

Research on Automated Post-Earthquake Building Damage Assessment Method

Siao Liu ^{1*}, Ye Tian ¹, Lei Chen ¹, Ning Wang ¹, Jianyu Wang ¹, Yihuan Li ¹

¹ Satellite Communications Branch China Telecom Co. Ltd, Beijing, China - lsa.wx@chinatelecom.cn, tianye6@chinatelecom.cn, chenlei.wx@chinatelecom.cn, wangning.wx@chinatelecom.cn, wangjy.wx@chinatelecom.cn, liyihuan@chinatelecom.cn

Keywords: Post-earthquake Damage Assessment, Segformer, Satellite Imagery, Grid-based Analysis, Large Language Model.

Abstract

This study addresses the challenges of low efficiency and limited scalability in traditional post-earthquake building damage assessment methods by proposing an automated assessment framework. The approach combines an improved Segformer model, a grid-based quantitative statistical method, and LLM-driven report generation for scalable and accurate structural damage assessment from high-resolution satellite imagery. First, an improved Segformer model is developed to extract and compare pre- and post-earthquake building footprints, with optimized feature fusion and training strategies tailored for post-disaster scenarios. The model effectively detects changes in building footprints under complex conditions. Second, the study introduces a grid-based quantitative statistical method that divides the affected area into uniform cells, within which damage is classified into four severity categories. To further streamline the process, the workflow is integrated into Dify, allowing for automated processing, interpretation, and report generation via LLMs. This integration enables quick and consistent delivery of actionable insights to decision-makers, reducing the need for human intervention. The method was validated using data from the 2025 Shigatse earthquake, where the model achieved a MIOU over 86% for building footprint extraction, and the damage classification showed strong alignment with ground-truth data. This study provides an efficient and scalable solution for post-earthquake building damage assessment, significantly enhancing disaster response capabilities and urban resilience planning.

1. Introduction

Earthquakes cause widespread destruction to buildings and infrastructure, posing significant challenges for urban resilience and disaster recovery. Prompt and accurate post-disaster damage assessment is critical to guide emergency response efforts, prioritize resources, and inform reconstruction plans. However, traditional field surveys, while highly accurate, are time consuming, labor intensive, and often impractical for large-scale or inaccessible areas (Wu et al., 2021; Ma et al., 2019; Nex et al., 2019).

As a result, remote sensing technologies, particularly high-resolution optical satellite imagery, have become powerful tools for post-earthquake damage analysis. In recent years, deep learning techniques, especially convolutional neural networks (CNNs), have been widely applied to building extraction and change detection tasks using remote sensing imagery (Kalantar et al., 2020; Wu et al., 2021). These models can identify building footprint changes by comparing pre- and post-disaster images. However, CNN-based approaches often struggle under complex post-disaster conditions due to their limited ability to capture long-range dependencies and contextual information (Ma et al., 2019; Xu et al., 2019).

Transformer-based architectures, such as SegFormer, have shown promising results by incorporating global attention mechanisms and multi-scale feature fusion, which enhance performance in complex scenarios (Xie et al., 2021). Other recent approaches, including DamFormer and Siamese Transformer frameworks, have demonstrated improved accuracy for multi-temporal change detection and damage assessment (Chen et al., 2022; Bandara and Patel, 2022; Mohammadian and Ghaderi, 2022). However, there remains a lack of research specifically optimizing these models for

detecting building footprint changes before and after earthquakes.

Another limitation of current methods is their focus on image segmentation or damage localization without embedding results into a structured and quantitative framework suitable for large-scale assessment. While many models can detect changes, they often do not generate interpretable or actionable insights for disaster management. Moreover, the reporting process in most existing systems remains manual, resulting in delays and inefficiencies during emergency response.

To address these challenges, this study proposes a novel end-to-end framework that integrates an improved SegFormer model for detecting building footprint changes, a grid-based approach for quantitative damage classification, and automated reporting through large language models (LLMs). The improved SegFormer model is optimized for post-disaster conditions, enabling accurate change detection. The grid-based approach allows scalable and interpretable classification of affected areas, and the LLM-driven reporting module converts evaluation outputs into structured disaster assessment reports. This integrated system enhances both the technical robustness and operational efficiency of post-earthquake building damage assessment, contributing practical value to emergency decision-making and urban recovery planning.

