Comparative Analysis for Post-Earthquake Road Debris Detection Based on Deep Neural Networks Using High-resolution Remote Sensing Imagery

Aydin Ebrahimi [1*], Ali Mohammadzadeh [1], Armin Moghimi [1,2], Saeed Homayouni [3]

1 Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran 19967-15433, Iran - a.ebrahimi3@email.kntu.ac.ir, a_mohammadzadeh@kntu.ac.ir

2 Ludwig-Franzius-Institute for Hydraulic, Estuarine and Coastal Engineering, Leibniz University Hannover, Nienburger Str. 4, 30167 Hanover, Germany - moghimi@lufi.uni- hannover.de

3 Centre Eau Terre Environnement, Institut National de la Recherche Scientifique, Québec, QC G1K 9A9, Canada – saeid.homayouni@inrs.ca

Keywords: Remote sensing, Deep learning, Road debris detection, Semantic segmentation, Disaster response

Abstract

Earthquakes can cause considerable damage to transportation infrastructure, affecting emergency response. Timely detection of road debris is important in such situations. The purpose of this research is to address the challenges related to the rapid detection of collapsed areas and road blockages. Deep learning has become an essential tool in remote sensing and image processing, offering improved capabilities for classification, and segmentation from high-resolution imagery. Its integration into post-disaster analysis enables more accurate assessments compared to traditional methods. This study examines the performance of three commonly used semantic segmentation models Unet, Attention Unet (AttUnet), and ResUnet++ for road debris detection. Two earthquake events were used as case studies: the 2023 earthquake in Osmaniye, Türkiye (Mw 7.8), captured by the Pleiades satellite, and the 2017 earthquake in Sarpol-e Zahab, Kermanshah, Iran (Mw 7.3), captured by a Phantom 4 Pro drone. The models were evaluated using three metrics: Intersection over Union (IoU), Recall, and Accuracy. The results show that ResUnet++ achieved higher performance compared to Unet and AttUnet in both cases. For the Osmaniye dataset, ResUnet++ reached an IoU of 80.81%, Recall of 78.88%, and Accuracy of 96.12%. On the Sarpol-e Zahab dataset, it obtained an IoU of 81.62%, Recall of 80.28%, and Accuracy of 97.24%. Unet and AttUnet performed at lower levels across all evaluated metrics. This comparison provides a clear assessment of model performance in post-earthquake debris detection tasks and contributes to ongoing work in the application of deep learning and high-resolution imagery for geospatial analysis in disaster response contexts.

1. Introduction

Earthquakes are among the most catastrophic natural disasters, often occurring with little or no warning and causing significant damage to buildings, infrastructure, and the environment (Huang & Li, 2009). Densely populated areas are particularly at risk, creating serious threats to human life. Although earthquakes cannot be prevented, the way humans respond can reduce damages. Quickly detecting the affected areas is important for directing relief supplies such as food, medicine, and shelter. Remote sensing (RS) technologies have proven valuable for assessing earthquake damage and supporting emergency response (Menderes et al., 2015; T. Zhang & Huang, 2018; X. Zhang et al., 2022). Various RS data sources are used to detect earthquake damage, including optical images, LiDAR point clouds (Erdogan & Yilmaz, 2019), synthetic aperture radar (SAR) (Shahzad et al., 2019), and aerial or UAV orthophotos (Kalantar et al., 2020). Very high-resolution (VHR) optical images from satellites or aircraft are particularly effective for detecting building damage when the resolution is finer than one meter (YAMAZAKI & MATSUOKA, 2007). However, manually inspecting these images can take a long time after a large earthquake. This has led to the use of machine learning (ML) methods for automatic damage detection (Menderes et al., 2015). Some studies have applied ML to find road debris and damage. For example, Jia et al. proposed a framework that segments road areas and classifies debris using a Support Vector Machine (SVM) (Jia & Ye, 2023). Although this method produced good results, it relies heavily on manual thresholds and features designed by hand, which limits its use in different environments (Kalantar et al., 2020). Deep learning (DL), especially convolutional neural networks (CNNs), can automatically extract complex features for detecting collapsed structures and debris (Hong et al., 2022; Nex et al., 2019; Xia et al., 2023). Most CNN-based methods, however, focus on classifying only a few damage levels (Arabameri et al., 2019). Pre-trained models like VGG and ResNet, combined with data augmentation, have improved detection of small and complex debris in images (Moskalenko et al., 2022), and integrating CNNs with GIS data can improve damage mapping after earthquakes (Jia & Ye, 2023). Despite this, advanced object detection methods such as YOLO (Diwan et al., 2023; Ma et al., 2020) and SSD (Li et al., 2019) still face challenges in dense urban areas. To overcome these issues, encoder–decoder networks like Unet and its variants have been widely implemented for pixel-level segmentation. For example, Bai et al. applied Unet to detect damage from the Tohoku tsunami (Bai et al., 2018). Studies have shown that encoder–decoder models can accurately detect multiple damage

