World Scientific Research Journal Volume 12 Issue 9, 2026
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

A Review on Building Earthquake Damage Assessment Using Multi-Source Remote Sensing Imagery

Liyi Han[1,*]

[1]Gansu Earthquake Agency, Lanzhou 730000, China

[*] Corresponding author

Abstract

The seismic damage assessment of buildings is an important basis for post-earthquake emergency response, deployment of search and rescue forces, disaster loss assessment and post-disaster reconstruction. With the advantages of large-scale coverage and non-contact observation, multi-source remote sensing images have become an important data source for rapid and automated evaluation of post-earthquake building damage. This paper focuses on the application of multi-source remote sensing images in the evaluation of earthquake damage of buildings, and systematically sorts out the data sets, task paradigms, damage levels, evaluation indicators and typical methods in this field. First, compare the four commonly used public data sets of xBD, QuickQuakeBuildings, BRIGHT and MEDBFS, and analyze their image modality, time-phase configuration, spatial resolution, task granularity and damage label system; then, according to the vein of technical evolution, the system Summarize the development process from spectral, texture and visual feature engineering to traditional machine learning, deep learning, two-time phase change detection, single-time phase evaluation, two-stage method, integrated learning and visual basic large models. On this basis, four major problems faced by the current research are summarized: insufficient cross-regional generalization ability, difficulty in identifying category imbalance and fine-grained damage, high cost of large-scale model adaptation and inference calculation, and ununiform destruction grading standards. In response to the above problems, we further look forward to the development direction of basic model and pre-training strategy, multi-modal and multi-phase information integration, efficient model adaptation and deployment, standardized evaluation system, and deep connection between seismic assessment and emergency decision-making system. This article aims to provide a systematic research context and reference for the follow-up research of the earthquake damage assessment method of buildings and its practical emergency application.

Keywords

Multi-source remote sensing imagery; Building earthquake damage assessment; Multimodal data fusion; Damage grading standards; Emergency decision support.

1. INTRODUCTION

Earthquake disasters have the characteristics of suddenness and a wide range of damage, which seriously threaten the safety of people's lives and property and social stability. Therefore, the rapid and accurate identification of building damage after the earthquake is of great significance for the deployment of search and rescue forces, disaster loss assessment and post-disaster reconstruction and other emergency decision-making work. However, the traditional earthquake damage assessment mainly relies on on-site manual survey and expert visual judgment. Not only is the manpower and time cost high, but it is also difficult to carry out under

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

extreme conditions such as traffic and communication interruption. Operators will also face the risk of secondary disasters and cannot meet the requirements of emergency response to timeliness and coverage.

With the advantages of large-scale coverage and non-contact observation, Earth Observation (EO) technology provides a safe and efficient alternative means for the rapid assessment of post-earthquake building damage [1]. The existing remote sensing data used for earthquake damage assessment mainly includes three categories: optical images, synthetic aperture radar (SAR) images and drone images: optical images are intuitive and easy to visually interpret, but they are passive imaging, which is easily limited by weather conditions and easily obscured by clouds and fog; SAR images use active imaging. It has all-day and all-weather imaging capabilities, and can penetrate clouds and rain to obtain surface information, but there are geometric distortions such as perspective contraction and shadows, and the difficulty of interpretation is relatively high. The UAV image deployment is flexible and the spatial resolution is high, which is suitable for detailed seismic assessment. Multi-source remote sensing data provides an effective data basis for the identification of damage to buildings after the earthquake, but how to quickly and accurately identify damaged buildings from it is still an urgent problem to improve the efficiency of emergency response [2].

Chapter 2 of this article is the definition of the problem, introduces four representative data sets, and elaborates on the task method of building seismic damage identification, damage classification and evaluation indicators; Chapter 3 sorts out the current situation of research in this field, covering early methods based on spectral texture characteristics, machine learning methods and current mainstream deep learning methods, and summarizes existing problems; Chapter 4 looks forward to the future development trend; Chapter 5 summarizes the full text.

2. PROBLEM DEFINITION

2.1. Datasets

Four representative publicly available datasets are introduced in this study: xBD [3], QuickQuakeBuildings [4], BRIGHT [5], and MEDBFS [6]. These datasets differ substantially in terms of imaging modality, temporal configuration, spatial resolution, damage-label granularity, and task formulation, thereby providing diverse benchmarks for building damage identification under different disaster-response scenarios.

xBD is one of the earliest and largest datasets developed for building damage assessment. It was jointly constructed by Carnegie Mellon University and several disaster-response organizations in the United States and served as the official dataset for the xView2 challenge. The dataset covers 19 natural disaster events of different types, including earthquakes, hurricanes, floods, volcanic eruptions, tornadoes, tsunamis, and wildfires. The imagery was obtained through the Maxar/DigitalGlobe Open Data Program, with a spatial resolution of less than 0.8 m, and more than 850,000 buildings were annotated across an area exceeding 45,000 km^2. For each disaster event, xBD provides paired pre- and post-disaster optical imagery together with building damage labels. Environmental factors, including fire, smoke, and water, are also annotated. Owing to its large scale and diverse disaster scenarios, xBD has become one of the most widely adopted benchmark datasets for building damage assessment.

