7th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Detection of collapsed buildings from post-earthquake imagery using mask region-based convolutional neural network

Esra Yildirim*[1], Taskin Kavzoglu [1]

[1] Gebze Technical University, Faculty of Engineering, Department of Geomatics Engineering, Kocaeli, Türkiye

Keywords
Earthquake
Remote sensing
Deep learning
Mask R-CNN
Building detection

Abstract

After large-scale natural disasters such as earthquakes, tsunamis, and floods, the rapid identification of collapsed buildings from high-resolution imagery plays a crucial role in post-disaster damage assessment, reconstruction, and emergency rescue operations. Deep learning (DL) architectures, widely applied across various scientific domains, have also been used for extracting damaged buildings from aerial and satellite images. This study is focused on identifying collapsed buildings using a DL algorithm applied to remotely sensed data collected after the February 6, 2023, Kahramanmaraş earthquake in Türkiye. To achieve this, post-earthquake WorldView-3 image with a spatial resolution of 0.3 m were obtained to establish a building dataset, from which the boundaries of collapsed and intact buildings were manually outlined. The Mask R-CNN model was then trained and validated using various hyperparameter combinations to optimize its performance. Experimental results revealed that the Mask R-CNN model with a ResNet-50 backbone yielded the most accurate results, successfully distinguishing between intact and collapsed buildings with an Average Precision (AP) of approximately 81% and 69%, respectively. The findings of the study illustrate the promising potential of using Mask R-CNN with high-resolution imagery for the detection and mapping of collapsed buildings following earthquake events. This application is particularly significant for post-disaster operations and mitigation studies.

1. Introduction

Earthquakes are considered among the most catastrophic natural calamities, causing extensive damage and resulting in significant loss of life and property. Even though earthquakes are unpreventable, the rapid detection and mapping of collapsed buildings after an earthquake is of utmost significance in emergency response and reconstruction efforts. Moreover, the extent, location, and degree of building damage, as well as the collapsed building rate, reflecting the magnitude of the earthquake, are essential information in supporting the evaluation processes of post-earthquake disasters (Song et al. 2020). While it is possible to obtain an accurate assessment of building damage through field surveys, this conventional method can be time-consuming and costly. It is also inefficient for the rapid evaluation of collapsed buildings during rescue operations (Turker and San 2004).

Due to progress in satellite and sensor technology, remote sensing methods can now capture Earth's surface with incredibly high spatial, spectral, and temporal resolutions. As a result, they have become a potent tool for identifying and tracking the impacts of natural disasters (Rathje and Adams 2008; Dell'Acqua and Gamba 2012; Dong and Shan 2013). A range of studies has been executed aiming to detect building damages triggered by earthquakes through the utilization of aerial and satellite imagery (Serifoglu Yilmaz et al. 2023; Turker and San 2004; Turker and Sumer 2008).

Recent research has explored the application of Convolutional Neural Networks (CNNs) in identifying building damage, displaying their effectiveness in automatically recognizing affected buildings within remotely sensed images. For instance, Moradi and Shah-Hosseini (2020) applied U-Net architecture to pre- and post-imagery of the Haiti earthquake images (WorldView-2) to identify damaged buildings and obtained 68.71% overall accuracy. Zhan et al. (2022) presented an adapted Mask R-CNN model designed for identifying impaired buildings and categorizing these structures based on the severity of their damage. Using the aerial images taken after the Kumamoto earthquake, their proposed model could detect buildings with about 90% accuracy and classified damage levels with about 80% accuracy.

* **Corresponding Author**

*(esrayildirim@gtu.edu.tr) ORCID ID 0000-0002-4951-0488
(kavzoglu@gtu.edu.tr) ORCID ID 0000-0002-9779-3443

Cite this study

Yildirim, E., & Kavzoglu, T. (2023). Detection of collapsed buildings from post-earthquake imagery using mask region-based convolutional neural network. Intercontinental Geoinformation Days (IGD), 7, 119-122, Peshawar, Pakistan

The aim of this research was to detect buildings that collapsed after the February 6, 2023, Kahramanmaraş Earthquake using remotely sensed imagery. To accomplish this, post-earthquake WorldView-3 image were utilized to create a dataset of buildings, and a DL-based Mask R-CNN model was trained and validated with this dataset. The study appraised the efficacy of the model in detecting collapsed buildings and employed the trained model to locate and map damaged structures in a different study zone.