* Corresponding author

classes, with reported accuracy between 70% and 90% (Alzubaidi et al., 2021; Khan et al., 2022). Other approaches, such as integrating pre-event building maps to improve training data, help reduce errors but can still miss debris regions (Ghaffarian et al., 2019; Zhao et al., 2022). These results highlight the growing role of deep learning and semantic segmentation for fast and accurate assessment of post-earthquake damage.

2. Material and Methods

2.1 Study Area and Dataset

Two sets of data, comprising images and their corresponding ground-truth labels, were utilized for the training and testing phases in networks. These datasets were derived from satellite imagery from the Pleiades satellite and UAV imagery that the KNTU LiDAR lab provided. A brief overview of the dataset characteristics is provided in the subsequent section, complemented by visual representations in Figure 1 and Table 1. Furthermore, there was no spatial overlap between the datasets used for both the training and testing phases to guarantee the accuracy of detecting debris on the road. Dataset 1 Pleiades includes a VHR Pleiades satellite image collected over Osmaniye, Türkiye, geographically limited [37°04′30″(N) 36°15′00″ (E)]. Dataset 2 comprises a VHR UAV orthophoto image obtained from a section of Sarpol-e Zahab, situated at [34° 27′32″ (N) 45°51′41″ (E)].

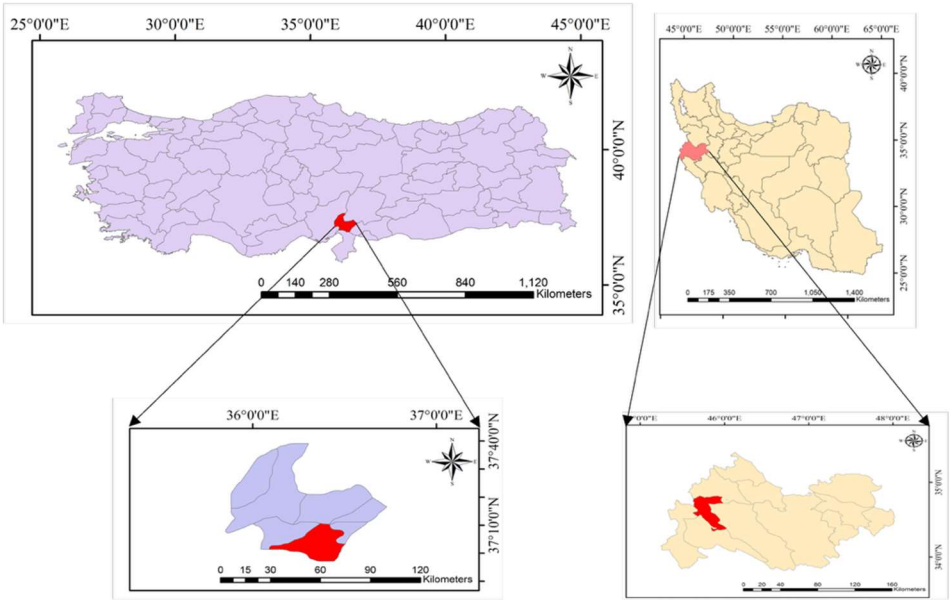

Figure 1. Location of Osmaniye on the map of Türkiye and the location of Sarpol-e Zahab on the map of Iran (Right), A part of the Satellite image of Osmaniye, and a section of the UAV image of Sarpol-e Zahab (Left).