QuickQuakeBuildings is designed for post-earthquake building damage detection using single-temporal SAR and optical imagery. The dataset focuses on the city of Islahiye, one of the areas severely affected by the 2023 Türkiye–Syria earthquake. It was constructed using spotlight-mode SAR imagery provided by Capella Space and optical imagery acquired by WorldView-3, together with building damage information annotated by the Humanitarian OpenStreetMap Team (HOTOSM). Through a detailed registration procedure involving pre-earthquake building footprints and SAR imagery, the dataset contains 4,029 buildings, including

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

169 damaged and 3,860 intact buildings. For each building, four types of data are provided: an SAR image patch, an SAR building-footprint mask, an optical image patch, and an optical building-footprint mask. Unlike datasets relying on paired pre- and post-earthquake imagery, QuickQuakeBuildings uses only post-earthquake single-temporal imagery, thereby reflecting the practical situation in which high-resolution pre-earthquake imagery from the same sensor or source may not be available after a disaster.

BRIGHT is a multimodal dataset developed for all-weather post-disaster building damage assessment by researchers from the University of Tokyo, RIKEN, and other institutions. It also served as the official dataset for the 2025 IEEE GRSS Data Fusion Contest. The dataset integrates high-resolution pre-disaster optical imagery acquired from sources such as Maxar and Google Earth with post-disaster SAR imagery from Capella Space and Umbra. The spatial resolution ranges from 0.3 to 1 m. BRIGHT covers 14 disaster events across 23 geographic regions, including five types of natural disasters and two types of human-induced disasters, and contains approximately 385,000 annotated buildings. Building damage is categorized into three classes: Intact, Damaged, and Destroyed. In addition to conventional supervised learning, BRIGHT supports a variety of research tasks, including unsupervised domain adaptation, semi-supervised learning, unsupervised multimodal change detection, and multimodal image registration, making it a versatile benchmark for multimodal disaster analysis.

MEDBFS (Multi-scale Earthquake-Damaged Building Feature Set) was developed by the Aerospace Information Research Institute, Chinese Academy of Sciences. It integrates data from five major earthquake events: the 2010 Yushu, China, Mw 7.1 earthquake, the 2017 Chiapas, Mexico, Mw 8.2 earthquake, the 2023 Türkiye, Mw 7.8 earthquake, the 2023 Morocco, Mw 6.9 earthquake, and the 2023 Afghanistan, Mw 6.2 earthquake. The dataset incorporates imagery from multiple satellite platforms, including QuickBird, Jilin-1, and WorldView, with spatial resolutions ranging from 0.3 to 0.75 m. It provides 7,062 pre- and post-earthquake image pairs, each with a size of 512 × 512 pixels. In addition to the original imagery, MEDBFS provides 16 shallow feature maps belonging to five feature categories: spectral features, texture features (GLCM and LBP), edge features (Prewitt and Laplacian), building-related indices (MBI), and temporal features (three-dimensional GLCM). These precomputed features can be directly reused in both traditional machine learning and deep learning approaches, facilitating comparative studies of different feature representations for earthquake damage identification.

2.2. Formulations of Building Earthquake Damage Assessment

According to the temporal configuration and imaging data used as inputs, building earthquake damage assessment can be broadly categorized into two major paradigms: (1) Bi-temporal change detection. Pre- and post-earthquake images are jointly utilized, and building damage is identified by analyzing changes in building appearance and structural characteristics between the two temporal observations [7]. (2) Single-temporal semantic segmentation. Only post-earthquake imagery is used to simultaneously locate buildings and determine their damage status or damage level [8].

The four datasets introduced in Section 2.1 differ in terms of their temporal configurations, imaging modalities, task granularity, and task paradigms, as summarized in Table 1.

2.3. Building Damage Levels

The European Macroseismic Scale 1998 (EMS-98) was among the earliest internationally adopted systems for classifying earthquake-induced building damage [9]. It categorizes damage into five levels: slight, moderate, substantial, very heavy, and destruction. However, because EMS-98 was originally developed for field-based damage surveys, distinguishing these levels accurately from remote sensing imagery can be challenging given the spatial resolution and observable information available in satellite data.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

Building on HAZUS, the FEMA disaster assessment guidelines, the Kelman scale, and EMS-98, the xBD dataset introduced a Joint Damage Scale specifically designed for satellite imagery. This scale groups building damage into four categories: no damage, minor damage, major damage, and destroyed.

