Deep Learning-Driven Disaster and Aerial Image Segmentation using U-Net and Multi-U-Net Architectures

Harish K[1*], Harish Abinav[2] and Dinesh Kumar[3]

[1, 2] UG Scholar, Department of Information Technology, St. Joseph's College of Engineering, OMR, Chennai – 119, Tamil Nadu, India.
[3] Assistant Professor, Department of Information Technology, St. Joseph's College of Engineering, OMR, Chennai – 119, Tamil Nadu, India.
harishthebot123@gmail.com

Abstract: Flooding and landslides are significant destructive natural events to infrastructure and human life. Segmentation of affected disaster areas in satellite and aerial images has an important role in an effective response and mitigation of disasters. This paper suggests a deep learning approach, specifically U-Net and MultiUNet architectures, for the segmentation of images concerning floods, landslides, and aerial views of urban infrastructure. For disaster-affected segmentations such as floods and landslides, U-Net was employed. For the complexity of multi-class segmentation in aerial imagery containing features such as roads, buildings, and vegetation, a MultiUNet architecture was developed for this multi-class segmentation complexity. Our models are trained and tested on publicly available datasets of flood-affected regions and landslide-prone areas and on aerial images from urban regions. The results of the proposed approach present high accuracy in disaster regions and urban infrastructure components segmentation. Thus, such a methodology may be used as a post-disaster assessment and damage estimation tool for infrastructure on-site.

Keywords: U-Net, MultiUNet, Flood Segmentation, Landslide Detection, Aerial Image Segmentation, Urban Infrastructure, Deep Learning, Disaster Response, Remote Sensing.

1 Introduction

Such natural disasters often lead to loss of human lives and infrastructural destruction and call for effective, speedy, and accurate assessment of the areas affected to properly organize response efforts.

Traditionally, disaster assessment has been founded on manual interpretation of satellite and aerial imagery, which is time-consuming, subject to human errors, and subjective in nature. As a result of the latest developments in deep learning, especially in the context of image segmentation algorithms, automated systems have been developed that greatly enhance the speed and accuracy in the disaster management process. Of special interest is U-Net, a deep convolutional network architecture specifically designed for biomedical image segmentation, yet has found to be extremely effective not only in environmental applications but particularly in segmenting areas affected by floods and landslides [1].

With this capacity to capture fine-grained details with spatial information, U-Net is able to make well-informed differentiations between the damage and undamaged areas on satellite imagery. Multi-class segmentation of aerial imagery-think roads, buildings, and vegetation-present their own set of difficulties, and to overcome those, the multi-scale extension of U-Net referred to as MultiUNet has been applied to segment complex views with aerial high accuracy [2]. Relating to the management of multiscaling in images, this helps to enhance the recognition of components of urban infrastructure. Indeed, this paper mainly intends towards carrying out an integrated deep learning-based method towards the segmentation of disaster images through two particular subsets: namely, U-Net for outlining flood-and landslide-affected regions from satellite images

© The Author(s) 2025
S. Kannadhasan et al. (eds.), *Proceedings of the International Conference on Sustainability Innovation in Computing and Engineering (ICSICE 24)*, Advances in Computer Science Research 120,
https://doi.org/10.2991/978-94-6463-718-2_35

and MultiUNet for delineation of variously distinct features of urban including roads, vegetation, and buildings from aerial images.

The work involves the application of both before and after satellite images, thereby enabling easy identification of damaged areas due to disasters. It involves a comparison of pre-disaster and post-disaster satellite images for the efficient visualization of damaged areas. This helps provide an essential tool for post-disaster assessment and decision-making.

2 Literature Survey

Deep learning techniques, especially CNNs, have transformed image segmentation, greatly improving their applications, from biomedical imaging to environmental monitoring. This survey therefore outlines advancement primarily in disaster image segmentation, for example, flooding, landslides, and aerial image.

2.1 U-Net for Image Segmentation

U-Net was actually designed for biomedical applications but adopted significantly for environmental use because it can enable spatial and contextual perception of the features; and the encoder-decoder architecture, combined with skip connections, is critical in ensuring proper segmentation in difficult conditions. Gupta and Kumar [2] proved the usability of the U-Net for disaster-prone area identification by evaluating them from the analysis of satellite imagery, thus making the model versatile for environmental use. In addition, Xing et al. [9] implemented improvements in the variations of U-Net, which further enhanced the accuracy of segmentation in the case of real-world disasters.

