A METHOD OF DISASTER-DAMAGED BUILDING SAMPLE LIBRARY CONSTRUCTION BASED ON TRANSFER LEARNING

Zeyu Wang [1 2], Feng Zhang [1 2 *], Chuyi Wu [1,2]

[1] **School of Earth Sciences, Zhejiang University, Hangzhou 310028, China - 3180104062@zju.edu.cn; zfcarnation@zju.edu.cn; wuchuyi@zju.edu.cn**

[2] **Zhejiang Provincial Key Laboratory of Geographic Information Science, Hangzhou 310028, China**

KEY WORDS: Remote Sensing, Disaster, Transfer Learning, Building Damage, Sample Library, U-Net

ABSTRACT:

In recent years, the analysis of satellite remote sensing data using deep learning methods has become an important solution in the field of disaster response. The sample library can provide knowledge support for deep learning tasks in a research area, and its sample size and labeling accuracy directly determine the model effectiveness. However, the traditional sample library is often constructed by manual annotation, which is time-consuming and less automated. This research explores a transfer learning method to automate the construction of a sample library of building damage. We used models instead of manual work to automatically annotate the damage information of disaster images, combined with an auxiliary manual inspection to achieve high quality and dynamic optimization of the sample library. Following are three main aspects of research work. (1) This study extensively collected remote sensing images related to disaster events, and carried out data pre-processing work. (2) This study built a building damage information identification model based on the U-Net framework, which applied pre-trained backbone model and was formally trained on xBD dataset. (3) With CycleGAN to implement color style transfer between RS images collected from two different data sources, we constructed a high-quality building damage sample library, mainly by automatic machine labeling and supplemented by manual sampling verification. Finally, we successfully constructed a sample library of disaster-damaged building images, and conducted several rounds of manual precision sampling work and dynamic optimization work. The number of image slices in the sample library reached 12,834, and the overall accuracy is higher than 0.80 under the manual inspection of 200 samples.

1. INRTODUCTION

According to a report published by the World Health Organization, there has been an alarming increase in the number of natural disasters worldwide over the past two decades, causing a global economic loss of approximately $2.97 trillion and affecting more than 4.2 billion people (WHO, 2020).

Remote sensing data can play an important role in all aspects of disaster emergency management: including monitoring and early warning before the occurrence of a disaster, emergency mapping of the disaster in the early stage of the disaster, and assistance program development after the disaster situation is stabilized. By tracking the whole process of disasters, remote sensing data can provide a series of decision support for disaster prevention, preparedness, and relief, effectively mitigate the negative impacts of disasters, and improve the accuracy and timeliness of disaster response (Boccardo et al., 2015; Said et al., 2019). Among them, 72 hours after the occurrence of a major natural disaster is the golden rescue time for emergency response, and emergency responders need remote sensing images based on the affected area to conduct rapid disaster emergency assessment and mapping. Compared with manual annotation and damage level calibration by image analysts, automatic extraction of damaged areas from images will be more efficient, and human errors can be avoided and mapping quality can be controlled.

The damage and collapse of buildings is the main form of damage caused by natural disasters to human settlement areas, and is an important basis for assessing the distribution of the affected population and the extent of damage in the area. Automated identification of disaster damage information from satellite remote sensing data requires the following conditions: (1) remote sensing images covering the scope of the affected area; (2) automated disaster damage labeling method with high accuracy rate; and (3) a complete sample library of disaster remote sensing images as the standard for damage location and level assessment.

There are several publicly available datasets for building change detection and damage detection, which are generally released along with research papers from universities and public AI competitions. (1) xBD dataset, contains more than 20,000 disaster slices under 7 disaster types and 19 natural disaster events acquired by Worldview-3 satellite photography images (Gupta et al., 2019); (2) the ABCD (AIST Building Change Detection) dataset for constructing and evaluating building damage detection models, containing approximately 8,500 pre- and post-disaster image pairs from the tsunami-affected region of Japan (Fujita et al., 2017); and (3) WHU building dataset and LEVIR-CD dataset are used for building change detection tasks (Ji et al., 2019; Chen and Shi., 2020).

At present, the publicly available datasets that can be applied to the task of building disaster damage information extraction have the following shortcomings.

