Tong et al. 2013; Turker and Sumer 2008). Furthermore, high-level semantic features based on bag-of-words models were also designed for the damage detection (Naito et al. 2020; Tu et al. 2016).

In order to overcome the shortcomings of artificial feature design, more recently, CNN-based methods have showed a growing potential in the task of automated building damage detection (Ali, Sultani, and Ali 2020; Bai et al. 2017; Ji et al. 2019; LeCun, Bengio, and Hinton 2015; Sun et al. 2017; Vetrivel et al. 2018). In this direction, various CNN architectures have been proposed, including VGGNet (Simonyan and Zisserman 2014), SqueezeNet (Iandola et al. 2016), ResNet (He et al. 2016), SEAM (Wang et al. 2020), and MDC (Wei et al. 2018). Under the framework of image classification network, Bai et al. (2017) employed a SqueezeNet network to distinguish between built-up and non-built-up regions. Subsequently, they developed a modified wide residual network to determine the damage level. Duarte et al. (2018) evaluated three multi-resolution CNN feature fusion approaches and two baseline CNN models for building damage assessment. Ali, Sultani, and Ali (2020) used a weakly supervised semantic segmentation (WSSS) method to generate class activation maps (CAMs) to assist in building damage extraction. Within the context of semantic segmentation network, Song et al. (2020) presented a DeepLab v2 model for the initial pixel-level identification of damaged building regions and then utilized a superpixel segmentation map to refine the boundaries of damaged buildings. In the realm of object detection networks, Ma et al. (2019) proposed an improved YOLOv3 model to identify collapsed buildings from post-earthquake HRRS images.

Furthermore, object-based image analysis (OBIA) has been also adopted in several approaches, which can improve the detection results through comprehensive utilization of spectral, texture, shape, and topological structure features (Bialas et al. 2016; Song et al. 2020; Zheng et al. 2021).

Overall, existing approaches using only post-disaster data may have the following shortcomings, i.e. (1) the lack of the pre-disaster building data as cues to locate the candidate damage area will result in decreased detection accuracy, and (2) semantic segmentation networks can obtain pixel-level identification results, but creating fine-grained training annotations are labor-intensive and time consuming. While image-level annotations may suffice for image classification models, they can only yield image-level labeling results.

2.2. Using both pre- and post-disaster data

The increased availability of optical remote sensing images and the frequently updated data archives makes it possible to obtain pre-disaster data to assist in the identification of building damage.

Compared to approaches that utilize only post-disaster data, more reliable and accurate results can be acquired by introducing prior knowledge from pre-disaster data (Vu and Ban 2010). For this reason, great attention was devoted to the development of methods using both pre- and post-disaster data for building damage detection (Dubois and Lepage 2014; Ghaffarian et al. 2019; Janalipour and Taleai 2017; Shen et al. 2020; Sublime and Kalinicheva 2019; Weber and Kané 2020).

Ancillary information, such as a pre-disaster building vector map, has been extensively exploited to improve the accuracy of building damage detection (Janalipour and Mohammadzadeh 2015; Ji, Liu, and Buchroithner 2018; Ji et al. 2019; Ma et al. 2020; Ye et al. 2017). Ye et al. (2017) and Janalipour and Mohammadzadeh (2015) applied statistical learning methods to classify building damage from post-disaster images based on multiple features, such as geometric, texture, and gradient features. Pre-disaster building maps were introduced into the methods to accurately locate building boundaries and eliminate the interference from other objects. Ji, Liu, and Buchroithner (2018) employed an improved SqueezeNet to identify damaged buildings from post-disaster image assisted by building footprint maps. Besides, due to the reason that the number of collapsed buildings is far less than other objects, three strategies including random over-sampling, random under-sampling, and cost-sensitive were used to deal with the sample imbalance problems (Buda, Maki, and Mazurowski 2018; Ji et al. 2019; Ma et al. 2020). The aforementioned studies prove that pre-disaster ancillary information does improve the accuracy of building damage recognition to some extent. Yet, the additional GIS vector data is not always available and require extra handcrafted processing, which limits the applicability in rapid disaster emergency response.

Therefore, using both pre- and post-disaster remote sensing images for building damage detection has become an alternative strategy. In the context of change detection, Ghaffarian et al. (2019) employed multitemporal satellite data for updating the building map after a disaster using automated deep learning techniques. Sublime and Kalinicheva (2019) applied an unsupervised deep neural network to develop a change detection method for mapping post-disaster damage. Shen et al. (2020) adopted a two-stage U-Net framework for building damage assessment. It mainly utilized a cross-directional feature fusion module to combine a pair of registered pre- and post-disaster images. However, even when pre-disaster images are available, obtaining reliable and accurate building damage results can still be challenging due to the obvious difference in color, spectral characteristics, and viewing angles between pre- and post-disaster images. So, in addition to change detection, several methods also attempted to initially extract the building regions from the pre-disaster image, and then compared the corresponding building areas between pre- and post-disaster images using various similarity measures to determine whether the building is damaged (Chini, Cinti, and Stramondo 2011; Korkmaz and Abualkibash 2018; Sarp et al. 2014). Moreover, Bai, Mas, and Koshimura (2018) explored to directly concatenate pre- and post-disaster images as input to train a U-Net model for mapping the damage.

All in all, it is crucial to initially delineate the building regions in the pre-disaster image as a guide for building damage detection in the post-disaster image. However, the design of an appropriate guiding methodology is likely the key factor in this endeavor.

3. Methodology

In this section, we will provide a detailed explanation of our method for the weakly supervised identification of building damage from HRRS images. The framework of the proposed method is depicted in Figure 1, which mainly includes three stages. Firstly, an FCN and a CNN model are designed respectively based on the proposed wide multi-scale block. In the second step, the FCN model, which is trained on pixel-level annotations, is used to extract building regions in the pre-disaster image. Subsequently, a superpixel segmentation algorithm is applied to divide the extracted building regions into a collection of small pure sub-regions, which are then mapped to the post-disaster image to generate destruction building candidate processing units. This step

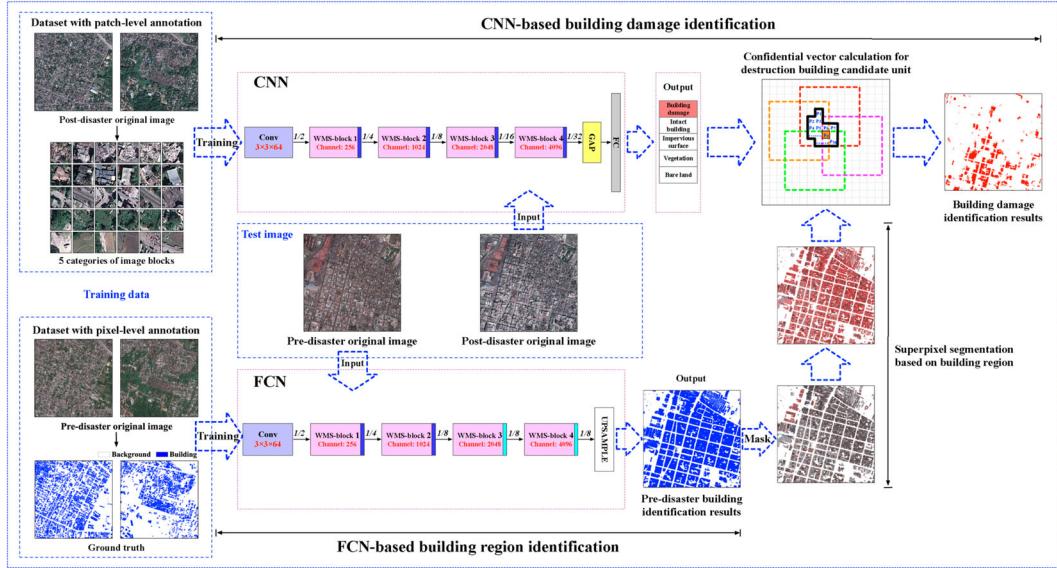


Figure 1. Overall framework of proposed methodology for building damage detection.

will help us to locate the damaged areas of buildings more precisely, because on many occasions, buildings are partially damaged. Finally, the CNN model that is trained on image-level annotations is used to perform pixel-wise predictions to locate building damage in the post-disaster image. Specifically, overlapping image patches generated from the post-disaster image are fed into the trained CNN to obtain the confidential vector for each image patch. The final pixel-wise building damage map can be obtained using the confidential vector through voting mechanism. The detailed explanation of the proposed method is described in the following three subsections.

3.1. Design of neural network architectures

As mentioned above, there is a weak distinction between the characteristics of damaged buildings and those of background in the post-disaster satellite images. In order to better serve the building damage detection task, it is crucial to achieve better image feature representation. To this end, a wide multi-scale (WMS) block, which aims to take advantage of multi-scale features, is designed as the key part of the neural networks, including a CNN model and an FCN model, as shown in Figures 2 and 3.

The WMS block replaces the 3×3 convolution kernel in the standard residual unit with a series of parallel dilated convolution layers, where dilation rates for the parallel paths are set to 1, 3, 6, 9 and 12, respectively. And then, multiple features outputted by the dilated convolution layers are concatenated together, as illustrated in Figure 2. It should be noted that the parameter N means the number of convolution kernels, which is also equivalent to the number of feature map channels. The use of dilated convolution allows for expanding the receptive field of the convolution kernel by employing various dilation rates, without increasing the total number of parameters and losing the resolution or coverage of the feature map. Moreover, benefiting from the way of cascaded connection, the dilated convolution layer's multi-scale output feature maps are fused. This makes it possible to learn both low-level and high-level multi-scale image features.

In addition, the study in Zagoruyko and Komodakis (2016) has confirmed that when different neural network models have approximate network complexity, a wider and shallower network

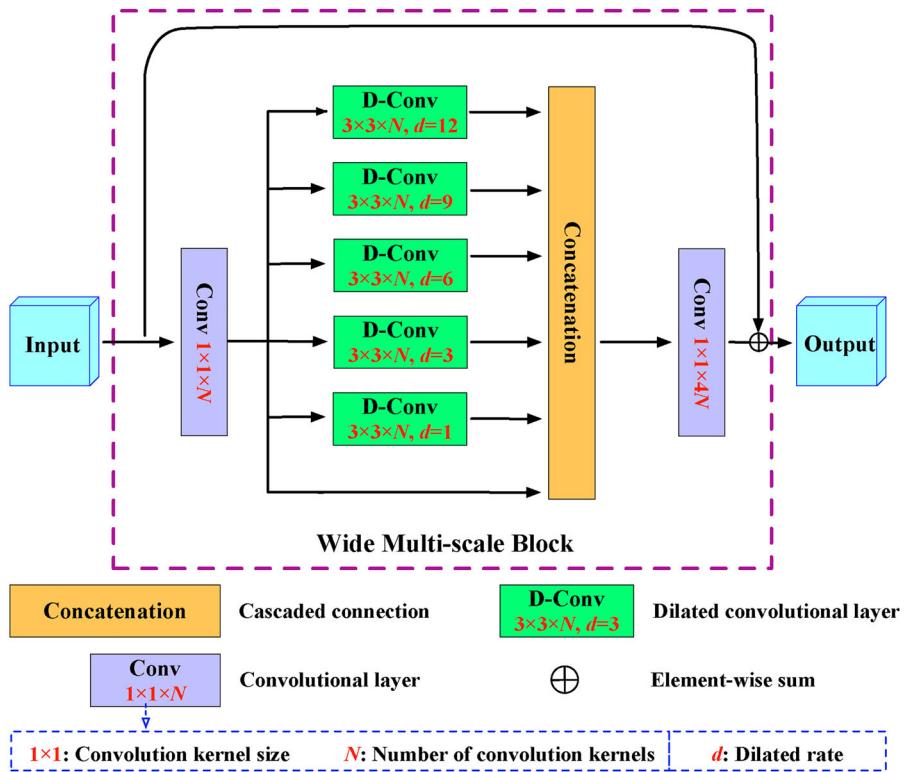


Figure 2. The proposed wide multi-scale block.