2. Building Footprint Change and Quantitative Evaluation Method

2.1 Improved SegFormer for Building Change Detection

Detecting building footprint changes in post-earthquake scenarios is crucial for accurate damage assessment and effective recovery planning. Traditional methods and CNN-based approaches often underperform in complex settings—

such as scenes with debris accumulation and variable lighting—because they lack the capacity to model long-range spatial dependencies and contextual variation (Chen et al., 2023; Yan et al., 2022). Transformer-based architectures like SegFormer, which integrate global attention and multi-scale feature fusion, have demonstrated superior ability to detect building footprint changes under such challenging conditions (Xie et al., 2021). Other transformer-driven models—such as Fully Transformer Network (FTN), SiamixFormer, and attention-based multiscale detector frameworks—have achieved state-of-the-art results in remote sensing change detection (Wan et al., 2022; Mohammadian and Ghaderi, 2022; Yan et al., 2022). However, few studies have focused explicitly on applying these architectures to building footprint change detection in real post-disaster earthquake contexts, indicating a need for more specialized adaptation within this domain.

2.1.1 Multi-Head Mixed Convolution (MHMC): A major challenge in traditional transformer models is their difficulty in effectively capturing fine-grained local features while processing large, high-resolution images. In the early stages of the SegFormer architecture, self-attention mechanisms are typically used, but they may struggle with capturing the necessary local details for accurate footprint change detection, especially in cluttered post-disaster imagery.

$$MHMC(X) = \text{Concat}(DSCn_1 \times n_1(x_1), \dots, DSCn_k \times n_k(x_k)), \quad (1)$$

where X = Input vector of MHMC layer
 Concat = Connection of network tensors
 DSC = Depthwise Separable Convolution
 x = Input features, $[x_1, x_2, \dots, x_k]$
 n_i = Convolution kernel size, $n_i \in \{3, 5, \dots, N\}$

To address this, the proposed model incorporates Multi-Head Mixed Convolution (MHMC) into the early layers of the encoder. Unlike traditional self-attention, which primarily focuses on long-range dependencies across the entire image, MHMC allows the model to prioritize the extraction of local features. This is crucial for post-earthquake scenes where fine details, such as changes in the outline of buildings or the presence of debris, are significant indicators of damage.

MHMC works by using depthwise separable convolutions for multiple convolutional heads, each capturing features at different scales. For example, one head might capture small-scale changes at the pixel level, such as the displacement of building boundaries due to partial collapse, while another head might focus on broader, large-scale changes, like the formation of new gaps or the collapse of entire structures. The outputs from each head are concatenated and passed to the next stage, allowing the model to leverage multi-scale features without incurring the computational overhead typically associated with multi-scale processing. This approach enhances the model's ability to detect changes across a variety of scales, ensuring that both small and large changes are captured with high accuracy.

2.1.2 Scale-Aware Aggregation (SAA): While MHMC improves the local feature extraction capabilities of the network, it is equally important to integrate these local features across different scales to capture the full context of the building footprint changes. To achieve this, Scale-Aware Aggregation (SAA) is introduced as an improvement to the network's feature fusion process. The specific structure of SAA is shown in Figure 1.

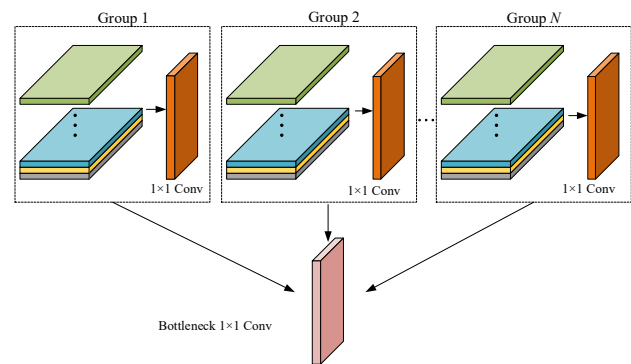


Figure 1. SAA module.

SAA aggregates features from multiple convolutional heads with varying receptive fields to improve the interaction between different scales of features. This helps the network better understand the spatial relationships between local and global changes in the building footprints. For example, changes at the local level need to be integrated with larger-scale information (e.g., the overall destruction or collapse of a building) to form a complete assessment of the damage.

