



Scalable approach to create annotated disaster image database supporting AI-driven damage assessment

Sun Ho Ro¹ · Jie Gong¹

Received: 17 September 2023 / Accepted: 20 April 2024
© The Author(s) 2024

Abstract

As coastal populations surge, the devastation caused by hurricanes becomes more catastrophic. Understanding the extent of the damage is essential as this knowledge helps shape our plans and decisions to reduce the effects of hurricanes. While community and property-level damage post-hurricane damage assessments are common, evaluations at the building component level, such as roofs, windows, and walls, are rarely conducted. This scarcity is attributed to the challenges inherent in automating precise object detections. Moreover, a significant disconnection exists between manual damage assessments, typically logged-in spreadsheets, and images of the damaged buildings. Extracting historical damage insights from these datasets becomes arduous without a digital linkage. This study introduces an innovative workflow anchored in state-of-the-art deep learning models to address these gaps. The methodology offers enhanced image annotation capabilities by leveraging large-scale pre-trained instance segmentation models and accurate damaged building component segmentation from transformer-based fine-tuning detection models. Coupled with a novel data repository structure, this study merges the segmentation mask of hurricane-affected components with manual damage assessment data, heralding a transformative approach to hurricane-induced building damage assessments and visualization.

Keywords Hurricane · Hurricane damage assessment · Component-level damage · Deep learning · Object detection

1 Introduction

1.1 General introduction

As coastal regions swell with increasing populations, the aftermath of hurricanes grows devastatingly apparent. Understanding and assessing hurricane damage is not just about

✉ Jie Gong
jg931@soe.rutgers.edu

Sun Ho Ro
sunhoro@scarletmail.rutgers.edu

¹ Department of Civil and Environmental Engineering, Rutgers University, Piscataway, NJ 08854, USA

directing immediate relief or aiding recovery. The broader spectrum encompasses preventive measures, urban planning, and long-term strategies to mitigate future damages, ensuring safer and more resilient infrastructure (Neumann et al. 2015; Nicholls and Small 2002; Ngo 2001; Berke et al. 2012).

1.2 Hurricane damage assessment from manual to digital via remote sensing

Post-hurricane evaluations, particularly of residential buildings, were manually conducted in the past. Field teams documented building details and associated imagery aligning with the guidelines (FEMA 2016). However, as technology progressed, so did the means of assessment. The shift from manual, time-consuming surveys to digital approaches marked a significant advancement, enabling faster responses and more detailed analyses.

Traditionally, the post-hurricane damage assessment, especially the evaluation of residential buildings, relied heavily on manual methods (Alzughabi 2018; Massarra 2012; Wengrowski 2019). Expert teams would precisely document the specifics of each building and capture relevant images. Manual damage assessment protocols were derived from the Rapid Needs Assessment (RNP) and Primary Damage Assessment (PDA) (FEMA 2016). The study and refinement of these protocols remain an active area of research, especially as data acquisition and processing methods evolve and the roles of various agencies shift (Pant 2019; Friedland 2009; Wilson et al. 2015). However, as we ventured further into the digital age, the landscape of damage assessment evolved dramatically. This evolution from labor-intensive manual surveys to innovative digital techniques represented a monumental leap. Not only did it streamline the entire process, making it more efficient, but it also enabled a deeper, more comprehensive analysis of the damage.

The rise of remote sensing technologies has transformed the world of damage assessment. Tools like aerial images (Kanistras et al. 2013; Zhong et al. 2020; Schaefer et al. 2020), sonar systems (Hayes and Gough 2009; Purser et al. 2018), light detection and ranging (LiDAR) (Zhou and Gong 2018; Gong and Maher 2014; Van Ackere et al. 2019), and satellite imagery (Gupta and Shah 2021; Kakooei and Baleghi 2017; Oludare et al. 2021) do not just enable more comprehensive data acquisition; they have paved the way for advanced methods like machine learning and deep learning to interpret this data. With these technologies, regions previously hard to access or evaluate can now be quickly assessed, providing insights that are crucial for immediate disaster response.

