DOI: 10.1111/mice.XXXXX

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

RESEARCH ARTICLE

Automated Building Damage Assessment and Large-Scale Mapping by Integrating Satellite Imagery, GIS, and Deep Learning

Abdullah M. Braik[1] | **Maria Koliou**[1]

[1] Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX, 77843, U.S.A.

Correspondence
Maria Koliou, Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX, 77843, U.S.A.
Email: maria.koliou@tamu.edu

Funding information
US National Science Foundation (NSF), Grant/Award Number: 2052930

ABSTRACT

Efficient and accurate building damage assessment is crucial for effective emergency response and resource allocation following natural hazards. However, traditional methods are often time-consuming and labor-intensive. Recent advancements in remote sensing and Artificial Intelligence (AI) have made it possible to automate the damage assessment process, and previous studies have made notable progress in machine learning classification. However, the application in post-disaster emergency response requires an end-to-end model that starts with satellite imagery as input and automates the generation of large-scale damage maps as output, which was rarely the focus of previous studies. Addressing this gap, this study integrates satellite imagery, Geographic Information Systems (GIS), and deep learning. This enables the creation of comprehensive, large-scale building damage assessment maps, providing valuable insights into the extent and spatial variation of damage. The effectiveness of this methodology is demonstrated in Galveston County following Hurricane Ike, where the classification of a large ensemble of buildings was automated using deep learning models trained on the xBD dataset. The results showed that utilizing GIS can automate the extraction of sub-images with high accuracy, while fine-tuning can enhance the robustness of the damage classification to generate highly accurate large-scale damage maps. Those damage maps were validated against historical reports.

1 | INTRODUCTION

1.1 | Motivation and problem statement

In the aftermath of disasters, the accurate assessment of building damage is crucial for guiding emergency responses, optimizing resource allocation, and expediting social and economic recovery. This assessment serves as the foundation for risk-informed decision-making, playing a vital role in enhancing community resilience against natural hazards (Koliou et al. 2020).

Current risk assessment methodologies rely on hazard and fragility analysis to predict the damage to structures and their economic and social consequences (e.g., Masoomi and van de Lindt (2016); Koliou and van de Lindt (2020); Nofal et al. (2021)). While effective for pre-hazard planning, preparedness, and hardening, their dependence on probabilistic predictions limits their post-disaster applicability due to the substantial aleatory uncertainties associated with natural hazards and their interactions with structures.

Methods based on post-disaster damage surveys (e.g., Highfield et al. (2014); Aghababaei et al. (2018); Aghababaei et al. (2020)) are important for continuous refinement of models and codes, estimating long-term recovery, and facilitating communities and insurance firms in submitting insurance claims. However, they are less effective in rapid emergency responses (Lozano et al. 2023). Traditional post-disaster assessments heavily depend on trained ground crews conducting visual inspections, leading to inefficiency, prolonged time consumption, and risks to crew members from hazardous materials and debris (Spencer Jr et al. 2019).

To address these challenges, there is a growing recognition of the potential presented by advanced technologies, such as remote sensing, artificial intelligence (AI), and data analytics. These are emerging as promising tools capable of revolutionizing the assessment process. Since the manual

inspection of the remote sensing data usually takes several weeks to complete (Lozano et al. 2023), recent studies have focused on the integration of AI, particularly deep learning, and remote sensing to automate the damage assessment process. However, current methods tend to concentrate solely on proposing models for machine learning classification, rarely extending their methodology to incorporate the full system automation starting from the acquisition of remote sensing images and ending with the generation of large-scale mapping of building damage (Matin and Pradhan 2022). This automation is essential to provide immediate post-disaster damage evaluation to be used in updating risk-assessment predictions of economic losses and recovery.

Hence, ongoing research in the three areas above; fragility-based risk assessment, surveys-based damage assessment, and AI-based damage classification, are currently running in parallel, and each, by itself, has limitations in applicability in the immediate and rapid post-disaster emergency response. Regional damage maps across a disaster-affected region are the main input of most risk-assessment methodologies that predict the losses and restoration in the short- and long-term recovery phases (Koliou et al. 2020). Hence, there is a potential to extend the application from pre-hazard mitigation and preparedness to post-disaster emergency response and adaptive decision-making. This requires a framework focusing on automation and data analytics to seamlessly transition from satellite images as inputs to the automated generation of large-scale damage maps as outputs.

However, previous studies, discussed in detail in Section 1.2, focused solely on developing machine learning algorithms for damage assessment rather than on the entire system from data acquisition to analysis. Hence, a gap lies in the transition from technical advancements in machine learning classification to the practical implementation of automated systems for post-disaster emergency response. Another gap pertains to the underutilization of GIS to identify buildings despite the heavy utilization of geotagged building databases in the risk assessment field (Amin Enderami et al. 2022).

To address these gaps, the current study integrates Geographic Information Systems (GIS) with satellite imagery and deep learning to automate the generation of comprehensive, large-scale building damage maps in the disaster region. These maps are intended to provide rapid and immediate insights into the extent and spatial variation of damage, laying the groundwork for future research to utilize them in post-hazard risk-informed decision-making. Moreover, there is potential for these maps to facilitate real-time updates of disaster management plans based on evolving on-site conditions, expanding the models' capabilities from pre-hazard offline learning to the development of digital twins for post-disaster online learning (Braik and Koliou 2023).

The research outlined in this paper introduces a methodology designed for the automated and rapid generation of large-scale and highly detailed regional damage assessment maps in the aftermath of natural hazards. Section 2 provides a comprehensive overview of the methodology, elucidating the integration of georeferenced satellite images, geotagged building data, and deep learning. The discussion within this section encompasses the validation of machine learning classifications, utilizing diverse metrics, and benchmarking against historical reports. Section 3 applies the methodology to a case study, focusing on Galveston County post-Hurricane Ike. The results are validated against historical reports, offering a robust demonstration of the methodology's effectiveness in post-disaster scenarios. Section 4 discusses the limitations of the current study, and potential works for future research are explored, building upon the groundwork established in this paper. Finally, Section 5 summarizes the conclusions drawn from this study.

1.2 | Literature review on the automation of building damage assessment

The application of deep learning for pattern recognition in civil engineering has a history spanning over three decades (Adeli and Yeh (1989); Adeli (2001)). In the last decade, deep learning found application in various civil engineering applications such as construction management (e.g., Rafiei and Adeli (2016)), finite element analysis (e.g., Pereira et al. (2020)), material engineering (e.g., Rafiei et al. (2017)), and modal parameter identification of smart structures (e.g., Perez-Ramirez et al. (2016); Pezeshki et al. (2023)). More recently, its application in structural health monitoring has also gained prominence (Javadinasab Hormozabad et al. (2021); Pezeshki et al. (2023)). Several studies focused on crack detection in concrete (e.g., Cha et al. (2017); Deng et al. (2020); Yang et al. (2018)), while others focused on automating the inspection of structural elements in buildings and bridges (e.g., Liang (2019); Zheng et al. (2022)). Chu et al. (2022) studied tiny crack segmentation using attention mechanisms, Ye et al. (2023) investigated the detection of slab crack track, Chen et al. (2023) studied the online monitoring of crack dynamic development, Shim (2023) utilized conditional generative adversarial networks for self-training of crack detection, and Yong et al. (2023) proposed a framework for zero-shot and few-shot defect detection. Also, Cha et al. (2018) utilized deep learning to detect various types of damage including steel and bolt corrosion. Moreover, researchers proposed frameworks to automate the updating of finite element models using deep-learning-based crack detection (e.g., Zhang and Lin (2022); Kong et al. (2023); Gao et al. (2024)). However, in the context of post-disaster regional damage assessment, there has been a shift towards the integration of remote sensing

technologies with deep learning. Among these, satellite imagery and aerial images obtained from unmanned aerial vehicles have emerged as primary tools.

Aerial remote sensing, distinguished by its relatively higher resolution, stands out for its ability to detect a wide range of damage states. Consequently, researchers have been actively exploring its potential for automating the damage assessment process. Liu et al. (2020) studied the crack assessment of bridge piers using unmanned aerial vehicles, while other studies focused on post-disaster damage assessment using aerial imagery (e.g., Cheng et al. (2021); Hong et al. (2022); Khajwal et al. (2023)). However, aerial imagery is often expensive and time-consuming, and hence, less suitable for rapid emergency response compared to satellite imagery (Matin and Pradhan 2022).

In contrast, satellite imagery, despite its comparatively lower resolution, offers the advantage of rapid accessibility and wider coverage. This makes it particularly well-suited for immediate large-scale mapping, providing crucial emergency support and serving as initial guidance for more in-depth follow-up assessments utilizing aerial and terrestrial imagery (Matin and Pradhan 2022). The significant challenge of requiring a large and comprehensive dataset for model training has been addressed to some extent by recent introductions of labeled datasets like xBD (Gupta et al. 2019), leading to notable progress.

Several studies (e.g., Li et al. (2019); Kaur et al. (2022)) proposed deep learning models to classify post-hazard satellite imagery into various damage states. However, no discussion was provided beyond machine learning classification on how to automate the extraction of the sub-images. Advancements were made by other studies, introducing models that first identify buildings from pre-disaster images and then classify the damage state from the difference between the pre- and post-disaster images. Hao et al. (2021) and Wu et al. (2021) proposed attention-based methodologies that utilize a U-Net for building segmentation in pre- and post-disaster images, coupled with a Siamese network to compare features and classify damage levels based on segmentation masks. Shen et al. (2021) employed a two-stage convolutional neural network, integrating a U-Net for building localization in the first stage and a two-branch multiscale U-Net with a cross-directional attention module in the second stage, along with data augmentation, to enhance building damage assessment from pre- and post-disaster satellite images by explicitly considering correlations between the images. Gupta and Shah (2021) proposed a localization-aware loss function for end-to-end training to enhance building segmentation and damage classification performance. Kaur et al. (2023) employed a hierarchical transformer architecture, utilizing four transformer blocks to map pre- and post-change features into a common domain, generating difference-features of varying resolutions, and constructing a hierarchical output from these features for building damage assessment.