BRIGHT further simplifies the damage taxonomy by integrating labeling systems from multiple sources, including the Copernicus Emergency Management Service, the United Nations Satellite Centre (UNOSAT), and FEMA. Building damage is ultimately categorized into three classes: Intact, Damaged, and Destroyed.

MEDBFS adopts a classification scheme that combines the principles of EMS-98 and xBD. Considering that certain relatively low-resolution satellite images, such as those from QuickBird, may not provide sufficient visual information to reliably distinguish minor damage, MEDBFS simplifies the damage categories into Intact/Minor Damage, Severe Damage, and Complete Collapse. QuickQuakeBuildings directly adopts the binary labels Damaged and Intact provided by HOTOSM.

Table 1. Comparison of building earthquake damage datasets

Dataset	Temporal configuration	Primary modality	Task	Task formulation
xBD	Bi-temporal (pre- and post-disaster)	Optical	Pixel-level	Building localization + damage classification (decoupled semantic segmentation)
BRIGHT	Bi-temporal (pre-disaster optical + post-disaster SAR; post-disaster optical available for some events)	Optical + SAR	Pixel-level	Supports both direct prediction and decoupled semantic segmentation
QuickQuakeBuildings	Single-temporal (post-disaster only)	SAR / Optical	Building-instance level	Binary classification (image-level classification based on known building footprints)
MEDBFS	Bi-temporal (pre- and post-earthquake, cropped into fixed-size image patches)	Optical	Image-patch / pixel level	Three-class feature extraction and classification

Considering that certain relatively low-resolution satellite images, such as those from QuickBird, may not provide sufficient visual information to reliably distinguish minor damage, MEDBFS simplifies the damage categories into Intact/Minor Damage, Severe Damage, and Complete Collapse. QuickQuakeBuildings directly adopts the binary labels Damaged and Intact provided by HOTOSM.

Although the damage taxonomies differ in granularity across datasets, their underlying semantic definitions are relatively consistent. Buildings that remain structurally intact and exhibit no obvious visual changes are generally assigned to the intact category. Buildings showing partial wall or roof collapse, visible cracks, or other localized damage while retaining their primary structural form are typically classified as damaged or an intermediate damage category. Buildings that have undergone extensive structural collapse, are buried by debris, or have lost their recognizable original footprint are generally classified as destroyed. Therefore,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

when integrating multiple datasets, careful cross-dataset semantic alignment is required to ensure consistency among damage categories.

2.4. Evaluation Metrics

The performance of building earthquake damage assessment models is commonly evaluated using Overall Accuracy (OA), F1-score, and Intersection over Union (IoU). Overall Accuracy measures the proportion of correctly classified samples or pixels across all categories and provides an intuitive assessment of overall prediction performance. However, in highly imbalanced datasets, OA can be dominated by the majority class and may therefore provide an overly optimistic assessment of model performance. Consequently, OA is generally used in conjunction with metrics such as F1-score and mean IoU (mIoU).

The F1-score is the harmonic mean of precision and recall and provides a balanced measure of classification performance, particularly when the positive and negative classes are unevenly distributed. The IoU measures the degree of spatial overlap between the predicted segmentation region and the corresponding ground-truth region and is therefore one of the fundamental metrics for pixel-level semantic segmentation tasks.

For class k, the metrics are defined as:

$$OA = \frac{\sum_{k=0}^{2} TP_k}{M} \tag{1}$$

$$\begin{aligned} Precision_k &= \frac{TP_k}{TP_k + FP_k} \\ Recall_k &= \frac{TP_k}{TP_k + FN_k} \\ F1_k &= \frac{2 \times Precisision_k \times Recall_k}{Precision_k + Recall_k} \end{aligned} \tag{2}$$

$$IoU_k = \frac{TP_k}{TP_k + FP_k + FN_k} \tag{3}$$

where M denotes the total number of pixels; TP_k, FP_k, FN_k, and TN_k denote true positive, false positive, false negative, and true negative counts for class k, respectively.

3. STATE OF THE ART

Research on building earthquake damage assessment has evolved from conventional approaches based on manually designed spectral and textural features to traditional machine learning methods and, more recently, deep learning and foundation-model-based approaches. This progression reflects a continuous shift from handcrafted feature representation toward automated and increasingly sophisticated semantic feature learning.

3.1. Traditional Methods Based on Spectral

Early-stage building seismic damage assessment primarily relies on shallow visual features—such as spectral, color, and texture characteristics of remote sensing imagery—to determine the extent of building seismic damage by applying empirical thresholds or simple statistical criteria. Liu Wei et al. [10] employed an object-oriented classification approach to

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

extract information on building seismic damage; they constructed classification rules based on the spectral, shape, texture, and other features of image objects, as well as their topological relationships. Through accuracy evaluation and visual comparison analysis, they found that the object oriented classification method demonstrated good consistency and higher classification accuracy.