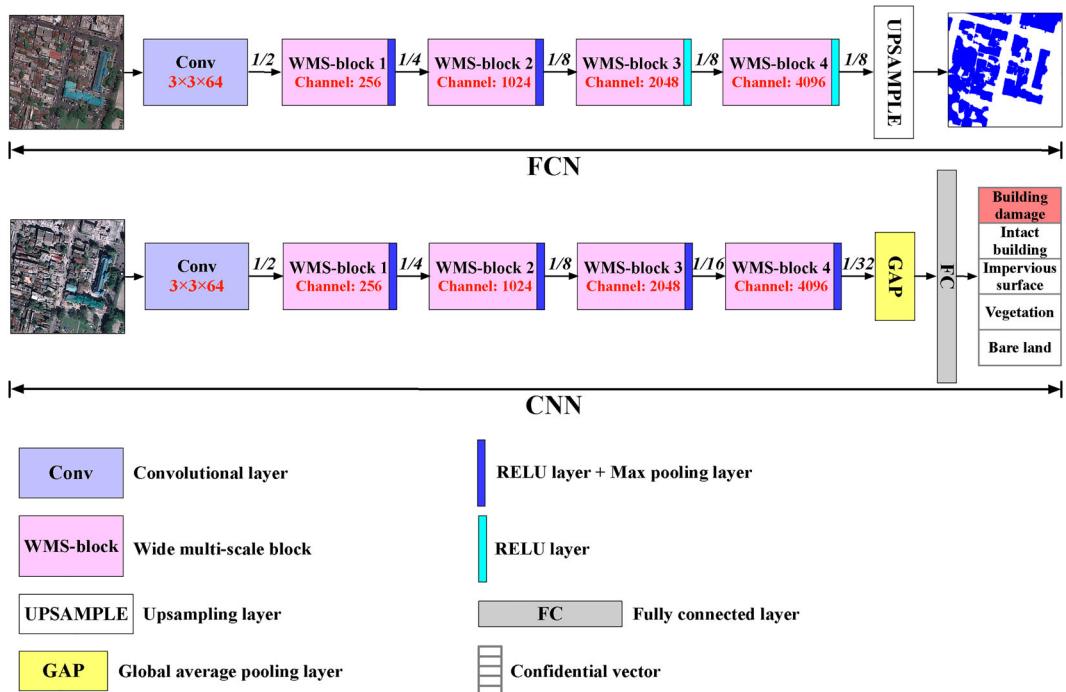


Figure 3. Architectures of the proposed WMSNet.

structure may utilize the GPU computing power more effectively and perform better. This finding inspires us to explore the value of increasing the network width. For this reason, the proposed WMS block uses five parallel dilated convolution layers to augment the width of network.

Based on the proposed WMS block, wide multi-scale networks (WMSNet), including an FCN model and a CNN model, are constructed, respectively. To be specific, the FCN model is used to extract building regions from pre-disaster images, and the CNN model is for detecting building damage in the post-disaster images. Both of the FCN and the CNN contain four connected WMS blocks. The detailed neural network architectures are illustrated in Figure 3. For the CNN model, in the neural network architecture, each WMS block is followed by a max pooling layer to reduce the resolution of the feature map. The size of the final extracted feature map will be 1/32 of the original input image. An global average pooling (GAP) layer, a fully connected (FC) layer and a softmax classifier are placed at the end of the network. Overlapping image patches generated from the post-disaster image will be fed into the CNN to obtain the confidential vector for each image patch. For the FCN model, the last two max pooling layers of the network are removed in order to reserve as much detail information as possible. Hence, the size of the feature map outputted by the WMS blocks will be 1/8 of the input image. Followed by a linear upsampling operation at the end, the FCN will recover the final feature map to the original image size.

3.2. Generation of destruction building candidate units

After an earthquake, in addition to the damage to buildings, the damage to roads can also take place. Moreover, debris from collapsed buildings might be dispersed onto the road. These elements are easily mistaken for building damage in the post-disaster HRRS images, posing a challenge for the task of building damage detection. However, since building damage can only occur within the building regions, it is therefore necessary to obtain the building regions in the pre-disaster image firstly to guide the detection of building damage in the post-disaster image.

In this paper, we use the previously designed FCN model to extract building regions in the pre-disaster satellite image. As shown in Figure 4, the FCN model can better extract the building areas at

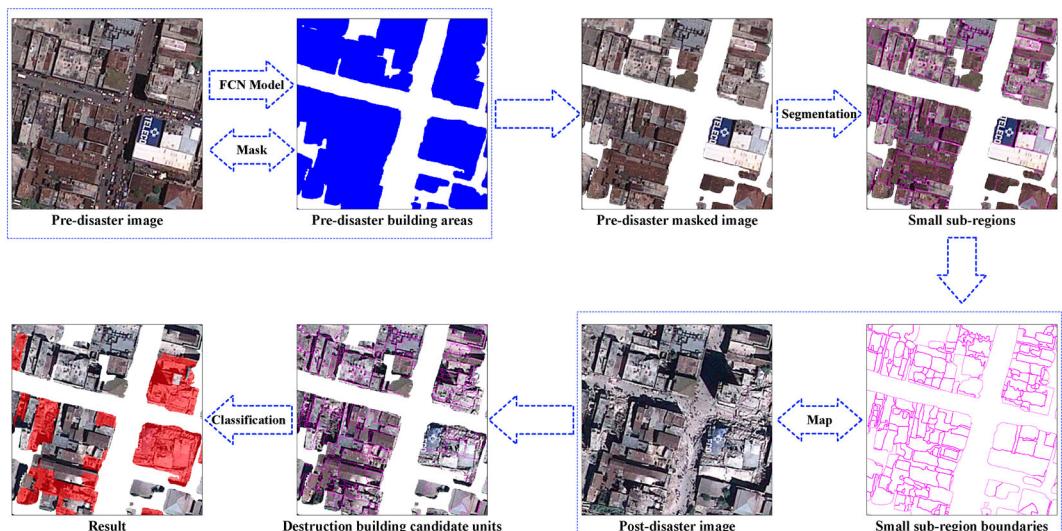


Figure 4. The generation process of destruction building candidate units.

the pixel level (Abriha and Szabó 2023; Qiao et al. 2019). But we also notice that the boundaries between different individual buildings are blurry in the semantic segmentation result, so it is inappropriate to directly use the obtained rough building areas as cues to detect building damages in the post-disaster image.

To address the above-mentioned issues, a novel strategy for the generation of destruction building candidate units is proposed, by conducting fine-grained segmentation of building areas. The purpose of this strategy is twofold. On the one hand, it helps to get clear building boundaries. On the other hand, it gains sub-building-level individuals, and will be beneficial to locate the local damaged areas of buildings more precisely, because buildings are partially damaged on many occasions.

As shown in [Figure 4](#), the proposed strategy includes the following steps. Firstly, the rough building regions are extracted using the designed FCN model from the pre-disaster satellite image. Secondly, we use the results of the building area segmentation to mask the original pre-disaster image. Afterward, a superpixel segmentation algorithm is applied to divide the masked image into a collection of smaller sub-regions. Finally, we map the acquired sub-regions to the post-disaster image to segment the image. Accordingly, destruction building candidate units, i.e. superpixel regions, are obtained.

Super-pixels are small regions consisting of a series of contiguous pixels with similar features, which can greatly keep the relationship between adjacent pixels for subsequent image processing (Song et al. 2020). It is import to point out that although the proposed strategy requires the use of a superpixel segmentation algorithm, it is not tied to one specific segmentation algorithm. Any approach that can generate reasonable superpixels could meet the requirement. In this paper, we choose Efficient graph-based image segmentation (EGIS) (Felzenszwalb and Huttenlocher 2004) algorithm for the superpixel segmentation task. EGIS is proven to be an effective segmentation method, which can preserve details in low-variability image regions while ignoring details in high-variability regions. It also should be noted that when using the superpixel segmentation method, we don't need to care about selecting the segmentation scale or whether the images are segmented at an optimal scale. Our primary focus should be on ensuring that each superpixel is a homogeneous building unit.

3.3. Pixel-wise building damage mapping

Once we have obtained each destruction building candidate unit in the post-disaster satellite image, the subsequent step is to assign a category label to each unit, indicating whether it is damaged or non-damage. In this stage, the previously designed CNN model trained on image-level annotations, which is regarded as weak supervision, is used to obtain the final pixel-wise building damage map from the post-disaster image.

To be specific, the designed CNN model is firstly trained based on image patch samples (labeled at the image level) covering C types of objects, i.e. building damage, intact building, etc. Secondly, the post-disaster satellite image is divided into a series of overlapping image patches for the test purpose. These patches are fed into the trained CNN, and then each image patch will be allocated a C -dimensional confidential vector I , which corresponds to the probability distribution that the image patch belongs to the pre-defined C object categories. Since the image patches are generated in an overlapping manner, each pixel in the post-disaster image will be covered by multiple image patches, as shown in [Figure 5](#). For a pixel P_n , suppose that there are K image patches covering it, its confidential vector can be mathematically given by:

$$I_{P_n} = \frac{1}{K} \sum_{i=1}^K I_i \quad (1)$$

Each destruction building candidate unit U may include M pixels, so the confidential vector for each

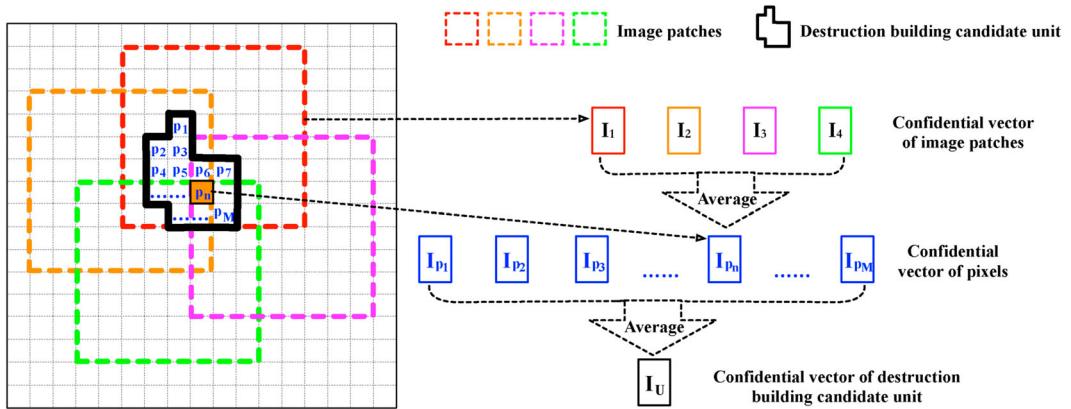


Figure 5. The illustration of calculating the confidential vector for each destruction building candidate unit.