(1) the number of datasets is small and the scale is small. The number of benchmark datasets that can support deep learning tasks in scenarios in the field of building disaster information extraction is small and the dataset size is also small. Several datasets related to change detection and disaster information extraction tasks have a big gap with the xBD dataset in terms of data scale and collection scope, while the xBD dataset also has a difference of hundreds of times in data

* Corresponding author

volume with ImageNet, the largest image recognition database in the field of machine learning (Deng et al., 2009).

(2) The degree of automation of dataset construction is low. Datasets in the domain are basically produced by manual annotation, so the construction and publication of large datasets, such as the xBD dataset, often require high costs. In contrast, a dataset or sample can be built more quickly by intelligent machine recognition, followed by manual review and fine-tuning, but there are few attempts to do so.

Although transfer learning methods have been widely used in the field of deep learning, most of them still serve for the transfer of models to application scenarios, and less for the construction of datasets. In this paper, we aim to automate the construction of a building damage sample library for typical natural disasters by using transfer learning methods. We mainly select images of hurricane disasters in North America for building damage samples, so that the sample library can be typical and transferable.

2. DATA

2.1 Data source

In this research, the data collected in the process of automating the construction of the building damage sample library consists of two main parts. One part is used to train the deep learning model for building damage extraction, which needs to use the officially released benchmark dataset; the other part is used as the raw materials for building the damage sample library, which needs to collect remote sensing images related with multiple natural disasters.

2.1.1 Benchmark Dataset for Training Models

The xBD dataset produced and released by MIT in collaboration with the U.S. Department of Defense in 2019 was adopted as the benchmark dataset for training the building damage information extraction model (Gupta et al., 2019). xBD dataset is derived from remote sensing images taken by Worldview-3 satellite with 0.3 m ground resolution, collecting 19 natural disaster events such as hurricanes, earthquakes, floods, wildfires, volcanic eruptions, and tsunamis. , volcanic eruptions, tsunamis, and other natural disaster types under 19 natural disaster events, a total of more than 20,000 pre- and post-disaster satellite images, and location markers and damage level calibration of 550,000 buildings. As the largest and highest quality dataset in the field of natural disaster information identification, the xBD dataset integrates the opinions of several official organizations and experts to propose certain criteria for building damage rating, as shown in Table 1 and Figure 1.

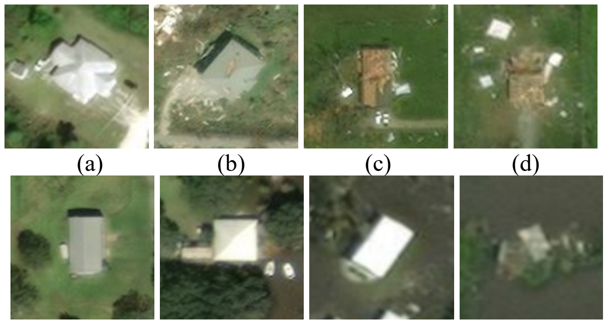
2.1.2 Materials for Sample Library Construction

In order to build a large sample library of building damage, it is necessary to collect a wide range of images of major natural disasters around the world, and to locate and rate buildings based on diachronic images of the affected areas.

The xBD dataset data sources, damage rating criteria, annotation methods and data organization methods were used to construct the sample library. Most of the source data was downloaded from the Open-Data project on the Maxar website, which publishes remote sensing imagery and layers related to multiple natural hazards under the Creative Commons 4.0 licence and allows organizations and people involved in disaster response to open source access, download data from the

Damage level	Description of the building damage appearance
No damage	Undisturbed. No water, structural damage, residual rubble or signs of burn marks.
Minor damage	The building is partially burned and surrounded by visible water structures, volcanic flows, roof debris or cracks, etc.
Major damage	Partially collapsed walls or roofs, eroded by volcanic flows, or surrounded by water/mud.
Destroyed	Burnt, completely collapsed, completely covered by water/mud, or no longer visible.

