UNET++ Convolution Neural Network for Building Damage Assessment: A Case Study of the 2023 Turkey–Syria Earthquakes

Nam H. Vu

Student number: 2723402

Vrije Universiteit Amsterdam

30/06/2024

Table of Contents

Abstract

On February 6, 2023, a devastating sequence of earthquakes struck Turkey and Syria, with magnitudes of 7.8 and 7.5, resulting in many deaths and much property damage. This calls for rapid and accurate building damage assessment to help with disaster response. This research studies the application of deep learning models, specifically UNet and UNet++, for post-earthquake building damage assessment. Using satellite imagery and ground truth data from the 2023 Turkey-Syria earthquakes, the author trained and evaluated the performance of these models. The study's final results indicate that while both UNet and UNet++ performed well on training data, UNet++ is slightly better than UNet. However, when applied to the Antakya case study, both models showed a significant decrease in performance, raising concerns about data quality, training dataset size, and model complexity. This research proves the potential of UNet and UNet++ for building damage assessment. It also stresses the need for higher-quality satellite data, larger training datasets for the building damage assessment task, and further optimisation of model architectures.

Introduction

On February 6, 2023, a destructive sequence of earthquakes and aftershocks struck Turkey and Syria, beginning with a magnitude 7.8 earthquake, followed closely by a magnitude 7.5 earthquake and thousands of aftershocks (United States Geological Survey, 2023). These earthquakes resulted in approximately 50,783 confirmed casualties in Turkey alone (Wang et al., 2023) and caused an estimated $84 billion in damages (Turkonfed, 2023). Earthquakes, among the most destructive natural disasters, have historically caused significant loss of life and economic disruption. Over a million earthquakes occur annually (Dong et al., 2013), accounting for 60% of all natural disaster-related casualties and significantly impacting affected countries' economies (Bartels et al., 2011). Urbanisation heightens the risk of earthquake damage, particularly in densely populated urban areas. Research indicates that building collapse and structural damage are the main causes of earthquake casualties (Shohet et al., 2016). Therefore, fast and accurate building damage assessment is essential for effective disaster response (Nex et al., 2019).

Challenges and Complexities of Earthquake Damage Assessment

Despite the essential need for fast damage assessment, the process is challenging. Although reliable, Manual interpretation of damage is costly and labour-intensive (Matin et al., 2021a). Automated and semi-automated methods, including machine learning and deep learning frameworks, offer possible solutions but are often viewed with doubt due to concerns about their reliability and lack of explainability (Ahmad et al., 2020). AI models, particularly in damage mapping, operate as "black boxes," making their decision-making processes unclear to users. This lack of transparency can restrict trust and adoption despite these tools' potential benefits in improving disaster response efforts. Also, since there is only a little tier-one training data for this particular task of damage mapping (Gupta et al., 2019), the generalisation of models can be affected. It might cause the models to perform poorly in unseen regions.

Role of Deep Learning in Earthquake Damage Assessment

Deep learning, particularly Convolutional Neural Networks (CNNs), has appeared as a promising approach for building damage assessment (Duarte et al., 2018). However, each method has its own set of pros and cons. Traditional manual processes, while accurate, are slow and resource-intensive. Automated methods, though faster, need help with issues of accuracy and explainability (Matin et al., 2021b). Deep learning models like UNET and its advanced variant, UNET++, offer potential improvements in this context.

UNet has been widely studied and applied in various image segmentation tasks, including biomedical tasks (Ronneberger et al., 2015) and earthquake damage assessment (Bai et al., 2018). However, UNet++, an enhanced version of UNet designed to improve segmentation accuracy through nested and dense skip connections and deep supervision capability, still needs to be explored in this domain.

Objective of the Study

To the author's knowledge, there has yet to be any research on using UNet++ for the building damage assessment task. Therefore, this study aims to better understand the capabilities and limitations of UNet++ by comparing it to its predecessor, UNet, for post-earthquake building damage assessment. Using the 2023 Turkey earthquake as a case study, the author will generate damage maps of Antakya using both UNet and UNet++ and compare these to ground truth data provided by the Humanitarian OpenStreetMap Team (HOTOSM). This comparative research aims to determine if Unet and UNet++ can be valuable tools in earthquake damage assessment.