2. Study Area and Dataset

On February 6, 2023, Türkiye was struck by two strong and consecutive earthquakes of magnitude 7.8 and 7.5, causing catastrophic damage to lifelines, facilities and buildings. The first earthquake occurred at 01:17:34 UTC in the Pazarcık district in the Kahramanmaraş province (southern Türkiye). About nine hours later, an aftershock occurred in the Elbistan district of Kahramanmaraş at 10:24:48 UTC (Goldberg et al. 2023).

In this study, the WorldView-3 image acquired on February 7, 2023, just one day after the earthquake, with 0.3 m spatial resolution covering the part of the Islahiye district of Gaziantep province, one of the mostly affected provinces, was employed in the analyses. Building boundaries were manually digitized in ArcGIS Pro 3.03 software in two categories (i.e., "Intact" and "Collapsed") using high-resolution base images of General Directorate of Mapping, a national mapping agency of Turkey, as a reference (Figure 1).

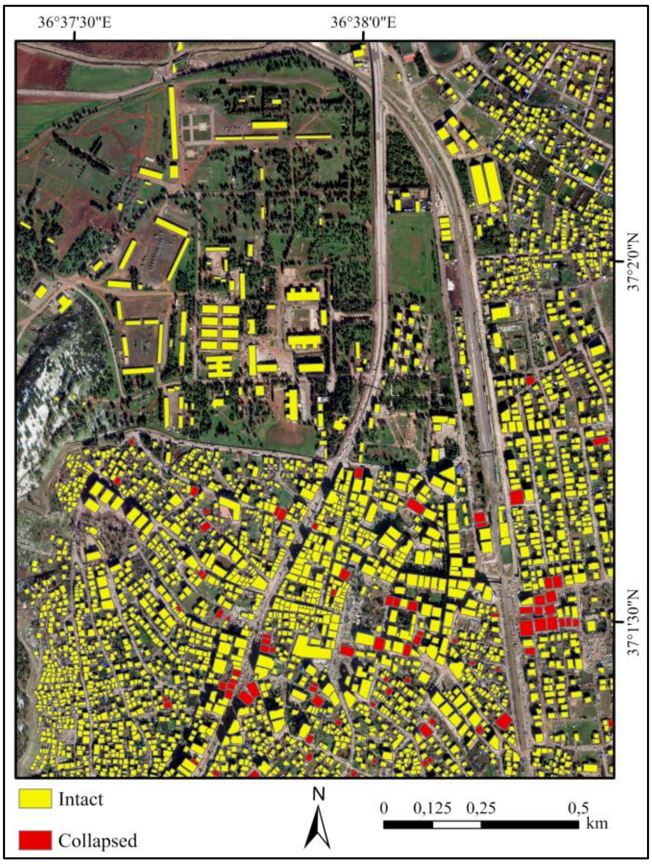


Figure 1. WorldView-3 image in the study area acquired after the 2023 Kahramanmaraş earthquake and locations of intact and collapsed building footprints

To provide training and validation datasets for the DL model, the WorldView-3 image was cropped into 256×256 pixel-sized image chips with a stride of 128×128 pixels (i.e., %50 overlap). The overlap was applied both to expand the dataset and to ensure that each image chip contained at least one building instance of two classes. Besides, the dataset was augmented using 180^0 rotation to increase the number of image samples thus improving the robustness of the model. Consequently, a building dataset containing a total of 3,792 image chips and 31,038 intact and 1,108 collapsed building features was obtained. In addition, corresponding label masks for each image chip were also generated (Figure 2).