Table 1. Building damage level and appearance description



(a) (b) (c) (d)
(e) (f) (g) (h)

Figure 1. Building morphology under each disaster damage level. (a)(e)No damage; (b)(f)Minor damage; (c)(g)Major damage; (d)(h)Destroyed

platform, and integrate with other existing disaster response technologies. The list of the natural disasters for which the original images were collected is shown in Table 2. After collecting hundreds of raw images, the images were screened and retained through manual visual inspection following the following criteria: (1) the images covered the affected area and the post-disaster images had obvious building damage information; (2) the affected area where the images were located had more complete pre-disaster and post-disaster dual time phase images; and (3) the images were RGB color maps with spatial resolution ≤ 0.5m, better color balance and lower cloud coverage.

Disaster Name	Time	Location
Hurricane Dorian	Sept. 01, 2019	The USA (Islands of Abaco and Grand Bahama)
Hurricane Delta	Oct. 08, 2020	The USA (Louisiana and Texas)
Hurricane Ida	Oct. 08, 2020	The USA (Louisiana and Mississippi)
Hurricane Laura	Aug. 26, 2020	The USA (Louisiana and Texas)
Hurricane Irma	Sept. 06, 2017	Caribbean
Mount Semeru Eruption	Dec. 04, 2021	Indonesia (Java Island)
Haiti Earthquake	Jan. 12, 2010	Haiti

Table 2. Disaster list of collecting images

2.2 Data pre-processing

Effective methods are needed to process the collected disaster images to better serve downstream tasks such as model training and sample library construction. These include image pre-processing, image slicing and sample processing.

(1) We adopt *ArcGIS* Desktop software to carry out image pre-processing on dual-temporal disaster images. When the satellite company released the remote sensing images, the basic radiation correction and atmospheric correction of the images had been completed, but the spatial resolution of the images differed from one frame to another and there were certain deviations in spatial location; after geographic calibration, the positions of the spatial entities in the pre-disaster and post-disaster dual-temporal images could be made consistent; after cropping and resampling, the spatial extent of the pre-disaster and post-disaster samples obtained by subsequent slicing could be made the same.

(2) Slicing of images. A large image with spatial information in tif format is sliced into png format with a pixel size of 512× 512, which can be input into the model in batches for convolution; at the same time, the image name needs to contain information such as disaster name, pre- and post-disaster temporal phase and image number.

(3) Some image processing is performed on the original samples and the prediction results output from the model to reduce image noise. A dark channel a priori-defogging algorithm was used to remove thin clouds from the images (He, K., 2009); median filtering and open-close operations were used to post-process the segmentation maps output by the model to fill in the small internal holes of the buildings and smooth out the broken contours.

In the accuracy verification session of the sample library, the manually annotated true value samples need to be compared with the machine annotated predicted samples. In this paper, the open source image labeling tool *Labelme* is used to label the locations of buildings in the form of polygons with different levels of damage as labels. In the manual labeling process, the building outlines are firstly labeled on the pre-disaster images, and the target objects in the pre-disaster images are all non-damaged by default; then the damage level labels are determined by comparing the pre-disaster and post-disaster images.

3. METHODS

3.1 U-Net for building damage information extraction model

The model training and testing used in this study was based on the open source code of the xView2 competition winner as the baseline (Weber and Kané, 2020). And the Attention U-Net from (Wu et al., 2021) was also used to train the building damage information extraction model.

3.1.1 U-Net

U-Net is one of the standard convolutional neural network structures for change detection tasks, which combines low-resolution information that can distinguish classes and high-resolution information that provides a basis for accurate segmentation localization (Ronneberger et al., 2015). In the task of disaster damage information recognition studied in this paper, the model needs to first determine whether each pixel point of the image belongs to a building for accurate segmentation based on the pre-disaster image, and then determine the damage level of the target object in the post-disaster image based on the dual-time phase images of the pre-disaster and post-disaster.

The sampling process of U-Net essentially uses convolutional kernels to learn image feature mappings, so a CNN structure with more learning capability can be used as the backbone network of the U-Net encoder. Based on the xView2-first-place code, four CNN network structures, ResNet34, ResNeXt50, ResNeXt154, and Dpn92, were mainly selected as the backbone network of U-Net in this study (He et al., 2016; Xie et al., 2017; Chen et al., 2017) , the channel attention mechanism module (Hu et al., 2019) was added to the network, and the target object localization model and the disaster damage level classification model were trained under the disaster damage information extraction task, respectively.

3.1.2 Building Damage Information Extraction Model

We trained a building damage information extraction model on the xBD dataset, so as to automatically extract building damage information from our collected images. The whole process is showed in Figure 2.