Related works

UNet

The paper "U-Net: Convolutional Networks for Biomedical Image Segmentation" by Ronneberger et al. (2015) introduced the UNet architecture, which revolutionised biomedical image segmentation due to its efficient use of limited training data and innovative network design. The UNet architecture includes a downsampling path to capture context and an upsampling path that enables precise localisation. This structure is notably suited for medical applications where training data are limited and high precision is required. The UNet enhanced the accuracy of biomedical image segmentation and showed superior performance in various competitions, setting a benchmark for future CNN architectures (Ronneberger et al., 2015).

UNet++

Zhou et al. research on "UNet++: A Nested U-Net Architecture for Medical Image Segmentation" (2018) built upon the UNet model with their UNet++ architecture, which introduces nested and densely connected skip pathways to reduce the semantic gap between the encoder and decoder sub-networks. This modification solves one of the critical limitations in traditional U-Net architectures by improving feature compatibility and enabling more effective learning. With its deep supervision mechanism, UNet++ has shown improved performance over both the original U-Net and its variants across different medical segmentation tasks, proving the effectiveness of its innovative architecture for detailed and accurate image segmentation (Zhou et al., 2018).

UNet in disaster response

Aside from medical imaging to remote sensing, another adaptation of the UNet model was explored by Bai et al. (2018), who used it to assess building damage after natural disasters. Their study "Towards Operational Satellite-Based Damage-Mapping Using U-Net

Convolutional Network: A Case Study of 2011 Tohoku Earthquake-Tsunami" used high-resolution satellite images to detect changes pre- and post-disaster, utilising a UNet-based model trained to identify varying damage grades: "washed away", "collapsed", and "survived". Although the model achieved high accuracy levels, its application was limited to the dataset from the 2011 Tohoku earthquake, raising concerns about its generalizability to other events or regions. This limitation indicates the need for further testing across diverse disaster scenarios to ensure more general usage (Bai et al., 2018).

Key points

These studies together highlight the versatility and strength of UNet and its derivatives in performing complex segmentation tasks across different domains. The ongoing developments in biomedical and geographical applications have shown the undeniable impact of convolution neural networks in image analysis. However, challenges such as data diversity and model generalizability remain relevant, requiring further research and testing across various scenarios.

Case Study and Dataset

Case study

The 2023 Kahramanmaras earthquake was a significant event that profoundly impacted the region. Antakya, located along the Asi River, was one of the most severely affected areas, experiencing a maximum Mercalli intensity of XII (extreme) during the earthquake (Wang et al., 2023). The earthquake resulted in Turkey's confirmed death count of 50,783 and 107,000 reported injuries as of May 18, 2023 (Wang et al., 2023).

The area along the Asi River is of particular interest due to the extent of the damage it sustained. The main shock and the largest aftershock produced significant ground deformation extending approximately 350 km and 150 km along the southern and northern strands, respectively, bifurcating west of the East Anatolian Fault (Yan et al., 2024). The earthquake also triggered a series of geohazards, such as fault ruptures, liquefaction, landslides, rockfalls, and lateral spreading along the major ruptured faults (Wang et al., 2023). These geohazards contributed to the extensive damage along the Asi River.

Dataset

One of the research's goals is to create a rapid disaster response tool, so we must refrain from introducing data from the case study to the training dataset. Therefore, the model is trained on xBD, a dataset for assessing building damage from satellite images, and the model's final performance on unseen data is tested with the HOTOSM building damage map of the 2023 Turkey Earthquake plus MAXAR satellite images.