1.3 Component-level damage assessment

When assessing hurricane damage through remote sensing, it is vital to align the evaluation technique with the type of information required. Hurricane damage can be broadly categorized into three levels:

1. Community-level: This assesses the extent of damage across large affected areas, giving an overview of the disaster's spread.
2. Property-level: Here, the focus narrows to individual structures, identifying their state post-disaster.
3. Component-level: This drills even deeper, evaluating specific elements like roofs, windows, walls, and doors.

This level of detail is not just about identifying present damages. It offers insights into structural weaknesses, aiding in better construction practices for the future. Additionally, by integrating deep learning and image segmentation, these damages' complexity and varied nature can be more accurately identified and classified, furthering the precision of such assessments.

While assessments at the community (Gupta et al. 2019a; Weber and Kané 2020; Gupta and Shah 2021; Gupta et al. 2019b) and property levels (Zhu et al. 2021; Yeom et al. 2019; Daud et al. 2022; Lindell and Prater 2003) are commonplace, component-level assessment is less frequent (Hatzikyriakou et al. 2016; Zhou and Gong 2018; Ou et al. 2021), and the reasons for this are multifaceted. The detailed and intricate nature of component-level assessment demands a coarser analysis of specific elements, and the depth of analysis often consumes more time and resources. Furthermore, automating such detailed assessments presents its own set of challenges, especially when trying to discern diverse and often subtle types of damage specific to building components. However, it is essential to understand that this coarse assessment transcends precise damage identification. It provides a deeper understanding of structural vulnerabilities, thereby informing improved construction standards and practices for future resilience. With the integration of advanced technologies like deep learning and image segmentation, we can enhance the accuracy and precision of these component-level evaluations, capturing the nuances of diverse damage types.

1.4 Deep learning, from CNN to transformers in damage assessment

Deep learning, especially within the realm of computer vision, has emerged as a transformative tool for hurricane damage assessment. Utilizing advanced neural networks, notably Convolutional Neural Networks (CNNs) such as You Only Look Once (YOLO) (Redmon et al. 2016) and regional convolutional neural networks (Mask R-CNN) (He et al. 2017), excel at analyzing image data to identify, classify, and assess damages inflicted by hurricanes. At the community and property level, CNNs are highly effective due to their ability to recognize patterns over large spatial scales. These networks can efficiently scan wide-area satellite or aerial images and differentiate between damaged and undamaged regions or structures based on color variations, textures, and spatial patterns that indicate large-scale damage. Numerous successful studies have employed CNNs for natural disaster object detection, particularly at community and property levels, including building damage assessment (Nex et al. 2019; Valentijn et al. 2020; Bhuyan et al. 2023), land cover change detection (Khan et al. 2017; Lv et al. 2022), landslide mapping (Kikuchi et al. 2023; Gao and Ding 2022), street-level change detection (JST 2015; Lenjani et al. 2020), and roof material classification (Kim et al. 2021).

In addressing component-level damage assessment, the precision required to analyze specific structural elements—like windows, doors, roofs, and walls—presents notable challenges. Traditional CNNs, including advanced variants like Mask-RCNN (Inc 2018), can effectively detect general damage to a building. However, identifying specific, intricate damages such as a cracked windowpane or dislodged roof tiles demands a level of detail and sensitivity beyond what general detection provides. Shadows, occlusions, and the heterogeneity in building design, materials, and orientation further complicate the task, making certain damages difficult to distinguish. Moreover, the feasibility of creating a comprehensive training dataset that covers the extensive range of possible damage patterns, especially in severely damaged buildings, becomes a significant limitation. Severely damaged structures often exhibit unique and irregular damage patterns, as depicted in Fig. 1,



Fig. 1 Complex and unique damage patterns from destroyed buildings, post-Hurricane Harvey

where buildings are near collapse with leaning walls and scattered roof beams. Such variability and unpredictability in damage patterns pose a substantial challenge for traditional CNN architectures, emphasizing the practical limitations in preparing datasets that can accurately represent every possible scenario of post-disaster damage.