The SAA method works by grouping features from different heads and performing intra-group and inter-group feature fusion. This process enables the model to combine fine-grained details from lower-level features with more global, contextual information from higher-level features. The SAA module uses a lightweight aggregation structure, where the features from different groups are fused through a reverse bottleneck structure, allowing for efficient computation while ensuring rich contextual information is preserved.

This feature fusion mechanism improves the model's ability to deal with complex post-disaster images by preserving both local details and global context, making it more robust to the challenges posed by debris, partial collapse, and other post-earthquake conditions.

Together, the MHMC and SAA improvements significantly enhance SegFormer's ability to detect building footprint changes in challenging post-earthquake scenarios. These innovations allow the model to efficiently capture both local and global features, ensuring high accuracy and computational efficiency even in complex, cluttered environments. This makes the model particularly effective for large-scale, real-time damage assessment in disaster-stricken areas.

2.2 Grid-based Quantitative Evaluation Method

To accurately assess the extent of building damage in post-earthquake scenarios, adopting a grid-based quantitative evaluation method has become increasingly prevalent. Such approaches divide the affected region into uniform spatial units, aggregating building footprint changes within each grid cell to support systematic and scalable damage assessment (Duarte et al., 2018; Corbane et al., 2011). Recent studies have demonstrated that grid-based frameworks, particularly when combined with change detection outputs, can effectively visualize damage distribution across urban areas and enable automated statistical summarization (Tu et al., 2016; Adriano et al., 2020). Within each grid cell, calculated building footprint change metrics are used to classify damage severity—such as area loss—providing interpretable data for resource allocation and recovery planning. Such quantification supports both spatial visualization and facility-level decision support.

2.2.1 Grid Division and Area Calculation: The first step in the grid-based evaluation method is to divide the earthquake-affected area into regular, uniformly-sized grids. The grid size is chosen based on the resolution of the satellite imagery and the scale of damage to ensure that each grid can capture relevant changes in building footprints. The size of each grid is selected to ensure that it corresponds to an appropriate area, neither too small to lose contextual information nor too large to miss smaller-scale changes.

Once the grid division is complete, the change in building footprint for each grid is calculated by comparing the pre- and post-earthquake satellite images. This comparison is facilitated by the improved SegFormer model, which accurately detects changes in building boundaries, including the displacement of buildings or areas of collapse. The area of change is quantified by measuring the difference in building footprint size before and after the earthquake.

In addition to simple area change detection, Weighted Area Change is introduced to enhance the accuracy and representativeness of the damage assessment. Instead of treating the overall grid area equally, the weighted area change method gives more significance to the proportion of the building footprint that has been altered. For instance, if only a small part of a building within a grid has been damaged (e.g., a corner or a small section of a wall), the total area change for that grid may be relatively small, but the damage could still be significant in terms of its impact on the structure. Therefore, grids with a higher proportion of the building footprint changed are weighted more heavily in the damage classification.

To implement this, the weighted area change is calculated by determining the ratio of the damaged building area within the grid to the total area of the building footprint in that grid. This ratio is then used to assign a weighted value to the grid's overall damage level. For example, a grid where 70% of the building area has been damaged would be weighted higher than a grid where only 10% of the building footprint has been affected. This ensures that grids with significant structural changes are given more importance in the overall damage assessment, providing a more accurate reflection of the extent of destruction.

2.2.2 Damage Classification and Quantification: After the area change within each grid is calculated, the grids are classified into four damage categories based on the magnitude of the weighted area change. The categories are as follows:

1. Severely Damaged: Grids where a significant portion of the building footprint has been altered, indicating complete collapse or destruction.
2. Moderately Damaged: Grids where the building footprint has been partially altered, but not entirely destroyed.
3. Slightly Damaged: Grids where only minor changes to the building footprint have occurred.
4. Unchanged: Grids where no noticeable change has occurred in the building footprint.

These classifications provide a quantitative distribution of damage across the affected area, helping to visualize and understand the extent of damage in various regions. By aggregating the grid statistics, the system generates a clear

report of the total area affected by each level of damage, along with the number of grids in each category.