The xBD dataset were divided into training set (90%) and validation set (10%) using random seeds 0,1,2 respectively. For each copy of data, four types of backbone-replaced U-Net model was trained and tested according to the following process: (1) training the building location model; (2) fine-tune the parameters of the localization model; (3) train the building damage classification model based on the previously trained localization model; (4) fine-tune the parameters of the damage classification model; and (5) evaluate the accuracy of the model based on the test set.

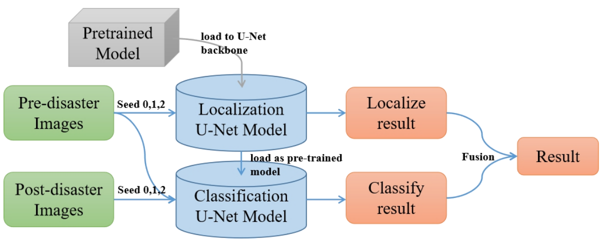


Figure 2. Overall framework of model training and prediction

3.2 Pre-trained Model and Fine-tuning

Transfer learning refers to the application of deep learning models in which an already trained model is transferred to a new application field. The following are two main types of fine-tuning-based transfer learning used in training the building damage information extraction model.

3.2.1 Loading pre-trained model on U-Net backbone network

The disaster damage information extraction models trained in this paper were loaded with the official pre-training models provided by Pytorch before training on the xBD dataset, and the backbone networks used included four types of ResNet34, ResNeXt50, ResNeXt154, and Dpn92. The model parameters were obtained by pre-training on large-scale images of ImageNet, which has the ability to extract the basic feature information of images.It can reduce the model's requirement for the sample size of the training set and speed up the model convergence.

As the pre-trained model is part of the overall U-Net structure, it is necessary to determine whether each network layer is common to both the target model and the pre-trained model

https://doi.org/10.5194/isprs-archives-XLVIII-3-W1-2022-65-2022

when loading, and only load the parameters when they are common, while the parameters of the remaining network layers are still randomly initialized. During the formal training of the model, images of the disaster samples and annotation masks are input, so the model is transferred to the disaster information extraction task.

3.2.2 Transfer the localization model to the damage classification model

Both the building localization task and the damage classification task are pixel-level semantic segmentation tasks, which require the recognition and extraction of semantic information of buildings in images. In this paper, we first train the building localization model U-Net to complete the image pixel binary classification to extract building location information, and then load the localization model onto Double U-Net as the initial weight of the damage level classification model to further identify the damage level of building structure and environment based on the pre-disaster and post-disaster dual-time images.

3.3 Sample color style transfer based on GAN

In the process of transfer learning to solve real-world problems, one of the common strategies is to achieve knowledge transfer through data adaptation and transformation. In this paper, the GAN method is applied to color style transfer of disaster-related images, thereby improving the overall recognition accuracy of the model for images from different data sources.

3.3.1 Generative Adversarial Networks

A GAN consists of a generator and a discriminator, with the generator generating fake images and the discriminator attempting to distinguish the difference between fake and real images (Goodfellow et al., 2014). And the cyclic generative adversarial network consists of a pair of generators and a pair of discriminators with the same structure (Zhu et al., 2017). Based on the cyclic structure, CycleGAN can automatically transform images from source domain X to target domain Y without the need for pairs of training data. Training data from different data sources or different regions varies greatly, and models trained on the source domain sample set may generalize poorly to the target domain sample set. CycleGAN can make the distribution of these two types of data as close as possible, thus improving the model's ability to generalize to new data without affecting the performance of existing models.

3.3.2 Style transfer between different satellite source image data

The essence of classifying building damage levels is to perform image change detection, which requires the preservation of the original features of the target object while minimizing differences in geographic environment, image color style, etc. in images of different time phases. The disaster-related images collected in this paper include the disaster image data provided by the Maxar website and the historical images downloaded from the Google Earth platform; among them, the downloaded historical Google Earth images are from the original images taken by different satellites, and their color styles differ significantly from the images obtained from the Worldview satellite on the Maxar website. There are significant differences between the color style of the downloaded Google Earth historical images and the Worldview satellite images from the Maxar website. Meanwhile, the building damage information extraction model in this paper is obtained based on the xBD dataset, which is also based on the original images from the Worldview series of satellites, and the model has been extensively learned on the Worldview satellite style images. As showed in Figure 3, we mainly use the CycleGAN method for data style transfer to convert the color style of the pre-disaster images collected from Google Earth to the post-disaster images collected from Maxar website, in order to reduce the non-target object differences between the pre-disaster and post-disaster images and improve the model's disaster damage recognition accuracy. The baseline code: https://github.com/junyanz/CycleGAN.