Nevertheless, the applicability of these is constrained without the georeferencing of buildings necessary for the generation of large-scale maps (Matin and Pradhan 2022). While Bai et al. (2018) employed the GeoAI platform to produce classification results suitable for integration within GIS software, the generated output lacks a direct link to any existing building database, which limits its utility to visualizations and restricts its capability for real-time updates of estimates or immediate integration with real-time statistical analyses. Additionally, all previous works depend on the availability of pre- and post-disaster images with nearly identical quality and angle, and the identification of objects before classification introduces a second layer of errors, constraining the robustness and generalizability to real-world applications. While the potential of GIS has been acknowledged in a few past studies (e.g., Cao and Choe (2020); Cao and Choe (2020)), these primarily relied on pre-labeled building pixel locations within satellite imagery, offering no discussion on automating GIS processing for new satellite data. Although Miyamoto and Yamamoto (2021) mentioned utilizing building databases to extract sub-images of buildings when preparing the training dataset, they provided minimal elaboration and restricted their methodology and application study to comparing various machine learning models. Consequently, prior studies integrating deep learning with satellite imagery have been limited in their applications beyond machine learning classification (Matin and Pradhan 2022).

2 | METHODOLOGY

The methodology presented in this section functions as a unified pipeline model, seamlessly transitioning from georeferenced satellite images as inputs to the generation of large-scale damage maps as outputs. The process initiates with satellite image processing, incorporating GIS information, followed by the extraction of sub-images for each building using geotagged data. These two steps allow for the automated extraction of building sub-images of the new unclassified post-disaster images, which differ from the training labeled ones used to train the deep learning model. At the core of the methodology is deep learning classification, employing a convolutional neural network (CNN) model for accurate damage classification. Hence, the pre-trained CNN can be used to automate the classification of the extracted sub-images into damage states. Finally, the integration of classification outcomes with the building database facilitates the automated creation of comprehensive large-scale damage maps, offering a spatial overview of the disaster-affected region. Therefore, the complete process could be automated, allowing for rapid and large-scale building damage assessment. Figure 1 visually illustrates the flowchart of the methodology, while

detailed discussions are provided in the following sections.

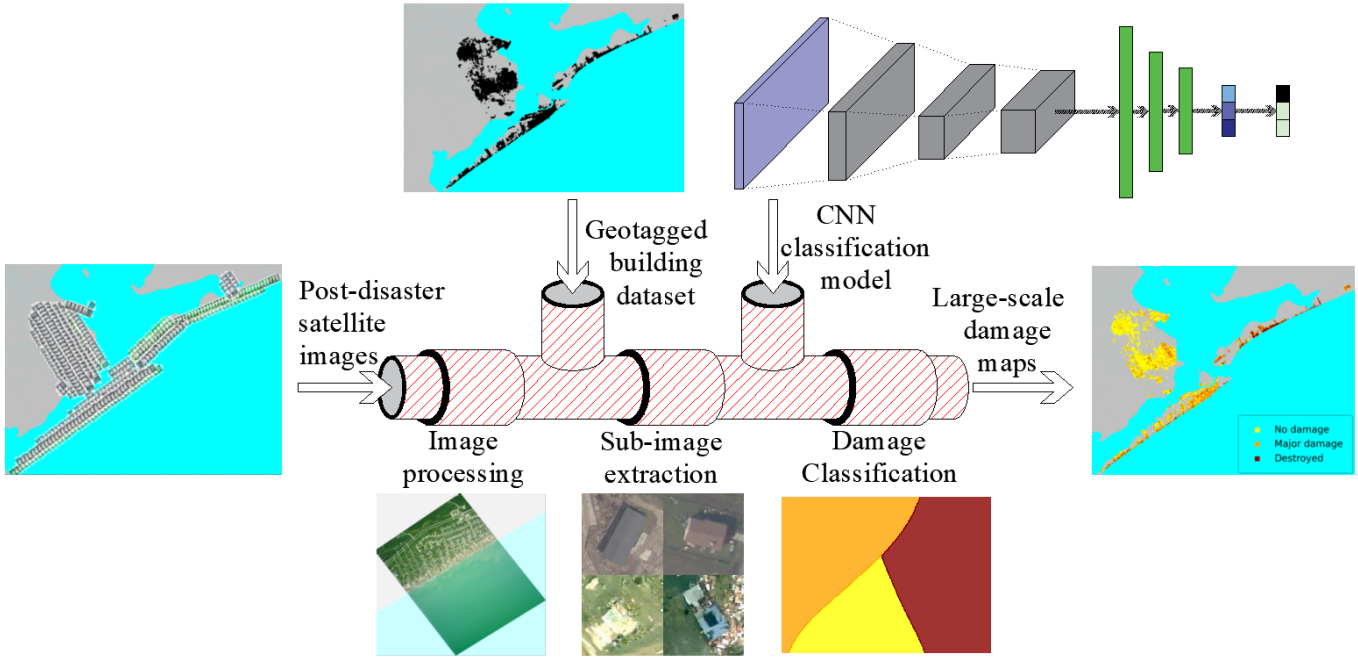

Figure 1 Flowchart showing key steps of the proposed methodology for large-scale building damage mapping

2.1 | Georeferenced satellite images

The methodology begins with the acquisition of satellite images, leveraging the rapid and wide coverage availability from sources such as the United States Geological Survey (USGS), the National Oceanic and Atmospheric Administration (NOAA), and the European Space Agency (ESA). For example, NOAA provides timely access to high-resolution satellite imagery after the occurrence of a natural hazard.

When using satellite imagery for assessing building damage, it is important to reference the images within a GIS framework. This involves providing details such as coordinate reference system, projection parameters, and transformation information. The satellite images are usually encoded using specialized GIS file formats like GeoTiff (Mahammad and Ramakrishnan 2003) or accompanied with separate world files (ESRI 2024), which allows them to be integrated into the broader GIS environment.

A standard world file stores 6 ordered transformation parameters: A, D, B, E, C, F, which are pixel size in the x-direction, negative pixel size in the y-direction, rotation about the y-axis, rotation about the x-axis, x-coordinate of the center of the upper left pixel, and y-coordinate of the center of the upper left pixel, respectively. Consequently, the coordinates (Lon_i, Lat_i) of the center of any pixel i with column numbers (x_i, y_i) can be calculated using Eq. (1) and Eq. (2), respectively (ESRI 2024).

$$Lon_i = Ax_i + By_i + C \qquad (1)$$

$$Lat_i = Dx_i + Ey_i + F \qquad (2)$$

2.2 | Sub-image extraction using geotagged building data

The utilization of deep learning for object detection is essential for objects that are typically not geotagged, such as post-hazard tree fall patterns (Rhee et al. 2021). When it comes to building identification, the necessity for deep learning arises only in the absence of geotagged building data; however, it becomes redundant when such data is readily available. Given that governments and counties usually maintain comprehensive datasets, and testbeds provide accessible data for research purposes (Amin Enderami et al. 2022), it becomes more efficient to leverage such existing data for extracting building sub-images. Employing highly accurate geotagged data ensures that the building identification process is virtually error-free and allows the deep learning model to focus exclusively on the damage classification task. By integrating them with georeferenced satellite imagery, sub-images of every building within the target region can be automatically extracted with the size of extraction proportional to the area or dimensions of the building. These sub-images represent the input of the deep learning classification model discussed in the next section.

2.3 | Damage classification using deep learning

2.3.1 | Training dataset

Sections 2.1 and 2.2 above discussed the GIS processing

and building sub-image extraction of the target post-disaster images. These differ than the training dataset needed to train the CNN, which is discussed in this section. A comprehensive dataset is essential for training a robust model to accurately classify buildings based on their damage states. It exposes the model to diverse examples of building damage, facilitating learning and generalization of associated patterns. Recent contributions, notably the xBD dataset (Gupta et al. 2019), help address this challenge and allow for the extraction of thousands of labeled sub-images of buildings for model training and testing. While xBD is the largest building damage assessment dataset to date, covering 16 different natural hazards, including earthquakes, hurricanes, floods, tornados, fires, and more, there is recognition that more thorough datasets may be required in the future (Matin and Pradhan 2022). This acknowledgment stems from potential differences between the terrain and damage patterns present in target satellite images and those covered in the training data. Fine-tuning, as discussed later, may be necessary to address this challenge.

The xBD dataset comprises both pre- and post-hazard satellite images, with post-hazard images labeled for four damage states: no damage, minor damage, major damage, and collapse. Identifying minor damage, relying on subtle features like missing roof elements and minor cracks, proves challenging in satellite imagery, yielding sub-optimal results in prior studies (e.g., Kaur et al. (2023)). Acknowledging this limitation, the current study focuses on the more distinguishable classes; no damage, major damage, and destroyed. Therefore, the immediate and rapid classification using satellite images can serve as an initial step for more comprehensive assessments, utilizing higher quality aerial and terrestrial imagery capable of capturing additional damage states. With the availability of higher-resolution satellite imagery, future research can explore their utility in detecting more nuanced damage states.

