Pixel-Level Identification of Post-Earthquake Building Collapses Based on Satellite Remote Sensing Images

Can Dong

Zaozhuang Engineering Quality and Safety Service Center, Jinan, China

zzjwdc@163.com

Wenyin Song, and Rui Liu

Shandong Quality Inspection and Testing Center of Construction Engineering Co., Ltd, Jinan, China

[b]389036081@qq.com, [c] lr0317liurui@163.com

Abstract: Due to the obvious diversity and complexity of damage patterns, geometries, and spatial scales of urban building complex earthquake hazards, conventional identification and assessment methods are less generalizable in real post-earthquake scenarios. Compared with time-range signals such as kinetic acceleration, image/video data provide a new source of perceptual information for accurately assessing the post-earthquake damage of urban building complexes. To realize the integrated, comprehensive, and rapid identification and assessment of the structural damage of urban buildings after an earthquake, this paper proposes a geometrically constrained deep learning framework for seismic damage identification and assessment of buildings based on computer vision. It systematically carries out a multi-scale seismic damage identification and assessment method that associates the “building group, building unit, and structural component" with the "building group, building unit, and structural component”. This paper proposes a geometrically constrained deep learning framework for seismic damage identification and assessment based on computer vision and systematically researches the identification and assessment of “building groups-building units-structural components”. A method for fine identification of densely distributed small-target buildings and rapid assessment of the collapse state after earthquake based on satellite remote sensing images at high altitudes is proposed. A semantic segmentation network for post-earthquake building cluster identification and assessment is built, the influence law of the weight coefficients of the GCE loss on the segmentation performance of the model is systematically investigated, and the geometric feature optimization performance of the GCE loss in the training process and the multi-level feature extraction ability are analyzed, which verifies the effectiveness and accuracy of the geometrically constrained deep learning method for multi-scale seismic damage identification of buildings proposed in this paper.

Keywords: post-earthquake building damage assessment, geometrically constrained deep learning, semantic segmentation, computer vision, damage patterns

1 Introduction

With the continuous innovation and development of image processing algorithms in the field of computer vision, researchers have begun to apply them to the task of building earthquake damage assessment based on remotely sensed optical images [1].

The building damage dataset of the real earthquake scene is very limited, and the feature information of the building data affected by the earthquake usually has obvious imbalance characteristics, and the collapsed or severely damaged buildings are generally much less than the slightly damaged or intact buildings. The imbalance of the training data makes it difficult for the model to learn the image features of the imbalanced categories, which ultimately leads to poor recognition performance of the model. Considering the limited data used for model training, three data balancing methods, random oversampling, random sampling and cost-sensitive sampling, are used to improve the accuracy of building collapse classification for the unbalanced building damage data. A two-stage evaluation method for building earthquake damage combining the improved YOLOv4 target detection algorithm and the support vector machine-based classification algorithm is proposed, as shown in Figure 1.