2.2.3 Visualization and Reporting: The grid-based quantitative evaluation system not only provides detailed statistical data but also facilitates visualization of the damage across the affected region. The damage distribution can be represented on color-coded maps, making it easier for decision-makers to understand the extent of destruction.

In addition to visual representations, the system generates quantitative reports, summarizing the total area affected and the severity of the damage in each grid. These reports can be used for resource allocation, recovery planning, and further analysis of the disaster's impact.

3. Automated Framework for Post-Earthquake Building Damage Assessment

To meet the urgent demand for timely, standardized, and scalable post-earthquake damage assessment, this study proposes a fully automated assessment framework built upon Dify, a large language model (LLM)-driven workflow orchestration platform. The framework is designed to streamline the entire process—from disaster input and satellite image acquisition to building footprint change detection, grid-based quantitative analysis, and automatic report generation—by integrating advanced deep learning techniques and remote sensing knowledge with intelligent task coordination and natural language generation. This end-to-end automation significantly reduces the reliance on manual interpretation and enables consistent disaster analysis that is both fast and scalable, even in vast or inaccessible regions. The overall process is shown in Figure 2.

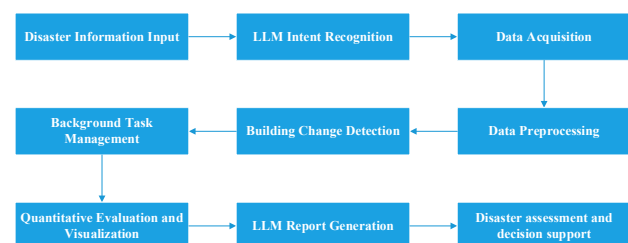


Figure 2. Automated damage assessment process.

The proposed system consists of eight tightly integrated stages that operate sequentially through a Dify-controlled pipeline. (1) Disaster Information Input begins the process, where users submit key parameters such as the earthquake's location, magnitude, and affected area. (2) In LLM Intent Recognition, the LLM parses this input to determine the appropriate data sources, analytical models, and spatial scopes required for processing. Based on the parsed intent, (3) Satellite Data Acquisition is triggered automatically. The system queries high-resolution pre- and post-event satellite images through APIs such as Google Earth Engine or Sentinel Hub, applying filters based on cloud cover and acquisition time to ensure suitable imagery. Once the data is collected, (4) Data Preprocessing follows, which includes atmospheric correction to normalize reflectance values, precise image co-registration to ensure spatial alignment, and clipping to the specified area of interest. These steps prepare the imagery for pixel-level comparison by minimizing radiometric and geometric distortions.

The core analysis takes place in stage (5), Building Footprint Change Detection, where the improved SegFormer model is

invoked. This transformer-based semantic segmentation model has been optimized to detect building footprint changes under complex post-disaster conditions, such as debris coverage, structural collapse, and lighting inconsistency. By leveraging multi-head mixed convolution (MHMC) and scale-aware aggregation (SAA), the model extracts robust multi-scale features that accurately capture differences in building shapes, sizes, and boundaries between the two time phases. The model outputs binary masks of changed vs. unchanged buildings, which serve as the basis for quantitative evaluation. Task execution across the pipeline is managed by (6) Celery, a distributed task queue framework integrated with Dify, which enables real-time monitoring, parallel computation, and fault-tolerant scheduling of long-running processes such as model inference and data processing.

In stage (7), Grid-Based Quantitative Evaluation and Visualization, the extracted change masks are spatially aggregated into uniform grids that segment the disaster area into fixed spatial units. Within each grid, the total changed building area is calculated and further refined using a weighted area change ratio, defined as the proportion of the changed footprint relative to the original building area. This approach avoids overestimation in sparsely built grids and allows for more representative classification of structural impact. Based on predefined thresholds, each grid is assigned one of four severity levels—severe, moderate, slight, or unchanged—forming the basis for damage heatmaps and statistical summaries. Finally, (8) LLM-Generated Damage Assessment Report is executed, where the LLM interprets both the statistical data and annotated imagery, referencing a curated remote sensing knowledge base. The output is a natural language report that includes damage summaries, quantitative tables, and spatial visualizations, formatted for direct use by emergency managers, government agencies, and urban planners.