Region	Platform	Date	Bands	Resolution
Osmaniye, Türkiye	Pleiades	9 Feb 2023	Blue: 450-520 nm Green: 530- Red: 620-690 nm	30 cm
Sarpol-e Zahab, Iran	DJI Phantom 4 Pro	15 Nov 2017	Blue: 450 nm ± 16 nm Green: 560 nm ± 16 nm Red: 650 nm ± 16 nm	10 cm

Table 1. Image data specifications

2.2 Methodology

2.2.1 Step 1: Image Preprocessing

In this step, VHR images are first divided into 256*256 patches to resize data and increase the capability of the segmentation model to learn more. The dataset is divided into three sections: training, validation, and testing. The annotation process involved identifying and defining road

debris within the images and providing textual descriptions of the specified debris.

The annotation process for road debris in very high-resolution (VHR) images involves manual labelling. Annotators identify debris and provide textual descriptions and use bounding boxes, polygons, or pixel masks to mark debris instances.

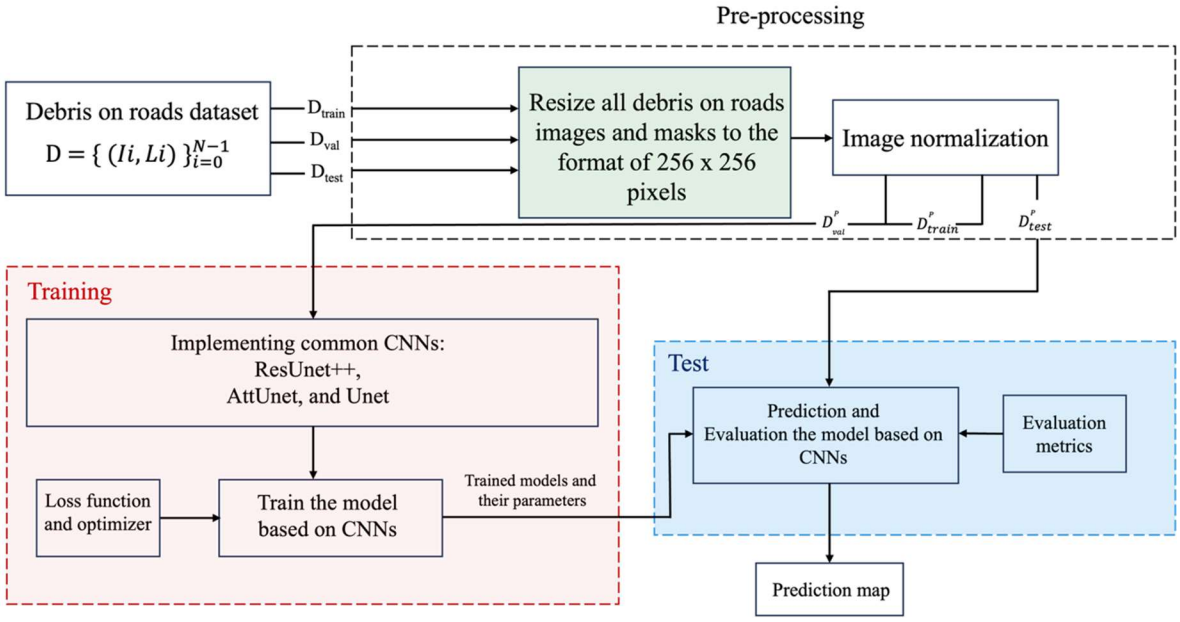

Figure 2. Flowchart of study

Dataset	Number of Images	Number of Debris in the image
Training	747	1102
Test	78	115
Validation	63	93

Table 2. Dataset Division for Osmaniye, Türkiye

Dataset	Number of Images	Number of Debris in the image
Training	672	910
Test	50	84
Validation	35	71

Table 3. Dataset Division for Sarpol-e Zahab, Iran

Given our experiments constrained availability of image samples, adopting image data augmentation techniques becomes indispensable. This methodology enriches the training dataset, thereby reducing the risk of overfitting. Specifically, our study leverages techniques including image rotation and flipping for data augmentation. Rotation operations spanning 90 and/or 180 degrees are applied alongside horizontal and vertical flipping. Details of data division can be seen in Table 2 and 3.

3. Experiments

In this part, we provide a concise overview of the evaluation metrics and technical implementation and demonstrate the effectiveness of the semantic segmentation models. The VHR images from both study areas generate segmentation results for considered models to evaluate their performance.