The xBD dataset consists of large-scale satellite images, with accompanying files of the pixel location of buildings within these images. Since the CNN used in this paper takes sub-images of buildings as input, the xBD sub-images were extracted using the pixel locations, with each image having a size of 120x120 pixels. Then, the images were manually filtered to include only the high-quality ones with distinct damage state, to ensure accurate training of the CNN. None of the selected sub-images included minor damage states as discussed earlier. These are used for the training and differ from the sub-images discussed in the previous section, which are automatically extracted from the satellite images using geolocations obtained from building databases. Since the images extracted automatically from the satellite ones could potentially have parts from neighboring buildings, especially in urban regions with closely spaced buildings, the training dataset was selected to include some images with neighboring buildings appearing. While the Hurricane Ike application dataset only includes post-disaster images, future research could be applied to other testbeds with both pre- and post-disaster satellite images available. Hence, advanced CNN models could be utilized to have such neighboring buildings artificially removed from the sub-images. This is acknowledged as a limitation of the current study.

2.3.2 | CNN training

CNN is a specialized deep learning model designed for image recognition and classification tasks, excelling in the analysis of intricate patterns within visual data (Goodfellow et al. 2016). This research employs a CNN architecture to automate the classification of building damage states in post-disaster satellite imagery. The proposed CNN model in this paper focuses on the classification of building sub-images, leveraging prior building identification through geotagging. This targeted approach allows for the use of a relatively shallow network.

The input to the CNN consists of labeled images, each represented as a 3D tensor with dimensions corresponding to pixels (height and width) and RGB normalized channels. The model architecture includes convolutional, pooling, fully connected, and SoftMax layers. Convolutional layers are integral for spatial feature extraction, employing convolution operations to identify patterns and detect local features. The application of ReLU activation functions within these layers introduces non-linearity, enabling the network to capture complex relationships in the data. Pooling layers systematically down sample spatial dimensions, contributing to dimensionality reduction and computational efficiency. Fully connected layers play a central role in comprehensive learning, connecting nodes between layers and integrating high-level features. The final SoftMax layer transforms the network's output into pseudo probabilities for each class, reflecting the model's confidence in its predictions. The classification is determined by selecting the class with the highest probability, representing the model's final classification output. Figure 2 visually represents the standard CNN architecture, illustrating the arrangement of convolutional (Conv), pooling (pool) layers, and fully connected (FC) layers, followed by a SoftMax layer and a classification layer (Goodfellow et al. (2016); Roberts et al. (2022); Prince (2023)).

The loss function serves as a critical component in the training of neural networks, guiding the optimization process by quantifying the dissimilarity between predicted values and ground-truth labels. In the context of multi-classification tasks, the cross-entropy loss stands out as a widely adopted and effective choice for minimizing the disparity between predicted and actual class distributions (Krizhevsky et al. 2017). The cross-entropy loss penalizes the model more severely when its predicted probabilities deviate from the true distribution. Mathematically, the

cross-entropy loss for a multi-class classification task is defined as shown in Eq. (3), where m is the number of classes, q_i is the ground-truth label, and p_i is the SoftMax probability.

$$CrossEntropy = -\sum_{i=1}^{m}[q_i \log(p_i)] \quad (3)$$

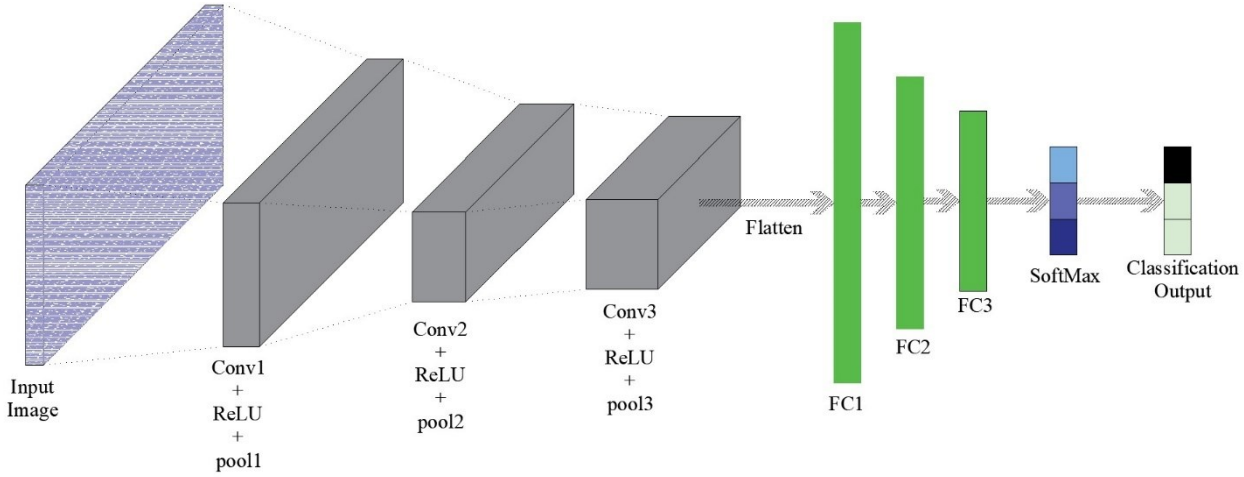

Figure 2 CNN architecture for image classification

2.3.3 | Performance evaluation and fine-tuning

Evaluating the model's performance is an important step in ensuring its efficacy in building damage state classification, employing key metrics such as accuracy, precision, recall, F1 score, and the confusion matrix (Powers 2020). Initially, the testing data, derived from a subset of the training dataset, serves as a benchmark to assess the model's proficiency in pattern recognition. However, success with the testing data doesn't guarantee generalization to new data. Therefore, for new target satellite images, an extracted subset is essential to evaluate the model's ability to generalize. When the model performs well on testing data but shows suboptimal results on the application data, it suggests challenges related to generalization, implying that the new satellite images don't align with the characteristics of the training images. In such cases, transfer learning through fine-tuning, using a small manually classified sample from the target satellite images, becomes necessary. This enhances the model's proficiency, ensuring it excels not only in controlled testing conditions but also demonstrates robust performance in real-world applications.

2.4 | Large-scale mapping

Following the CNN training, the model is applied to classify every geotagged sub-image extracted from the satellite imagery into distinct damage states. This automated mapping enables the creation of comprehensive damage maps, providing a spatial overview of damage variation across the affected disaster region. The geotagged sub-images, being connected to the building database, offer additional utility beyond mapping. They can serve as inputs for statistical analyses, contributing valuable insights into damage patterns. Moreover, they provide a means to update pre-hazard prediction models, enhancing the predictive capabilities of future assessments.

3 | APPLICATION STUDY

3.1 | Hurricane Ike satellite images and Galveston Testbed

The application of the proposed methodology is demonstrated using the Galveston testbed and Hurricane Ike satellite images. In contrast to the xBD dataset utilized for model training, the Hurricane Ike images were independently sourced from NOAA (NOAA 2008). Additionally, the geotagged building data was acquired from Incore (Incore 2023). This diversification of data sources ensures a comprehensive validation of the model's robustness in real-world scenarios.

Figure 3 (a) provides a detailed overview of the building map of Galveston, incorporating a total of 84,024 buildings. This breakdown includes 49,060 buildings in Mainland Galveston, 29,480 buildings in Galveston Island, and 5,484 buildings in Bolivar Peninsula. Complementing this, Figure 3 (b) shows the projection of satellite images onto Galveston County, totaling 527 images, each with dimensions ranging between 2.75 km to 3.75 km. The integration of these maps is presented in Figure 4, showcasing the projection of geotagged buildings over sample georeferenced satellite images in Mainland Galveston, Galveston Island, and Bolivar Peninsula. The seamless alignment between the two maps ensures error-free extraction of building sub-images.

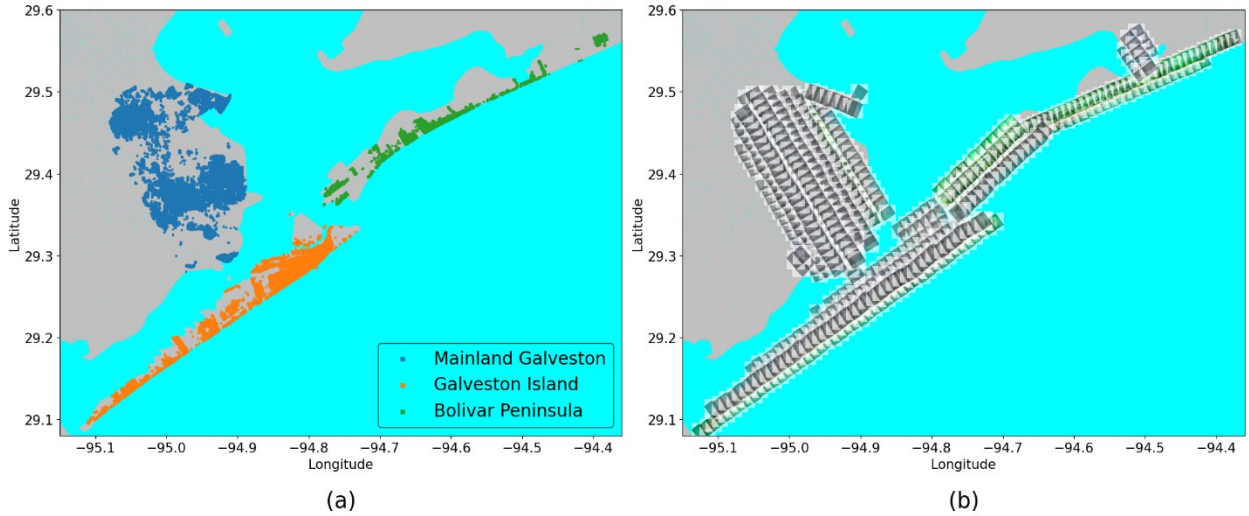

Figure 3 Graphic view of Galveston Island's: (a) Building map (b) Post-Ike satellite images