The framework offers multiple advantages over traditional or semi-automated approaches. It eliminates subjective variability in damage interpretation, accelerates processing from hours to minutes, and ensures repeatable assessment results across different disaster events. Moreover, the LLM's integration with domain knowledge allows it not only to summarize results but also to highlight patterns, identify high-impact zones, and generate narrative insights previously dependent on expert intervention. The modular design also ensures extensibility: the workflow can be adapted to different disaster types (e.g., floods, landslides) or alternate data sources by substituting specific modules without changing the overall architecture. In conclusion, this automated framework demonstrates a practical and intelligent solution for post-earthquake damage assessment, bridging advanced remote sensing analytics with AI-driven decision support to enable efficient, informed, and large-scale disaster response.

4. Experiment

4.1 Experimental Environment

This study's experiments were conducted on a high-performance deep learning workstation using the PyTorch 2.1.0 framework. The workstation is equipped with two NVIDIA RTX A6000 GPUs, each providing 48GB of video memory, enabling the training of transformer-based models with high-resolution satellite imagery. All experiments were executed under the CUDA 11.8 platform to ensure computational efficiency and stability.

4.2 Study Area and Data Sources

This study uses the earthquake event that struck Dingri County, Shigatse, Tibet on January 7, 2025, as a case study to evaluate the performance of the proposed automated building damage assessment framework. The selected experimental area is a local urban region within Dingri County, bounded by coordinates (87.43°, 28.53°) in the northwest and (87.48°, 28.48°) in the southeast. This area was chosen due to its relatively high building density and visible post-disaster damage signals.

Two high-resolution (approximately 0.8m) GF-2 optical satellite images were used: one acquired before the earthquake on December 24, 2024, and the other after the earthquake on January 8, 2025. As shown in Figure 3 and Figure 4. The images were obtained from the national remote sensing data platform. Prior to analysis, the imagery underwent standard remote sensing preprocessing, including radiometric calibration, atmospheric correction, and orthorectification, to ensure accurate spatial and spectral consistency between the two datasets.

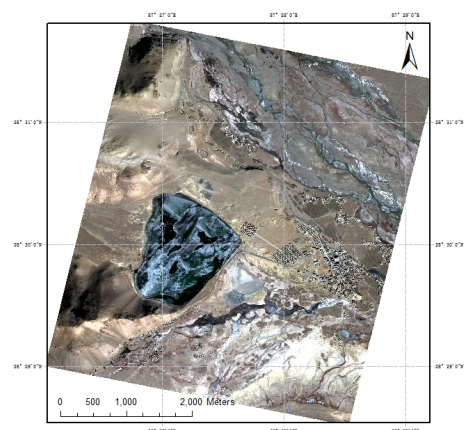


Figure 3. Pre-earthquake satellite image.

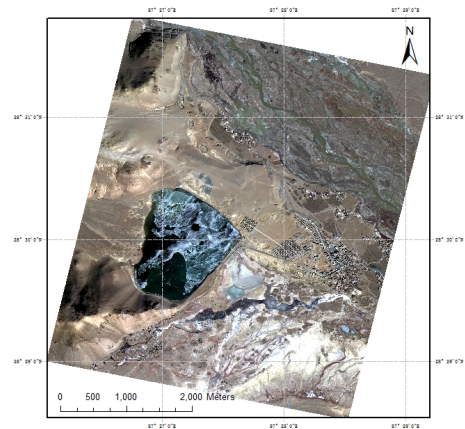


Figure 4. Post-earthquake satellite image.

4.3 Experimental Procedure

The processed imagery was used as input to the improved SegFormer model, which was configured to perform building footprint extraction for both time points. The model outputs binary segmentation maps representing detected building regions in each image. Subsequently, a pixel-wise comparison

of pre- and post-event building footprints was performed to determine areas of change, indicating possible building damage, collapse, or removal.

To evaluate the accuracy of the model, a reference dataset was manually constructed. Experts performed detailed digitization of pre- and post-earthquake building outlines within the study area using visual interpretation. These manually annotated vector maps were then used to generate three reference categories: increased, decreased, and unchanged building areas. As shown in Figure 5. The automatic footprint change detection results were compared to this ground-truth dataset to verify consistency and reliability.