3.1 Evaluation metrics

The semantic segmentation model is assessed using the pixel-based confusion matrix, which records true and predicted values. The rows of the matrix depict the true

values, and the predicted values are represented in the columns. True Positives (TP) and True Negatives (TN) are located at the main diagonal, while False Positives (FP) sit in each column (excluding the main diagonal), and False Negatives (FN) lie along the rows. We employ three evaluation metrics to evaluate the semantic segmentation is: mean intersection over union (mIoU), recall, and overall accuracy (OA). The definitions for these metrics are described as follows:

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (1)$$

$$IoU = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \ , mIoU = \frac{\sum_{i=1}^{N} \mathrm{IoU}}{\mathrm{N}} \quad (2)$$

$$OA = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FN} + \mathrm{FP}} \quad (3)$$

The Overall Accuracy (OA) evaluates the complete image, with N indicating the total number of categories, whereas the Intersection over Union (IoU) evaluates specific classes. mIoU is calculated by taking the average of IoU for each category.

3.1.1 Implementation details

We initiated training for all models from basics, without using any pre-trained parameters or models. The parameter settings for both the Osmaniye and Sarpol-e Zahab datasets were kept consistent throughout the training process: a batch size of 8, a learning rate of 10^-3, 100 epochs, and utilizing the Adam optimizer. The experiments were conducted using the Keras platform. As a result, we selected binary cross-entropy loss (Seydi & Hasanlou, 2020) as our loss function, which is a weighted sum of the modified cross-entropy to facilitate effective training.

Model	mIoU (%)	Recall (%)	Accuracy (%)
ResUnet++	80.81	78.88	96.12
AttUnet	80.22	78.62	95.86
Unet	78.22	77.72	94.68

Table 4. Results from experiments on the Osmaniye, Türkiye dataset show the segmentation of original categories. The mIoU/Recall is used to evaluate the accuracy of each category. The bold signifies the top performance.

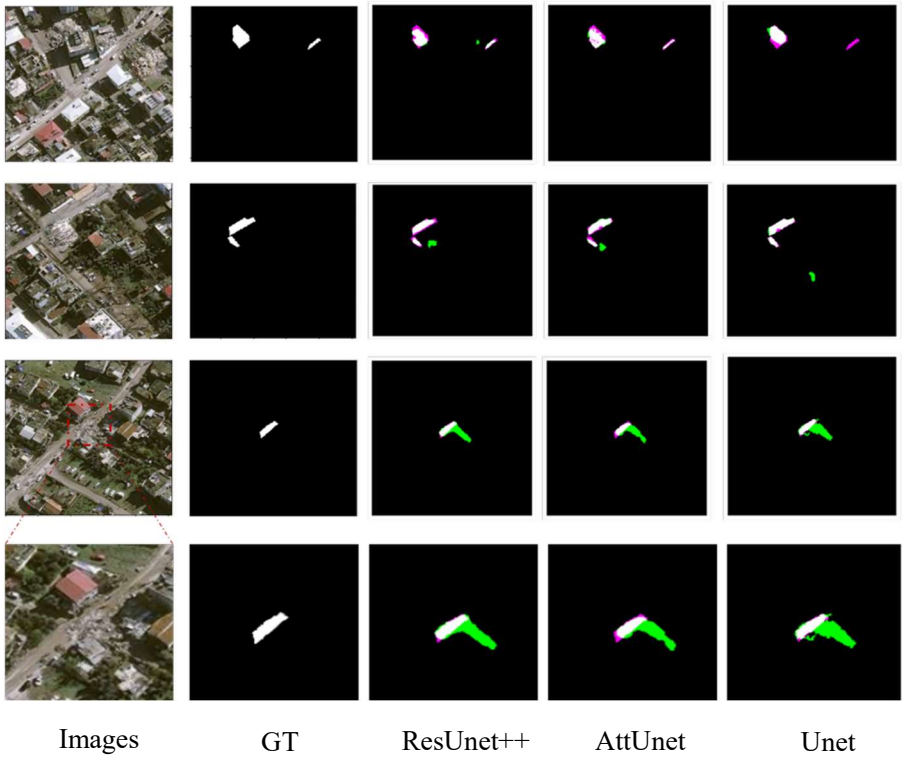

Figure 3. Comparison of the predicted masks with the ground truth masks and their superimposed maps for different networks in Osmaniye, Türkiye. Purple colours indicate areas of unpredicted debris (True positive), while green colours highlight regions where predictions deviated (False negative).