Training data.

xBD is one of the largest public datasets of high-resolution satellite images with building locations and damage rated before and after natural disasters. The xView2 challenge provides the dataset to push research on building damage assessment after a natural disaster. Due to the limited time and resources of the study, the author has decided only to use the highest-quality tier 1 data in the dataset, which includes six different disaster types: earthquake, tsunami, flood, volcano eruption, wildfire and wind. In addition, their damage rating contains "no damage", "minor damage", "major damage" and "destroyed". The original xBD tier 1 dataset contains 4665 sets of pre-disaster, post-disaster satellite, and ground-truth images, and every image has a pixel size of 1024 x 1024.

Case study data.

Creating a dataset to evaluate our model requires satellite images before and after the 2023 Turkey Earthquake of the Antakya region and ground truth data that contain information about all buildings destroyed and not destroyed in our area of interest. Luckily for the research, MAXAR has an open program that provides pre- and post-disaster satellite images of the 2023 Turkish Earthquake. In addition, the Humanitarian Open Street Map Team provided us with GIS data of all buildings and GIS data of all destroyed buildings due to the 2023 Earthquake in Antakya.

Methodology

Software

Google Colab.

Google Colab was the primary tool used to train and deploy the models in this research. It provides an easy-to-use platform that lets users write and execute Python code via Jupyter Notebook, run behind the scenes by Google's computing resources (GPU or CPU). Our research utilised Google Colab mainly because more computing power is needed to train a deep-learning model that processes many images.

Google Drive.

Google Drive was used to store all of the research's data on the cloud so that Google Colab could easily access it.

QGIS.

QGIS loads the GIS ground truth data provided by HOTOSM and satellite images by MAXAR. After verifying the data, the damage maps are exported with matching dimensions and coordinated with the satellite images.

Visual Studio Code.

Visual Studio Code is the code editor the project used to write minor Python scripts that can help preprocess and check the quality of the research's datasets.

GitHub.

GitHub stores every Jupyter Notebook, python scripts, and the parameters of the trained models.

Data processing

The training data was downloaded from the xView2 challenge page. Every image was PNG with a dimension of 1024 x1024 pixels. After that, they are split into three folders containing pre-disaster images, post-disaster images, pre-disaster building maps and post-disaster damage maps. According to the paper by Gupta et al. (2019), their baseline model, ResNet50, did not perform well in predicting buildings with minor and major damage. We have decided to process the post-disaster damage map to turn pixels representing "minor damage" into "no damage" and "major damage" into "destroyed". The post-disaster folder will go through another processing step to create a new folder called the post-disaster building map with all pixels marked "destroyed" turned into "non-building". Because GPU memory is required to store the feature map of each image during training, the research chose to crop the image into a 256 x 256-pixel dimension. The last step to finish the training dataset is to crop every image in every folder to 256 x 256. The final dataset should have five folders, each with the same number of images. The whole process is shown in Figure 1.