U can be represented as the following:

$$I_U = \frac{1}{M} \sum_{n=1}^M I_{p_n} \quad (2)$$

Obviously, I_U is a C -dimensional confidential vector. Each destruction building candidate unit U can be classified as a pre-defined object label c with the maximum confidential rule, i.e.

$$c = \text{argmax}(I_U) \quad (3)$$

It should be noted that we can aggregate all the non-damage categories as a single class, i.e. non-damage. Then, the pixel-wise building damage map including damage or non-damage category can be finally obtained.

4. Experiments

4.1. Study area and experiment data

To demonstrate the effectiveness of the proposed method, the region of Port-au-Prince, Haiti, which was severely damaged by 2010 Haiti earthquake occurred on 12 January 2010, is selected as the study area. The disaster event caused an extensive destruction for buildings and other man-made facilities. Due to the dense distribution of buildings, Port-au-Prince area was heavily affected, with more than 100,000 buildings being completely damaged.

Figure 6(a) shows a pair of HRRS images covering Port-au-Prince area of Haiti. The images were taken on 26 August 2009 (pre-disaster) and on 13 January 2010 (post-disaster). Both images were captured from Google Earth Pro with a spatial resolution of 0.28 m, including three channels: red (R), green (G), and blue (B).

Two distinct datasets have been separately prepared to accommodate the requirements of both the FCN model and the CNN model. Each dataset is then divided into three groups, i.e. the training set for training the model, the validation set for validating the model during the training stage, and the test set for evaluating the model. For the FCN model, the building dataset includes twelve pre-disaster image patches of 3500×3500 pixels and their corresponding pixel-level annotations. As shown in Figure 6(a), the image patches for the training set and validation set are marked with yellow boxes, while the image patches for the test set are marked with red boxes. An example image patch and its corresponding ground truth mask annotation can be illustrated in Figure 6(b). Moreover, Inria aerial image labeling dataset is used as pre-training data for the building extraction model (Maggiori et al. 2017).

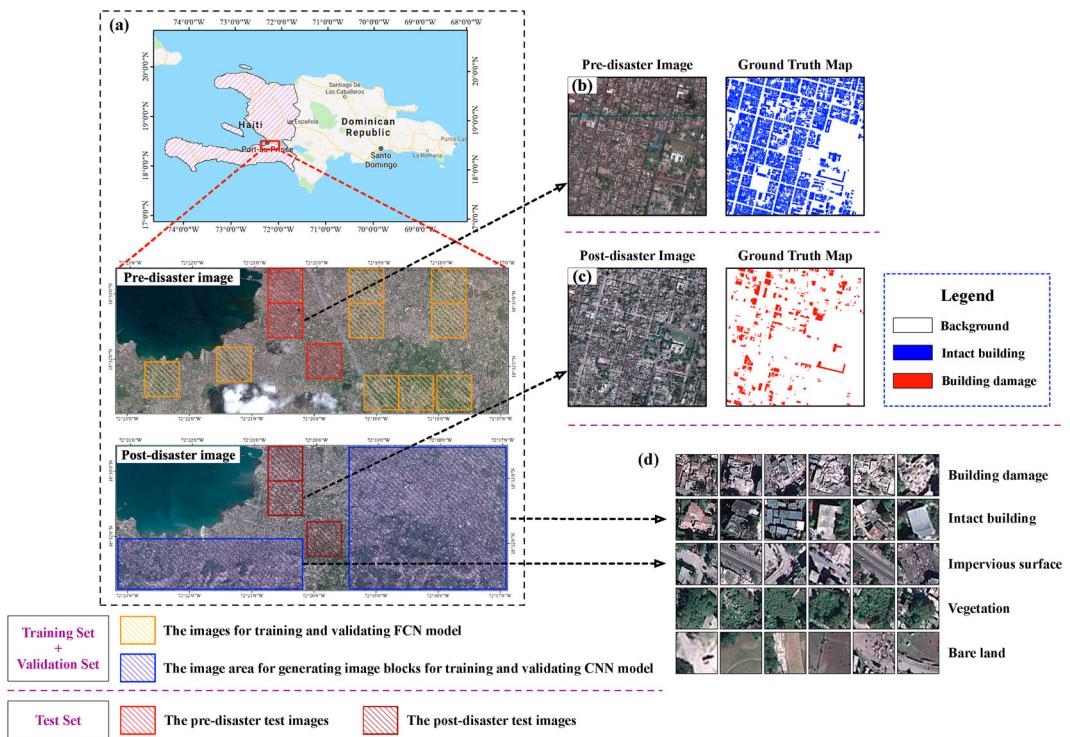


Figure 6. Study area and image data.

For the CNN model, as shown in Figure 6(a), the areas marked by blue boxes in the post-disaster image correspond to the training set and validation set, while the red areas represent the test set. A series of overlapping square image blocks is produced from these areas to construct the dataset. As shown in Figure 6(d), each image block is annotated at the image level, with one of five categories, i.e. building damage, intact building, impervious surface, vegetation, and bare land. The side lengths of generated square image blocks are set between 50 and 500 pixels, and all image blocks are resampled to 256×256 pixels as the input of the CNN model. Table 1 shows the dataset information for both the FCN model and the CNN model. To alleviate the problem of insufficient training samples and sample imbalance, we employ a few basic image data augmentation techniques, such as mirroring, rotating (90° , 180° and 270°) and Gaussian blur.

Three test images of 3500×3500 pixels are used to evaluate the performance of the proposed method for building damage detection. The sites of three images are marked with red rectangular

Table 1. Dataset information for both the FCN model and the CNN model.

	Category	Number of image blocks			Image size (pixels)
		Training	Validation	Test	
FCN	Building and Background (Pre-training)	120	30	30	5000×5000
	Building and Background (Fine-tuning)	6	3	3	3500×3500
CNN	Building damage	1400	300	300	256×256
	Intact building	1400	300	300	256×256
	Impervious surface	1400	300	300	256×256
	Vegetation	1400	300	300	256×256
	Bare land	1400	300	300	256×256



Figure 7. Annotation details of the ground truth maps for buildings and building damage. (a) Ground truth maps of buildings. (b) Ground truth maps of building damage. (c) Comparison of pre- and post-disaster images for building damage.

boxes in Figure 6(a), including two typical scenarios, i.e. areas with dense building distribution and sparse building distribution. Figure 6(c) shows an example of a post-disaster image and its corresponding ground truth map of building damage. To generate high-quality ground truth for buildings and building damage, we select experienced annotators and establish rigorous guidelines for annotation standards. Regular meetings are also organized to facilitate real-time discussions and problem-solving during the annotation process. For the ground truth maps of buildings, we refer to existing building semantic segmentation datasets (Ji, Wei, and Lu 2018; Maggiori et al. 2017) as a reference for manually annotating building footprints. Regarding building damage, the ground truth maps are manually interpreted by combining both pre- and post-disaster images. In addition, satellite and aerial images with higher resolutions, along with other supplementary data, are utilized to assist in interpretation. We meticulously annotate the pre-disaster building outlines for buildings that have experienced damage after the disaster, serving as the ground truth for building damage. Our aim is to annotate all building damages that can be identified with the naked eye. Furthermore, a triple cross-checking process is implemented to minimize the risk of false judgments. Figure 7 illustrates the annotation details of the ground truth maps for buildings and building damage.

It should be noted that the categorization of damage to buildings consists of five grades according to the European Macroseismic Scale (EMS-98) (Grünthal 1998). Examples for each grade are illustrated in Figure 8. For both G1 and G2 grades, building structures are basically complete. The roofs and walls of buildings are not seriously damaged, with only a few cracks on the surface or perhaps losing roof tiles. This level of building damage cannot be identified accurately even by manual interpretation from HRRS images (Yamazaki, Yano, and Matsuoka 2005). Obviously, it is also almost impossible to identify these slight damages using automated detection methods. Moreover, the losses caused by these damage situations are relatively small. Therefore, in this paper, we define the identifiable building damage by HRRS images as the damages correspond to G3, G4 and G5 grades (Kerle 2010).

4.2. Experimental setup

4.2.1. Methods for comparison

To evaluate the effectiveness, the performance of the proposed approach is compared with that of ten state-of-the-art methods. These approaches can be divided into two categories according to the

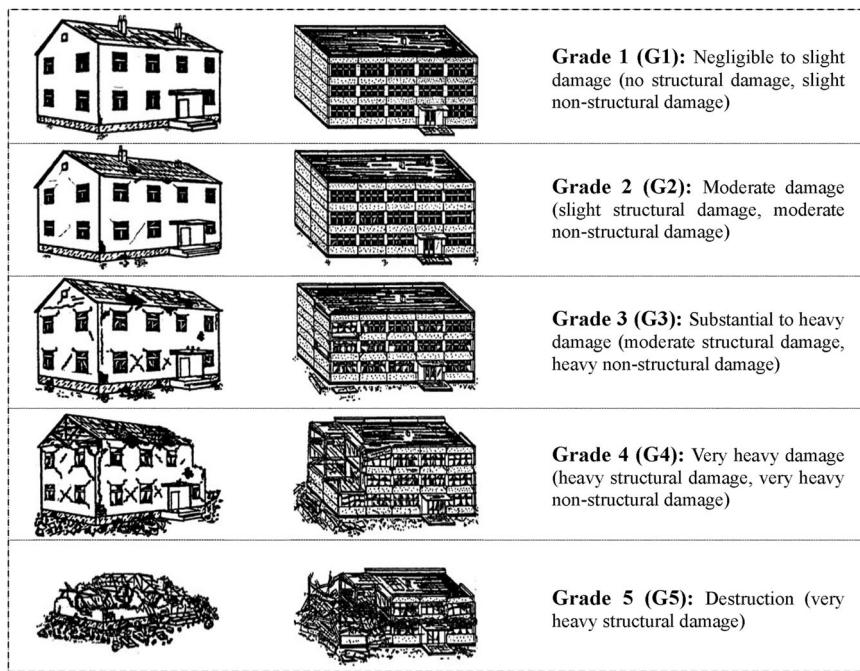


Figure 8. Different grades of damage to buildings.

difference of the data utilized, i.e. employing only post-disaster HRRS images and applying both pre- and post-disaster HRRS images.