3.2 Analysis

3.2.1 Segmentation based on the trained models

To evaluate model performance, we compared three semantic segmentation architectures Unet, ResUnet++, and AttUnet on the Osmaniye and Sarpol-e Zahab datasets. To assess the relative performance of the three semantic segmentation models Unet, AttUnet, and ResUnet++ quantitative results are detailed in Tables 4 and 5. Table 4 displays the results for the Osmaniye, Türkiye dataset, indicating that the metrics mIoU, Recall, and Overall Accuracy (OA) reached 80.81%, 78.88%, and 96.12%, respectively. Among the models assessed, ResUnet++ achieved the best segmentation accuracy, surpassing Unet by 2.59% in mIoU, 1.16% in Recall, and 1.44% in OA, illustrating its enhanced ability to extract features specific to categories.

Model	mIoU (%)	Recall (%)	Accuracy (%)
ResUnet++	81.62	80.28	97.24
AttUnet	80.23	80.16	97.02
Unet	79.28	78.72	96.21

Table 5. Results from experiments on the Sarpol-e Zahab, Iran dataset show the segmentation of original categories. The mIoU/Recall is used to evaluate the accuracy of each category. The bold signifies the top performance.

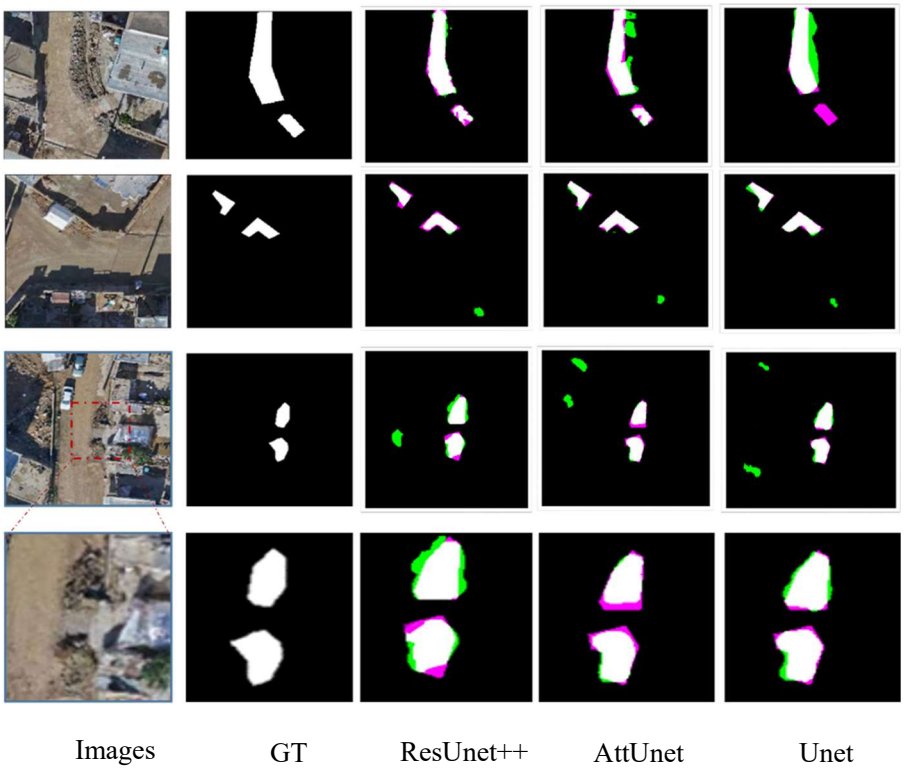

Figure 4. Comparison of the predicted masks with the ground truth masks and their superimposed maps for different networks in Sarpol-e Zahab, Iran. Purple colours indicate areas of unpredicted debris (True positive), while green colours highlight regions where predictions deviated (False negative).

Likewise, Table 5 outlines the results for the Sarpol-e Zahab dataset, where ResUnet++ secured 81.62% mIoU, 80.28% Recall, and 97.24% OA. When compared with AttUnet, which is the second-best model, ResUnet++ demonstrated advancements of 1.37% in mIoU, 0.12% in Recall, and 0.22% in OA. In summary, the results verify that ResUnet++ consistently achieves superior segmentation accuracy for detecting road debris after earthquakes. Visual comparisons in Figures. 3 and 4 further reinforce this conclusion, showing that ResUnet++ delivers clearer boundaries, smoother edge delineations, and more reliable detection of small or irregular debris areas across both datasets.