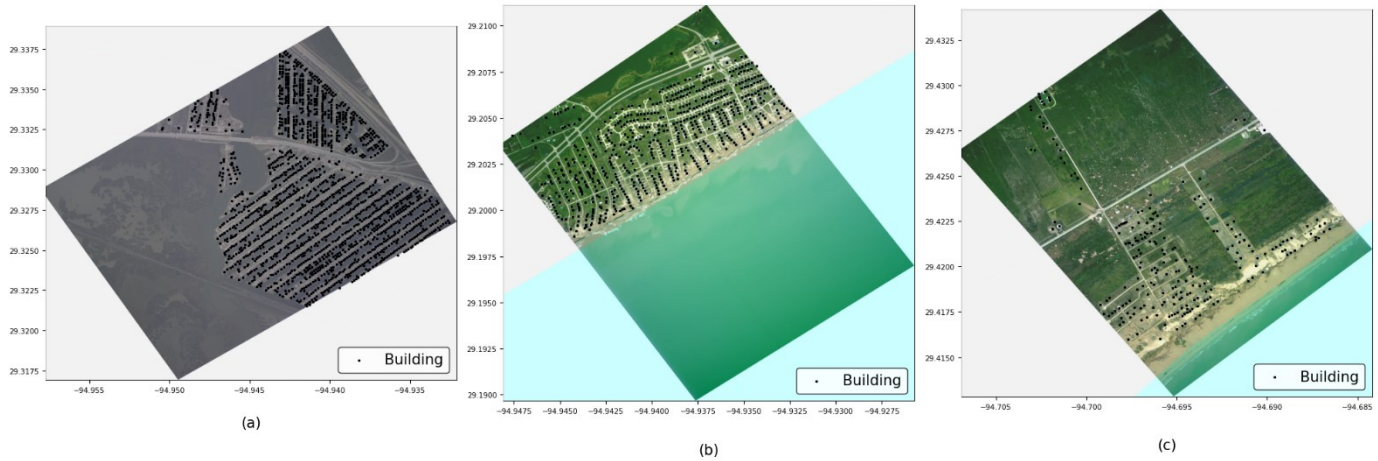

Figure 4 Projection of geotagged buildings over sample georeferenced satellite images in: (a) Mainland Galveston (b) Galveston Island (c) Bolivar Peninsula

3.2 | CNN training and testing

Approximately 5,000 high-quality labeled sub-images were extracted from the xBD dataset, representing samples from four hazards, namely Hurricane Harvey, Hurricane Michael, Joplin Tornado, and Tuscaloosa Tornado. The dataset was partitioned, allocating 80% for training and the remaining 20% split equally between validation and testing subsets. To address class imbalance, the training dataset underwent oversampling using the SMOTE technique (Chawla et al. 2002). Additionally, data augmentation techniques, such as random contrast and brightness adjustments, were employed to enhance variability, resulting in an augmented dataset comprising approximately 29,000 sub-images for training.

The architecture of the CNN encompasses five convolutional layers (of 3x3 filters with sizes 32, 64, 128,128, 128), five max-pooling layers (2x2), and five fully connected layers (of sizes 128, 128, 64, 64, 32). The model, concluding with a Softmax layer for classification, is characterized by a total of 485,187 trainable parameters. This architecture was selected based on manual experiments. Then, hyperparameter tuning was performed using the validation data. Therefore, the training of the model underwent 15 epochs with a batch size of 64, utilizing the Adam optimizer with a 0.001 learning rate and the cross-entropy loss function.

The model's performance was evaluated on the testing dataset, achieving accuracy, precision, Recall, and F1 scores of 93%, 90%, 86%, and 88%, respectively. Figure 5 presents the confusion matrix for the testing dataset, illustrating the CNN's high accuracy in classifying the non-damaged and destroyed damage states (98% and 96%, respectively), with sufficient accuracy for the major damage state (65%). Additionally, Table 1 compares the model's performance with previous published papers. While the table demonstrates the current methodology's strong performance compared to state-of-the-art studies, any further direct conclusions should consider the major differences between studies, such as the variations in the number of damage classes, training datasets, use of pre-hazard images, and consideration of segmentation tasks before classification. For example, while Kaur et al. (2022) achieved a higher F1 score, the current study would achieve

an even higher F1 score of 98% if both major damage and destroyed damage states are merged since most of the error shown in Figure 5 stems from the confusion between these two damage states.

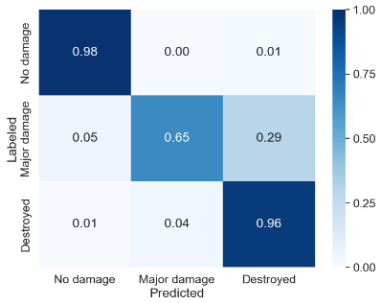

Figure 5 Confusion matrix of testing data

Table 1 Comparision of the testing results of the current study with previously published papers

Reference	Number of classes	Training dataset	Satellite images	CNN task	F1 score (%)
Li et al. (2019)	3	Sandy	Post only	Classification	68
Dotel et al. (2020)	4	Harvey	Pre and post	Classification & segmentation	68
Kaur et al. (2022)	2	Harvey	Post only	Classification	96
Kaur et al. (2023)	4	xBD	Pre and post	Classification & segmentation	80
Current paper	3	xBD	Post only	Classification	88

Moreover, it should be noted that the high performance on the testing data indicates the CNN's efficacy on images with characteristics similar to the xBD dataset, encompassing comparable damage, building, and terrain patterns. The training sub-images, chosen for high quality and distinctive damage patterns, allowed the CNN to effectively learn features, especially as it could focus exclusively on this task, given successful building identification using GIS. The testing subset served to protect against overfitting. However, it's crucial to acknowledge that this doesn't guarantee the model's generalizability to new hazards with different patterns. While most previous studies conclude at this stage, this study goes further by applying the methodology to the Galveston testbed when subjected to Hurricane Ike, as discussed in the subsequent sections.

3.3 | CNN evaluation on real-world data

The CNN model's performance was evaluated using approximately 600 sub-images extracted from Hurricane Ike satellite imagery. To establish a ground truth for comparison, these sub-images underwent manual classification. Figure 6 (a) shows the confusion matrix when applying the trained CNN model on the real-world data. Despite the model's success in classifying "no damage" and "destroyed" classes, its performance was suboptimal for "major damage". This suggests that the damage patterns of Hurricane Ike significantly differ from those in the xBD dataset, necessitating transfer learning via fine-tuning.

Figure 6 (b), Figure 6 (c), and Figure 6 (d) show the confusion matrices after fine-tuning the model with 0.5%, 1.0%, and 1.5% of Hurricane Ike sub-images, respectively. The fine-tuning subsets excluded any evaluation data to ensure accurate assessment. Moreover, Table 2 summarizes the evaluation metrics for all 4 cases. These highlight the positive impact of fine-tuning on the model's robustness, demonstrating a significant enhancement in performance with an increased fine-tuning sample size. Given that real-world applications often cover areas with diverse damage patterns and terrains compared to the training data, as evident in this application study, fine-tuning becomes important to ensure the methodology's robustness in varied scenarios.

Figure 7 shows a sample of classified sub-images using the fine-tuned model. These sub-images were automatically extracted from the post-disaster satellite imagery using the methodology proposed in Section 2.2, and then fed into the pre-trained CNN model. While most of the images were correctly classified as demonstrated in Figure 6 (d), the model still misclassified some images. It is also worth noting that the extraction included some parts of neighboring buildings, which is acknowledged as a limitation of utilizing GIS and building area features in extracting the sub-images.

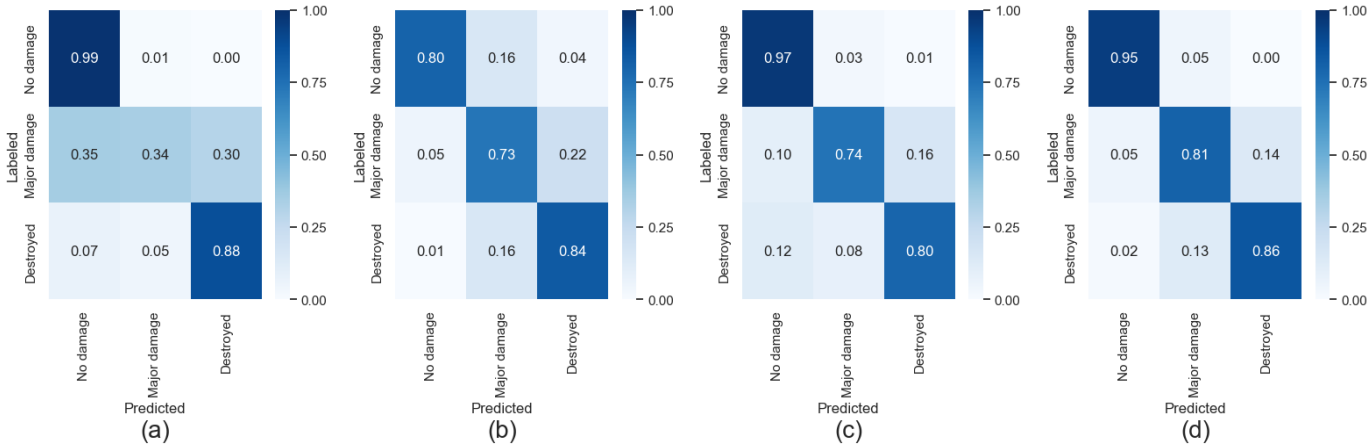

Figure 6 Confusion matrix of Ike data: (a) Without fine-tuning (b) Fine-tuned with 0.5% of Ike images (c) Fine-tuned with 1.0% of Ike images(d) Fine-tuned with 1.5% of Ike images