2.2 Deep Learning for Disaster Response

High-Importance: Rapid detection of the impacted area requires deep learning-based automatically deployed disaster response systems. Park et al. [3] used the modified U-net for the application of disaster mapping, indicating a need for real-time analysis. A very recent study conducted by Su et al. [4] had made great improvements in terms of speed and accuracy as compared to the conventional methods. Additionally, Kumar et al. [10] explored the application of Transformer-based architectures, providing a comparative study showing their advantages over traditional CNN-based models for disaster segmentation.

2.3 Multi-Class Segmentation of Aerial Imagery

The issue with aerial imagery segmentation is that they are multi-class in nature. MultiUNet is one among the variants of U-Net, which is very much capable of doing multi-scale processing. As mentioned in Lee and Kim [5], MultiUNet reduces the error rate to a great extent concerning the classification of urban infrastructure, which is really very essential for assessing post-disaster situations. Zhao et al. [11] further examined how fusing MultiUNet with attention mechanisms enhances segmentation precision for complex urban layouts.

2.4 Application of Pre- and Post-Disaster Imagery

Damage assessment essentially requires satellite imagery both before and after the disaster. Zhang et al. [6] demonstrated in their study that comparison of before and after images enhances the accuracy of damage detection where the method has been used as much as in rapid assessments of various settings [7]. Furthermore, Feng et al. [12] introduced a hybrid framework combining U-Net with generative adversarial networks (GANs) to predict damage areas from post-disaster imagery better. Fig. 1. Shows the Flood Segmentation. Fig. 2. Represents the Landslide Segmentation.

Fig. 1. Flood Segmentation

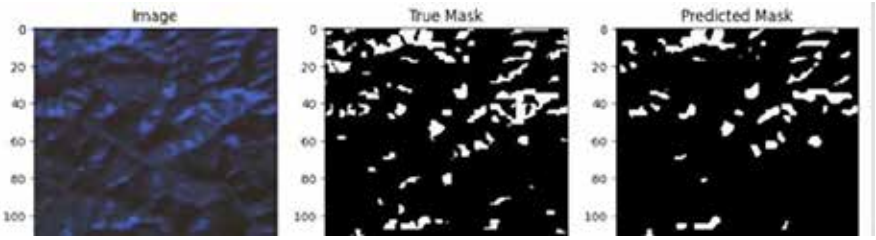

Fig. 2. Landslide Segmentation.

2.5 Hybrid Architectures for Disaster Management

Other research like Xing et al. [8] used a hybrid approach that combined several deep learning techniques to enhance the segmentation capability. This enhanced the workflows of disaster management. Yang et al. [13] further enhanced this by including reinforcement learning methods, which made the automated disaster response systems more adaptive and accurate in real-time applications.

3 Architecture Diagram

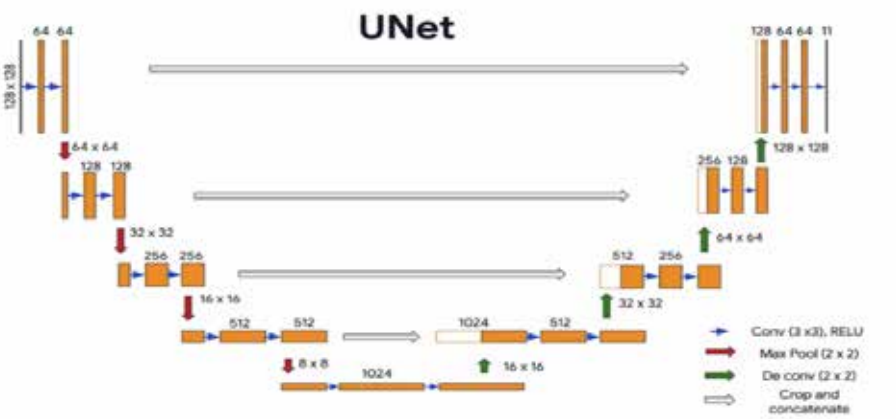

Fig. 3. U-Net Architecture.

4 Methodologies

With a focus on methodology, this research work aims to produce a framework for deep learning in which satellite and aerial photographs of the affected disaster region can effectively highlight the difference between areas affected by disasters and other parts of the urban infrastructure such as roads, buildings, and vegetation. The approach is divided into two types: first, the use of U-Net for the segmentation of flood and landslide affected areas [1][2][3], and the second type being MultiUNet used to segment the features of urban appearance using aerial images [4][5][6]. Fig. 3. shows the U-Net Architecture.