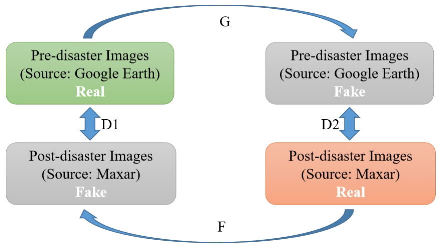


Figure 3. The CycleGAN structure to implement color style transfer. G and F: generators. D1and D2: discriminators.

3.4 Accuracy Assessment

3.4.1 Model accuracy assessment

In this paper, common semantic segmentation accuracy evaluation metrics such as Dice coefficient, F1 score, IoU and MIoU are used to evaluate the accuracy of building disaster damage models on the xBD dataset.

In the application scenario of building damage information recognition in this paper, the accuracy assessment consists of three components: the binary classification accuracy of the localization model, the multi-classification accuracy of the damage classification model, and the combined accuracy of the model. (i)Localization model accuracy: take F1 score to evaluate the binary accuracy of the model. (ii) Classification model accuracy: take the summed average of F1 scores (Equation 1), and the summed mean of MIoU (Equation 2) to comprehensively evaluate the accuracy performance of the model under each damage class. (iii)Overall accuracy: assign a weight of 0.3 and 0.7 to localization accuracy and damage classification accuracy respectively and then sum.

$$F1_{\text{total}} = 4/(\frac{1}{F1_{\text{no\_damage}}} + \frac{1}{F1_{\text{minor\_damage}}} + \frac{1}{F1_{\text{major\_damage}}} + \frac{1}{F1_{\text{destroyed}}}) \quad (1)$$

$$MIoU_{\text{total}} = 4/(\frac{1}{MIoU_{\text{no\_damage}}} + \frac{1}{MIoU_{\text{minor\_damage}}} + \frac{1}{MIoU_{\text{major\_damage}}} + \frac{1}{MIoU_{\text{destroyed}}}) \quad (2)$$

3.4.2 Sample Library Accuracy Assessment and Optimization

For the sample library constructed by automatic machine annotation, there are several methods of accuracy verification as follows: (1) External sampling verification: according to the proportion of the number of samples under each disaster, a random sample of samples for each disaster is taken, and the true value of manual annotation is used as the benchmark to calculate accuracy; (2) Internal cross-validation: to check the balance of internal samples, the training set and validation set are randomly divided into internal multiple times based on the machine annotated samples to train. (3) External test set validation: Based on the model trained by the machine labeled samples, the xBD dataset is used as the validation set to observe the accuracy of the model and indirectly test the quality of the sample library.

The overall process of dynamic data replenishment and dynamic model optimization is used to construct a large-scale building disaster damage sample library. The samples obtained through automatic machine annotation need to be verified, and if the accuracy meets the requirements, the annotated batch of samples will be stored; if the accuracy is lower than the requirements, the samples with poor annotation accuracy need to be optimized for bad case analysis, and the model annotation method is optimized for the analysis results, thus continuously cycling the process of sample annotation, accuracy verification and method optimization.

4. RESULTS

4.1 Performance of disaster damage information extraction model

The experiments were conducted based on the xBD dataset to train the U-Net model for building localization and the U-Net model for disaster classification under each of the four backbone networks, and three random seeds were used for each model to divide the training and validation sets, and the performance of the F1 scores of the obtained models is shown in Table 3. The model using ResNeXt50 as the backbone network obtained the best results for the building localization and disaster classification tasks, with F1 scores of 0.888 and 0.774, respectively; ResNeXt154 and Dpn92 also performed better, and the model using ResNet34 as the backbone network obtained the lowest F1 scores. Therefore, ResNeXt50, which has the best accuracy, was mainly used as the backbone network of the U-Net model in the machine labeling stage of the sample library.