4. Discussion

In this study, we compared three deep learning models, Unet, AttUnet, and ResUnet++, to check how well they can find road debris after earthquakes. The datasets we used had different resolutions about 30 cm for Pleiades satellite images and 10 cm for UAV photos so we divided all images into 256×256-pixel patches. This size worked reasonably well because it kept the processing time manageable while still capturing small debris details. Detecting road debris from high-resolution images is still quite tricky. One problem is that debris usually takes up only a small part of the image, which makes the models more likely to focus on the background. This imbalance often lowers detection accuracy. Debris also varies a lot in how it looks different shapes, colours, and textures which makes generalization difficult. On top of that, lighting, shadows, and other environmental conditions can hide some of the details and cause the predictions to change between images. We tried to reduce confusion between road debris and nearby collapsed buildings by focusing the annotation on debris found on or next to the roads. The labelling process used high-resolution base maps and manual checking to keep it consistent. In real emergencies, both precision and recall matter because false alarms waste time, while missed detections can delay rescue work. The results suggest that models like ResUnet++ could be helpful tools for quick damage assessment after earthquakes. They might support emergency planners and field teams by showing where roads are blocked or partially damaged. With some adjustments, these models could become part of real disaster management systems to speed up response and help coordinate resources better. These findings highlight the potential of ResUnet++ for integration into automated disaster management systems.

5. Conclusion

This study focuses on post-earthquake damage assessment, particularly the detection of road debris using very high-resolution (VHR) images from Osmaniye, Türkiye (2023), and Sarpol-e Zahab, Iran (2017). Three semantic segmentation models Unet, Attention Unet (AttUnet), and ResUnet++ were tested and compared. The results showed that ResUnet++ gave the highest overall accuracy, with mean IoU, Recall, and Accuracy of 80.81%, 78.88%, and 96.12% for Osmaniye, and 81.62%, 80.28%, and 97.24% for Sarpol-e Zahab. The model performed consistently well with both datasets, showing good robustness in detecting road debris under different imaging conditions. For future work, we plan to include transformer-based modules to speed up training and improve classification of different debris categories. Adding more satellite and UAV images with varying resolutions is also expected to make the model more reliable and accurate for real post-disaster applications.

References

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., Farhan, L., 2021: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. In *Journal of Big Data* (Vol. 8, Issue 1). Springer International Publishing. https://doi.org/10.1186/s40537-021-00444-8

Arabameri, A., Pradhan, B., Rezaei, K., Lee, C. W., 2019: Assessment of landslide susceptibility using statistical- and artificial intelligence-based FR-RF integrated model and multiresolution DEMs. *Remote Sensing*, *11*(9). https://doi.org/10.3390/rs11090999

Bai, Y., Mas, E., Koshimura, S., 2018: Towards operational satellite-based damage-mapping using U-net convolutional network: A case study of 2011 Tohoku Earthquake-Tsunami. *Remote Sensing*, *10*(10). https://doi.org/10.3390/rs10101626

Diwan, T., Anirudh, G., Tembhurne, J. V., 2023: Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications*, *82*(6), 9243–9275. https://doi.org/10.1007/s11042-022-13644-y

Erdogan, M., Yilmaz, A., 2019: Detection of building damage caused by Van Earthquake using image and Digital Surface Model (DSM) difference. *International Journal of Remote Sensing*, *40*(10), 3772–3786. https://doi.org/10.1080/01431161.2018.1552816

Ghaffarian, S., Kerle, N., Pasolli, E., Arsanjani, J. J., 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), 1–20. https://doi.org/10.3390/rs11202427

Hong, Z., Zhong, H., Pan, H., Liu, J., Zhou, R., Zhang, Y., Han, Y., Wang, J., Yang, S., Zhong, C., 2022: Classification of Building Damage Using a Novel Convolutional Neural Network Based on Post-Disaster Aerial Images. *Sensors*, *22*(15). https://doi.org/10.3390/s22155920

Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E., 2020: Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(8), 2011–2023. https://doi.org/10.1109/TPAMI.2019.2913372

Huang, R. Q., Li, W. L., 2009: Analysis of the geo-hazards triggered by the May 12 2008 Wenchuan Earthquake, China.