Table 2 Accuracy, precision, recall, and F1 scores for various fine-tuning levels

Evaluation Metric	Without Fine-tuning	Fine-tuning with 0.5% of Ike sub-images	Fine-tuning with 1.0% of Ike sub-images	Fine-tuning with 1.5% of Ike sub-images
Accuracy (%)	85	80	88	90
Precision (%)	82	75	85	85
Recall (%)	74	79	83	87
F1 (%)	75	76	84	86

Figure 7 Sample of classified sub-images

3.4 | Classification outcomes

The fine-tuned model was applied to classify all extracted sub-images, and subsequently, the classification results were geotagged and projected back onto the satellite images. In Figure 8, Figure 9, and Figure 10, sample classification maps for Mainland Galveston, Galveston Island, and Bolivar Peninsula are presented, respectively. These maps visually show the model's classification outcomes across various geographic areas, providing valuable insights into the spatial distribution of building damage states following Hurricane Ike. The buildings were precisely identified and accurately classified using the proposed methodology. Figure 8 illustrates no damage classification for all buildings in the satellite image, representing the situation in most areas on Mainland Galveston. Figure 9 shows a substantial amount of major damage and some destruction in Palm Beach, while Figure 10 showcases significant damage and destruction on the Bolivar Peninsula.

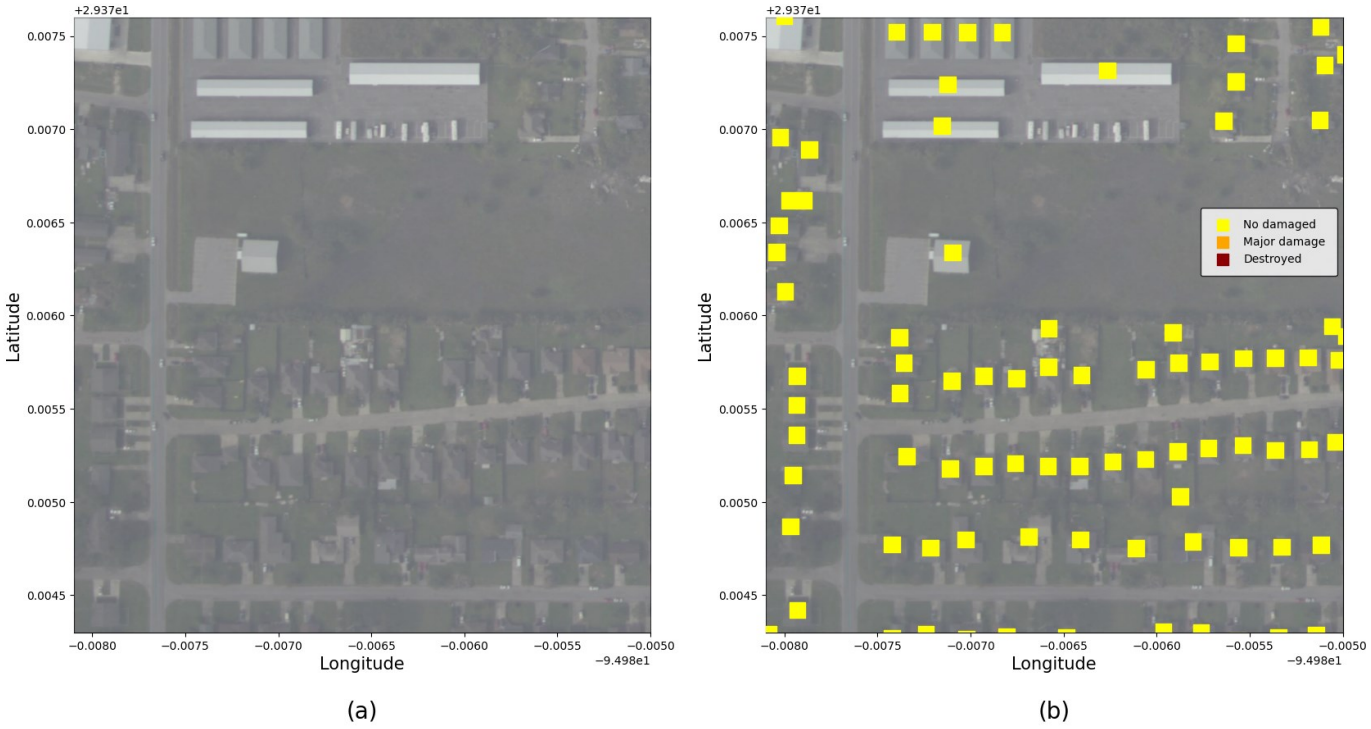

Figure 8 Sample map of Galveston Mainland's: (a) Satellite image (b) Detection of major damage and destruction

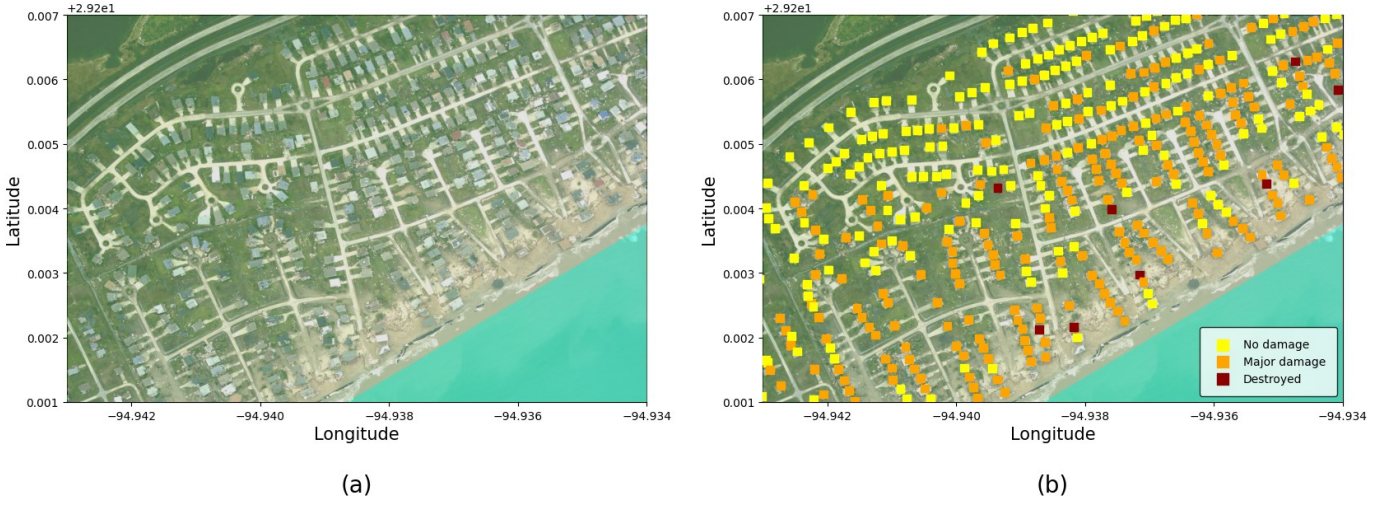

Figure 9 Sample map of Galveston Island's: (a) Satellite image (b) Detection of major damage and destruction

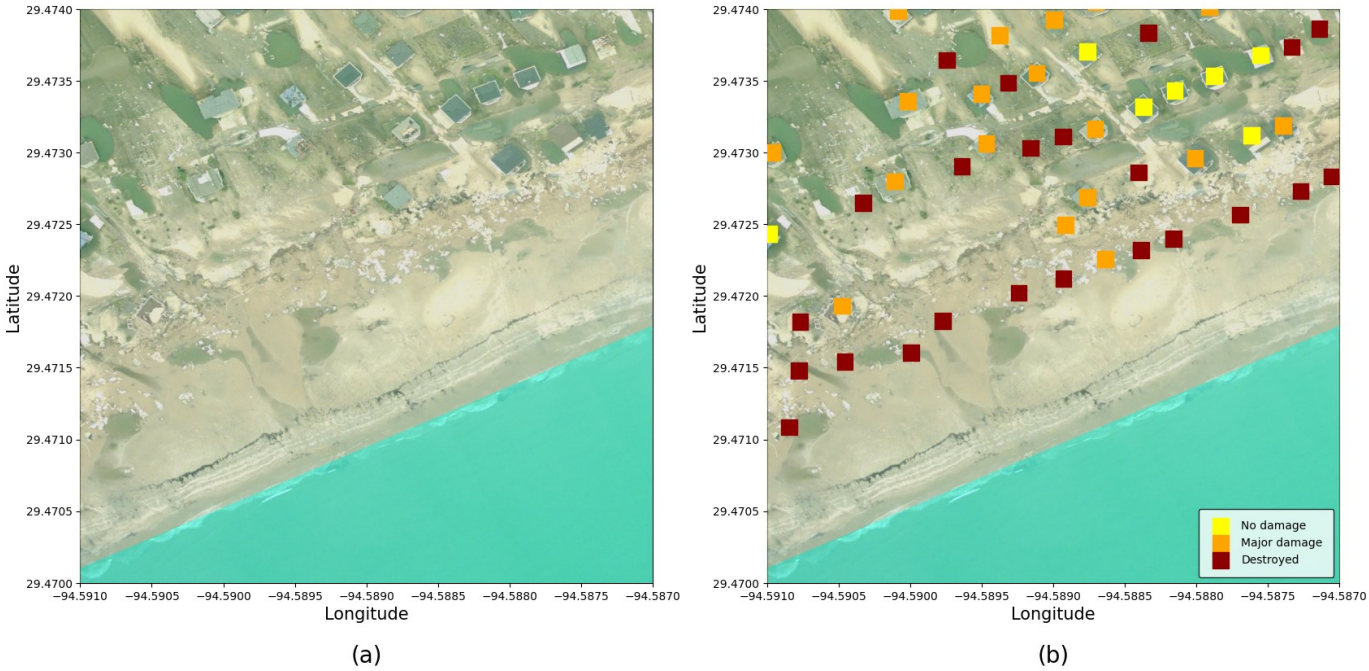

Figure 10 Sample map of Bolivar Peninsula's: (a) Satellite image (b) Detection of major damage and destruction