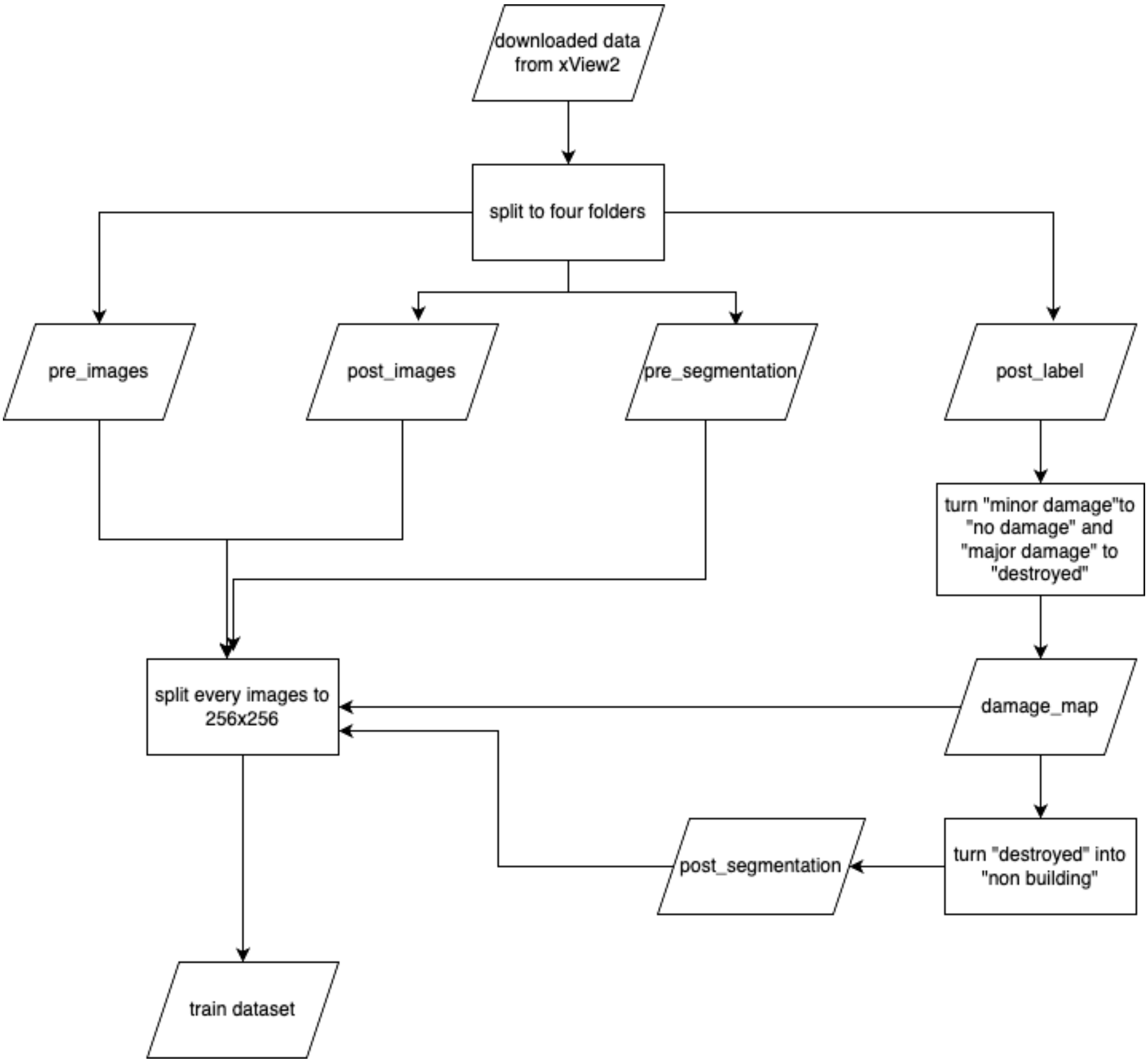

Figure 1: process of making the training dataset

The case study data first include pre- and post-disaster TIFF images of Antakya downloaded from MAXAR. The GIS building and damage data were downloaded from the HOTOSM and loaded into QGIS. After that, a PNG damage map of each TIFF image will be generated. After the last step, we will have sets of 17408 x 17408 pixels images of pre-disaster, post-disaster, and damage images. The final step will be splitting all these images into 256 x 256 and feeding them into our models. The whole process of creating the case study dataset is shown in Figure 2. The output of the models will then be merged to form the predicted damage map of Antakya.