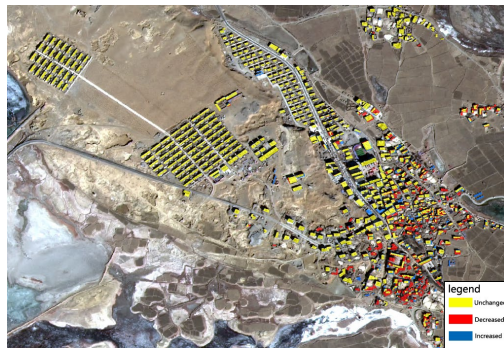


Figure 5. Plotting results of building changes.

In addition to pixel-level validation, the study also implemented grid-based quantitative analysis. The entire study area was divided into uniform spatial units using a grid resolution of $0.0003^\circ \times 0.0003^\circ$ (longitude $\times$ latitude). Within each grid, the total building footprint area loss was computed, and a weighted area change ratio was applied to account for the relative magnitude of structural loss with respect to pre-disaster building coverage. This method enabled a more representative classification of damage severity, particularly in mixed or low-density construction zones. Each grid was then classified into one of four categories: severely damaged, moderately damaged, slightly damaged, or unchanged.

The final phase of the workflow was conducted via the Dify-based orchestration platform, which automatically invoked each module in sequence, monitored task status via Celery, and triggered the LLM-based report generation component. The large language model synthesized statistical indicators and change maps into a natural language report, complete with embedded tables, summary descriptions, and thematic visualizations. The entire process, from data input to report output, was executed without manual intervention, demonstrating the viability of full automation in real disaster assessment scenarios.

4.4 Accuracy Evaluation

To verify the consistency between the extracted results and the manually annotated ground truth, three evaluation metrics were employed: Overall Accuracy (OA), Intersection over Union (IOU), and Mean Intersection over Union (MIOU).

Overall Accuracy (OA) refers to the ratio of correctly classified pixels to the total number of pixels. It serves as a global indicator for evaluating the overall performance of the classification task.

Intersection over Union (IOU) is defined as the ratio of the intersection to the union between the predicted segmentation and the ground truth for a given class. It can be used to assess

the segmentation accuracy of individual categories, such as increased area, decreased area, and unchanged area.

Mean Intersection over Union (MIOU) is the average of the IOU values computed for each class, and it reflects the overall segmentation accuracy of the model across all categories.

The formulas for these evaluation metrics are defined as follows:

$$OA = \frac{TP + TN}{TP + FP + FN + TN}, \quad (2)$$

$$IOU = \frac{TP}{TP + FP + FN}, \quad (3)$$

$$MIOU = \frac{IOU_{increased} + IOU_{decreased} + IOU_{unchanged}}{3}, \quad (4)$$

where TP, FP = true positive and false positive
 TN, FN = true negative and false negative

4.5 Results and Analysis

The framework produced accurate and reliable results. To maintain the clarity of the diagram, some key areas are intercepted for visual display. Compared with manually annotated ground truth, the model achieved an Overall Accuracy (OA) of 93.16% and a mean Intersection over Union (MIOU) of 86%. The IOU scores for each class were 83.13% for increased areas, 84.63% for decreased areas, and 90.24% for unchanged areas. As shown in the following figure, Figure 6 shows the real results, and Figure 7 shows the automatic extraction results, which maintain a high degree of consistency, in which the red area reveals the key areas of concern for disaster research and judgment. These results confirm the effectiveness of the improved SegFormer model in detecting building footprint changes under complex post-earthquake conditions.

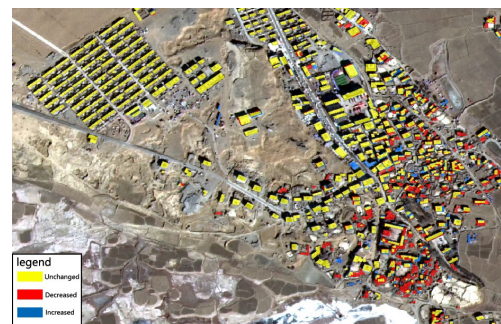


Figure 6. Plotted real results.



Figure 7. Automatic extraction results.