3.5 | Large-scale mapping and validation against historical reports

Galveston County, spanning Galveston Island, Bolivar Peninsula, and Mainland Galveston, provides a diverse landscape that showcases distinctive damage patterns resulting from Hurricane Ike. This diversity allows for a comprehensive evaluation of the methodology's effectiveness in capturing and categorizing unique damage and destruction features in each region.

As per historical reports (FCD (2008); Berg (2009); Highfield et al. (2014)), the Bolivar Peninsula emerged as the most severely affected area, with a majority of its buildings either destroyed or suffering major damage. Approximately 3,266 buildings on the peninsula faced destruction or severe damage in the aftermath of the hurricane. Galveston Island, on the other hand, displayed a more varied pattern of damage. Higher damage levels were observed on the bay side, attributed to surge washing back from the bay. The main city on the right portion of Galveston Island, closer to the seawall, experienced comparatively less destruction, but considerable damage, nonetheless. In contrast, the western part of the island suffered the most extensive destruction. Mainland Galveston, like Houston-surrounding area, was fortunate to avoid significant damage, sparing most of its buildings from the major impacts of the hurricane.

Figure 11 illustrates a large-scale damage map of Galveston County, covering tens of thousands of buildings across the entire region. Generated using the proposed methodology, the map aligns well with the historical reports, depicting the Bolivar Peninsula as severely affected, with thousands of buildings extensively damaged or destroyed. Galveston Island exhibits concentrated damage on the bay side, while Galveston Mainland appears to be the least affected, with very few buildings damaged and even fewer destroyed. Complementing the map, Figure 12 presents bar plots of damage classification results for Galveston County as a whole, Mainland Galveston, Galveston Island, and Bolivar Peninsula. The plots align well with the historical records and further support the conclusions made based on Figure 11 regarding the spatial variability of damage between the different regions. The statistics generated in Figure 12 can be automatically and rapidly generated using the proposed methodology, which holds great potential to gain immediate insight into the spatial distribution of damage. This can be used to update emergency response plans and direct rescue crews. Moreover, there is a potential to integrate social science with the results of the current methodology to identify socially vulnerable areas within the disaster affected region.

In contrast to historical reports that took weeks to prepare, such as Highfield et al. (2014), which stated that assessing only 1,500 buildings took 2,000 hours of fieldwork, the map and statistics presented here can be generated automatically and almost immediately after the hazard, as soon as satellite images become available.

It is important to acknowledge that while satellite images prove effective in detecting major damage and destruction, their efficacy in identifying minor damages may be limited. In the case of Galveston County, where around 24,000 buildings experienced some level of damage (FCD 2008), the proposed methodology concentrates specifically on detecting the buildings that either faced destruction or suffered major damage. Consequently, the damage assessment derived from satellite images provides a comprehensive overview, serving as an initial step for a

more thorough evaluation utilizing additional data sources, such as aerial and terrestrial imagery. The potential integration of these diverse data sources could be explored in a future study, where a unified methodology combines all available sources of data-driven damage detection.

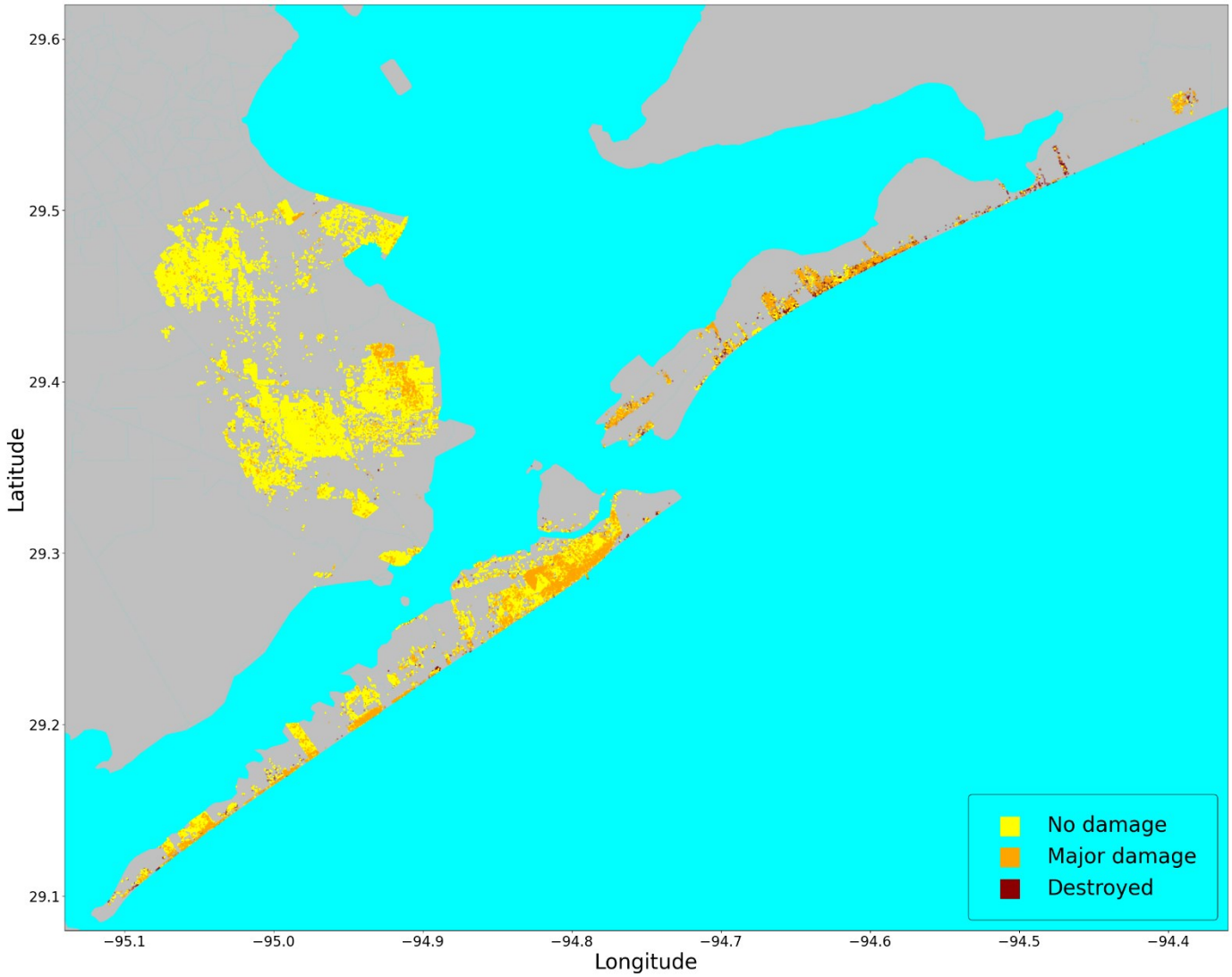

Figure 11 Galveston County large-scale map for the detection of major damage and destruction

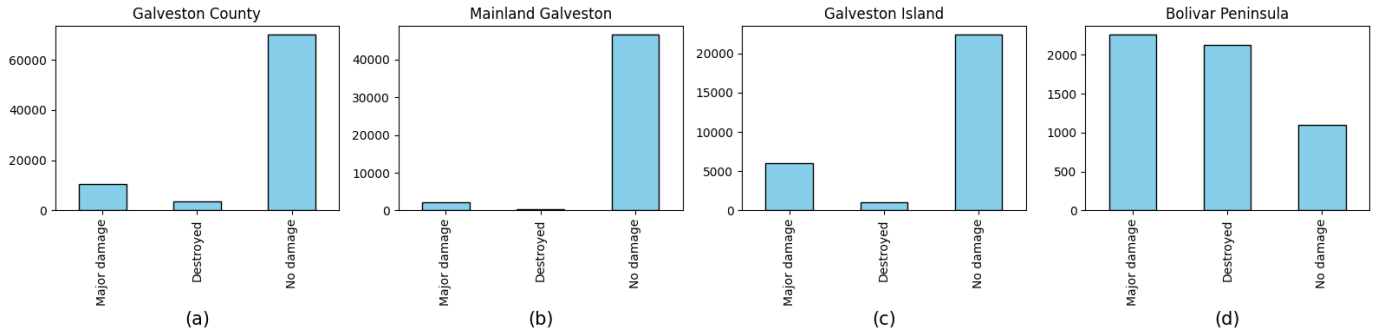

Figure 12 Bar plots of damage classification results for (a) Galveston County (all parts) (b) Mainland Galveston (c) Galveston Island (d) Bolivar Peninsula

4 | DISCUSSION, LIMITATIONS, AND FUTURE WORK

4.1 | Strengths and weaknesses of the proposed methodology compared to state-of-the-art methodologies

The primary strength of the proposed methodology lies in its full automation, spanning from initial satellite image acquisition to large-scale damage mapping. This comprehensive approach distinguishes it from prior studies, offering a valuable tool for future disaster risk management methodologies. Importantly, it eliminates the need for pre-hazard images with identical quality and angles to post-disaster imagery, a requirement in previous segmentation-focused approaches. Leveraging highly accurate geo-tagged building data enables precise extraction of building sub-images, reducing potential errors introduced by machine learning techniques. Nevertheless, the reliance on geotagged building databases presents a limitation, restricting the methodology's applicability in scenarios where such data is unavailable. Furthermore, while comparison between pre- and post-hazard images could enhance CNN's performance, the current methodology solely relies on post-disaster imagery, missing out on potential improvements afforded by pre-disaster data integration. These comparisons are summarized in Table 3.