However, this kind of method has a small calculation volume and strong interpretability, but it has limited ability to express features. It is difficult to cope with complex and changeable post-earthquake scenes. It is more sensitive to changes in lighting, shooting angle and image resolution, and has poor generalization ability.

3.2. Machine Learning Methods

Therefore, some researchers began to introduce traditional machine learning algorithms such as support vector machine (SVM), Random Forest, and decision tree into the task of building earthquake damage assessment. Compared with the early threshold discrimination method, machine learning algorithms can automatically learn the mapping relationship between features and damage categories, which improves the recognition accuracy and robustness to a certain extent. Chenguang Wang et al. [11] proposed a Bayesian network method based on causality, which uses the overall causal Bayesian network to damage and interfere with building damage synthetic aperture radar (Interferometric Synthetic Aperture The complex causal dependence between Radar, InSAR) images is encoded to infer large-scale unobserved building damage without the need for real ground labels. Valentina Macchiarulo et al. [12] introduced a multi-case joint learning strategy to explore the generalization ability of the model in cross-earthquake event scenarios. Finally, the experimental results showed that the overall classification accuracy of the method in unsighted areas can reach up to 72%. Although the machine learning method is superior to the traditional method in terms of feature expression ability and classification accuracy, the machine learning method relies on manually designed features. The upper limit of model performance is limited by the quality of feature engineering, and it is difficult to fully exploit the deep semantic information in remote sensing images.

3.3. Deep Learning Methods

With the rapid development of deep learning, architectures such as convolutional neural networks (CNNs) and Transformers have achieved substantial advances in remote sensing image interpretation. Deep learning has therefore gradually become the dominant paradigm for building earthquake damage assessment, owing to its ability to automatically learn hierarchical representations directly from image data.

Bai et al. [13] introduced U-Net into earthquake damage assessment and used paired pre- and post-earthquake imagery from the area affected by the 2011 Tohoku earthquake and tsunami in Japan as input. The method performed pixel-level three-class building damage classification and achieved an overall accuracy of 70.9%. Ablation experiments further demonstrated that the use of bi-temporal imagery outperformed the use of post-earthquake imagery alone, highlighting the importance of temporal information for building damage identification.

Chen et al. [14] employed DeepLabv3+, which uses atrous spatial pyramid pooling (ASPP) to capture multi-scale contextual information and integrates low-level features in the decoder to recover fine boundary details. This architecture provides an effective framework for simultaneously exploiting high-level semantic information and low-level spatial structures, which is particularly important for building-level damage delineation.

Chen et al. [15] subsequently proposed ChangeMamba (MambaBDA), which employs VMamba as a Siamese encoder to model global spatial context with linear computational

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

complexity. Three bi-temporal feature fusion mechanisms were further designed to effectively exploit information from pre- and post-earthquake observations. On the xBD dataset, ChangeMamba improved the overall F1-score by 2.64%–4.39% compared with DamFormer and exhibited stronger robustness to noise and scale variations.

In contrast to ChangeMamba, Gebre et al. [16] developed a multimodal attention framework using ConvNeXt-Tiny as a shared backbone. The method introduces learnable change tokens and employs cross-attention to integrate pre-disaster imagery, post-disaster imagery, and building masks. On the xBD dataset, the approach achieved an overall accuracy of 94.90%, with an F1-score of 94.10% for the destroyed category. This study demonstrates the potential of attention-based multimodal fusion for exploiting complementary information from heterogeneous disaster-related observations.

Although bi-temporal inputs generally provide advantages in terms of damage identification accuracy, the acquisition of suitable pre-earthquake imagery can be challenging during the early stages of emergency response because of the sudden occurrence of earthquakes and the limited availability of timely remote sensing data. In particular, high-resolution pre-earthquake imagery from the same sensor or acquisition source may not always be accessible. Therefore, building damage assessment based solely on post-earthquake single-temporal imagery remains practically important and continues to attract considerable research attention.

Gomroki et al. [17] proposed EMYNet-BDD, in which two pretrained networks, EfficientViT-B and YOLOv8, are employed in parallel as encoders. The decoder adopts an improved U-Net architecture and integrates multi-level encoder features through feature concatenation. The method achieved overall accuracies of 97.62%, 98.63%, and 96.43% on datasets from the Türkiye, Morocco, and Libya earthquakes, respectively. Nevertheless, because no pre-earthquake imagery is available as a reference, the reconstruction of boundaries for completely destroyed buildings remains challenging. Gomroki et al. [18] further proposed UNet-GCViT, which incorporates Global Context Vision Transformer (GCViT) blocks into a U-Net encoder. Local and global self-attention mechanisms are alternately employed to model short- and long-range spatial dependencies. The proposed architecture reduces the number of parameters from 19.9 M to 12.4 M. On datasets corresponding to the Türkiye earthquake, the Batna explosion, and the Haiti earthquake, the method achieved overall accuracies exceeding 96%. Its training speed was comparable to that of FastViT, while its reconstruction of building boundaries in densely built-up areas outperformed competing methods such as TinyViT. These results further demonstrate that single-temporal approaches can achieve a practical balance between computational efficiency and prediction accuracy.