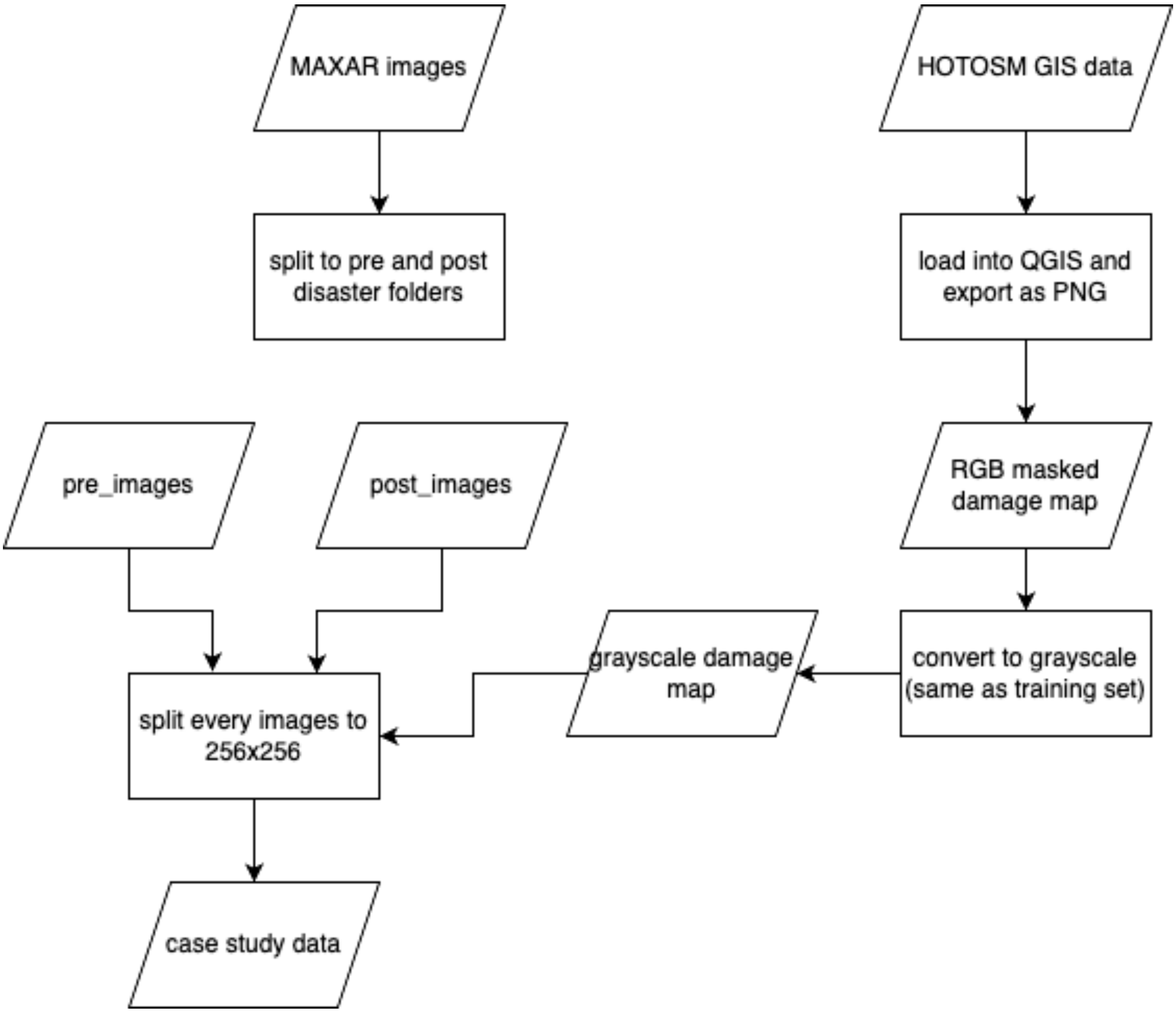

Figure 2: process of making the case study dataset

The GitHub link in the appendix allows access to every Python script used in the data processing step.

UNet neural network architecture

Our baseline model is based on UNet architecture by Ronneberger et al. (2015) in this study. Although Ronneberger et al. (2015) use the original model for biomedical imaging tasks, Bai et al. (2018) show it can also adapt to building damage assessment. Our model has the same architecture as Ronneberger et al. (2015), containing four down-sampling layers (encoder), a bottleneck layer, and four up-sampling layers (decoder). However, the main difference is that every convolution layer's size is less than that of the original model four times. The author has to reduce the size of the layers due to limited GPU memory and computing resources. Each downsampling layer contains two pairs of 2D convolution and ReLU (DoubleConv) stacked with another 2D max pooling layer. The downsampling layer returns its output and a copy of the DoubleConv layer output. The downsampling layer output is the input of the next downsampling layer, and its DoubleConv output is the skip connection to the upsampling layer, which is the main idea of UNet. The upsampling layer, which contains a 2D transposed convolution and a DoubleConv layer, will take these skip connections and the layer's output before it as input. In other words, UNet uses skip connections to utilise data from earlier stages to help the model understand the overall context and finer data detail.