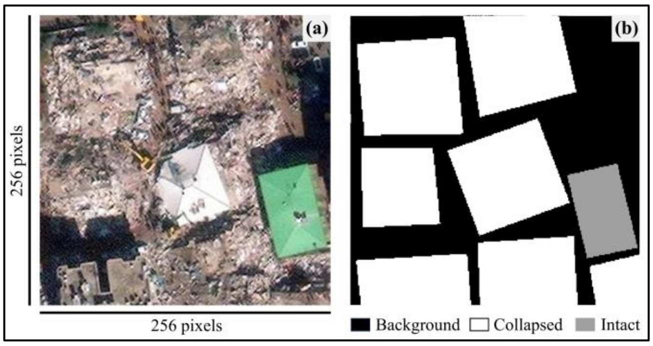


Figure 2. Building dataset, (a) sample of image chip and (b) corresponding ground-truth mask

3. Methods

3.1. Mask Region-Based Convolutional Neural Network (Mask R-CNN)

Mask R-CNN, developed by He et al. (2017), is an enhanced version of Faster R-CNN, capable of predicting both bounding boxes and detailed pixel-wise masks of objects (Figure 3). Mask R-CNN, similar to the Faster R-CNN algorithm, employs a two-stage detection pipeline that starts with the same initial phase, scanning the entire image and generating proposals. In the subsequent phase, while predicting the class and bounding box offsets, Mask R-CNN also generates a segmentation mask for each Region of Interest (RoI). Considering the network architecture of the Mask R-CNN model, it involves: (i) a backbone network (ResNet) responsible for extracting features over a whole image and creating feature maps; (ii) a Region Proposal Network (RPN) for generating regions (RoIs) for areas where objects can be found from feature maps; (iii) a RoI classifier and a bounding box regressor for classifying RoI and refining the bounding box; (iv) a Fully Convolutional Network (FCN) to generate a pixel-wise segmentation mask (Potlapally et al. 2019).

3.1. Design and implementation

The training and validation of the Mask R-CNN model was implemented in ArcGIS API for Python. In order to obtain the best-performing model, different hyperparameter combinations, shown in Table 1, were utilized in the training of the model. Thus, four experiments were conducted using ResNet-50 and ResNet-100 backbone architectures, batch sizes, epochs, and the experimented hyperparameters were adjusted according to the hardware configuration. Additionally, in each model, 90% of the created buildings were used for training and 10% for validation. All experiments were carried out on a Windows 11 laptop with an Intel® Core™ i7-10870H CPU, and a NVIDIA GeForce 3060 RTX GPU, with 32 GB RAM memory.

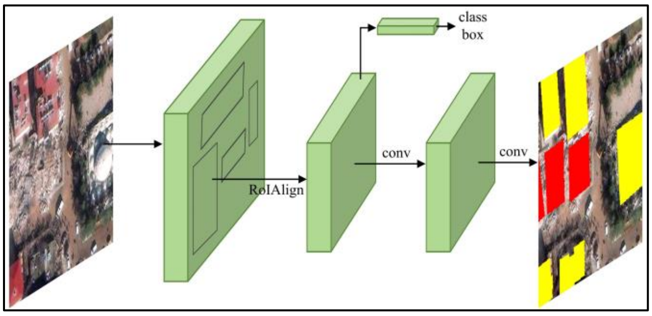


Figure 3. Architecture of the Mask R-CNN algorithm

Table 1. Hyperparameter configuration for Mask-RCNN

	Hyperparameters		
Experiment	**Backbone**	**Batch size**	**Epoch**
1	ResNet-50	2	100
2	ResNet-50	4	100
3	ResNet-100	4	100
4	ResNet-50	2	200

4. Results

The performances of the four trained models were evaluated using the Average Precision (AP) metric calculated for intact and collapsed building classes. The overall accuracy assessments indicated that using Mask R-CNN with a ResNet-50 backbone, trained for 200 epochs across two batch sizes, resulted in the highest AP scores for both intact and collapsed building categories (Table 2). More precisely, it demonstrated improved performance, attaining an AP score of 81.28% for intact or undamaged buildings and 69.26% for collapsed structures. Considering the computational cost of the models, the best-performing model (Experiment 4) required the longest training time of 32 hours and 50 minutes. The model was trained for more than twice the duration of the other models experimented with, owing to the extended training epoch.