The above studies indicate that single-temporal damage assessment can substantially reduce data-access requirements while maintaining competitive performance. However, the absence of pre-earthquake references fundamentally limits the ability of these approaches to distinguish pre-existing building characteristics from earthquake-induced structural changes. This issue is particularly pronounced for buildings that have undergone complete collapse, where the original building footprint may no longer be clearly observable.

In addition to direct pixel-level damage classification, some studies have adopted a two-stage strategy, in which buildings are first localized or segmented and their damage states are subsequently classified. This task decomposition separates building extraction from damage assessment and allows specialized models to be designed for the two subtasks.

Benedict et al. [19] proposed a two-stage CNN-based framework using the xBD dataset. In the first stage, TransUNet is employed for building segmentation. In the second stage, the segmented buildings are classified using a lightweight EfficientNet network through transfer learning. The study also demonstrated the effectiveness of Focal Loss in mitigating class imbalance. Zhao et al. [20] proposed a two-stage network termed ECADS-CNN. In the first stage,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

SE-ResNeXt50 is used as the encoder, while the second stage employs a Siamese segmentation branch to compare the feature similarity between pre- and post-earthquake observations. Compared with the baseline approach, the proposed method improved the F1-score for damage categories by 5.2% while reducing inference time by 7.3%. These results suggest that explicitly modeling the building localization and damage classification processes can improve both recognition performance and computational efficiency. Sirma et al. [21] constructed the D'RespNeT dataset specifically for search-and-rescue scenarios. The dataset provides instance-level annotations for 28 categories, including building structural conditions and accessibility of building entrances. Using YOLOv8-seg, the study achieved an mAP@50 of 92.7% at an inference speed of 27 FPS. These results demonstrate the potential of instance segmentation for supporting post-earthquake search-and-rescue operations, particularly in applications involving rescue-route planning and the identification of accessible building entrances.

Beyond individual deep learning architectures, ensemble learning has also been explored to improve the robustness and generalization capability of post-earthquake damage assessment models. Soleimani-Babakamali et al. [22] proposed RAPID-A (Rapid Post-Earthquake Aerial Imagery Damage Assessment), an ensemble framework for rapid large-scale post-earthquake damage assessment. The framework integrates six deep segmentation models, including EfficientNet-B7 and SegFormer. Experiments on remote sensing imagery from the 2023 Türkiye earthquakes demonstrated that the ensemble strategy can effectively reduce estimation variance compared with individual constituent models. Bhardwaj et al. [23] proposed a ResNet-U-Net ensemble architecture, in which ResNet-50 is employed to extract deep features and U-Net performs pixel-level localization. An attention module is further introduced to adaptively weight feature maps and emphasize damaged regions. For binary building damage classification using satellite imagery acquired after Hurricane Iota, the method achieved overall accuracies of 96.91% and 89.35% on imbalanced and balanced test sets, respectively. These results indicate that ensemble architectures can provide improved robustness, although their performance may remain sensitive to dataset composition and class distribution.

More recently, visual foundation models represented by Segment Anything Model (SAM) and SAM2 have increasingly been introduced into remote sensing building damage assessment because of their powerful zero-shot and prompt-driven segmentation capabilities. These models provide a new paradigm for extracting building-level information with reduced dependence on task-specific annotation.Chen et al. [24] proposed RSPrompter to address two key limitations of SAM in remote sensing applications: its dependence on manually specified prompts and its inherently category-agnostic segmentation mechanism. Specifically, semantic features are extracted from intermediate layers of the frozen SAM image encoder, while a prompt generator is trained to produce prompt embeddings automatically. Consequently, the SAM decoder can generate category-aware instance masks without requiring manual interaction.On the WHU Building, NWPU VHR-10, and SSDD datasets, RSPrompter outperformed competing approaches including Mask R-CNN, Mask2Former, and SAM-det. This study demonstrates the feasibility of adapting general-purpose visual foundation models to remote sensing instance segmentation through task-specific prompt learning. Siva et al. [25] focused specifically on the label noise and class imbalance problems in the xBD dataset. A building-centroid-based tiling strategy was introduced during preprocessing to suppress background noise, and different fine-tuning strategies based on DINOv2-small and DeiT-Ti were investigated. Experimental results showed that end-to-end fine-tuning of DeiT achieved a macro-averaged F1-score of 0.599 and outperformed existing CNN baselines for the severe-damage category.