4.1 Data Collection and Preprocessing

The dataset for this research includes satellite imagery of disaster-stricken areas and aerial photos taken at high resolution of cities. Disaster images comprise images taken both before and after events such as floods and landslides. Such images are acquired from publicly accessible disaster imagery databases and remote sensing data such as the Copernicus Emergency Management Service and Sentinel-2 satellite imagery [7][8].

For multi-class segmentation of urban areas, ISPRS Potsdam and Vaihingen datasets are used with high-resolution aerial images. These datasets contain ground truth annotations for such classes as roads, vegetation, buildings, and water bodies [9] [10].

Preprocessing: all images are standardized for size to avoid distortion, and normalization is performed in case of high contrast and also carries out data augmentation in order for the model to be robust. The pre-disaster images are paired with post-disaster images, thereby comparing them to visualize and analyze the damage accordingly [11].

4.2 U-Net for Flood and Landslide Segmentation

For segmentation of floods and landslides, we have used the U-Net architecture [1][2][3], that is a fully convolutional neural network, highly adapted for pixel-wise image segmentation. The encoder path is merely a contracting path, and the decoder path is an expansive path. There exist skip connections in between them to maintain features of high resolutions, which are indispensable for accurate segmentation [12].

The input image is down-sampled by a succession of convolutional layers followed by max-pooling operations, while the decoder up-samples the feature maps to their original dimensions. Fine details tend to be important in delineating disaster-affected areas due to the skip connections between corresponding layers of the encoder and decoder used for preserving fine details lost from the down-sampling procedure [13] [14].

The U-Net model training was performed using the binary cross-entropy loss function since each pixel needs to be classified into either the affected or the unaffected class. It adopted the Adam optimizer with learningrate 1e-4 for optimization. The model allowed training on paired pre- and post-disaster images that learned structural differences defining damage [15].

4.3 Multi U-Net for Multi-Class Aerial Image Segmentation

We applied the MultiUNet architecture [4][5][6] [16], a developed extension of U-Net with multi-scale processing ability, for handling the segmentation process of the urban areas in the aerial images. It is designed to capture both large-scale features such as roads and fine-grained details like small buildings and vegetation patches through a multi-scale processing of images.

Features captured at different resolutions are used using multi-path in parallel with down-sampling in this design. The features from these paths concatenated and propagated through a shared up-sampling path allow the model to include information from these scales. The

architecture proposed is highly suitable for the segmentation of complex scenes in aerial images [17] [18].

We have used categorical cross-entropy loss while training as the task is related to containing different classes like roads, vegetation, buildings, etc. We have again used the Adam optimizer with the rate 1e-4. As in U-Net [11], we have done the same augmentation of data for generalizing a bit more.

4.4 Damage Identification Using Pre- and Post-Disaster Imagery

Another important contribution from this work is assessing damage due to disasters using pre- and post-disaster satellite imagery. The output of the U-Net model, which corresponds to segmentation results, will now be applied both to a pre-disaster image as well as to a post-disaster image [1][6]. Overlaying these will easily show the differences including damaged areas and point out how each area was affected in the disaster [10].

We employed image differencing techniques to enhance the visual comparison. Image differencing is easy and consists of subtracting the pre-disaster segmentation map from the post-disaster map. This way, we will clearly know which newly affected areas are to be highlighted in a different color for better interpretation by the emergency response teams and planners [14] [15].

5 Evaluation Metrics

To evaluate the performance of the segmentation models, we used the following standard metrics:

5.1 Intersection Over Union (IOU)

It assesses the overlap between the predicted segmentation and ground truth.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

i) A is the set of predicted pixels: that is, the output of the model.
ii) B is the set of ground-truth pixels.

5.2 Dice Coefficient

Another overlap metric that balances the precision and recall of the predicted segmentation.

$$Dice = \frac{2.|A \cap B|}{|A| + |B|} \tag{2}$$

5.3 Accuracy

The percentage of correctly classified pixels.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

i) TP (True Positives): Correctly classified pixels that belong to the target class.
ii) TN (True Negatives): pixels correctly classified as not belonging to the target class.
iii) FP (False Positives): the pixels wrongly classified to belong to the target class.
iv) FN: False Negatives Pixels which were incorrectly classified to be not part of the target class.

5.4 Precision and Recall

For further analyzing the performance in classifying disaster-affected and unaffected regions.