For the category using only post-disaster HRRS images, firstly, three typical types of patch-based CNN classification networks, i.e. SqueezeNet (Iandola et al. 2016), VGG-16 (Simonyan and Zisserman 2014), and ResNet-101 (He et al. 2016), are selected. Since image blocks with image-level labels generated from post-disaster HRRS image are regarded as input to these networks for both training and testing, sliding window strategies are employed to obtain pixel-level predictions (Cha, Choi, and Büyüköztürk 2017). For simplicity of presentation, these methods are simply referred to as sw_SqueezeNet, sw_VGG-16, and sw_ResNet-101. Moreover, three WSSS methods (i.e. SEAM (Wang et al. 2020), MDC (Wei et al. 2018), and WSAN (Ali, Sultani, and Ali 2020)), which use CAMs as the cornerstone to generate pseudo masks for segmentation, are also used for comparison.

For the category using both pre- and post-disaster HRRS images, four representative approaches with different tactics are chosen. Firstly, to verify the effectiveness of super-pixel segmentation on the result, a method that combines a CNN model and the super-pixel segmentation from pre-disaster HRRS images, termed as CSSCNN, is added for comparison. Originating from the method in (Song et al. 2020), the CSSCNN replaces the DeepLab v2 semantic segmentation network with a patch-based ResNet-101 network. The second one is the two-stage CNN model, termed as TSCNN (Sun, Tang, and Zhang 2017). It uses a proposed CNN network to firstly identify the building areas from pre-disaster HRRS images, which are treated as cues to aid in locating the candidate building damage. Thus, in the second stage, by paying attention on the above cues, the proposed CNN network is used again to detect building damage from post-disaster HRRS images. In order to improve the accuracy of extracting building areas from pre-disaster HRRS images, we use FCN instead of CNN to identify building areas. The third method is an approach that uses a CNN model and building footprint maps, termed as FMCNN (Ji, Liu, and Buchroithner 2018). The building footprint maps manually interpreted from pre-disaster HRRS images are used to assist the identification of building damage from post-disaster HRRS images based on an improved

SqueezeNet model. At last, particularly, a fully supervised semantic segmentation method utilizing concatenated pre- and post-disaster HRRS images, termed as U-Net (Bai, Mas, and Koshimura 2018), is also added for comparison. It is chosen to evaluate that the accuracy gap between our weakly supervised method and the fully supervised method.

It should be noted that, with the exception of the U-Net model, the above-mentioned methods for comparison all use the image-level annotations for model training. In other words, same as our method, they are all weakly supervised. For the fully supervised U-Net model, nine 3500×3500 pixel concatenated pre- and post-disaster HRRS images, along with their corresponding pixel-level labels, are used for training and validation.

4.2.2. Implementation details

All neural network models are implemented using Pytorch on a server with Intel® Xeon (R) 8-core CPU @ 3.4 GHz processor and NVIDIA GTX 1080Ti GPU (11.5 GB RAM). For various methods based on neural networks, the parameter settings are configured as the same as possible for a fair comparison. All models are trained using stochastic gradient descent (SGD) with momentum and cross-entropy loss function. We set the momentum to 0.9 and the weight decay to 0.0005. The batch size and the total number of iterations are empirically set to 5 and 100,000, respectively. The learning rate is initially set to 0.01, and then the poly-like learning rate rule is applied to update it. The rule is mathematically given by:

$$lr' = lr \times \left(1 - \frac{iter}{Max_iter}\right)^{power}. \quad (4)$$

Here, *iter* and *Max_iter* denote the current iteration and maximum iteration, respectively. *power* represents the power variable, which controls the decay rate of the learning rate. We set *power* to 0.9. *lr* and *lr'* stand for current learning rate and updated learning rate.

As previously stated, for the CNN model, all image blocks are resampled to dimensions of 256×256 pixels to serve as the model's input. For the FCN model, due to the limitation of GPU memory, the HRRS images with large size cannot be directly fed into the network. Thus, we also divide the original image into square image blocks as input. To be specific, at each training step, image blocks of 320×320 pixels are sampled from original training images randomly. During the inference stage for the test images, a fixed stride of 100 pixels is set to acquire the overlapping image blocks of 320×320 pixels as input. The final predictions on the overlapping regions are then averaged, which can reduce the border effects and improve the accuracy.

4.2.3. Evaluation metrics

Five quantitative criteria, including precision (*P*), recall (*R*), F_1 -score (F_1), mean intersection over union (MIoU) and overall accuracy (OA), are used in the experiments to evaluate the performance of building damage extraction for all the methods. The damage building detection task can be regarded as a binary classification problem, i.e. distinguishing between building damage (positive class) and non-building damage (negative class). Thus, we use four statistics, i.e. true positive (*TP*), true negative (*TN*), false positive (*FP*) and false negative (*FN*), to calculate the above criteria. Here, a *TP* denotes a positive sample correctly classified as positive, while a *TN* denotes a negative sample correctly classified as negative. A *FP* represents a negative sample incorrectly classified as positive, while a *FN* represents a positive sample incorrectly classified as negative. As given by formula (5)–(9), *P* is calculated as the number of *TP* divided by the total number of *TP* and *FP*, and *R* is calculated as the number of *TP* divided by the total number of *TP* and *FN*. *P* measures the percentage of predictions made by the model that are correct, while *R* measures the percentage of relevant samples that are correctly identified by the model. F_1 provides a single metric that weights both the concerns of *P* and *R* in a balanced way. MIoU represents the area of intersection between ground truth mask and predicted mask over their union. Moreover, OA is a metric that is generally used to

describe the model performance across all classes.

$$P = \frac{TP_n}{TP_n + FP_n} \quad (5)$$

$$R = \frac{TP_n}{TP_n + FN_n} \quad (6)$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (7)$$

$$MIoU = \frac{1}{2} \left(\frac{TP_n}{TP_n + FP_n + FN_n} + \frac{TN_n}{TN_n + FN_n + FP_n} \right) \quad (8)$$

$$OA = \frac{TP_n + TN_n}{TP_n + FP_n + TN_n + FN_n} \quad (9)$$

Here, the subscript n indicates the total number of TP , TN , FP or FN .

4.3. Accuracy assessment of proposed method

The visual extraction results of building damage on test HRRS images of three scenarios using our proposed approach are shown in Figure 9. The first and second columns represent pre- and post-

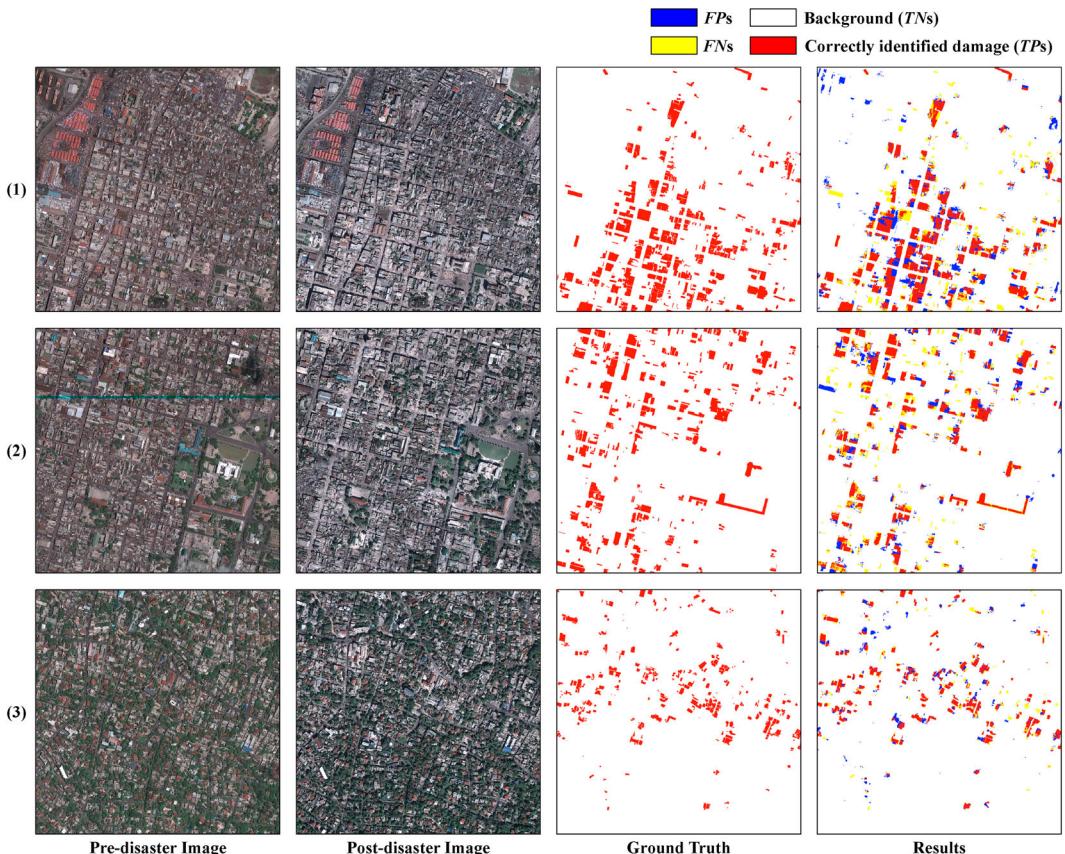


Figure 9. Building damage identification results in three test images by our proposed method.

Table 2. Precision (%), recall (%), F_1 (%), MIoU (%), and OA (%) of our proposed method for test images of three scenarios.

Test Region	Precision (%)	Recall (%)	F_1 (%)	MIoU (%)	OA (%)
1	70.18	73.71	71.90	74.60	93.64
2	71.92	69.16	70.51	74.30	94.52
3	73.43	78.27	75.77	79.35	97.79
Three test regions as a whole	71.41	72.79	72.09	75.69	95.31

disaster HRRS images to be tested. The third column corresponds to extraction results of building damage. As can be observed, the red area stands for correctly identified building damage (*TPs*), while the blue, yellow, and white areas denote *FPs*, *FNs*, and background (*TNs*), respectively. From visual inspection, in general the proposed method yields satisfactory results in building damage detection, with a quite high *TPs* ratio. From quantitative evaluation, four quantitative criteria are calculated, as reported in **Table 2**. Compared with the other two test regions, all accuracy indicators of test region 3 are higher, indicating that better recognition results in this region can be achieved. This phenomenon can be explained by the fact that in the test region 3, buildings are sparsely distributed, and the space between buildings is larger. For this reason, individual building objects are more likely to be correctly segmented and identified, and thus to better guide the extraction of building damage. On the contrary, for the areas with dense building distribution such as test region 1 and test region 2, buildings are very close to each other and even are connected. In this occasion, it is not easy to distinguish and interpret between individual buildings, resulting in decreased accuracy of building damage extraction.