These studies suggest that foundation models and large-scale pretrained visual representations may provide a promising direction for building earthquake damage assessment. Nevertheless, the transferability of these models from general-purpose visual tasks to fine-grained earthquake damage assessment remains an open problem, particularly when damage categories are subtle, labels are noisy, and post-earthquake scenes exhibit substantial domain shifts.

3.4. Emerging Exploration of Quantum Computing

In addition to conventional deep learning and foundation models, recent studies have begun to investigate the potential of quantum computing for building seismic damage identification. Bhatta et al. [26] proposed a hybrid quantum-classical convolutional neural network (QCNN) by incorporating quantum convolutional layers into a conventional CNN architecture. Quantum superposition and entanglement are exploited to encode and map local image features. When evaluated on imagery from the 2023 Türkiye earthquake that was not included during training, the QCNN achieved an overall recognition accuracy of 61.2%, representing an improvement of approximately 11% over the CNN baseline. These preliminary results indicate that quantum convolutional operations may enhance the discriminative representation of image features and provide a potential new direction for earthquake damage assessment.

In summary, although the accuracy of building earthquake damage assessment has continued to improve, several issues remain prevalent:

(1) Limited cross-regional generalization.

Differences in architectural styles and structural types across regions can lead to a significant degradation in model performance outside the geographic areas represented in the training data. Although existing studies have attempted to improve model generalization through transfer learning, pretrained models, and training with multi-regional datasets, robust transferability across different earthquake events, geographic regions, and even sensing modalities remains to be further improved.

(2) Class imbalance.

Earthquake damage datasets, such as xBD, generally exhibit a highly imbalanced class distribution, with undamaged buildings accounting for the overwhelming majority of samples, while samples from other damage levels are relatively scarce, resulting in a long-tailed distribution. Although weighted loss functions such as Focal Loss can partially alleviate this problem, the visual differences between adjacent damage levels, particularly minor and moderate damage, are often subtle. Their discrimination therefore relies heavily on local structural changes and textural differences in buildings, resulting in considerably lower recognition accuracy than that achieved for intact and completely destroyed buildings.

(3) Limited real-time efficiency of large-model fine-tuning and inference.

Large-scale models provide stronger feature representation capabilities; however, compared with conventional lightweight networks, they generally involve larger numbers of parameters, higher GPU memory consumption, and more complex computational procedures, resulting in substantial fine-tuning and inference costs. In post-earthquake emergency response scenarios, remote sensing data acquired after an earthquake often need to be processed within a limited time frame to assess building damage over large areas. Consequently, a trade-off remains between model accuracy and computational efficiency.

(4) Inconsistent damage grading standards.

Current datasets differ substantially in terms of data sources, image spatial resolution, and damage-level definitions. Such inconsistencies make it difficult to conduct rigorous and direct comparisons of performance across different datasets, while also increasing the difficulty of benchmark dataset selection, experimental reproduction, and validation of methodological effectiveness in subsequent studies.

4. FUTURE TRENDS

4.1. More Generalizable Foundation Models and Pretraining Strategies

To address the limited cross-regional generalization of existing approaches, future building earthquake damage assessment is expected to gradually shift from specialized models trained for individual earthquake events or geographic regions toward more general visual representation learning with stronger transferability. In addition, disaster-oriented pretraining, cross-event representation learning, domain adaptation, and parameter-efficient fine-tuning strategies can be further explored to enable a single model to rapidly adapt to different earthquake events, geographic regions, and sensor conditions.

4.2. Multimodal and Multitemporal Data Fusion

Optical, SAR, and UAV imagery exhibit strong complementarity in terms of information content, imaging conditions, and spatial resolution, while pre- and post-earthquake imagery provide direct temporal information on structural and appearance changes. Therefore, multimodal and multitemporal data fusion is expected to become a key direction for improving the reliability of building earthquake damage assessment. Future research may focus on cross-modal feature alignment, adaptive modality weighting, temporal attention, and joint representation learning. These techniques can exploit complementary information from heterogeneous data sources while reducing the dependence on co-registered pre-earthquake imagery from the same source, ideal weather conditions, or a single sensor, thereby enhancing the ability to identify minor and fine-grained damage patterns.

4.3. Efficient Model Adaptation and Real-Time Deployment

The substantial computational and memory requirements of large-scale models remain important factors limiting their broader application to post-earthquake emergency response. Future research should therefore place greater emphasis on balancing model representation capacity and computational efficiency. Techniques such as parameter-efficient fine-tuning, knowledge distillation, and model pruning are expected to reduce memory consumption and inference latency while preserving strong feature extraction capabilities. Furthermore, model architectures and deployment strategies should be jointly designed to accommodate practical constraints in post-earthquake emergency scenarios, including limited computational resources, rapid data acquisition, large-scale image processing, and robust inference across heterogeneous environments. Such efforts could facilitate the transition of building earthquake damage assessment from offline experimental evaluation toward near-real-time operational applications.