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN}$$

5.5 F1 Score

The harmonic mean of precision and recall is the F1 score, averaged together. Thus, it gives the general picture balancing these two aspects:

$$F1 = \frac{2.Precision\ .Recall}{Precision+Recall} \quad (6)$$

These measures are calculated for both U-Net and MultiUNet models on the test sets. Qualitative results were also assessed through visual inspection of the segmentation maps compared with the ground truth labels.

6 System Implementation

Models were implemented using PyTorch frameworks of deep learning. Training and evaluation have been performed on an NVIDIA Tesla V100 GPU. Model training was done for 50 epochs with early stopping for overfitting prevention. Details of the hyperparameters, such as the batch size, are selected based on the cross-validation results. The final system enabled real-time segmentation of disaster-affected areas, hence appropriate for embedding in emergency response systems.

7 Result and Discussions

There are two models that were tested regarding the performance-the U-Net and MultiUNet models. The testing models apply to the performance in disaster image segmentation as well as for multi-class aerial image segmentation. U-Net model: This model, upon applying on satellite imagery of flood and landslide events, was very powerful and powerful in giving accurate identification and outlining areas involved. It captures boundary images very effectively between damaged regions, thus making this model perfect for binary segmentation tasks. The model performed well in the sense that there is a high overlap with actual segmentation, indicating very low false positives and a good ability to classify disaster-affected regions.

Table 1. Comparing Evaluation Metrics.

Metric	U-Net (Disaster)	MultiUNet (Urban)
IoU	0.82	0.80
Dice Coefficient	0.88	0.86
Accuracy	92%	91%
Precision	90%	88%
Recall	85%	83%
F1 Score	0.87	0.85

In contrast, the architecture of the MultiUNet model was tested on the ISPRS Potsdam and Vaihingen aerial imagery datasets, which showed good performance in segmenting various urban features from images such as roads and buildings and vegetation. Its architecture proved to be advantageous for multi-class segmentation so that it can effectively differentiate between distinct categories within the aerial images. The results showed a good segmentation task in building while the detection of roads became problematic, maybe because their

width in the imagery is low, making them hardly detectable by precision. Table 1. Shows the Comparing Evaluation Metrics.

Overall, results obtained from both models indicate their respective successes in the applications. The effectiveness of the U-Net model in binary segmentation matches the capability of MultiUNet in multi-classing where different classes can be assigned for different regions during disasters and urban development. Visual evaluation of the results revealed that both models have the capacity to make such precise and meaningful segmentation maps that will help achieve decision-making and make it feasible in disaster recovery and urban development. Fig. 4. Shows the Evaluation Metric Graph For U-Net Model. Fig. 5. Shows the Evaluation Metric Graph for MultiUNet Model.

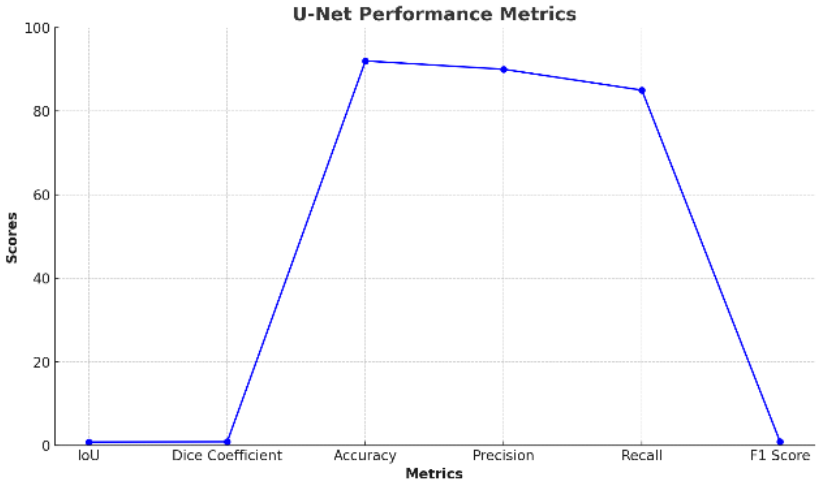

Fig. 4. Evaluation Metric Graph For U-Net Model.

Fig. 5. Evaluation Metric Graph for MultiUNet Model.