Backbone	**Localization (Single U-Net）**		
	seed = 0	seed = 1	seed = 2
ResNet34	0.863	0.856	0.843
ResNeXt50	0.888	0.867	0.873
Dpn92	0.885	0.863	0.873
ResNeXt154	0.879	0.859	0.865
Backbone	**Classification (Double U-Net)**		
	seed = 0	seed = 1	seed = 2
ResNet34	0.727	0.691	0.709
ResNeXt50	0.757	0.774	0.763
Dpn92	0.753	0.721	0.752
ResNeXt154	0.755	0.749	0.765

Table 3. F1 score of U-Net model under four backbone models

We further test the generalization ability of the mode. The Test test set and Holdout retention set of the xBD dataset were used as test samples to train the ResNeXt50 U-Net model with three random seeds, respectively, and the prediction results output from the three models were fused to obtain the final output. According to the test results in Table 4, the F1 score of the disaster damage information extraction model obtained from the training is 0.869 on the building localization task and 0.764 on the building damage class classification task, which is consistent with the accuracy performance of the model on the validation set during the training process, and there is no over-fitting problem. Among the four damage classes, the model has better recognition ability for buildings under the categories of no damage and total damage, and the worst recognition ability for buildings under the category of light damage.

In order to check the model's ability to extract disaster damage information for new natural disasters, samples of Dorian hurricane disasters outside the xBD dataset were selected and input to the model for building extraction and damage class classification, and the model could effectively output prediction results. Taking Figure 4 as an example, the building localization model can better identify the building outlines in the sample and distinguish them from objects such as roads and boats; the damage level classification model can better detect the changes based on the pre-disaster and post-disaster dual-time phase image samples and identify the damage levels of different buildings.

Metrics		**F1 Score**	**MIoU**
Localization		0.869	0.769
Damage	Overall	0.764	0.618
	No damage (Level 1)	0.918	0.849
	Minor damage (Level 2)	0.607	0.436
	Major damage (Level 3)	0.756	0.608
	Destroyed (Level 4)	0.849	0.741

Table 4. F1 score and MIoU of the model on test set

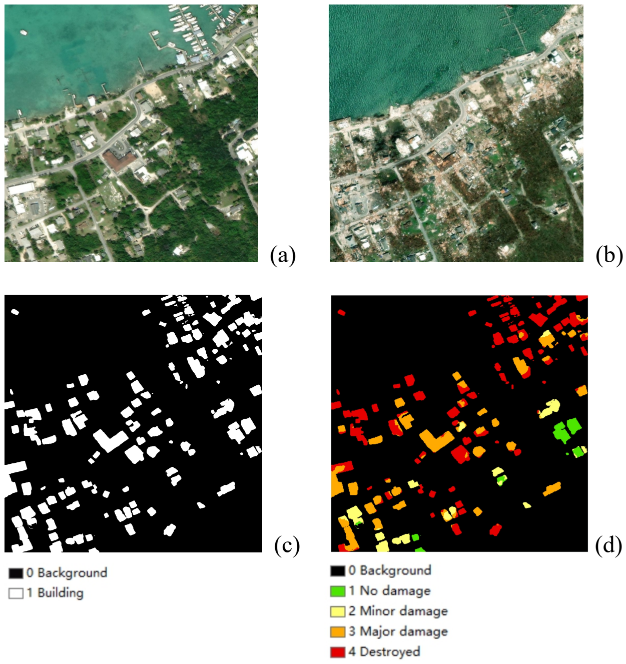


Figure 4. Model prediction results of Hurricane Dorian sample. (a) pre-disaster image; (b) post-disaster image; (c) predicted building localization map; (d) predicted building damage level result map.

4.2 Sample accuracy improvement due to GAN method

In this experiment, the pre-disaster and post-disaster images of Hurricane Delta are selected as data to test the accuracy improvement of disaster damage recognition task brought by CycleGAN method after style transfer of data from different satellite remote sensing sources. The disaster occurred in the North American Gulf Coast region, and the post-disaster images were obtained from the Open-data project of Maxar website, with greenish tones; the pre-disaster images were historical images downloaded from Google Earth platform, with red-brown tones. The pre-disaster image is used as the source domain and transferred to the target domain image style of the post-disaster image; 500 images are selected for each of the pre-disaster and post-disaster, and U-Net256 is used as the generator

https://doi.org/10.5194/isprs-archives-XLVIII-3-W1-2022-65-2022

of CycleGAN, and 80 epochs are trained to obtain the results as shown in Figure 5.