4.4. Standardized Damage Grading Systems

To address the inconsistency of damage grading standards across different datasets, future research should establish standardized and reproducible benchmarks for building earthquake damage assessment, with explicit definitions of key factors including data sources, spatial resolutions, temporal configurations, annotation protocols, class semantics, and evaluation metrics. In particular, a unified semantic mapping mechanism for damage levels should be developed to enable reasonable alignment among binary, three-class, and four-class damage-label schemes adopted by different datasets.

4.5. Deeper Integration of Damage Assessment with Emergency Decision-Support Systems

The ultimate objective of building earthquake damage assessment is not merely to generate damage segmentation maps, but to provide actionable information for search-and-rescue operations and emergency response. Existing studies have begun to investigate applications involving building structural conditions, accessibility of building entrances, and rescue-route planning in search-and-rescue scenarios [21]. Future research could further integrate building-level damage assessment results with geospatial information, population distributions, road accessibility, and critical infrastructure data to generate structured outputs for emergency briefing generation, search-and-rescue force allocation, and dynamic resource dispatch.

5. CONCLUSION

This paper presented a systematic review of building earthquake damage assessment using multi-source remote sensing imagery. Four widely used datasets, namely xBD, QuickQuakeBuildings, BRIGHT, and MEDBFS, were systematically examined, with their characteristics analyzed from the perspectives of imaging modalities, temporal configurations, task paradigms, and damage-level definitions. On this basis, the technical evolution of the field was reviewed in chronological and methodological terms, covering the progression from conventional spectral and textural features and machine learning to deep learning, bi-temporal change detection, single-temporal assessment, two-stage approaches, ensemble learning, and vision foundation models. Existing studies have substantially improved the accuracy of automated building earthquake damage assessment; however, several challenges remain, including limited cross-regional generalization, class imbalance and difficulties in fine-grained damage discrimination, the relatively high computational cost of large-model adaptation and inference, and inconsistent damage grading standards. In response to these challenges, future research should focus on more generalizable foundation models and pretraining strategies, multimodal and multitemporal information fusion, efficient model adaptation and deployment, standardized evaluation frameworks, and deeper integration of building earthquake damage assessment with emergency decision-support systems.

ACKNOWLEDGEMENTS

This work was supported by the Earthquake and Geological Disaster Rescue Department of the Ministry of Emergency Management of the People's Republic of China under the project “Research on AI Large Language Model-based Decision Support for Earthquake Disaster Response and Deployment” (Grant No. 2025- DizhenDizhiSi -Rescue21), and the project "Research on Key Technologies for Establishing Life-Saving Corridors in Major Earthquake Disaster Relief Operations in Northwest China" (Grant No. 2025-DizhenDizhiSi-Rescue22).

REFERENCES

[1] Al Shafian, S., & Hu, D. (2024). Integrating machine learning and remote sensing in disaster management: A decadal review of post-disaster building damage assessment. Buildings, 14(8), 2344. https://doi.org/10.3390/buildings14082344.

[2] Nia, K. R., & Mori, G. (2017). Building damage assessment using deep learning and ground-level image data. In 2017 14th Conference on Computer and Robot Vision (CRV) (pp. 95–102). IEEE. https://doi.org/10.1109/CRV.2017.19

[3] Gupta, R., Hosfelt, R., Sajeev, S., & others. (2019). xbd: A dataset for assessing building damage from satellite imagery. arXiv preprint arXiv:1911.09296.

[4] Sun, Y., Wang, Y., & Eineder, M. (2024). Quickquakebuildings: Post-earthquake SAR-optical dataset for quick damaged-building detection. IEEE Geoscience and Remote Sensing Letters, 21, 1–5. https://doi.org/10.1109/LGRS.2024.3370827

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

[5] Chen, H., Song, J., Dietrich, O., & others. (2025). BRIGHT: A globally distributed multimodal building damage assessment dataset with very-high-resolution for all-weather disaster response. Earth System Science Data, 17(11), 6217–6253. https://doi.org/10.5194/essd-17-6217-2025

[6] Gao, G., Wang, F., Wang, Z., & others. (2024). Multi-scale earthquake damaged building feature set. Data, 9(7), 88. https://doi.org/10.3390/data9070088

[7] Lu, W., Wei, L., & Nguyen, M. (2024). Bitemporal attention transformer for building change detection and building damage assessment. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 4917–4935. https://doi.org/10.1109/JSTARS.2024.3380488

[8] Deng, L., & Wang, Y. (2022). Post-disaster building damage assessment based on improved U-Net. Scientific Reports, 12(1), 15862. https://doi.org/10.1038/s41598-022-19985-4

[9] Del Mese, S., Graziani, L., Meroni, F., & others. (2023). Considerations on using MCS and EMS-98 macroseismic scales for the intensity assessment of contemporary Italian earthquakes. Bulletin of Earthquake Engineering, 21(9), 4167–4189. https://doi.org/10.1007/s10518-023-01654-7

[10] Wang, L., Li, J., & Liu, Y. (2015). Comparative study on extraction methods of earthquake damage information of buildings from high-resolution remote sensing images. Journal of Heilongjiang Institute of Technology, 29(4), 11–13.