Quantitatively, the total building footprint in the study area decreased from 230,071.74 m² (pre-event) to 212,419.85 m²

(post-event), indicating a net loss of 17,651.89 m², which reflects the extent of physical damage. The grid-based classification yielded the following results: 238 grids were classified as severely damaged, 202 as moderately damaged, and 867 as slightly damaged. These results were visualized in a series of maps showing the spatial distribution of damage intensity across the region. The pre-earthquake building overlay map (Figure 8) and the post-earthquake building overlay map (Figure 9) illustrate the effectiveness of the improved Segformer in capturing structural changes. The building damage heatmap (Figure 10) clearly highlights severely affected zones, which closely align with field observations and government-reported impact areas. The overlaid change maps and damage heatmaps, along with extracted statistical charts, were integrated into a structured LLM-generated disaster assessment report.

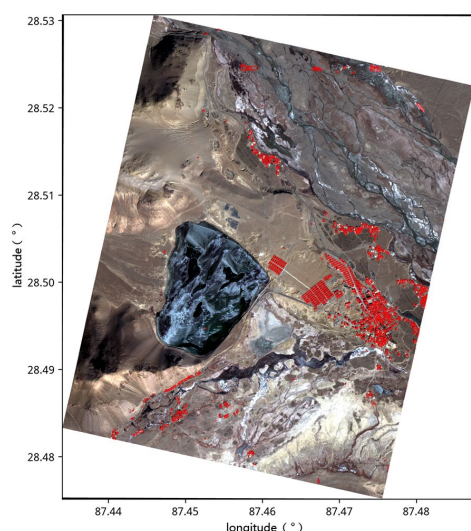


Figure 8. Pre-earthquake building overlay map.

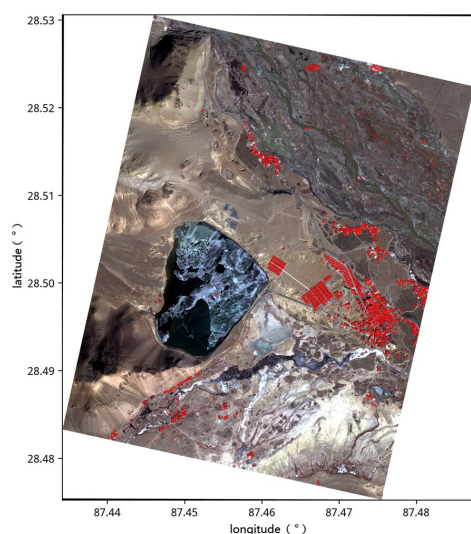


Figure 9. Post-earthquake building overlay map.

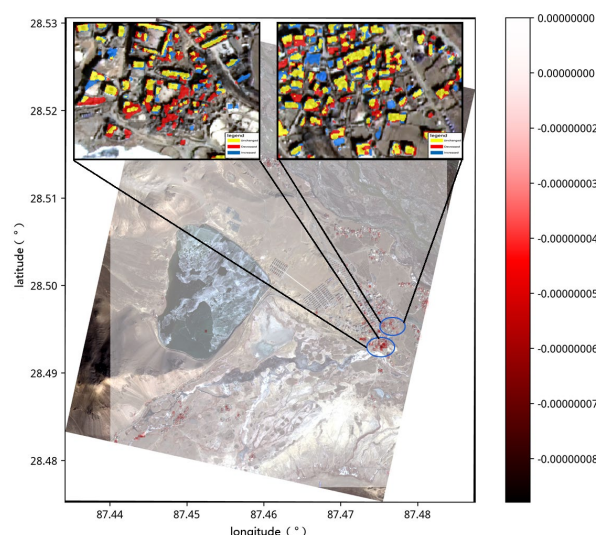


Figure 10. Heatmap of building damage severity.

It is important to note that all results are confined to the selected test region and based on satellite observations. The findings are intended for scientific research purposes only and should not be interpreted as official disaster loss statistics. Nonetheless, the high degree of automation, spatial consistency, and accuracy demonstrated by the proposed method confirms its potential as a valuable tool for emergency response and post-disaster decision-making.