Table 3 Comparing the advantages and disadvantages of the proposed methodology other to state-of-the-art methodologies

References	Method for building identification	Building identification depends on the availability of	Automates the generation of large-scale damage maps
Li et al. (2019), Kaur et al. (2022)	N/A	N/A	No
Hao et al. (2021), Wu et al. (2021), Shen et al. (2021), Gupta and Shah (2021), Kaur et al. (2023)	Deep learning	Pre-hazard satellite imagery	No
Current paper	GIS	Geotagged building database	Yes

4.2 | Limitations and future works

Despite the promising results for the automation of large-scale damage mapping, the proposed methodology still has limitations in its current form. The study relies heavily on the xBD dataset. While being one of the most comprehensive datasets, xBD suffers from a relatively low resolution and doesn't fully represent the diversity of building damage patterns and archetypes worldwide. Hence, future work could enhance the generalization across different disaster scenarios by expanding the training datasets to include a more diverse range of hazards and regions. Consequently, the robustness of the methodology could be further validated by applying it to other types of natural hazards, such as tornados, earthquakes, and wildfires.

Moreover, damage detection using satellite imagery is limited to major roof damage and complete destruction of walls. Therefore, future work could integrate geolocation features (as suggested by Cao and Choe (2020)) and engineering models to predict the minor and moderate damage states. These features associated with different hazards could include the distance from the earthquake epicenter and earthquake magnitude, distance from the shoreline and hurricane scale, or distance from the tornado center path and tornado intensity. This will necessitate a more complex deep learning model compared to the simple CNN architecture utilized in the current study. Therefore, exploring advanced deep learning architectures and techniques would further augment the model's capacity to capture damage patterns. In addition to the U-net and Siamese structures discussed earlier, future research could explore the application of other sophisticated deep learning algorithms such as neural dynamic classification and self-supervised learning algorithms (Rafiei and Adeli (2017); Alam et al. (2020); Rafiei et al. (2022)).

Integrating the proposed methodology with building databases holds significant potential, allowing the combination of pre-hazard risk assessment models with post-hazard immediate response strategies, representing a significant step towards achieving a disaster management digital twin. Hence, restoration and recovery models could be updated using real-time-generated damage classifications. This could also facilitate continuous learning and model improvement by updating physics-based prediction models.

Moreover, the large-scale maps generated by the methodology lay the groundwork for comprehensive damage assessment, incorporating terrestrial and aerial imagery alongside field observations. Therefore, the temporal integration of satellite imagery and real-time data streams could introduce a dynamic element to damage assessment, allowing for more responsive and up-to-date analyses.

5 | CONCLUSIONS

This research introduces a methodology to automate the large-scale damage assessment following natural hazards. The novelty lies in the utilization of geotagged building datasets to extract sub-images from georeferenced satellite imagery, resulting in error-free building identification, and enabling deep learning to focus on damage classification. Hence, the entire system can be automated starting from the acquisition of satellite imagery and ending with the generation of large-scale damage maps connected to

building databases, as explained in the methodology introduced in Section 2.

The application study presented in Section 3, using the Galveston Testbed and Hurricane Ike satellite images, showcased the methodology's efficiency in identifying buildings using GIS, and the training of the CNN using the xBD dataset showed high performance on testing datasets, with an F1 score of 88%. Fine-tuning further ensured the model's robust performance when applied to a new dataset with diverse landscapes and damage patterns, with an F1 score of 86%. The validation against historical reports provided evidence of the model's high accuracy in detecting damage patterns and spatial variability.

This research serves to automate the rapid generation of large-scale damage assessment maps, offering valuable support for immediate emergency response, and establishing a foundation for future advancements in disaster management and recovery efforts. The next step is to apply the current methodology in disaster management by updating the risk-assessment predictions using post-disaster damage assessment, hence, extending the risk-assessment models toward disaster management digital twins.

6 | ACKNOWLEDGMENT

Financial support for this work was provided by the US National Science Foundation (NSF) under Award Number 2052930. This financial support is gratefully acknowledged. Any opinions, findings, conclusions, and recommendations presented in this paper are those of the authors and do not necessarily reflect the views of NSF.

7 | REFERENCES

Adeli, H. (2001). "Neural networks in civil engineering: 1989–2000." *Computer-Aided Civil and Infrastructure Engineering*, 16(2), 126-142.

Adeli, H., and Yeh, C. (1989). "Perceptron learning in engineering design." *Computer-Aided Civil and Infrastructure Engineering*, 4(4), 247-256.

Aghababaei, M., Koliou, M., and Paal, S. G. (2018). "Performance assessment of building infrastructure impacted by the 2017 Hurricane Harvey in the Port Aransas region." *Journal of Performance of Constructed Facilities*, 32(5), 04018069.

Aghababaei, M., Koliou, M., Pilkington, S., Mahmoud, H., van de Lindt, J. W., Curtis, A., Smith, S., Ajayakumar, J., and Watson, M. (2020). "Validation of time-dependent repair recovery of the building stock following the 2011 Joplin Tornado." *Natural Hazards Review*, 21(4), 04020038.

Alam, K. M. R., Siddique, N., and Adeli, H. (2020). "A dynamic ensemble learning algorithm for neural networks." *Neural Computing and Applications*, 32, 8675-8690.

Amin Enderami, S., Mazumder, R. K., Dumler, M., and Sutley, E. J. (2022). "Virtual Testbeds for Community Resilience Analysis: State-of-the-Art Review, Consensus Study, and Recommendations." *Natural Hazards Review*, 23(4), 03122001.

Bai, Y., Mas, E., and 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.

Berg, R. (2009). *Tropical cyclone report: Hurricane ike (al092008), 1-14 september 2008*, National Hurricane, Center.

Braik, A. M., and Koliou, M. (2023). "A novel digital twin framework of electric power infrastructure systems subjected to hurricanes." *International Journal of Disaster Risk Reduction*, 104020.

Cao, Q. D., and Choe, Y. (2020). "Building damage annotation on post-hurricane satellite imagery based on convolutional neural networks." *Natural Hazards*, 103(3), 3357-3376.

Cao, Q. D., and Choe, Y. (2020). "Post-hurricane damage assessment using satellite imagery and geolocation features." *arXiv preprint arXiv:2012.08624*.

Cha, Y. J., Choi, W., and Büyüköztürk, O. (2017). "Deep learning-based crack damage detection using convolutional neural networks." *Computer-Aided Civil and Infrastructure Engineering*, 32(5), 361-378.

Cha, Y. J., Choi, W., Suh, G., Mahmoudkhani, S., and Büyüköztürk, O. (2018). "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types." *Computer-Aided Civil and Infrastructure Engineering*, 33(9), 731-747.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). "SMOTE: synthetic minority over-sampling technique." *Journal of artificial intelligence research*, 16, 321-357.

Chen, W., He, Z., and Zhang, J. (2023). "Online monitoring of crack dynamic development using attention-based deep networks." *Automation in Construction*, 154, 105022.

Cheng, C. S., Behzadan, A. H., and Noshadravan, A. (2021). "Deep learning for post-hurricane aerial damage assessment of buildings." *Computer-Aided Civil and Infrastructure Engineering*, 36(6), 695-710.

Chu, H., Wang, W., and Deng, L. (2022). "Tiny-Crack-Net: A multiscale feature fusion network with attention mechanisms for segmentation of tiny cracks." *Computer-Aided Civil and Infrastructure Engineering*, 37(14), 1914-1931.

Deng, J., Lu, Y., and Lee, V. C. S. (2020). "Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network." *Computer-Aided Civil and Infrastructure Engineering*, 35(4), 373-388.

Dotel, S., Shrestha, A., Bhusal, A., Pathak, R., Shakya, A., and Panday, S. P. "Disaster assessment from satellite imagery by analysing topographical features using deep learning." *Proc., Proceedings of the 2020 2nd International Conference on Image, Video and Signal Processing*, 86-92.

ESRI (2024). "Understanding World Files." <https://webhelp.esri.com/arcims/9.3/General/topics/author_world_files.htm>. (2024).

FCD (2008). "Harris County Flood Control District. Hurricane Ike 2008." <https://www.hcfcd.org/About/Harris-Countys-Flooding-History/Hurricane-Ike-2008>. (2023).