4.4. Comparison with existing methods

In order to verify the effectiveness of the proposed method, we perform quantitative and qualitative comparisons with existing methods. The quantitative evaluation results of various methods are shown in **Table 3**. The best results are marked in bold, and the second best results are underlined. In order to reflect the quantitative results of different methods more intuitively, the corresponding bar chart (**Figure 10**) is drawn according to **Table 3**. As can be seen in **Table 3** and **Figure 10**, our proposed method outperforms all of the previous weakly supervised semantic segmentation approaches in all evaluation metrics. The precision, recall, F_1 -score, MIoU, and OA in the test data are 71.41%, 72.79%, 72.09%, 75.69%, and 95.31%, respectively, confirming that the proposed method is an effective approach for HRRS imagery building damage extraction. Compared with sw_SqueezeNet, sw_VGG-16, and sw_ResNet-101, the SEAM, MDC, and WSAN achieve relative improvements in precision, F_1 , MIoU, and OA, respectively, indicating that CAMs can improve the performance of accurate localization of building damage. Although sw_SqueezeNet, sw_VGG-16, and sw_ResNet-101 can achieve relatively high recall values, the precision values are low, which shows that the methods based on sliding window classification are rough and

Table 3. Precision (%), recall (%), F_1 (%), MIoU (%), and OA (%) of various methods.

Methods	Precision (%)	Recall (%)	F_1 (%)	MIoU (%)	OA (%)
sw_SqueezeNet	29.66	71.75	41.97	54.51	83.50
sw_VGG-16	35.79	64.66	46.07	58.32	87.42
sw_ResNet-101	33.41	62.12	43.45	56.79	86.55
SEAM	32.62	72.00	44.90	56.66	85.31
MDC	40.71	60.32	48.61	60.47	89.40
WSAN	38.25	69.24	49.27	60.06	88.15
CSSCNN	33.90	63.23	44.14	57.14	86.69
TSCNN	57.54	58.99	58.25	66.85	92.97
FMCNN	68.81	55.94	61.71	69.29	94.23
U-Net	78.52	74.60	76.51	78.95	96.19
Ours	<u>71.41</u>	<u>72.79</u>	<u>72.09</u>	<u>75.69</u>	<u>95.31</u>

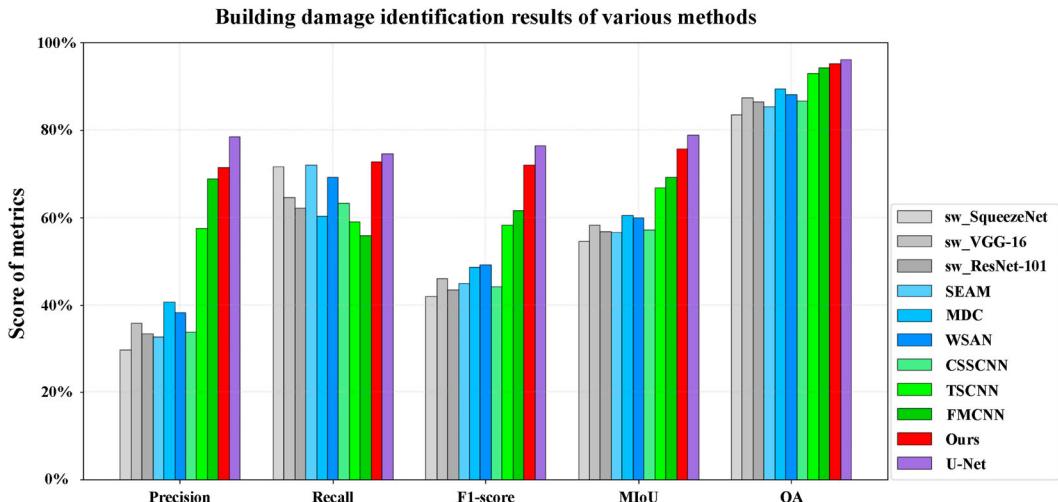


Figure 10. Bar chart of precision (%), recall (%), F_1 (%), MIoU (%), and OA (%) of various methods.

inaccurate for building damage extraction. In addition, it can be observed that adding pre-disaster image information prior to the process of building damage extraction significantly improves accuracy of damage area and boundary extraction for both low-density and high-density buildings. Compared with the sw_ResNet-101, the CSSCNN achieves approximately 0.49%, 1.11%, 0.69%, 0.35%, and 0.14% relative increases in the values of precision, recall, F_1 -score, MIoU, and OA, respectively. The results demonstrate the effectiveness of extracting building damage from post-disaster images in combination with pre-disaster image segmentation objects. TSCNN and FMCNN achieve an obvious improvement in the value of precision. It is worth noting that the reason for the main difference depends on the building area constraints. This fact further demonstrates that combining pre-disaster building areas with building damage extraction can help reduce confusion with surrounding objects.

Furthermore, it should be noted that the quantitative results of the proposed weakly supervised semantic segmentation method are only slightly lower than that of the FCN-based U-Net semantic segmentation method, which indicates that the proposed method is comparable to the fully supervised semantic segmentation method.

Figure 11 displays part of the test image patches and their building damage extraction results using different methods across different scenes. Among all the compared methods, sw_SqueezeNet, sw_VGG-16, and sw_ResNet-101 perform the worst visually. The three sliding window classification methods are obviously rough and inaccurate in the localization of building damage. The building damage extraction results of SEAM, MDC, and WSAN are relatively better than sw_SqueezeNet, sw_VGG-16, and sw_ResNet-101 in terms of smoothness and accuracy visually, which may be mainly attributed to the object localization ability of CAMs. However, it can be observed that there are still numerous false positive predictions for surrounding intact buildings and roads, which share similarities with building damage objects in terms of material and spectrum. This underscores the challenge of achieving high-quality building damage extraction results using only post-disaster HRRS images. Moreover, the damaged roads and scattered rubble from collapsed buildings are also easily confused with building damage objects, which may also significantly affect the performance of building damage extraction from post-disaster HRRS images. Therefore, it is necessary to utilize pre-disaster image information to guide building damage extraction from post-disaster HRRS images. CSSCNN can slightly refine the boundaries of building damage, but does not fundamentally alleviate the problem of building damage being confused with other categories. And the building damage extraction results of CSSCNN are not

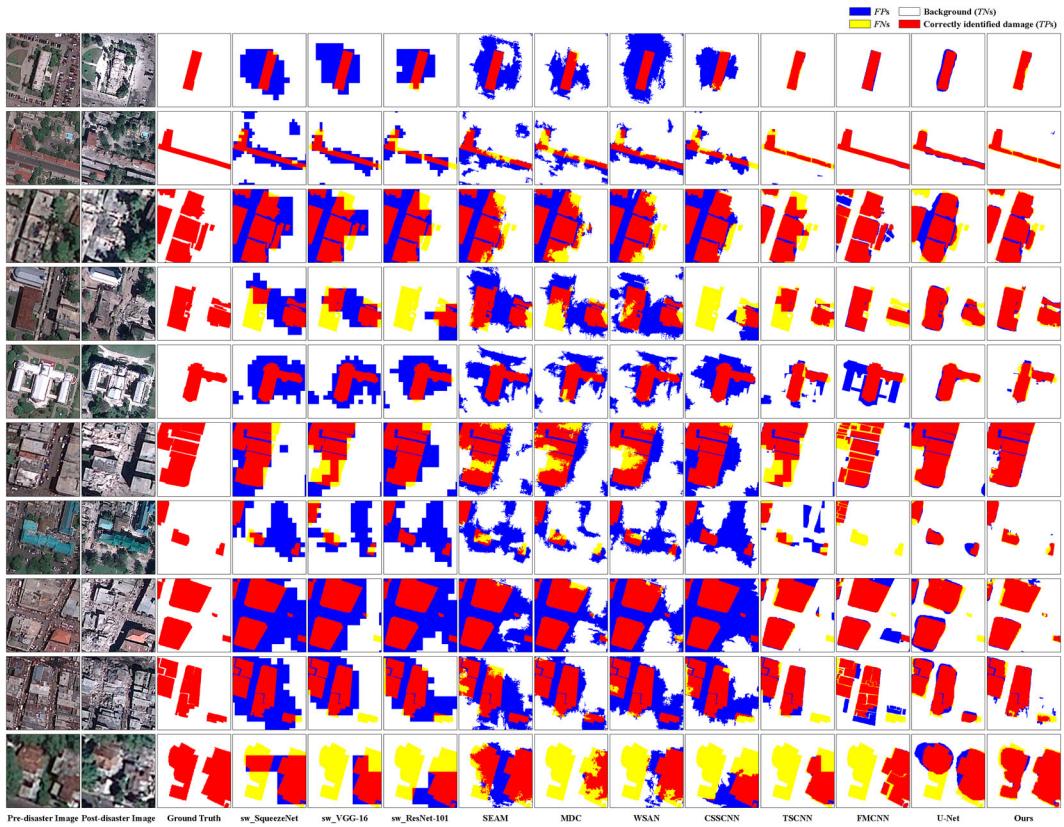


Figure 11. Qualitative comparison of building damage identification results of various methods.

robust enough in different complex scenes and are easily affected by segmentation effects. As can be seen in Figure 11, TSCNN, FMCNN, U-Net, and our proposed method obtain better building damage extraction results compared with other methods visually. The improvements mainly lie in the more accurate identification and localization of building damage and a significant reduction in false positives, all while maintaining high accuracy among different scenes. These improvements are attributed to the building location information acquired from pre-disaster HRRS images. Given building damage can only occur within the building regions, it is an effective strategy to initially acquire building regions in the pre-disaster image to guide building damage discrimination and reduce confusion. Compared with FMCNN, TSCNN performs poorly in extracting building damage with complete areas and smooth boundaries, which may be related to not considering object-based predictions. FMCNN performs well for independent and complete damaged buildings but does not perform well in extracting building damage for partially damaged buildings or in high-density building areas. For example, for test image patches in Figure 11 row 5, FMCNN recognizes the entire large size building as building damage, and for test image patches in Figure 11 row 7, FMCNN shows false negative for the partially damaged building. In addition, the building damage results extracted by such methods are largely affected by the quality of building footprint maps manually interpreted from pre-disaster HRRS images.

U-Net performs the best among the ten comparison results, with rather smooth boundaries and high accuracy in extracting building damage in both low-density and high-density building areas. In comparison, the results of our proposed method are not visually inferior to U-Net and slightly outperform in some scenes, such as Figure 11 row 10.

5. Discussion

5.1. Scene classification

Scene classification performance provides the basic guarantee for the proposed method to accurately identify building damage. To demonstrate the advantage of the proposed method, we also assess and compare the performance of image-level scene classification among SqueezeNet, VGG-16, ResNet-101, WResNet, and WMSNet using F_1 -score and OA. WResNet represents the model that the dilation rates for the parallel paths in the WMS block of WMSNet are all set to 1. The quantitative results are shown in [Table 4](#), with bold values representing the highest values for the different evaluation metrics among the five models. The confusion matrices of different models are shown in [Figure 12](#). It can be observed that the proposed WMSNet obtains the highest values of F_1 -score and OA when compared to other models, achieving the best overall scene classification performance. In particular, WResNet still achieves the second best overall scene classification performance despite the lack of the multi-scale component (parallel dilation convolutions), indicating that our wider network structure is effective.