The code implementation of UNet will be in the thesisUNET.ipynb file from the GitHub link in the appendix.

UNet++ Neural Network Architecture

Like its predecessor, UNet, UNet++ has the same encoder and decoder parts. However, the main difference is that it has nested skip pathways to fill the semantic feature gap between the encoder and decoder, which can improve its accuracy on image segmentation tasks (Zhou et al., 2018). One way to easily comprehend its architecture is to imagine UNet inside a bigger UNet, and so on. This nested architecture allows the decoders to have input from both low and high-level encoders, allowing a better understanding of the detail and context of images (Zhou et al., 2018).

The code implementation of UNet++ will be in the thesisUNETPP.ipynb file from the GitHub link in the appendix.

Some final adjustment

In their paper, Bai et al. (2018) apply pre- and post-disaster images by concatenating them and feeding them to UNet. In another study by Malmgren et al. (2023), they feed the pre- and post-disaster images separately into UNet and feed their respective output into another simple convolution neural network. Although Malmgren et al. (2023) did not mention it, the general idea was to make a pre- and post-disaster building map, and the difference between the two maps would be the damaged building. Since it could improve the final performance, we have implemented both UNet and UNet++ with this minor adjustment.

Training the models

UNet and UNet++ have the same hyperparameter for fairness in comparison. Due to limited resources, we trained the models for 15 epochs and a batch size of 32. The learning rate was 0.001, reduced by five times if the loss value did not reduce for three consecutive epochs. We chose the Adam optimiser on Pytorch and the cross-entropy loss function.

Evaluating the models

The trained UNet and UNet++ models would then use the case study data to create damage map predictions for the Antakya region. The f1 score of both models will be calculated by comparing the pixels' values of the predicted and the ground-truth damage map.

Visualisation

The output of the UNet and UNet++ will be processed to form two damage maps for Antakya, respectively.

Results and Discussion

UNet and UNet++ results

Training performance.

Figures 3 and 4 show the relationship between cross-entropy loss and the number of epochs trained for the two models. The line graphs show a gradual downward trend of the loss values as the epoch count increases. This shows that our UNet and UNet++ models are learning something, and our training dataset is decent. The training duration of UNet is around 5 hours for 15 epochs, and UNet++ is around 10 hours for 15 epochs.

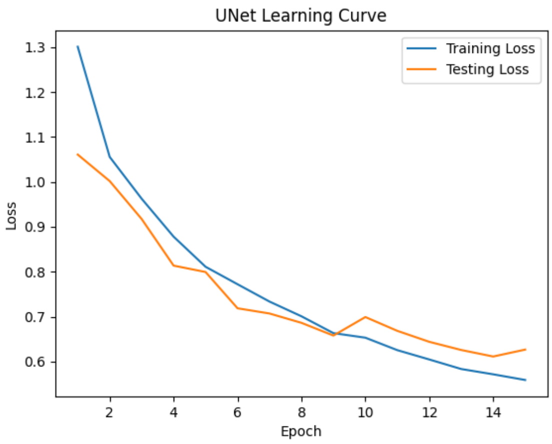

Figure 3

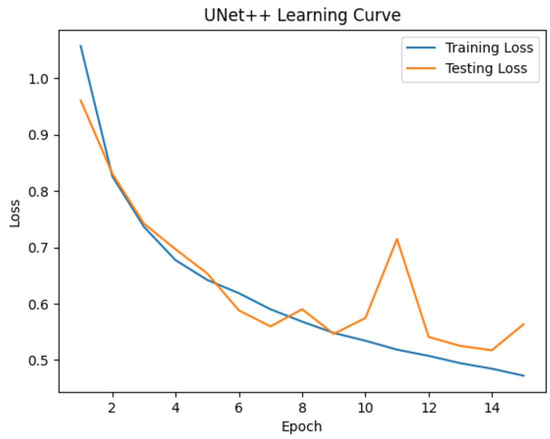

Figure 4

Table 1 and Table 2 show the f1 score for the validation data of the train data set of UNet and UNet++, respectively. Both models' f1 scores of "no damage" and "destroyed" have beaten those of Gupta et al.'s (2019) ResNet 50 baseline model of the xBD dataset. This proves that UNet and UNet++ architecture can be used for the building damage assessment task.