Gao, T., Yuanzhou, Z., Ji, B., and Xie, Z. (2024). "Vision-based fatigue crack automatic perception and geometric updating of finite element model for welded joint in steel structures." *Computer-Aided Civil and Infrastructure Engineering*.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*, MIT press.

Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J., Heim, E., Choset, H., and Gaston, M. (2019). "xbd: A dataset for assessing building damage from satellite imagery." *arXiv preprint arXiv:1911.09296*.

Gupta, R., and Shah, M. "Rescuenet: Joint building segmentation and damage assessment from satellite imagery." *Proc., 2020 25th International Conference on Pattern Recognition (ICPR)*, IEEE, 4405-4411.

Hao, H., Baireddy, S., Bartusiak, E. R., Konz, L., LaTourette, K., Gribbons, M., Chan, M., Delp, E. J., and Comer, M. L. "An attention-based system for damage assessment using satellite imagery." *Proc., 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, 4396-4399.

Highfield, W. E., Peacock, W. G., and Van Zandt, S. (2014). "Mitigation planning: Why hazard exposure, structural vulnerability, and social vulnerability matter." *Journal of Planning Education and Research*, 34(3), 287-300.

Hong, Z., Zhong, H., Pan, H., Liu, J., Zhou, R., Zhang, Y., Han, Y., Wang, J., Yang, S., and Zhong, C. (2022). "Classification of building damage using a novel convolutional neural network based on post-disaster aerial images." *Sensors*, 22(15), 5920.

Incore (2023). "Galveston Testbed." <https://incore.ncsa.illinois.edu/>.

Javadinasab Hormozabad, S., Gutierrez Soto, M., and Adeli, H. (2021). "Integrating structural control, health monitoring, and energy harvesting for smart cities." *Expert Systems*, 38(8), e12845.

Kaur, N., Lee, C. C., Mostafavi, A., and Mahdavi-Amiri, A. (2023). "Large-scale building damage assessment using a novel hierarchical transformer architecture on satellite images." *Computer-Aided Civil and Infrastructure Engineering*.

Kaur, S., Gupta, S., Singh, S., Koundal, D., and Zaguia, A. (2022). "Convolutional neural network based hurricane damage detection using satellite images." *Soft Computing*, 26(16), 7831-7845.

Khajwal, A. B., Cheng, C. S., and Noshadravan, A. (2023). "Post-disaster damage classification based on deep multi-view image fusion." *Computer-Aided Civil and Infrastructure Engineering*, 38(4), 528-544.

Koliou, M., and van de Lindt, J. W. (2020). "Development of building restoration functions for use in community recovery planning to tornadoes." *Natural Hazards Review*, 21(2), 04020004.

Koliou, M., van de Lindt, J. W., McAllister, T. P., Ellingwood, B. R., Dillard, M., and Cutler, H. (2020). "State of the research in community resilience: Progress and challenges." *Sustainable and resilient infrastructure*, 5(3), 131-151.

Kong, Q., Gu, J., Xiong, B., and Yuan, C. (2023). "Vision-aided three-dimensional damage quantification and finite element model geometric updating for reinforced concrete structures." *Computer-Aided Civil and Infrastructure Engineering*.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). "ImageNet classification with deep convolutional neural networks." *Communications of the ACM*, 60(6), 84-90.

Li, Y., Hu, W., Dong, H., and Zhang, X. (2019). "Building damage detection from post-event aerial imagery using single shot multibox detector." *Applied Sciences*, 9(6), 1128.

Liang, X. (2019). "Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization." *Computer-Aided Civil and Infrastructure Engineering*, 34(5), 415-430.

Liu, Y. F., Nie, X., Fan, J. S., and Liu, X. G. (2020). "Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction." *Computer-Aided Civil and Infrastructure Engineering*, 35(5), 511-529.

Lozano, J. M., Nichols, E., Frost, J. D., and Tien, I. (2023). "Spatial and temporal evolution of post-disaster data for damage assessment of civil infrastructure systems." *Geomatics, Natural Hazards and Risk*, 14(1), 2250531.

Mahammad, S. S., and Ramakrishnan, R. (2003). "GeoTIFF-A standard image file format for GIS applications." *Map India*, 28-31.

Masoomi, H., and van de Lindt, J. W. (2016). "Tornado fragility and risk assessment of an archetype masonry school building." *Engineering Structures*, 128, 26-43.

Matin, S. S., and Pradhan, B. (2022). "Challenges and limitations of earthquake-induced building damage mapping techniques using remote sensing images-A systematic review." *Geocarto International*, 37(21), 6186-6212.

Miyamoto, T., and Yamamoto, Y. (2021). "Using 3-D convolution and multimodal architecture for earthquake damage detection based on satellite imagery and digital urban data." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 8606-8613.

NOAA (2008). "National Oceanic and Atmospheric Administration. Hurricane Ike Images." <https://www.noaa.gov/>. (2023).

Nofal, O. M., Van De Lindt, J. W., Do, T. Q., Yan, G., Hamideh, S., Cox, D. T., and Dietrich, J. C. (2021). "Methodology for regional multihazard hurricane damage and risk assessment." *Journal of Structural Engineering*, 147(11), 04021185.

Pereira, D. R., Piteri, M. A., Souza, A. N., Papa, J. P., and Adeli, H. (2020). "FEMa: A finite element machine for fast learning." *Neural Computing and Applications*, 32, 6393-6404.

Perez-Ramirez, C. A., Amezquita-Sanchez, J. P., Adeli, H., Valtierra-Rodriguez, M., Camarena-Martinez, D., and Romero-Troncoso, R. J. (2016). "New methodology for modal parameters identification of smart civil structures using ambient vibrations and synchrosqueezed wavelet transform." *Engineering Applications of Artificial Intelligence*, 48, 1-12.

Pezeshki, H., Adeli, H., Pavlou, D., and Siriwardane, S. C. "State of the art in structural health monitoring of offshore and marine structures." *Proc., Proceedings of the Institution of Civil Engineers-Maritime Engineering*, Thomas Telford Ltd, 89-108.

Pezeshki, H., Pavlou, D., Adeli, H., and Siriwardane, S. C. (2023). "Modal analysis of offshore monopile wind turbine: An analytical solution." *Journal of Offshore Mechanics and Arctic Engineering*, 145(1), 010907.

Powers, D. M. (2020). "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." *arXiv preprint arXiv:2010.16061*.

Prince, S. J. (2023). *Understanding Deep Learning*, MIT press.

Rafiei, M. H., and Adeli, H. (2016). "A novel machine learning model for estimation of sale prices of real estate units." *Journal of Construction Engineering and Management*, 142(2), 04015066.

Rafiei, M. H., and Adeli, H. (2017). "A new neural dynamic classification algorithm." *IEEE transactions on neural networks and learning systems*, 28(12), 3074-3083.

Rafiei, M. H., Gauthier, L. V., Adeli, H., and Takabi, D. (2022). "Self-supervised learning for electroencephalography." *IEEE Transactions on Neural Networks and Learning Systems*.

Rafiei, M. H., Khushefati, W. H., Demirboga, R., and Adeli, H. (2017). "Supervised deep restricted Boltzmann machine for estimation of concrete." *ACI Materials Journal*, 114(2), 237.

Rhee, D. M., Lombardo, F. T., and Kadowaki, J. (2021). "Semi-automated tree-fall pattern identification using image processing technique: Application to Alonsa, MB tornado." *Journal of Wind Engineering and Industrial Aerodynamics*, 208, 104399.

Roberts, D. A., Yaida, S., and Hanin, B. (2022). *The principles of deep learning theory*, Cambridge University Press Cambridge, MA, USA.

Shen, Y., Zhu, S., Yang, T., Chen, C., Pan, D., Chen, J., Xiao, L., and Du, Q. (2021). "Bdanet: Multiscale convolutional neural network with cross-directional attention for building damage assessment from satellite images." *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-14.

Shim, S. (2023). "Self-training approach for crack detection using synthesized crack images based on conditional generative adversarial network." *Computer-Aided Civil and Infrastructure Engineering*.

Spencer Jr, B. F., Hoskere, V., and Narazaki, Y. (2019). "Advances in computer vision-based civil infrastructure inspection and monitoring." *Engineering*, 5(2), 199-222.

Wu, C., Zhang, F., Xia, J., Xu, Y., Li, G., Xie, J., Du, Z., and Liu, R. (2021). "Building damage detection using U-Net with attention mechanism from pre-and post-disaster remote sensing datasets." *Remote Sensing*, 13(5), 905.

Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., and Yang, X. (2018). "Automatic pixel-level crack detection and measurement using fully convolutional network." *Computer-Aided Civil and Infrastructure Engineering*, 33(12), 1090-1109.

Ye, W., Ren, J., Zhang, A. A., and Lu, C. (2023). "Automatic pixel-level crack detection with multi-scale feature fusion for slab tracks." *Computer-Aided Civil and Infrastructure Engineering*.

Yong, G., Jeon, K., Gil, D., and Lee, G. (2023). "Prompt engineering for zero-shot and few-shot defect detection and classification using a visual-language pretrained model." *Computer-Aided Civil and Infrastructure Engineering*, 38(11), 1536-1554.

Zhang, Y., and Lin, W. (2022). "Computer-vision-based differential remeshing for updating the geometry of finite element model." *Computer-Aided Civil and Infrastructure Engineering*, 37(2), 185-203.

Zheng, Y., Gao, Y., Lu, S., and Mosalam, K. M. (2022). "Multistage semisupervised active learning framework for crack identification, segmentation, and measurement of bridges." *Computer-Aided Civil and Infrastructure Engineering*, 37(9), 1089-1108.