Image annotation at the component-level for damage assessment is significantly more challenging than at community or property-levels. This added complexity comes from the detailed attention needed for individual building parts. Annotators have to spot and label specific damages, which can appear in many subtle forms. For example, a roof might exhibit a spectrum of issues, from missing shingles to barely noticeable punctures. Furthermore, images often present overlapping or adjoining building components. For instance, an image might capture both a damaged awning and the window beneath it. Separating and annotating these overlapping components becomes a meticulous task. The angle and distance from which an image is captured can further complicate the assessment. Varying perspectives might distort or hide essential details, demanding a keen eye and a deep familiarity with architectural nuances. This mix of required precision, expertise, and the need for distinct component visibility makes component-level assessment time-consuming and intricate, emphasizing its intricacy compared to broader assessments.

Transformers (Vaswani et al. 2017) have recently demonstrated remarkable capabilities in diverse areas, including computer vision (Xie et al. 2021; Zhao et al. 2021). Unlike traditional CNNs that process data sequentially, transformers can simultaneously attend to different parts of the input data, capturing intricate relationships. When applied to computer vision, transformers, such as Vision Transformers (ViTs) (OpenAI), can capture long-range dependencies and relationships in images, offering a global perspective that the localized view of CNNs may overlook. Several recent studies have demonstrated the effectiveness of transformers in natural disaster damage assessment (Da et al. 2022; Kaur et al. 2023; Asad et al. 2023; Tounsi and Temimi 2023). However, these investigations primarily focus on community and property-level evaluations, with component-level assessment still largely unexplored.

For component-level damage assessment, utilizing fine-tuning with transformers can be particularly impactful. Fine-tuning is a technique often employed in deep learning to adapt a pre-trained model to a new but related task. Fine-tuning leverages initially trained knowledge by taking a large pre-trained model and continuing the training on a smaller, task-specific dataset (typically a smaller custom dataset). Transformers trained on large datasets capture extensive information, and by fine-tuning them on specific tasks like hurricane damage assessment, they can be tailored to recognize intricate patterns and details that are vital for accurate results, potentially outperforming models like CNNs.

1.5 Manual data and image disconnection

Over the past decades, extensive post-hurricane damage data has been meticulously compiled. Historically, these manual assessments from hurricane sites have been detailed and archived in spreadsheets, often accompanied by images of the affected buildings. A major issue is that the detailed spreadsheet data is not directly connected to its matching images. For example, to correlate spreadsheet data with its corresponding images, one must laboriously align the damage data with images using side-by-side comparisons. This task becomes even more daunting considering the unstructured and varied nature of damage patterns that each building component can exhibit. The challenge lies not only in interpreting these intricate patterns but also in the absence of platforms that adeptly merge these images with manual damage assessment data.

Thus, while technology has greatly advanced the depth and precision of modern damage assessments, a significant void remains. Despite their potential wealth of insights, these archived manual assessments are gathering dust due to the technical challenges of integration. It is imperative to not only harness modern methodologies for ongoing and future assessments but also tap into this historical data, bridging the past with the present for a more comprehensive understanding of hurricane impacts over time. Finally, integrating historical data may be the sole approach to capture diverse building details like physical address, GPS, structure type, and other information that could evolve or change over time.

1.6 Conclusion and contribution

Addressing the complexities of component-level hurricane building damage assessment and manual assessments digital disconnection, this study introduces a new workflow, leveraging state-of-the-art deep learning models for refined semi-automated analysis. Specifically, it utilizes large-scale pre-trained instance segmentation models for efficient and precise image annotation and transformer-based fine-tuning for object detection. The pre-trained instance segmentation model is adopted for precise image annotation. This capability is pivotal for assessing hurricane-caused building damage, where complex damage patterns often require detailed polygon-shaped segmentation masks. Manually creating such masks can be an arduous and time-consuming process. Then, fine-tuning allows a pre-trained model to be tailored using a specialized dataset specific to the detection task, such as damaged building components. This approach enables the model to comprehend the nuances of the targeted dataset while building upon the knowledge from its comprehensive initial training. It does not lose the foundational knowledge and capabilities accumulated during its broader-scale initial training, which is crucial for holistic understanding. Moreover, this study recommends a new natural disaster data repository structure designed to visualize segmented images of hurricane-affected building components, seamlessly integrating them with manual damage assessment data.