Table 2. Performance comparison of Mask R-CNN models trained with different hyperparameters

	Average Precision (AP) (%)		
Experiment	**Intact**	**Collapsed**	**Training Time**
1	80.28	66.61	14 h 2 m
2	79.45	59.24	14 h 45 m
3	74.49	53.72	15 h 35 m
4	81.28	69.26	32 h 50 m

After the training and validation process of the Mask R-CNN model, total loss graphs were generated. It was evident from the loss curves of the most effective Mask R-CNN model that both the training and validation curves exhibited a decreasing trend as the number of epochs increased, reaching their minimum values without overfitting by the end of the process (Figure 4).

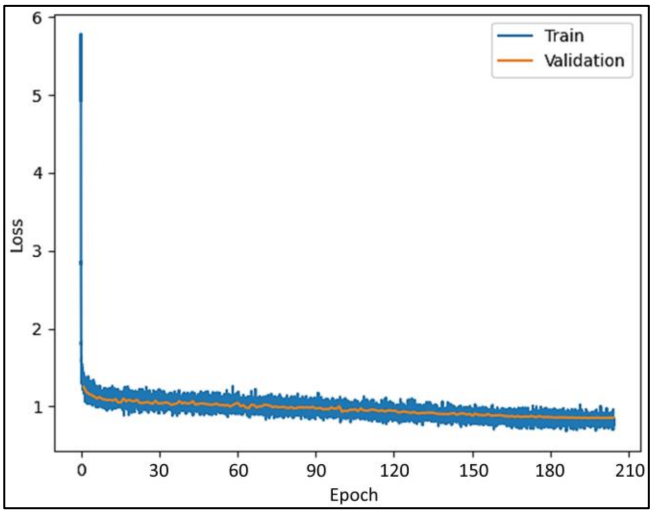
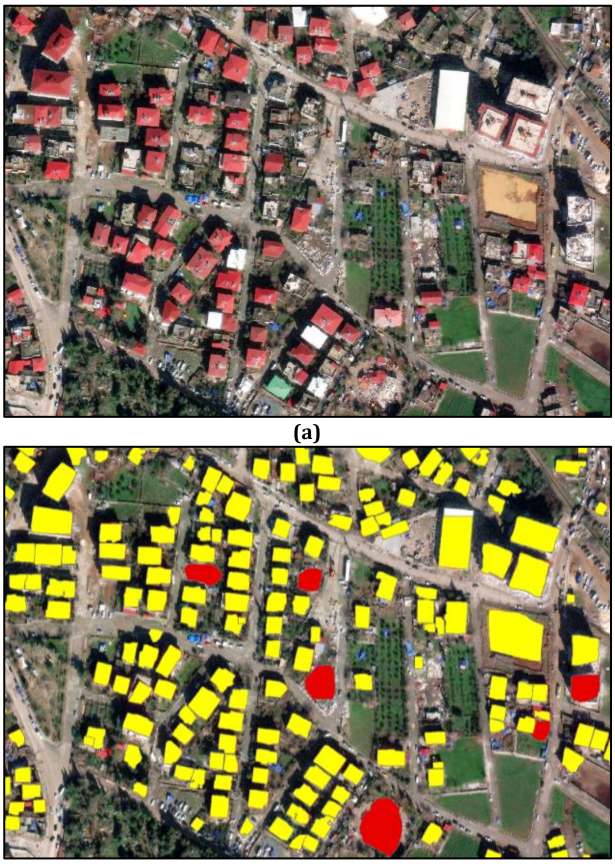


Figure 4. Training and validation loss curves for the Mask R-CNN model

To investigate the transferability of Mask R-CNN combined with best-performing hyperparameters, it was used for the detection and mapping of intact and collapsed buildings on the independent WorldView-3 image, which covers different part of the Islahiye district of Gaziantep province (Figure 5). It was observed that the model was able to accurately distinguish collapsed and intact buildings. However, due to the limited number of collapsed building samples in the training dataset, the model was more robust in identifying and locating intact buildings. Besides, detection errors were observed in collapses due to the viewing angle of the nadir satellite images when the roof of the building did not collapse but the floor collapsed.