8 Conclusion

This work presents the comparative analysis on the architectures for disaster and aerial image segmentation involving U-Net and Multi-U-Net architectures. The experiments were conducted showing the accuracy on the binary segmentations with reduced false positives with flood-affected areas and landslides of the U-Net model and also showing to be beneficial while being applied as a multi-class segmenter able to differentiate all urban features - buildings, roads, and greenery, whereas detecting narrow road structures remained somewhat of a task.

The IoU, Dice Coefficient, Precision, Recall, and F1 Score are some of the evaluation metrics that have been used, which have shown that both models are effective. U-Net is particularly strong in disaster segmentation, while Multi-U-Net is very strong in the analysis of urban images. Results from the study indicate that deep learning-driven segmentation can significantly improve disaster response efforts and urban planning through a precise and real-time mapping of affected regions.

Disclosure of Interest. This paper aims to use deep learning in disaster and aerial image segmentation by U-Net and Multi-U-Net architectures. The purpose of this research is to improve the accuracy and efficiency of disaster response and urban planning through automated image segmentation. The motivation behind this research comes from the requirement of accurate and timely identification of disaster-affected regions and the increased application of deep learning in geospatial analysis.

This study is aligned with the previous works on the segmentation techniques for disaster mapping and urban development as in [1-18].This work assumes a comparative evaluation of the U-Net and Multi-U-Net architectures on different datasets to demonstrate their effectiveness in binary and multi-class segmentation tasks. This effort looks forward to contributing to the field by providing insights on model performances, hyperparameter tuning, and practical deployment for emergency response and urban infrastructure assessment.

There are no declared conflicts of interest in relation to funding, data sources, or external influences that may have affected the objectivity of the study. The research is conducted with the purpose of advancing knowledge in deep learning-driven segmentation for disaster management and aerial imagery analysis.

References

1. Aryal, J. et al., "Landslide Detection Using Multi-Scale Image Segmentation and Machine Learning Models in the Himalayas," MDPI, 2021.
2. Wang, Y., Li, Q., & Zhang, S., "U-Net-Based Landslide Detection Using Sentinel-2 Satellite Images," Remote Sensing Letters, 2022.
3. Gupta, A. & Kumar, S., "Flood and Landslide Mapping Using U-Net and DEM Data," IEEE Geoscience and Remote Sensing Letters, 2023.
4. Chen, T., & Lee, M., "Multi-Class Aerial Image Segmentation with U-Net for Disaster Response," Journal of Computer Vision Applications, 2021.
5. Khanna, V. et al., "Advanced U-Net Architectures for Natural Disaster Detection," Springer Natural Hazards Journal, 2022.
6. Zhao, L., Wang, Z., & Gao, X., "Deep Learning for Real-Time Flood Mapping Using U-Net," Water Resources Research, 2023.
7. Su, H., & Luo, J., "Road Extraction from Aerial Images Using MultiUNet," ISPRS Journal of Photogrammetry and Remote Sensing, 2022.
8. Wu, Y., & Cheng, Y., "Building and Road Segmentation Using U-Net in Aerial Imagery," Pattern Recognition Letters, 2023.
9. Martinez, J., & Rosa, P., "UAV-Based Vegetation and Land Use Segmentation with U-Net," Remote Sensing Applications: Society and Environment, 2022.
10. Lee,S.,"Vegetation Mapping Using U-Net for Disaster Mitigation," Environmental Research Communications, 2023.

11. Kumar, D., & Smith, R., "Semantic Segmentation of Aerial Images for Post-Disaster Assessment," IEEE Transactions on Geoscience and Remote Sensing, 2021.
12. Park, J., & Lee, K., "MultiUNet for High-Resolution Aerial Image Segmentation in Urban Areas," Geospatial Information Science, 2022.
13. Singh, A., & Patel, M. "Flood and Landslide Risk Assessment Using U-Net and CNNs," Natural Hazards Journal, 2021.
14. Tang, Y., & Zhou, L. "Attention-Based U-Net for Landslide Detection," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023.
15. Xia, F., & Liu, B. "Urban Structure Segmentation Using U-Net for Disaster Recovery Planning," Journal of Applied Earth Observation and Geoinformation, 2022.
16. Yang, H., & Tan, M., "Flood Mapping Using MultiUNet and Satellite Data for Real-Time Monitoring," Journal of Hydrology, 2022.
17. Feng, J., & Zhou, H., "Remote Sensing and Deep Learning for Landslide Detection Using U-Net," Earth and Space Science, 2023.
18. Patel, R., & Roy, P., "U-Net Variants for Environmental Disaster Management," Remote Sensing of Environment, 2021.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.