This study harnesses state-of-the-art deep-learning models to streamline the evaluation of component-level hurricane damages. By digitally combining them with manual damage assessment data, this study transforms how we assess and understand the impact of these natural disasters on building components. The contribution of this work lies in the utilization of sophisticated models which collectively introduce a transformative approach to damage assessment practices. The rapid and precise image labeling offered by large-scale pre-trained instance segmentation models expedites the identification of intricate

damage patterns, while transformer-network-based fine-tuning refined predictions under limited training data enhances the precision of damage evaluation. Through these innovative methods, the study enhances our understanding of the multifaceted variables impacting hurricane-induced damage and furnishes practical tools to expedite post-disaster decision-making processes. Furthermore, integrating segmented image merging with manual damage assessment is a novel concept presenting a synergistic approach that combines the precision of advanced computer vision with the reliability of human expertise. This fusion not only refines accuracy but also paves the way for more comprehensive damage analyses, solidifying it as a promising innovation in disaster management and assessment.

2 Proposed methodology

As illustrated in Fig. 2, the workflow begins with preparing training image data. This involves the collection of a diverse array of RGB images showcasing hurricane-induced building damage, each sized at 1080×810 pixels. These images form the bedrock for training the model to identify and outline various objects precisely. After data collection, the Segment Anything Model (SAM), a pre-trained instance segmentation model, is utilized for annotating the images, thereby ensuring efficient and high-precision labeling in readiness for the next phase. The model's core is powered by DETR (Detection Transformer), integrated with a ResNet-50 backbone, chosen for its robust object detection capabilities, blending the strengths of transformer models with deep residual networks. Training and validation data loaders are prepared to supply the model with annotated images, formatted in alignment with the COCO dataset standards, a prevalent format for object detection tasks. The DETR model undergoes fine-tuning to detect building components damaged by hurricanes. This process adjusts the pre-trained DETR model, initially trained on the COCO dataset, to accommodate the unique dataset and labels pertaining to hurricane damage assessment. The final step integrates the processed images with manual damage assessment data. A simple label-matching script facilitates the overlay of manual damage assessments onto the segmentation masks, achieving a comprehensive visualization. This

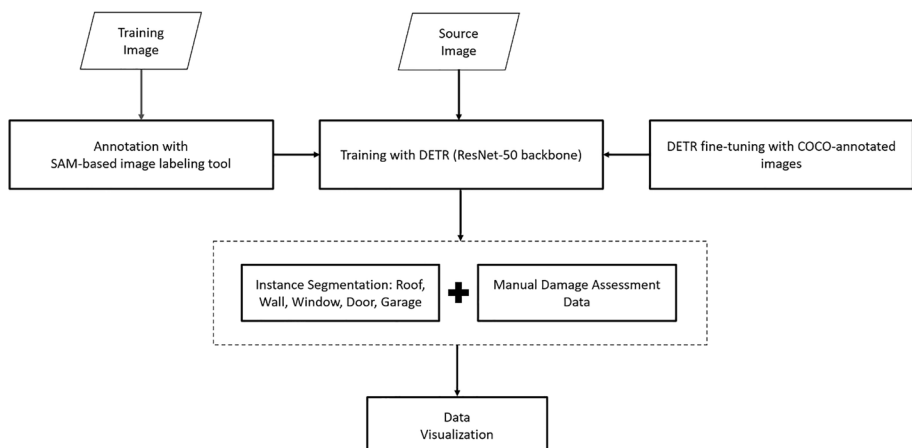


Fig. 2 Overall workflow

integration is carried out to produce a holistic visualization, with the output layer dimensions specified, ensuring a detailed representation of the damage assessment.

2.1 Image annotation with pre-trained segmentation model

Foundation models in the domain of natural language processing (NLP) have become immensely popular and transformative (Min et al. 2021). OpenAI's GPT (Generative Pre-trained Transformer) (Brown et al. 2020) stands out as one of the pioneering models in this area, leveraging vast datasets with trillions of tokens. These models are trained to predict the next word in a sentence and are distinguished by their massive scale and diverse training data. As a testament to their success in NLP, similar concepts began to emerge in other domains, notably computer vision. The field of computer vision, specifically image segmentation, involves extensive specialization. Traditionally, tasks like biomedical image analysis (Yang et al. 2023; Bloice et al. 2019; Kwon et al. 2020), photo editing (Lee et al. 2020; Zhu et al. 2020; Elharrouss et al. 2020), or autonomous driving (Huang et al. 2018; Chen et al. 2015; Song et al. 2019) required models trained for specific tasks, demanding domain expertise, specialized data collection, and lengthy training.