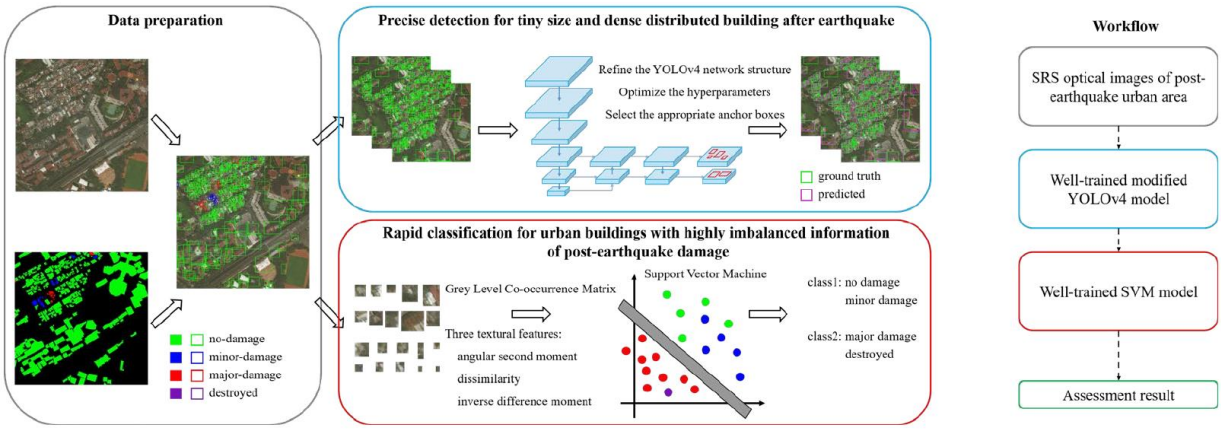

Fig. 1 Framework for two-stage assessment of post-earthquake building damage

2 Remote sensing image dataset establishment and hyperparameter setting

In this study, the seismic damage dataset of urban building complexes used includes 34 remote sensing optical images of damaged urban areas after typical earthquakes in two places in China, with the image resolution of 4800×2748. 24 images in the dataset were obtained from Yushu, Qinghai Province, China (Yushu Earthquake, 2010), and 10 images were obtained from Beichuan Old County, Sichuan Province, China (Hanchuan Earthquake, 2008) [2]. For the corresponding original image, real labelling production in the two places is based on the results of post-earthquake field investigation and visual interpretation of remote sensing images. The damage status of the post-earthquake buildings in the two cities was classified into two categories: collapsed and unclasped, and the post-earthquake building targets were labeled using polygonal boxes. Among them, buildings with complete or partial collapse, serious damage to structural components, and buildings with destructive forms of roofs were labeled as the collapsed category, and the rest of the buildings were labeled as the non-collapsed category. The representative original images and visualized label maps of urban building complexes after the Yushu and Beichuan earthquakes are shown in Figure 2.

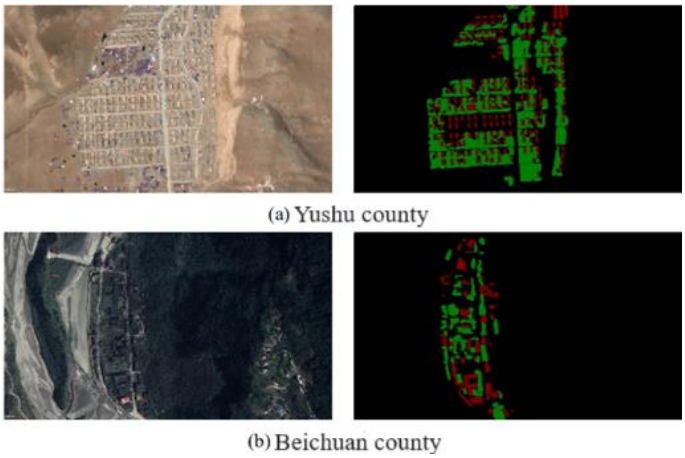

Fig. 2 Representative original image and corresponding visualization label map

Twenty-eight images were randomly selected for model network training, and the remaining six were used for model recognition ability testing. Since directly inputting high-resolution original images into the model faces challenges in terms of training cost and computer hardware, while down-sampling the original high-resolution large image (4800 × 2748) to a smaller size image inevitably loses most of the information. In order to meet the requirements of model training image size, effectively retain the key feature information of the image, and comprehensively consider the computer configuration conditions, model training efficiency and detection speed, 256 × 256 is selected as the model image input size. Eventually, the 28 original training images formed 18,620 sub-image sets, which were divided into the training set (13,034) and validation set (5,586) according to the ratio of 7:3. The six original images used for testing were cropped without overlap using a sliding step of 256 × 256 sub-images, and stitched to an initial size of 4608 × 2560 after the sub-images were input to the model for prediction [3].

3 Post-earthquake building cluster identification assessment results and validation study

3.1 GCE loss sensitivity analysis and impact law study

in order to obtain the optimal weight coefficients of the three subitems inside the GC loss, the GC loss term a was selected as 0.05, and the weight coefficients of the three subitems inside it were increased or decreased by 10 times, respectively, and a total of six sets of experimental optimization analyses were conducted. As shown in Figure 3, when the weight coefficients of each subitem are set to 1, each subitem presents a synergistic optimization ability, which in turn obtains the optimal validation accuracy (mIoU=86.98%).