Overall, this experiment validates the framework's capacity to deliver fast, scalable, and reliable building damage assessments, combining deep learning segmentation, grid-based spatial analytics, and LLM-driven interpretation. The integration of satellite data, transformer-based detection models, and language-driven reporting offers a practical solution with strong potential for deployment in future disaster scenarios across diverse geographic regions.

5. Conclusion

This study proposed a fully automated framework for post-earthquake building damage assessment by integrating high-resolution satellite imagery, an improved SegFormer-based footprint change detection model, grid-based quantitative evaluation, and large language model (LLM)-driven reporting. The framework addresses key limitations in traditional assessment workflows by enabling end-to-end automation—from data acquisition and preprocessing to damage classification and report generation—within a Dify-based orchestration environment. The improved SegFormer model incorporates multi-head mixed convolution and scale-aware aggregation mechanisms, significantly enhancing its ability to extract building footprint changes under complex post-disaster conditions.

A case study using the 2025 Shigtse earthquake demonstrated the framework's effectiveness. The proposed method achieved an overall accuracy of 93.16% and a MIOU of 86% when validated against manually annotated data. Grid-based damage classification, based on weighted area change, revealed clear spatial patterns of structural impact, supporting visual interpretation and statistical analysis. The automatically generated LLM report provided comprehensive summaries and intuitive visualizations, enabling direct support for decision-making.

Overall, the framework offers a scalable, interpretable, and operationally practical solution for post-earthquake building damage assessment. By integrating advanced deep learning and natural language generation technologies, it contributes to the development of intelligent remote sensing systems and provides valuable support for rapid disaster response and recovery planning.

References

- Adriano, B., Yokoya, N., Xia, J., Miura, H., Liu, W., Matsuoka, M., Koshimura, S., 2020. Learning from multimodal and multitemporal earth observation data for building damage mapping. *arXiv preprint* arXiv:2009.06200.
- Bandara, W. G. C., Patel, V. M., 2022. A transformer-based Siamese network for change detection. *arXiv preprint* arXiv:2201.01293.
- Chen, H., Song, J., Yokoya, N., Xia, J., Han, C., 2022. Dual-tasks Siamese transformer framework for building damage assessment. *arXiv preprint* arXiv:2201.10953.
- Chen, M., Zhang, Q., Ge, X., Xu, B., Hu, H., Zhu, Q., Zhang, X., 2023. A full-scale connected CNN–Transformer network for remote sensing image change detection. *Remote Sens.*, 15(22), 5383. doi:10.3390/rs15225383
- Corbane, C., Carrion, D., Lemoine, G., Broglia, M., 2011. Comparison of damage assessment maps derived from very high spatial resolution satellite and aerial imagery produced for the Haiti 2010 earthquake. *IEEE Trans. Geosci. Remote Sens.*, 48, 2403–2419. doi:10.1193/1.3630223
- 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 Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2, 89–96. doi:10.5194/isprs-annals-IV-2-89-2018
- Kalantar, B., Al-Najjar, H. A. H., Ueda, N., 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 Sens.*, 12(21), 3529. doi:10.3390/rs12213529
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B. A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.*, 152, 166–177. doi:10.1016/j.isprsjprs.2019.04.015
- Mohammadian, A., Ghaderi, F., 2022. SiamixFormer: A fully-transformer Siamese network with temporal fusion for accurate building detection and change detection in bi-temporal remote sensing images. *Int. J. Remote Sens.*, in press. doi:10.1080/01431161.2023.2225228
- 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 Sens.*, 11(23), 2765. doi:10.3390/rs11232765
- Tu, J., Sui, H., Feng, W., Song, Z., 2016. Automatic building damage detection method using high-resolution remote sensing images and 3D GIS model. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-8, 43–50. doi:10.5194/isprs-annals-III-8-43-2016
- Wan, Z., Yan, T., Zhang, P., 2022. Fully Transformer Network for change detection of remote sensing images. *arXiv preprint* arXiv:2210.00757
- Wu, C., Zhang, F., Xia, J., Xu, Y., Li, G., Xie, J., Du, Z., Liu, R., 2021. Building damage detection using U-Net with attention mechanism from pre- and post-disaster remote sensing datasets. *Remote Sens.*, 13(5), 905. doi:10.3390/rs13050905
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., Luo, P., 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Adv. Neural Inf. Process. Syst.*, 34, 12077–12090