5. Conclusion

DL-based algorithms have shown great potential to automatically detect damaged buildings after natural disasters using remotely sensed imagery. In this study, the DL-based Mask R-CNN model was utilized for the identification of collapsed buildings from post-disaster remotely sensed imagery. To meet the objective of the study, a building detection dataset was created using WorldView-3 imagery acquired one day after the February 6, 2023, Kahramanmaraş earthquake. Then, the Mask R-CNN model was trained and validated with the created dataset using different hyperparameter combinations. Experimental results revealed that Mask R-CNN combined with the ResNet-50 backbone and trained with two batch sizes for 200 epochs produced the most accurate results (AP=81.28% for intact buildings, AP=69.26% for collapsed buildings). These results highlighted that the Mask R-CNN model could be an effective solution for detecting and mapping collapsed buildings, which is particularly important for post-earthquake operations. It contributes to the accurate and rapid evaluation of collapsed buildings during emergency rescue operations. However, it should be noted that the main limitation of this study may be attributed to the lack of a high-quality large dataset. To be more specific, the unbalanced instances of collapsed and intact buildings, the constrained spatial resolution, and the use of a single image source were critical issues that influenced the accuracy and generalization ability of the DL model. Given these requirements, the use of an expanded dataset would significantly increase the accuracy and transferability of the study. From this perspective, future research endeavors could emphasize generating a superior building dataset and enhancing the resilience of DL algorithms.



(a)

(b)

Figure 5. Test result, (a) WorldView-3 image of the test site (b) collapsed (red) and intact (blue) buildings detected by the model

References

Dell'Acqua, F., & Gamba, P. (2012). Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. Proceedings of the IEEE, 100(10), 2876-2890. https://doi.org/10.1109/JPROC.2012.2196404

Dong, L., & Shan, J. (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

Goldberg, D. E., Taymaz, T., Reitman, N. G., Hatem, A. E., Yolsal-Çevikbilen, S., Barnhart, W. D., et al. (2023). Rapid characterization of the February 2023 Kahramanmaraş, Türkiye, Earthquake sequence. The Seismic Record, 3(2), 156-167. https://doi.org/10.1785/0320230009

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. IEEE International Conference on Computer Vision, 2961-2969, Venice, Italy. https://doi.org/10.1109/ICCV.2017.322

Moradi, M., & Shah-Hosseini, R. (2020). Earthquake damage assessment based on deep learning method using VHR images. Environmental Sciences Proceedings, 5(1), 16. https://doi.org/10.3390/IECG2020-08545

Potlapally, A., Chowdary, P. S. R., Shekhar, S. R., Mishra, N., Madhuri, C. S. V. D., & Prasad, A. V. V. (2019). Instance segmentation in remote sensing imagery using deep convolutional neural networks. International Conference on Contemporary Computing and Informatics (IC3I), 117-120. https://doi.org/10.1109/IC3I46837.2019.9055569

Rathje, E. M., & Adams, B. J. (2008). The role of remote sensing in earthquake science and engineering: Opportunities and challenges. Earthquake Spectra, 24(2), 471-492. https://doi.org/10.1193/1.2923922

Serifoglu Yilmaz, C., Yilmaz, V., Tansey, K., & Aljehani, N. S. (2023). Automated detection of damaged buildings in post-disaster scenarios: A case study of Kahramanmaraş (Türkiye) earthquakes on February 6, 2023. Natural Hazards, 1-25. https://doi.org/10.1007/s11069-023-06154-z

Song, D., Tan, X., Wang, B., Zhang, L., Shan, X., & Cui, J. (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

Turker, M., & San, B. T. (2004). Detection of collapsed buildings caused by the 1999 Izmit, Turkey earthquake through digital analysis of post-event aerial photographs. International Journal of Remote Sensing, 25(21), 4701-4714. https://doi.org/10.1080/01431160410001709976

Turker, M., & Sumer, E. (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

Zhan, Y., Liu, W., & Maruyama, Y. (2022). Damaged building extraction using modified Mask R-CNN model using post-event aerial images of the 2016 Kumamoto earthquake. Remote Sensing, 14(4), 1002. https://doi.org/10.3390/rs14041002