Table 1: UNet train dataset f1 score

Damage type	no damage	destroyed
F1 score	0.9278	0.6329

Table 2: UNet++ train dataset f1 score

Damage type	no damage	destroyed
F1 score	0.9325	0.6547

Antakya prediction.

Tables 3 and 4 show the f1 scores of the two models when predicting the damage map of Antakya. The low f1 scores of both models show that while the models perform well with the training set, their generalisation capability is still low. This can be due to the lack of diversity of the training datasets or the model being overfit.

Table 3: UNet Antakya f1 score

Damage type	no damage	destroyed
F1 score	0.9101	0.1770

Table 4: UNet++ Antakya f1 score

Damage type	no damage	destroyed
F1 score	0.8793	0.1817

Figures 6 and 7 illustrate the damage map produced by the two models. The maps show that compared to the ground truth data in Figure 5, where destroyed buildings are often clustered together, the predicted destroyed buildings are spread out, and there are many false positive instances.

Figure 5: Antakya damage map

Figure 6: UNet damage map

Figure 7: UNet++ damage map

Discussion

The results of our research showed that UNet++ performs slightly better in damage assessment than UNet. This is proved by UNet++'s higher f1 score for predicted destroyed buildings. Although both models perform well **in training**, even better than the ResNet 50 baseline model of Gupta et al. (2019), their performance in predicting Antakya destroyed buildings decreases significantly, with f1 scores of 0.1770 and 0.1817.

Several factors may cause the poor performance of both models. First is the quality of the data. Since the satellite images provided by MAXAR were intended to aid rapid disaster response, some properties of the before-and-after images can be sub-optimal—for example, the overall colour of the images, as in Figure 8. Not only is there a colour difference, but the two images also have different angles, which means the position of the roofs before and after

Figure 8

is changed. Another noticeable thing in Figure 8 is that due to the nature of satellite images, we rarely have images of buildings that align perfectly perpendicular to the ground; they are always skewed. The second factor is the size of the training dataset. For a tool that deals with imagery of a whole city or country, our model was trained with not enough data. This may have made the model unable to adapt to the Antakya scene. The third factor is that training the model for only 15 epochs may not have been enough. As shown in Figures 3 and 4, the trend of the lines shows that the loss value can still go down more, so training for more iterations may still be beneficial. Fourthly, compared to the original models proposed by Ronneberger et al. (2015) and Zhou et al. (2018), my UNet and UNet++ models have much smaller size, so this may have affected something.

Conclusion and future studies

The comparative study of UNet and UNet++ in the task of building damage assessment for the 2023 Turkey-Syria earthquakes provides valuable understandings into the capabilities and limitations of these deep learning models. The study demonstrated that UNet++ slightly outperforms UNet regarding f1 scores for destroyed buildings, indicating its potential for more accurate damage assessment. Despite the promising results during training, the models exhibited a significant decrease in performance when applied to the case study of Antakya, with f1 scores of 0.1770 and 0.1817 for UNet and UNet++, respectively. This in turn call for some further improvements like improving the data quality, expanding the training dataset, increase the number of training epochs, and optimizing the models architecture in future studies.