5.2. Ablation study

In order to further explore the contribution of different processing strategies (i.e. pre-disaster image super-pixel segmentation (SPS) and pre-disaster building semantic segmentation (BSS)) in improving the performance of building damage extraction, we conduct ablation experiments on our dataset. For ease of comparison, we firstly construct the baseline on the basis of WMSNet. In the baseline, we utilize WMSNet to extract building damage from post-disaster HRSS images through sliding window strategy. On the basis of the baseline, combining the results of super-pixel segmentation of pre-disaster images is called baseline + SPS. The auxiliary prediction that is combined with pre-disaster building semantic segmentation results is called baseline + BSS.

[Table 5](#) reports quantitative results for method with different strategy combinations, including the baseline, baseline + SPS, baseline + BSS, and baseline + SPS + BSS. Values in bold represent the highest values for the different evaluation metrics in the ablation study. On the basis of the baseline, the quantitative evaluation of the contributions of different strategies can be reflected in [Table 5](#). As can be observed, both the SPS and the BSS can produce a positive impact on the performance, and the baseline + SPS + BSS, i.e. the proposed approach, achieves the best precision, recall, F_1 -score,

Table 4. F_1 (%) and OA (%) of SqueezeNet, VGG-16, ResNet-101, WResNet, and WMSNet on image-level scene classification.

Methods	Building damage F_1 (%)	Intact building F_1 (%)	Impervious surface F_1 (%)	Vegetation F_1 (%)	Bare land F_1 (%)	Mean F_1 (%)	OA (%)
SqueezeNet	86.36	85.95	96.84	97.32	93.79	92.05	92.00
VGG-16	86.81	87.65	96.83	97.68	94.59	92.71	92.67
ResNet-101	84.97	85.08	95.97	96.43	88.28	90.15	90.13
WResNet	88.40	87.04	96.30	98.16	95.06	92.99	92.93
WMSNet	92.52	92.38	97.68	98.50	96.62	95.54	95.53

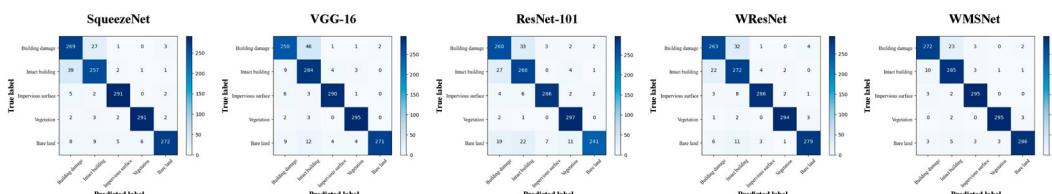


Figure 12. Confusion matrices of SqueezeNet, VGG-16, ResNet-101, WResNet, and WMSNet on image-level scene classification.

Table 5. Precision (%), recall (%), F_1 (%), MIoU (%), and OA (%) of various processing setting.

Methods	Precision (%)	Recall (%)	F_1 (%)	MIoU (%)	OA (%)
baseline	46.73	65.32	54.48	63.92	90.93
baseline + SPS	47.66	66.60	55.56	64.54	91.14
baseline + BSS	70.44	60.11	64.87	71.15	94.59
baseline + SPS + BSS	71.41	72.79	72.09	75.69	95.31

MIoU, and OA. However, it is worth noting that adding only SPS can slightly improve the performance, while adding only BSS has a significant improvement in evaluation metrics. It can be inferred that the primary difficulty of building damage extraction task is to solve the problem of confusion with surrounding objects. Moreover, compared with baseline + BSS, baseline + SPS + BSS contributes 0.97%, 12.68%, 7.22%, 4.54%, and 0.72% to the increment in precision, recall, F_1 -score, MIoU, and OA, respectively, which indicates that the addition of both SPS and BSS can achieve the best building damage extraction performance.

Some experimental results, including the test image patches and their corresponding pixel-wise building damage extraction maps, are shown in Figure 13. Compared with the baseline, baseline + SPS is more effective in optimizing the damage boundary of buildings, and the results are smoother and more complete. Nevertheless, it has no significant effect on reducing *FP*. Baseline + BSS is much better at distinguishing building damage from similar non-building areas and obtains relatively accurate building damage extraction results. However, there are many *FNs* in the building damage extraction results. We note that by benefiting from the advantages of incorporating SPS and BSS, the building damage extraction results of our proposed method can perform the best.

6. Conclusion

The intricate visual characteristics exhibited by building damage areas in HRRS imageries poses a significant challenge for extracting such information from the images. On one hand, approaches relying exclusively on post-disaster images for damage localization lack the necessary building boundary guidance, leading to diminished accuracy. Furthermore, achieving pixel-level interpretation results hinges on the availability of meticulous annotations. On the other hand, change detection methods utilizing dual-temporal imagery are susceptible to introducing spurious changes. To address these challenges, this paper revolutionizes building damage detection from two aspects. Firstly, this paper introduces a novel framework to leverage both pre- and post-disaster images. An FCN-based method is used to extract architectural features from pre-disaster images, while a CNN-based technique is employed for the extraction of building damage information from post-disaster images. The knowledge derived from the former is subsequently integrated to enhance and complement the latter process. Secondly, the proposed approach operates under weak supervision, primarily demonstrated by our sole reliance on image-level annotations to label building damage information for CNN model training. Nonetheless, it ultimately produces pixel-level building damage results from post-disaster imagery. Extensive experiments on the 2010 Haiti earthquake datasets demonstrate the effectiveness and robustness of the proposed methods.

Furthermore, we believe that the method framework we have established possesses a certain level of generality and scalability. In other words, the FCN model, superpixel algorithm, and CNN approach featured in our method could be substituted with similar alternatives. However, the manner in which we utilize pre-event and post-event images, as well as the strategy of weak supervision, exhibit transferability, ensuring that changes to local methods do not render them ineffective.

Due to geographical and temporal heterogeneity, diverse disaster scenarios, and the uniqueness of available post-disaster HRRS images for different regions and times, the practical target domain for building damage detection often exhibits significant data distribution disparities compared to

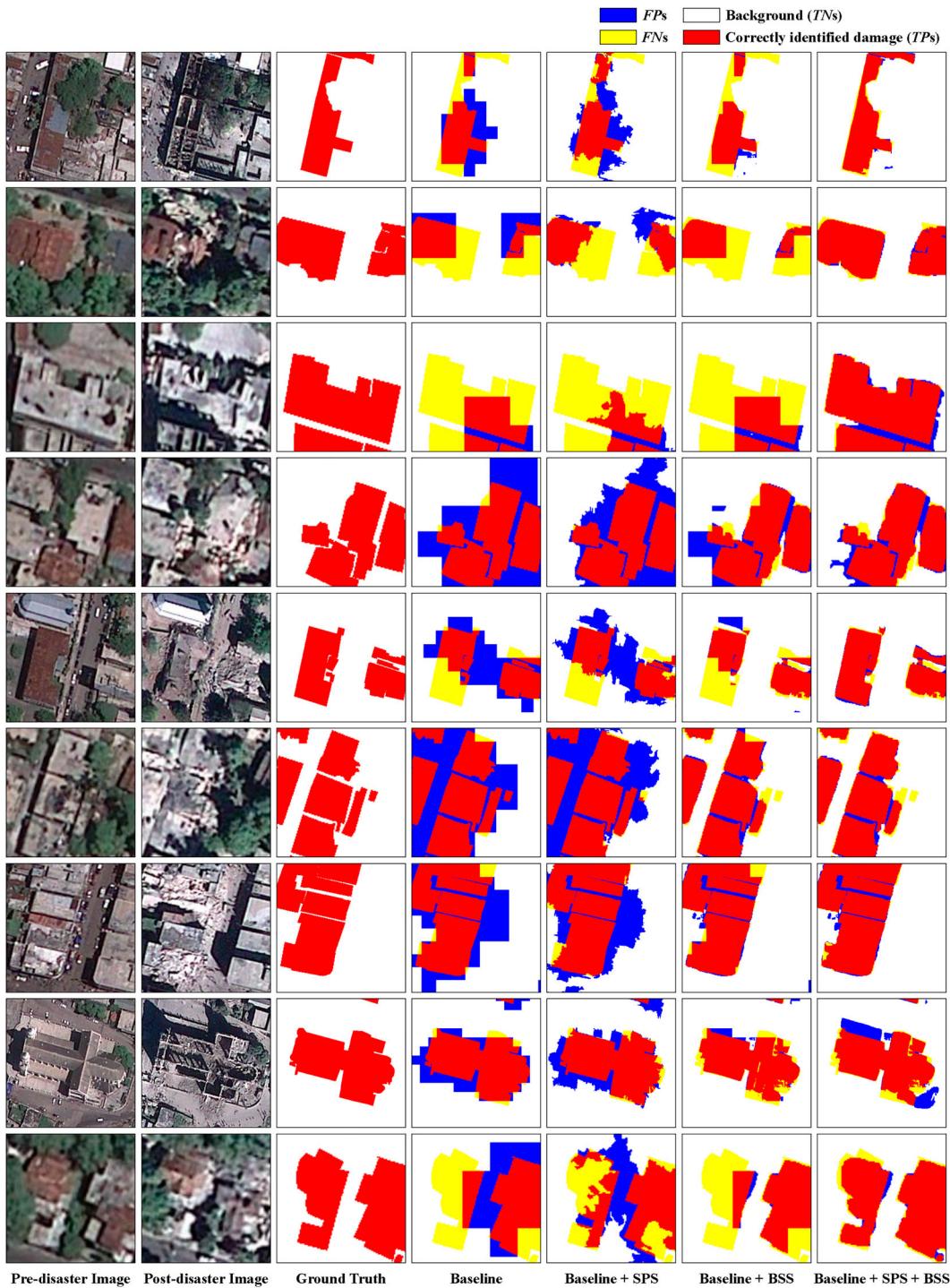


Figure 13. Qualitative comparison of building damage identification results of various processing setting.

the source domain of constructed sample sets. As a result, models learned from the training sample set struggle to be directly transferred and adapted to real-world target tasks. In our future work, we will be dedicated to studying the domain adaptation challenges involved in building damage

detection from HRRS images. The proposed approach and its future enhancements are anticipated to enhance the capacity for building damage assessment, disaster prevention, and mitigation, contributing towards the fulfillment of Sustainable Development Goal 11.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [Grant Number 42071386, 41930104 and 41871283]; the National Key Research and Development Program of China [Grant Number 2016YFB0501403]; the MiaoZi Project of Sichuan Province in China [Grant Number 2019013].

Data availability statement

The data used in this study are available by contacting the corresponding author.