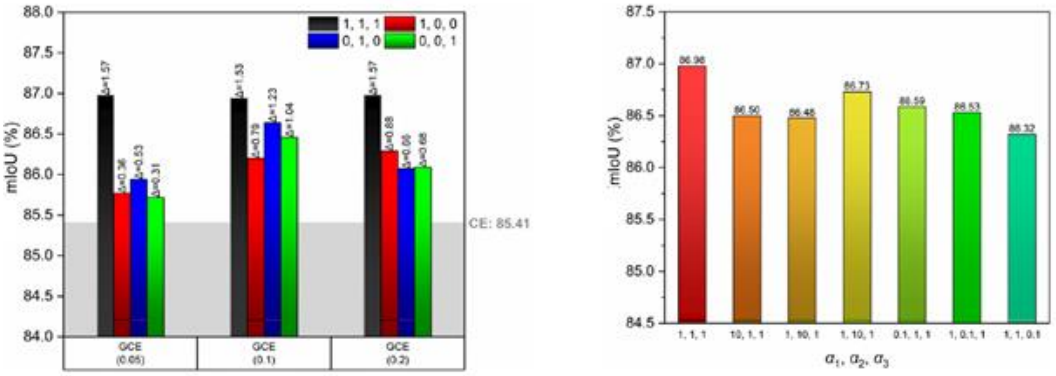

Fig. 3 Comparison of validation set accuracy with different GCE loss weight coefficients

The results show that the training mIoU and validation mIoU of the model show a gradual increase with the increase in the number of training rounds and finally converge to a stable trend, the training loss and validation loss in the early training of the 50 rounds of a sharp decline, followed by the downward trend of the curves tends to flatten out; the loss is no longer decreasing and reaches convergence after 100 rounds of training, the training loss and the validation loss are 0.0666 and 0.1060, respectively [4].

3.2Actual earthquake scenario modeling test results

In order to deeply reveal the superiority of the network model developed in this chapter for post-earthquake dense small-target building recognition and collapse state assessment, Figure 4 shows the local zoomed sub-images of the high-resolution test image (size 256 × 256. The prediction results of the model using the three sub-terms of GC loss alone in GCE loss are also shown as a comparative reference. The results show that the synergistic contribution of the three subcomponents of GC loss endows the optimized model with excellent segmentation capability for fine detail recognition compared to the baseline model using CE loss and the optimized model using the GC loss subcomponents alone [5]. In particular, for collapsed buildings with complex boundary shapes and poor geometric integrity, the three GC loss sub-terms work together to reduce misrecognition of the background predicted as a building, as well as building miss-detection for image boundary locations, promote edge closure and smoothing of the target building, and suppress misrecognition of the interior of the target building region. The results show that by using the geometric consistency enhancement loss function to guide the model optimization in the training phase, the model's boundary and region segmentation performance for earthquake-damaged building targets with complex geometric features is significantly improved, and accurate and reliable pixel-level identification and collapse state assessment of dense small targets in post-earthquake wide-area building clusters can be achieved.