[11] Wang, C., & others. (2024). Scalable and rapid building damage detection after hurricane Ian using causal Bayesian networks and InSAR imagery. International Journal of Disaster Risk Reduction, 104, 371–387. https://doi.org/10.1016/j.ijdrr.2024.104371

[12] Macchiarulo, V., Giardina, G., Milillo, P., & others. (2025). Integrating post-event very high resolution SAR imagery and machine learning for building-level earthquake damage assessment. Bulletin of Earthquake Engineering, 23(12), 5021–5047. https://doi.org/10.1007/s10518-025-02112-1

[13] 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

[14] Chen, L. C., Zhu, Y., Papandreou, G., & others. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In European Conference on Computer Vision (pp. 833–851). Springer International Publishing. https://doi.org/10.1007/978-3-030-01234-2_49

[15] Chen, H., Song, J., Han, C., & others. (2024). ChangeMamba: Remote sensing change detection with spatiotemporal state space model. IEEE Transactions on Geoscience and Remote Sensing, 62, 1–20. https://doi.org/10.1109/TGRS.2024.3364758

[16] Gebre, T. S., Talreja, J., & Hashemi-Beni, L. (2026). Multi-Modal Attention for Automated Disaster Damage Assessment Using Remote Sensing Imagery and Deep Learning. arXiv preprint arXiv:2606.14963.

[17] Gomroki, M., Hasanlou, M., Chanussot, J., & others. (2024). EMYNet-BDD: EfficientViTB Meets Yolov8 in the encoder–decoder architecture for building damage detection using postevent remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 13120–13134. https://doi.org/10.1109/JSTARS.2024.3414684

[18] Gomroki, M., Hasanlou, M., Chanussot, J., & others. (2025). UNet-GCViT: a UNet-based framework with global context vision transformer blocks for building damage detection. International Journal of Remote Sensing, 46(6), 2587–2610. https://doi.org/10.1080/01431161.2025.2444761

[19] Benedict, R., Winartio, R. B., Adinata, M. F., & others. (2024). Comparison on difference deep learning models for building damage assessment using xbd dataset. In 2024 Arab ICT Conference (AICTC) (pp. 181–186). IEEE. https://doi.org/10.1109/AICTC62075.2024.10710337

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202609_12(9).0006

[20] Zhao, Z., Wang, F., Chen, S., & others. (2024). Deep object segmentation and classification networks for building damage detection using the xBD dataset. International Journal of Digital Earth, 17(1), 2302577. https://doi.org/10.1080/17538947.2024.2302577

[21] Sirma, A., Plastropoulos, A., Tang, G., & others. (2025). DRespNeT: A UAV Dataset and YOLOv8-DRN Model for Aerial Instance Segmentation of Building Access Points for Post-Earthquake Search-and-Rescue Missions. arXiv preprint arXiv:2508.16016.

[22] Soleimani-Babakamali, M. H., Askari, M., Heravi, M. A., & others. (2025). Deep ensemble learning for rapid large-scale postearthquake damage assessment: Application to satellite images from the 2023 Türkiye earthquakes. ASCE OPEN: Multidisciplinary Journal of Civil Engineering, 3(1), 04025003. https://doi.org/10.1061/AO.MJCE.0000158

[23] Bhardwaj, D., Nagabhooshanam, N., Singh, A., & others. (2025). Enhanced satellite imagery analysis for post-disaster building damage assessment using integrated ResNet-U-Net model. Multimedia Tools and Applications, 84(5), 2689–2714. https://doi.org/10.1007/s11042-024-19221-1

[24] Chen, K., Liu, C., Chen, H., & others. (2024). RSPrompter: Learning to prompt for remote sensing instance segmentation based on visual foundation model. IEEE Transactions on Geoscience and Remote Sensing, 62, 1–17. https://doi.org/10.1109/TGRS.2024.3370103

[25] Siva, S., & Cross-Zamirski, J. (2026). Building Damage Detection using Satellite Images and Patch-Based Transformer Methods. arXiv preprint arXiv:2602.08117.

[26] Bhatta, S., & Dang, J. (2024). Multiclass seismic damage detection of buildings using quantum convolutional neural network. Computer-Aided Civil and Infrastructure Engineering, 39(3), 406–423. https://doi.org/10.1111/mice.13061