To verify the accuracy improvement effect of the style-converted images for the disaster damage information recognition task, the original pre-disaster images and the style-converted pre-disaster images, respectively, were fed into the disaster damage level classification model together with the post-disaster images. 56 pairs (512×512) of these images were randomly selected for manual truth annotation, and the accuracy of the prediction results before and after the style conversion was calculated. The results show that the prediction accuracy of the model is **0.622** for the post-styling sample and **0.603** for the original sample; the generalization ability of the model on the new sample is effectively improved by the style transfer between different satellite source images through CycleGAN network.


(a)


(b)

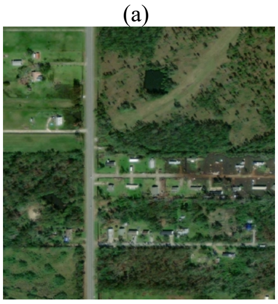

(c)

Figure 5. Example graph of the results of image style transfer by CycleGAN. (a) pre-disaster image before transfer(Real); (b) pre-disaster image transferred to post-disaster style (Fake); (c) post-disaster image of the same region (Real)

4.3 Overall accuracy check and optimization of the sample library

A U-Net model with the backbone network of ResNeXt50 was used, and image slice samples of five hurricane disasters were input. The model was run for about 1.5 hours, and the disaster damage sample library constructed automatically contained 12,384 512×512 image slices. The accuracy of the constructed sample library was mainly verified by external sampling: a certain percentage of samples were taken under each disaster for manual annotation and accuracy checking, with a total of 200 samples. The accuracy of building location labeling for this batch of samples was obtained as 0.885, the accuracy of building damage level classification labeling was 0.769, and the overall accuracy was 0.804, as shown in Table 5. Further sample replenishment and multidimensional accuracy verification are needed to improve the overall accuracy of the sample database.

Disaster	Total number	Inspected number	Localization accuracy	Classification accuracy
Hurricane Dorian	1080	24	0.878	0.741
Hurricane Delta	2760	32	0.874	0.749
Hurricane Ida	5320	80	0.897	0.846
Hurricane Laura	1600	32	0.901	0.717
Hurricane Irma	1624	32	0.834	0.626
Total	**12384**	**200**	**0.885**	**0.769**

Table 5. Sample library accuracy inspection

5. DISCUSSION

5.1 The role of transfer learning in sample library construction

We loaded the model obtained by pre-training on ImageNet as a backbone network into the house localization model and loaded the building localization model as a pre-trained model to the building damage level classification model. This approach avoids training the model from scratch and speeds up the convergence of the model, while reducing the required sample size and the training cost of the model. The test results show that the model trained on the xBD dataset can effectively achieve the task of house contour extraction.

Meanwhile, CycleGAN is used for color style transfer between different data sources. The disaster damage extraction model was trained on xBD dataset which has the color style of images from Maxar platform. After the color style transfer of historical images collected from Google Earth platform, the model showed a higher accuracy on the newly collected samples.

5.2 Automatic labeling with manual inspection

The building damage sample library constructed by automatic machine annotation contains more than 12,000 sliced images, and the overall accuracy is above 0.8 after manual sampling. In the process of dynamically constructing a large-scale sample library, we continuously cycle the process of automatic machine annotation, manual sampling accuracy, bad case analysis, and optimization of annotation methods, so as to continuously improve the quality and usability of the sample library. In fact, the automatic construction of the sample database is still inseparable from the necessary manual intervention. After the model has completed the task of extracting building damage information from disaster images, it unavoidably contains some mislabels of building damage levels, which needs to be verified by manual sampling to verify its usability.