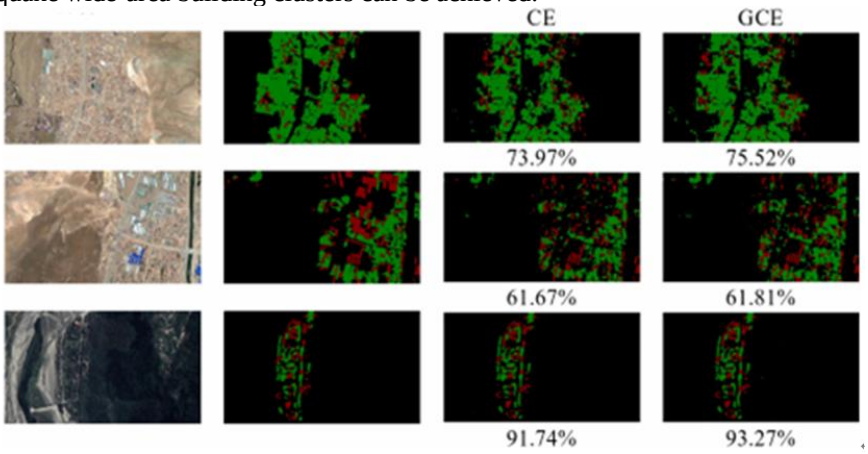

Fig. 4 Prediction results of the optimal and baseline models for the original high-resolution image

3.3 GCE loss feature extraction capability analysis

In order to further explore the feature extraction capability of GCE loss in the model training phase, Figure 5 shows the results of visualizing the feature maps for six subgraphs during the training process. The feature layers of all channels in the last convolutional and activation layer of each level of the coding structure in the network after 100 rounds of training of the optimal model and the baseline model are extracted, and then these five sets of channel feature layers are fused into a single channel feature layer respectively [6]. The analysis results show that the trained model mainly extracts shape, edge and other underlying features at the low level of the network; with the depth and deepening of the network, the model can extract abstract semantic feature information that is difficult for humans to understand and capture intuitively. Compared with the baseline model trained with CE loss only, the optimal model trained with GCE loss shows a higher degree of feature map activation at the deepest coding network level (layer 5), indicating that its feature extraction ability for damaged buildings in earthquake scenarios is significantly improved [7].

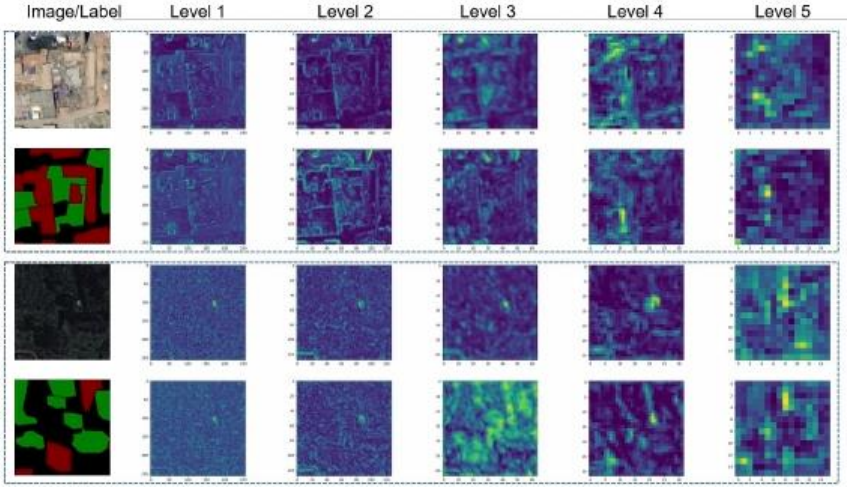

Figure. 5 Visualization of feature maps at different levels of the optimal

The results show that the GCE loss optimized by the weight coefficients can collaboratively use the geometric features of the segmentation line and segmentation region to constrain the model training optimization and achieved a validation accuracy of 86.98%, which has a 1.57% enhancement compared to the use of CE loss alone.

4 Post-earthquake building identification assessment and localization methods

4.1Adaptive Encoder and Decoder Network Architecture

In this paper, QuakeCityNet-M-N (QCNet), a semantic segmentation network with flexibly configurable encoding and decoding structures, is designed based on the traditional U-Net encoder-one-decoder architecture. Where, M and N represent the number of encoder layers in the encoding structure and the number of convolutional layers from its third to the bottom layer, respectively. This chapter focuses on nine QCNet-M-N networks with different network volumes, three types of encoder hierarchies including 4, 5, and 6, and three types of convolutional layers from the third layer to the bottom layer including 2, 3, and 4. The structure of the QCNet-M-N network with nine configurations is shown in Figure 6 [8].