ORCID

Wenfan Qiao  <http://orcid.org/0000-0003-4555-5731>

Li Shen  <http://orcid.org/0000-0003-4638-3329>

References

- Abriha, Dávid, and Szilárd Szabó. 2023. "Strategies in Training Deep Learning Models to Extract Building from Multisource Images with Small Training Sample Sizes." *International Journal of Digital Earth* 16 (1): 1707–1724. <https://doi.org/10.1080/17538947.2023.2210312>.
- Ajmar, Andrea, Simone Balbo, Piero Boccardo, Fabio Giulio Tonolo, Marco Piras, and Jan Princic. 2013. "A Low-Cost Mobile Mapping System (LCMMS) for Field Data Acquisition: a Potential Use to Validate Aerial/satellite Building Damage Assessment." *International Journal of Digital Earth* 6 (sup2): 103–123. <https://doi.org/10.1080/17538947.2011.638991>.
- Ali, Muhammad Usman, Waqas Sultani, and Mohsen Ali. 2020. "Destruction from Sky: Weakly Supervised Approach for Destruction Detection in Satellite Imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 162:115–124. <https://doi.org/10.1016/j.isprsjprs.2020.02.002>.
- Amirebrahimi, Sam, Abbas Rajabifard, Priyan Mendis, and Tuan Ngo. 2016. "A Framework for a Microscale Flood Damage Assessment and Visualization for a Building Using BIM–GIS Integration." *International Journal of Digital Earth* 9 (4): 363–386. <https://doi.org/10.1080/17538947.2015.1034201>.
- Bai, Yanbing, Chang Gao, Sameer Singh, Magaly Koch, Bruno Adriano, Erick Mas, and Shunichi Koshimura. 2017. "A Framework of Rapid Regional Tsunami Damage Recognition from Post-Event TerraSAR-X Imagery Using Deep Neural Networks." *IEEE Geoscience and Remote Sensing Letters* 15 (1): 43–47. <https://doi.org/10.1109/LGRS.2017.2772349>.
- Bai, Yanbing, Erick Mas, and Shunichi Koshimura. 2018. "Towards Operational Satellite-Based Damage-Mapping Using U-net Convolutional Network: A Case Study of 2011 Tohoku Earthquake-Tsunami." *Remote Sensing* 10 (10): 1626. <https://doi.org/10.3390/rs10101626>.
- Bialas, James, Thomas Oommen, Umaa Rebbaapragada, and Eugene Levin. 2016. "Object-Based Classification of Earthquake Damage from High-Resolution Optical Imagery Using Machine Learning." *Journal of Applied Remote Sensing* 10 (3): 036025. <https://doi.org/10.1117/1.JRS.10.036025>.
- Buda, Mateusz, Atsuto Maki, and Maciej A. Mazurowski. 2018. "A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks." *Neural Networks* 106:249–259. <https://doi.org/10.1016/j.neunet.2018.07.011>.
- Cha, Young-Jin, Wooram Choi, and Oral Büyüköztürk. 2017. "Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks." *Computer-Aided Civil and Infrastructure Engineering* 32 (5): 361–378. <https://doi.org/10.1111/mice.12263>.
- Chini, M, F. R. Cinti, and S. Stramondo. 2011. "Co-seismic Surface Effects from Very High Resolution Panchromatic Images: the Case of the 2005 Kashmir (Pakistan) Earthquake." *Natural Hazards and Earth System Sciences* 11 (3): 931–943. <https://doi.org/10.5194/nhess-11-931-2011>.

- Dong, Laigen, and Jie Shan. 2013. "A Comprehensive Review of Earthquake-induced Building Damage Detection with Remote Sensing Techniques." *ISPRS Journal of Photogrammetry and Remote Sensing* 84:85–99. <https://doi.org/10.1016/j.isprsjprs.2013.06.011>.
- Duarte, Diogo, Francesco Nex, Norman Kerle, and George Vosselman. 2018. "Multi-resolution Feature Fusion for Image Classification of Building Damages with Convolutional Neural Networks." *Remote Sensing* 10 (10): 1636. <https://doi.org/10.3390/rs10101636>.
- Dubois, David, and Richard Lepage. 2014. "Fast and Efficient Evaluation of Building Damage from Very High Resolution Optical Satellite Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (10): 4167–4176. <https://doi.org/10.1109/JSTARS.2014.2336236>.
- Ehrlich, Daniele, H. D. Guo, Katrin Molch, J. W. Ma, and Martino Pesaresi. 2009. "Identifying Damage Caused by the 2008 Wenchuan Earthquake from VHR Remote Sensing Data." *International Journal of Digital Earth* 2 (4): 309–326. <https://doi.org/10.1080/17538940902767401>.
- Fan, Yida, Qi Wen, Wei Wang, Ping Wang, Lingling Li, and Peng Zhang. 2017. "Quantifying Disaster Physical Damage Using Remote Sensing Data—A Technical Work Flow and Case Study of the 2014 Ludian Earthquake in China." *International Journal of Disaster Risk Science* 8 (4): 471–488. <https://doi.org/10.1007/s13753-017-0143-8>.
- Felzenswalb, Pedro F., and Daniel P. Huttenlocher. 2004. "Efficient Graph-Based Image Segmentation." *International Journal of Computer Vision* 59 (2): 167–181. <https://doi.org/10.1023/B:VISI.0000022288.19776.77>.
- Fiedrich, Frank, Fritz Gehbauer, and Uwe Rickers. 2000. "Optimized Resource Allocation for Emergency Response After Earthquake Disasters." *Safety Science* 35 (1-3): 41–57. [https://doi.org/10.1016/S0925-7535\(00\)00021-7](https://doi.org/10.1016/S0925-7535(00)00021-7).
- Ge, Pinglan, Hideomi Gokon, and Kimiro Meguro. 2020. "A Review on Synthetic Aperture Radar-Based Building Damage Assessment in Disasters." *Remote Sensing of Environment* 240:111693. <https://doi.org/10.1016/j.rse.2020.111693>.
- Ghaffarian, Saman, Norman Kerle, Edoardo Pasolli, and Jamal Jokar Arsanjani. 2019. "Post-Disaster Building Database Updating Using Automated Deep Learning: An Integration of Pre-disaster OpenStreetMap and Multi-temporal Satellite Data." *Remote Sensing* 11 (20): 2427. <https://doi.org/10.3390/rs11202427>.
- Grünthal, Gottfried. 1998. *European macroseismic scale 1998*. Technical Report. European Seismological Commission (ESC).
- Gu, Jiuxiang, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, et al. 2018. "Recent Advances in Convolutional Neural Networks." *Pattern Recognition* 77:354–377. <https://doi.org/10.1016/j.patcog.2017.10.013>.
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. "Deep Residual Learning for Image Recognition." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.
- Iandola, Forrest N., Song Han, Matthew W Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer. 2016. "SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and 0.5 MB Model Size." Preprint, arXiv:1602.07360. <https://doi.org/10.48550/arXiv.1602.07360>.
- Janalipour, Milad, and Ali Mohammadzadeh. 2015. "Building Damage Detection Using Object-Based Image Analysis and ANFIS from High-Resolution Image (Case Study: BAM Earthquake, Iran)." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (5): 1937–1945. <https://doi.org/10.1109/JSTARS.2015.2458582>.
- Janalipour, Milad, and Ali Mohammadzadeh. 2018. "Evaluation of Effectiveness of Three Fuzzy Systems and Three Texture Extraction Methods for Building Damage Detection from Post-Event LiDAR Data." *International Journal of Digital Earth* 11 (12): 1241–1268. <https://doi.org/10.1080/17538947.2017.1387818>.
- Janalipour, Milad, and Mohammad Taleai. 2017. "Building Change Detection After Earthquake Using Multi-criteria Decision Analysis Based on Extracted Information from High Spatial Resolution Satellite Images." *International Journal of Remote Sensing* 38 (1): 82–99. <https://doi.org/10.1080/01431161.2016.1259673>.
- Ji, Min, Lanfa Liu, and Manfred Buchroithner. 2018. "Identifying Collapsed Buildings Using Post-earthquake Satellite Imagery and Convolutional Neural Networks: A Case Study of the 2010 Haiti Earthquake." *Remote Sensing* 10 (11): 1689. <https://doi.org/10.3390/rs10111689>.
- Ji, Min, Lanfa Liu, Runlin Du, and Manfred F. Buchroithner. 2019. "A Comparative Study of Texture and Convolutional Neural Network Features for Detecting Collapsed Buildings After Earthquakes Using Pre-and Post-Event Satellite Imagery." *Remote Sensing* 11 (10): 1202. <https://doi.org/10.3390/rs1101202>.
- Ji, Shunping, Shiqing Wei, and Meng Lu. 2018. "Fully Convolutional Networks for Multisource Building Extraction from an Open Aerial and Satellite Imagery Data Set." *IEEE Transactions on Geoscience and Remote Sensing* 57 (1): 574–586. <https://doi.org/10.1109/TGRS.2018.2858817>.
- Kerle, Norman. 2010. "Satellite-Based Damage Mapping Following the 2006 Indonesia Earthquake—How Accurate was it?." *International Journal of Applied Earth Observation and Geoinformation* 12 (6): 466–476. <https://doi.org/10.1016/j.jag.2010.07.004>.
- Korkmaz, Kasim A., and Munther Abualkibash. 2018. "Earthquake Damage Detection Using Before and After Earthquake Satellite Images." In *Proceedings of the 2018 IEEE International Conference on Electro/Information Technology (EIT)*, 0615–0619. <https://doi.org/10.1109/EIT.2018.8500225>.

- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2012. "Imagenet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems* 25: 1097–1105.
- LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep Learning." *Nature* 521 (7553): 436–444. <https://doi.org/10.1038/nature14539>.
- Li, Liwei, Bing Zhang, and Yuanfeng Wu. 2012. "Fusing Spectral and Texture Information for Collapsed Buildings Detection in Airborne Image." In *Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 186–189. <https://doi.org/10.1109/IGARSS.2012.6351606>.
- Lin, Chen, Yundong Li, Yi Liu, Xiang Wang, and Shuo Geng. 2021. "Building Damage Assessment from Post-hurricane Imageries Using Unsupervised Domain Adaptation with Enhanced Feature Discrimination." *IEEE Transactions on Geoscience and Remote Sensing* 60:1–10. <https://doi.org/10.1109/TGRS.2021.3054869>.
- Long, Jonathan, Evan Shelhamer, and Trevor Darrell. 2015. "Fully Convolutional Networks for Semantic Segmentation." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431–3440.
- Ma, Haojie, Yalan Liu, Yuhuan Ren, Dacheng Wang, Linjun Yu, and Jingxian Yu. 2020. "Improved CNN Classification Method for Groups of Buildings Damaged by Earthquake, Based on High Resolution Remote Sensing Images." *Remote Sensing* 12 (2): 260. <https://doi.org/10.3390/rs12020260>.
- Ma, Haojie, Yalan Liu, Yuhuan Ren, and Jingxian Yu. 2019. "Detection of Collapsed Buildings in Post-earthquake Remote Sensing Images Based on the Improved YOLOv3." *Remote Sensing* 12 (1): 44. <https://doi.org/10.3390/rs12010044>.
- Maggiori, Emmanuel, Yuliya Tarabalka, Guillaume Charpiat, and Pierre Alliez. 2017. "Can Semantic Labeling Methods Generalize to any City? The Inria Aerial Image Labeling Benchmark." In *Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 3226–3229. <https://doi.org/10.1109/IGARSS.2017.8127684>.
- Minh, Hoang-Le, Samir Khatir, R. Venkata Rao, Magd Abdel Wahab, and Thanh Cuong-Le. 2023. "A Variable Velocity Strategy Particle Swarm Optimization Algorithm (VVS-PSO) for Damage Assessment in Structures." *Engineering with Computers* 39 (2): 1055–1084. <https://doi.org/10.1007/s00366-021-01451-2>.
- Minh, Hoang-Le, Thanh Sang-To, Samir Khatir, Magd Abdel Wahab, and Thanh Cuong-Le. 2023. "Damage Identification in High-rise Concrete Structures Using a Bio-inspired Meta-heuristic Optimization Algorithm." *Advances in Engineering Software* 176:103399. <https://doi.org/10.1016/j.advengsoft.2022.103399>.
- Moya, Luis, Abdul Muhari, Bruno Adriano, Shunichi Koshimura, Erick Mas, Luis R. Marval-Perez, and Naoto Yokoya. 2020. "Detecting Urban Changes Using Phase Correlation and L1-Based Sparse Model for Early Disaster Response: A Case Study of the 2018 Sulawesi Indonesia Earthquake-Tsunami." *Remote Sensing of Environment* 242:111743. <https://doi.org/10.1016/j.rse.2020.111743>.
- Moya, Luis, Homa Zakeri, Fumio Yamazaki, Wen Liu, Erick Mas, and Shunichi Koshimura. 2019. "3D Gray Level Co-occurrence Matrix and Its Application to Identifying Collapsed Buildings." *ISPRS Journal of Photogrammetry and Remote Sensing* 149:14–28. <https://doi.org/10.1016/j.isprsjprs.2019.01.008>.
- Naito, Shohei, Hiromitsu Tomozawa, Yuji Mori, Takeshi Nagata, Naokazu Monma, Hiromitsu Nakamura, Hiroyuki Fujiwara, and Gaku Shoji. 2020. "Building-Damage Detection Method Based on Machine Learning Utilizing Aerial Photographs of the Kumamoto Earthquake." *Earthquake Spectra* 36 (3): 1166–1187. <https://doi.org/10.1177/8755293019901309>.
- Nex, Francesco, Diogo Duarte, Anne Steenbeek, and Norman Kerle. 2019. "Towards Real-Time Building Damage Mapping with Low-Cost UAV Solutions." *Remote Sensing* 11 (3): 287. <https://doi.org/10.3390/rs11030287>.
- Qiao, Wenfan, Li Shen, Jicheng Wang, Yungang Cao, Shi He, and Yanshuai Dai. 2019. "A Fine-Grained Fully Convolutional Network for Extraction of Building Along High-Speed Rail Lines from VHR Remote Sensing Image." In *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 1244–1247. <https://doi.org/10.1109/IGARSS.2019.8899095>.
- Qiao, Wenfan, Li Shen, Jicheng Wang, Xiaotian Yang, and Zhilin Li. 2023. "A Weakly Supervised Semantic Segmentation Approach for Damaged Building Extraction from Postearthquake High-Resolution Remote-Sensing Images." *IEEE Geoscience and Remote Sensing Letters* 20:1–5. <https://doi.org/10.1109/LGRS.2023.3243575>.
- Sang-To, Thanh, Hoang Le-Minh, Magd Abdel Wahab, and Cuong-Le Thanh. 2023. "A New Metaheuristic Algorithm: Shrimp and Goby Association Search Algorithm and Its Application for Damage Identification in Large-scale and Complex Structures." *Advances in Engineering Software* 176:103363. <https://doi.org/10.1016/j.advengsoft.2022.103363>.
- Sarp, Gulcan, Arzu Erener, Sebnem Duzgun, and Kemal Sahin. 2014. "An Approach for Detection of Buildings and Changes in Buildings Using Orthophotos and Point Clouds: A Case Study of Van Erriş Earthquake." *European Journal of Remote Sensing* 47 (1): 627–642. <https://doi.org/10.5721/EuJRS20144735>.
- Shen, Yu, Sijie Zhu, Taojiannan Yang, and Chen Chen. 2020. "Cross-Directional Feature Fusion Network for Building Damage Assessment from Satellite Imagery." Preprint, arXiv:2010.14014. <https://doi.org/10.48550/arXiv.2010.14014>.
- Simonyan, Karen, and Andrew Zisserman. 2014. "Very Deep Convolutional Networks for Large-scale Image Recognition." Preprint, arXiv:1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>.

- Song, Dongmei, Xuan Tan, Bin Wang, Ling Zhang, Xinjian Shan, and Jianyong Cui. 2020. "Integration of Super-pixel Segmentation and Deep-learning Methods for Evaluating Earthquake-damaged Buildings Using Single-phase Remote Sensing Imagery." *International Journal of Remote Sensing* 41 (3): 1040–1066. <https://doi.org/10.1080/01431161.2019.1655175>.
- Sublime, Jérémie, and Ekaterina Kalinicheva. 2019. "Automatic Post-Disaster Damage Mapping Using Deep-learning Techniques for Change Detection: Case Study of the Tohoku Tsunami." *Remote Sensing* 11 (9): 1123. <https://doi.org/10.3390/rs11091123>.
- Sun, Genyun, Yanling Hao, Jun Rong, Shuna Shi, and Jinchang Ren. 2017. "Combined Deep Learning and Multiscale Segmentation for Rapid High Resolution Damage Mapping." In *Proceedings of the 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 1101–1105. <https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2017.238>.
- Sun, Li, Yuqi Tang, and Liangpei Zhang. 2017. "Rural Building Detection in High-Resolution Imagery Based on a Two-Stage CNN Model." *IEEE Geoscience and Remote Sensing Letters* 14 (11): 1998–2002. <https://doi.org/10.1109/LGRS.2017.2745900>.
- Tiachacht, Samir, Samir Khatir, Cuong Le Thanh, Ravipudi Venkata Rao, Seyedali Mirjalili, and Magd Abdel Wahab. 2021. "Inverse Problem for Dynamic Structural Health Monitoring Based on Slime Mould Algorithm." *Engineering with Computers*, 1–24. <https://doi.org/10.1007/s00366-021-01378-8>.
- Tong, Xiaohua, Xiaofei Lin, Tiantian Feng, Huan Xie, Shijie Liu, Zhonghua Hong, and Peng Chen. 2013. "Use of Shadows for Detection of Earthquake-Induced Collapsed Buildings in High-Resolution Satellite Imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 79:53–67. <https://doi.org/10.1016/j.isprsjprs.2013.01.012>.
- Tu, Jihui, Haigang Sui, Wenqing Feng, Kaimin Sun, and Li Hua. 2016. "Detection of Damaged Rooftop Areas from High-Resolution Aerial Images Based on Visual Bag-of-words Model." *IEEE Geoscience and Remote Sensing Letters* 13 (12): 1817–1821. <https://doi.org/10.1109/LGRS.2016.2614298>.
- Turker, Mustafa, and Emre Sumer. 2008. "Building-Based Damage Detection Due to Earthquake Using the Watershed Segmentation of the Post-Event Aerial Images." *International Journal of Remote Sensing* 29 (11): 3073–3089. <https://doi.org/10.1080/01431160701442096>.
- Vettrivel, Anand, Markus Gerke, Norman Kerle, Francesco Nex, and George Vosselman. 2018. "Disaster Damage Detection Through Synergistic Use of Deep Learning and 3D Point Cloud Features Derived from Very High Resolution Oblique Aerial Images, and Multiple-Kernel-Learning." *ISPRS Journal of Photogrammetry and Remote Sensing* 140:45–59. <https://doi.org/10.1016/j.isprsjprs.2017.03.001>.
- Vu, Tuong Thuy, and Yifang Ban. 2010. "Context-Based Mapping of Damaged Buildings from High-Resolution Optical Satellite Images." *International Journal of Remote Sensing* 31 (13): 3411–3425. <https://doi.org/10.1080/01431161003727697>.
- Wang, Xue, and Peijun Li. 2020. "Extraction of Urban Building Damage Using Spectral, Height and Corner Information from VHR Satellite Images and Airborne LiDAR Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 159:322–336. <https://doi.org/10.1016/j.isprsjprs.2019.11.028>.
- Wang, Yude, Jie Zhang, Meina Kan, Shiguang Shan, and Xilin Chen. 2020. "Self-supervised Equivariant Attention Mechanism for Weakly Supervised Semantic Segmentation." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 12275–12284.
- Weber, Ethan, and Hassan Kané. 2020. "Building Disaster Damage Assessment in Satellite Imagery with Multi-temporal Fusion." Preprint, arXiv:2004.05525. <https://doi.org/10.48550/arXiv.2004.05525>.
- Wei, Yunchao, Huaxin Xiao, Honghui Shi, Zequn Jie, Jiashi Feng, and Thomas S. Huang. 2018. "Revisiting Dilated Convolution: A Simple Approach for Weakly-and Semi-supervised Semantic Segmentation." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7268–7277.
- Yamazaki, Fumio, Yoshihisa Yano, and Masashi Matsuoka. 2005. "Visual Damage Interpretation of Buildings in Bam City Using QuickBird Images Following the 2003 Bam, Iran, Earthquake." *Earthquake Spectra* 21 (1_suppl): 329–336. <https://doi.org/10.1193/1.2101807>.
- Ye, Xin, Mingchao Liu, Jun Wang, Qiming Qin, Huazhong Ren, Jianhua Wang, and Jian Hui. 2017. "Building-Based Damage Detection from Postquake Image Using Multiple-feature Analysis." *IEEE Geoscience and Remote Sensing Letters* 14 (4): 499–503. <https://doi.org/10.1109/LGRS.2017.2651050>.
- YiFei, Li, Hoang-Le Minh, Samir Khatir, Thanh Sang-To, Thanh Cuong-Le, Cao MaoSen, and Magd Abdel Wahab. 2023. "Structure Damage Identification in Dams Using Sparse Polynomial Chaos Expansion Combined with Hybrid K-means Clustering Optimizer and Genetic Algorithm." *Engineering Structures* 283:115891. <https://doi.org/10.1016/j.engstruct.2023.115891>.
- Zagoruyko, Sergey, and Nikos Komodakis. 2016. "Wide Residual Networks." Preprint, arXiv:1605.07146. <https://doi.org/10.48550/arXiv.1605.07146>.
- Zheng, Zhuo, Yanfei Zhong, Junjue Wang, Ailong Ma, and Liangpei Zhang. 2021. "Building Damage Assessment for Rapid Disaster Response with a Deep Object-Based Semantic Change Detection Framework: From Natural Disasters to Man-made Disasters." *Remote Sensing of Environment* 265:112636. <https://doi.org/10.1016/j.rse.2021.112636>.