Bulletin of Engineering Geology and the Environment, *68*(3), 363–371. https://doi.org/10.1007/s10064-009-0207-0

Jia, J., Ye, W., 2023: Deep Learning for Earthquake Disaster Assessment: Objects, Data, Models, Stages, Challenges, and Opportunities. *Remote Sensing*, *15*(16). https://doi.org/10.3390/rs15164098

Kalantar, B., Ueda, N., Al-Najjar, H. A. H., Halin, A. A., 2020: Assessment of convolutional neural network architectures for earthquake-induced building damage detection based on pre-and post-event orthophoto images. *Remote Sensing*, *12*(21), 1–20. https://doi.org/10.3390/rs12213529

Khan, M. S. I., Rahman, A., Debnath, T., Karim, M. R., Nasir, M. K., Band, S. S., Mosavi, A., Dehzangi, I., 2022: Accurate brain tumor detection using deep convolutional neural network. *Computational and Structural Biotechnology Journal*, *20*, 4733–4745. https://doi.org/10.1016/j.csbj.2022.08.039

Li, Y., Hu, W., Dong, H., Zhang, X., 2019: Building damage detection from post-event aerial imagery using single shot multibox detector. *Applied Sciences (Switzerland)*, *9*(6). https://doi.org/10.3390/app9061128

Ma, H., Liu, Y., Ren, Y., Yu, J., 2020: Detection of collapsed buildings in post-earthquake remote sensing images based on the improved YOLOv3. *Remote Sensing*, *12*(1). https://doi.org/10.3390/RS12010044

Menderes, A., Erener, A., Sarp, G., 2015: Automatic Detection of Damaged Buildings after Earthquake Hazard by Using Remote Sensing and Information Technologies. *Procedia Earth and Planetary Science*, *15*, 257–262. https://doi.org/10.1016/j.proeps.2015.08.063

Moskalenko, V., Kharchenko, V., Moskalenko, A., Petrov, S., 2022: Model and Training Method of the Resilient Image Classifier Considering Faults, Concept Drift, and Adversarial Attacks. *Algorithms*, *15*(10), 1–24. https://doi.org/10.3390/a15100384

Nex, F., Duarte, D., Tonolo, F. G., Kerle, N., 2019: Structural building damage detection with deep learning: Assessment of a state-of-the-art CNN in operational conditions. *Remote Sensing*, *11*(23). https://doi.org/10.3390/rs11232765

Seydi, S. T., Hasanlou, M., 2020: Binary hyperspectral change detection based on 3D convolution deep learning. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *43*(B3), 1629–1633. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1629-2020

Shahzad, M., Maurer, M., Fraundorfer, F., Wang, Y., Zhu, X. X., 2019: Buildings detection in VHR SAR images using fully convolution neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 1100–1116. https://doi.org/10.1109/TGRS.2018.2864716

Turkey-Earthquake: Emergency Situation Report (24.02.2023) - Türkiye | ReliefWeb. (n.d.). Retrieved April 12, 2024, from https://reliefweb.int/report/turkiye/turkey-earthquake-emergency-situation-report-24022023

Xia, H., Wu, J., Yao, J., Zhu, H., Gong, A., Yang, J., Hu, L., Mo, F., 2023: A Deep Learning Application for Building Damage Assessment Using Ultra-High-Resolution Remote Sensing Imagery in Turkey Earthquake. *International Journal of Disaster Risk Science*, *14*(6), 947–962. https://doi.org/10.1007/s13753-023-00526-6

YAMAZAKI, F., MATSUOKA, M., 2007: Remote Sensing Technologies in Post-Disaster Damage Assessment. *Journal of Earthquake and Tsunami*, *01*(03), 193–210. https://doi.org/10.1142/s1793431107000122

Zhang, T., Huang, X., 2018: Monitoring of Urban Impervious Surfaces Using Time Series of High-Resolution Remote Sensing Images in Rapidly Urbanized Areas: A Case Study of Shenzhen. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(8), 2692–2708. https://doi.org/10.1109/JSTARS.2018.2804440

Zhang, X., Zhou, Y., Luo, J., 2022: Deep learning for processing and analysis of remote sensing big data: a technical review. *Big Earth Data*, *6*(4), 527–560. https://doi.org/10.1080/20964471.2021.1964879

Zhao, G., Yao, P., Fu, L., Zhang, Z., Lu, S., Long, T., 2022: A Deep Learning Method Based on Two-Stage CNN Framework for Recognition of Chinese Reservoirs with Sentinel-2 Images. *Water (Switzerland)*, *14*(22), 1–21. https://doi.org/10.3390/w14223755