References

Ahmad, K., Maabreh, M., Ghaly, M., Khan, K., Qadir, J., & Al-Fuqaha, A. (2020). Developing future human-centered smart cities: Critical analysis of smart city security, interpretability, and ethical challenges. Retrieved from https://researchgate.net/publication/347398615_Developing_Future_Human-Centered_Smart_Cities_Critical_Analysis_of_Smart_City_Security_Interpretability_and_Ethical_Challenges

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), 1626. https://doi.org/10.3390/rs10101626

Bartels, S. A., & VanRooyen, M. J. (2011). Medical complications associated with earthquakes. *The Lancet, 379*(9817), 748-757. https://doi.org/10.1016/S0140-6736(11)60887-8

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

Duarte, D., Nex, F., Kerle, N., & Vosselman, G. (2018). Satellite image classification of building damages using airborne and satellite image samples in a deep learning approach. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2*, 89-96. https://doi.org/10.5194/isprs-annals-IV-2-89-2018

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*. https://arxiv.org/abs/1911.09296

Humanitarian Open Street Map Team. (2023). HOTOSM Turkey destroyed buildings (2023 Kahramanmaras Earthquake). Retrieved from https://data.humdata.org/dataset/hotosm_tur_destroyed_buildings

Malmgren, J., & Karlbrg, T. (2023). *Deep learning for building damage assessment of the 2023 Turkey earthquakes: A comparison of two remote sensing methods* (Dissertation). Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:kth

Matin, S. S., & Pradhan, B. (2021a). Challenges and limitations of earthquake-induced building damage mapping techniques using remote sensing images-A systematic review. *Geocarto International*, *37*(21), 6186–6212. https://doi.org/10.1080/10106049.2021.1933213

Matin, S. S., & Pradhan, B. (2021). Earthquake-induced building-damage mapping using explainable AI (XAI). *Sensors, 21*(13), 4489. https://doi.org/10.3390/s21134489

Maxar Technologies. (2023). Turkey earthquake 2023: Open data. Retrieved from https://www.maxar.com/open-data/turkey-earthquake-2023

Nex, F., Duarte, D., Steenbeek, A., & Kerle, N. (2019). Towards real-time building damage mapping with low-cost UAV solutions. *Remote Sensing, 11*(3), 287. https://doi.org/10.3390/rs11030287

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *arXiv*. https://arxiv.org/abs/1505.04597

Shohet, I. M., Aharonson-Daniel, L., Levi, T., Levy, R., Salamon, A., Vilnay, O., Ornai, D., Adini, B., Bar-Dayan, Y., Wei, H.-H., Shapira, S., & Levi, O. (2016). Analytical-empirical model for the assessment of earthquake casualties and injuries in a major city in Israel - The case of Tiberias. Retrieved from https://cris.bgu.ac.il/en/publications/analytical-empirical-model-for-the-assessment-of-earthquake-casua

Turkish Minute. (2023). Turkey's Feb. 6 earthquake death toll reaches 50,500: minister. https://www.turkishminute.com/2023/04/14/turkey-feb-earthquake-death-toll-reach-50500-minister/

Turkonfed. (2023). 2023 Kahramanmaras earthquake pre-assessment & status report. https://turkonfed.org/Files/ContentFile/turkonfed2023kahramanmarasearthquakepre-assessmentstatusreport021223-9583.pdf

United State Geological Survey. (2023). M 7.5 - Elbistan earthquake, Kahramanmaras earthquake sequence. https://earthquake.usgs.gov/earthquakes/eventpage/us6000jlqa/executive

Wang, X., Feng, G., He, L., An, Q., Xiong, Z., Lu, H., Wang, W., Li, N., Zhao, Y., Wang, Y., & et al. (2023). Evaluating urban building damage of 2023 Kahramanmaras, Turkey earthquake sequence using SAR change detection. *Sensors, 23*(14), 6342. https://doi.org/10.3390/s23146342

Yan, K., Miyajima, M., Kumsar, H., & et al. (2024). Preliminary report of field reconnaissance on the 6 February 2023 Kahramanmaras earthquakes in Türkiye. *Geoenvironmental Disasters, 11*(1), 11. https://doi.org/10.1186/s40677-024-00272-x

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. Retrieved from https://www.researchgate.net/publication/324574642_UNet_A_Nested_U-Net_Architecture_for_Medical_Image_Segmentation

Appendix

GitHub link to the models' code:

https://github.com/nanotnam/thesis/tree/main