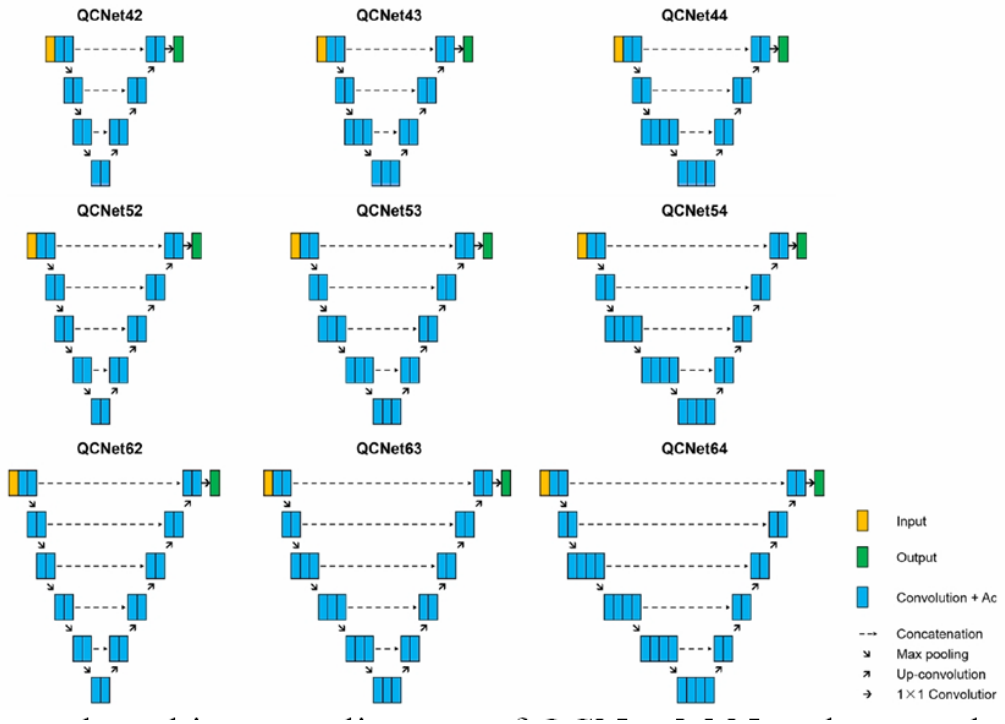

Fig. 6 General network architecture diagram of QCNet-M-N and network structure diagram

4.3Algorithm for building localization in post-earthquake region based on UAV images

In this experiment, four paths parallel to the facade are preset as scanning paths, and a one-transmitter-two-receiver ultra-wideband stepping-frequency radar system deployed in two unmanned vehicles is used to scan the scene [9]. The actual geographic locations of the buildings in the post-earthquake region in the images are obtained by using the Gaode Map API to obtain the key and calling the Python dictionary to construct the parameter request parameters, which are solved using geographic/inverse geographic coding. Table 1 shows the localization results of this method for four actual post-earthquake region buildings (taking the center position of each image as an example).

Table 1 Results of building localization in post-earthquake region based on UAV images

Image	Calculate Location	Actual Location	Geographic Location
	104.4577, 31.8302	104.4587, 31. 8300	Xiqiang Shangjie, Qushan Town, Beichuan Qiang Autonomous County, Mianyang City, Sichuan Province, China
	105.3585, 29.2128	105.3579, 29.2129	National Highway 321, Fuji Town, Hu County, Huzhou City, Sichuan Province, China
	102.9932, 30.3393	102.9953, 30.3381	Ludeng Road, Taiping Town, Lushan County, Ya'an City, Sichuan Province, China
	102.8104, 30.3714	102.8114, 30.3710	Yanjiang Road, Muping Town, Baoxing County, Ya'an City, Sichuan Province, China

This chapter proposes a post-earthquake building identification and damage state assessment method for complex scenes based on UAV images, and builds an adaptive encoder-one-decoder network architecture (QCNet-M-N) for post-earthquake building identification and assessment parameters; by fusing the UAV flight attitude information and the internal and external parameters of the on-board camera, and utilizing the pixel-image-camera-earth multi-coordinate system transformation and the geographic/inverse geolocation coding and decoding algorithm The accurate positioning of buildings in the post-earthquake region is achieved, which can provide a spatial location determination method for visualizing and marking the distribution of post-earthquake building complexes [10].