6. CONCLUSION

In this paper, transfer learning approaches are used to automate the construction of a building damage sample library. The pre-trained model were loaded to U-Net backbone network and formally trained on the xBD dataset, allowing the model to perform the disaster damage information extraction task. ResNeXt50 works best as the backbone network among all backbones, and the model can have good generalization ability on hurricane disaster samples outside the xBD dataset. Meanwhile, the CycleGAN method is used to make Google Earth's color style consistent with Maxar. and the accuracy validation work is performed on the obtained disaster damage sample library. The F1 Score and MIoU are used as the accuracy evaluation indexes of the model. The disaster damage sample library constructed by automation includes 12,384 512 × 512 image slices, and the accuracy of localization and classification of disaster damage level are 0.885 and 0.769 respectively after manual accuracy sampling.

Our sample library can be applied to automatic damage information identification tasks as base dataset. Also, in the face of new natural disasters, the model transfer and sample transfer methods proposed in this research can improve the semantic recognition of remote sensing images by the disaster damage information extraction model. For developing countries which lack damage analysis capability, this library can effectively serve in the natural disaster mitigation response activities.

In subsequent study, we will continue to collect RS images of disasters from multiple platforms. Also, we will explore the methods of transferring existing model to other disaster damage types, such as forest fires, road inundation, landslides, etc. The sample library will be public and accessible together with the publication of paper.

REFERENCES

Boccardo, P., Giulio Tonolo, F., 2015. Remote Sensing Role in Emergency Mapping for Disaster Response, in: Lollino, G., Manconi, A., Guzzetti, F., Culshaw, M., Bobrowsky, P., Luino, F. (Eds.), Engineering Geology for Society and Territory - Volume 5. Springer International Publishing, Cham, pp. 17–24. https://doi.org/10.1007/978-3-319-09048-1_3

Chen, H., Shi, Z., 2020. A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. Remote Sensing 12, 1662. https://doi.org/10.3390/rs12101662

Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S., Feng, J., 2017. Dual Path Networks. arXiv:1707.01629 [cs].

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, Li Fei-Fei, 2009. ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL, pp. 248 – 255. https://doi.org/10.1109/CVPR.2009.5206848

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Networks. arXiv:1406.2661 [cs, stat].

Fujita, A., Sakurada, K., Imaizumi, T., Ito, R., Hikosaka, S., Nakamura, R., 2017. Damage detection from aerial images via convolutional neural networks, in: 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). IEEE, Nagoya, Japan, pp. 5 – 8. https://doi.org/10.23919/MVA.2017.7986759

Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J., Heim, E., Choset, H., Gaston, M., 2019. xBD: A Dataset for Assessing Building Damage from Satellite Imagery. arXiv:1911.09296 [cs].

He, K., Sun, J., Tang, X., 2011. Single Image Haze Removal Using Dark Channel Prior. IEEE Trans. Pattern Anal. Mach. Intell. 33, 2341–2353. https://doi.org/10.1109/TPAMI.2010.168

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition. arXiv:1512.03385 [cs].

Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E., 2019. Squeeze-and-Excitation Networks. arXiv:1709.01507 [cs].

Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q., 2018. Densely Connected Convolutional Networks. arXiv:1608.06993 [cs].

Ji, S., Wei, S., Lu, M., 2019. Fully Convolutional Networks for Multisource Building Extraction From an Open Aerial and Satellite Imagery Data Set. IEEE Trans. Geosci. Remote Sensing 57, 574 – 586. https://doi.org/10.1109/TGRS.2018.2858817

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 234 – 241. https://doi.org/10.1007/978-3-319-24574-4_28

Said, N., Ahmad, K., Riegler, M., Pogorelov, K., Hassan, L., Ahmad, N., Conci, N., 2019. Natural disasters detection in social media and satellite imagery: a survey. Multimed Tools Appl 78, 31267 – 31302. https://doi.org/10.1007/s11042-019-07942-1

Weber, E., Kané, H., 2020. Building Disaster Damage Assessment in Satellite Imagery with Multi-Temporal Fusion. arXiv:2004.05525 [cs].

Wu, C., Zhang, F., Xia, J., Xu, Y., Li, G., Xie, J., Du, Z., Liu, R., 2021. Building Damage Detection Using U-Net with Attention Mechanism from Pre- and Post-Disaster Remote Sensing Datasets. Remote Sensing 13, 905. https://doi.org/10.3390/rs13050905

Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. arXiv:1703.10593 [cs].

World Health Organization 2020: The human cost of disasters: an overview of the last 20 years (2000-2019).