5 Conclusion

In this paper, we carry out the research of geometrically constrained deep learning-enhanced seismic damage refinement identification and assessment methods for multi-scale seismic damage of urban buildings. A semantic segmentation network for post-earthquake densely distributed small-target building identification and collapse state assessment is constructed, and the influence of the GCE loss weight coefficients on the model identification and segmentation performance is systematically investigated, the geometric feature optimization performance and multilevel feature extraction ability of GCE loss in the training process are analyzed, and the generalization performance of GCE loss for other different loss functions and semantic segmentation networks is verified. The framework of the method for pixel-level refinement identification and rapid assessment of building multiscale seismic damage is proposed. Based on the medium-scale remote sensing images from a drone, the three classifications of building seismic damage status (destroyed, severely damaged, and other) are realized; the small-scale evaluation realizes the pixel-level identification of building components and damage assessment after the earthquake by collecting detailed images of external components from the drone approaching the building, and then the cross-scale seismic damage assessment is realized by the building seismic damage assessment method based on the number of components, and the results can be integrated into the assessment results of each scale to realize the refinement of the building seismic damage assessment. Finally, the results can realize the refined building seismic damage assessment.

References

[1] Y, Bao, H. Li, “Machine learning paradigm for structural health monitoring” Structural Health Monitoring, 2021, 20(4): 1353-1372.

[2] B. Spencer Jr , V. Hoskere Y Narazaki, Advances in computer vision-based civil infrastructure inspection and monitoring [J]. Engineering, 2019, 5 (2) 199-222.

[3] Y, Bao, Z. Chen, S, Wei, “The state of the art of data science and engineering in structural health monitoring,” Engineering, 2019, 5(2) 234-242.

[4] L. Sun, Z. Shang, Y. Xia, “Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection”, Journal of Structural Engineering, 2020, 146(5) 04020073.

[5] Y, Xu W, Qian N Li, “Typical advances of artificial intelligence in civil engineering” Advances in Structural Engineering, 2022, 25(16) 3405-3424.

[6] Spencer Jr B, Ruiz-Sandoval M, Kurata N. Smart sensing technology: Opportunities and challenges[J]. Structural Control and Health Monitoring, 2004, 11(4): 349-368.

[7] P. Chang, A. Flatau, S. Liu, “Health monitoring of civil infrastructure” Structural Health exp Monitoring, 2003, 2(3): 257-267.

[8] J. Ou, H. Li, “Structural health monitoring in mainland China: Review and future trends” Structural Health Monitoring, 2010, 9(3): 219-231.

[9] L Zhou, G Yan, L. Wang, “Review of benchmark studies and guidelines for structural health monitoring” . Advances in Structural Engineering, 2013, 16 (7): 1187-1206.

[10] L, Zhou, G. Yan, L. Wang, “Review of benchmark studies and guidelines for structural health monitoring” Advances in Structural Engineering, 2013, 16(7): 1187-1206.

Image	Calculate Location	Actual Location	Geographic Location
	104.4577, 31.8302	104.4587, 31. 8300	Xiqiang Shangjie, Qushan Town, Beichuan Qiang Autonomous County, Mianyang City, Sichuan Province, China
	105.3585, 29.2128	105.3579, 29.2129	National Highway 321, Fuji Town, Hu County, Huzhou City, Sichuan Province, China
	102.9932, 30.3393	102.9953, 30.3381	Ludeng Road, Taiping Town, Lushan County, Ya'an City, Sichuan Province, China

	102.8104, 30.3714	102.8114, 30.3710	Yanjiang Road, Muping Town, Baoxing County, Ya'an City, Sichuan Province, China