The Segment Anything project, inspired by the success of foundation models in NLP, is one of the most recognized pre-trained models that sought to revolutionize this domain by democratizing image segmentation (Brown et al. 2020). The segment anything model (SAM) is influenced by the success of NLP foundation models, aiming to democratize image segmentation. SAM is an automatic segmentation model that requires minimal human involvement and bypasses individual dataset training. SAM uses deep learning and has been trained on a staggering 1 billion masks across 11 million images. With a simple Python inference, users can prompt SAM by various methods, including clicking on image points or drawing bounding boxes. SAM's utility in handling unstructured data is particularly evident in the assessment of the aftermath of natural disasters like hurricanes. Post-hurricane building damage presents intricate patterns that are challenging to identify and annotate, especially given the vast amounts of visual data that must be processed swiftly for timely response and rehabilitation efforts. Manual annotation, although meticulous, is time-consuming and susceptible to human error, particularly when delineating the multifaceted damage patterns onto polygon-shaped segmentation masks. SAM, however, swiftly and accurately interprets these complex patterns, minimizing manual labeling effort.

A Python-based script utilized with LabelMe (Wada 2021) was developed for the SAM-based image labeling inspired by Roboflow (Skalski 2023). Once an image is loaded, the user can draw a rectangle around any object they are interested in. The tool then automatically masks the area inside that rectangle and emphasizes the one critical identifiable object. Finally, it displays the original photo with the highlighted area and another version focusing on just the shaded object, allowing users to see the details clearly. If necessary, the generated mask can be edited before assigning a label. This process can be repeated until the mask has been generated and labeled to all desired objects in the image (Fig. 3).

2.2 DETR with fine-tuning

Before the advent of transformers in object detection, the realm of computer vision was predominantly influenced by Convolutional Neural Network (CNN) models. Pioneering models like Mask-RCNN (Inc 2018) and YOLO (Redmon et al. 2016) were instrumental in driving advancements in this field. However, as technology evolved, recent



Fig. 3 Example of SAM-based image labeling

transformer models have begun to surpass these traditional CNNs, heralding a new era in object detection. One of the standout contributors to this shift has been the Detection Transformer (DETR) (Carion et al. 2020).

DETR was conceived in response to the limitations associated with traditional CNN-based object detection. CNN relied heavily on mechanisms such as anchor boxes and region proposals, which often added complexity and limited efficiency. With an aspiration to streamline object detection and overcome the restrictions of earlier methods, DETR was developed to integrate the transformative capabilities of transformer architectures into the world of visual data.

DETR is a model that carries transformer-based structures, typically seen in NLP, with object detection paradigms. DETR uniquely sidesteps the conventional reliance on anchor boxes and region proposals. Instead, it utilizes a fixed set of learned object queries, which are then passed through its decoder to generate predictions. The model envisions object detection as a direct set prediction challenge, obviating the need for procedures such as non-maximum suppression and streamlining post-processing.

Fine-tuning in the context of deep learning and object detection involves adapting a pre-trained neural network to cater to a specific task or domain. It involves leveraging the foundational knowledge embedded in a model—gained from training on a large dataset—and refining it further using a smaller, specialized dataset to hone its proficiency in a particular domain. In DETR, fine-tuning begins with a model that's already been trained on expansive datasets such as COCO (Lin et al. 2014). This model, equipped with a broad understanding of various object features, is subjected to further training on a smaller targeted dataset, the SAM-labeled training data of building components in this context. The subsequent training narrows down the model's focus, adjusting its internal parameters and learned representations to specialize in the intricacies of the specific task at hand.

DETR, by design, brings to the table a unique set-based global loss and a transformer encoder-decoder architecture. This combination allows it to holistically reason about the relations of objects within the broader image context, making it adept at discerning intricate details. When this ability is paired with fine-tuning, DETR becomes highly specialized in identifying even the most unstructured and intricate patterns, such as those seen in post-hurricane building damage. By utilizing custom annotated building

component datasets, the fine-tuning process meticulously directs the model's attention, enabling it to accurately discern the disordered and complex aftermaths of hurricanes.

3 Data sets

Hurricane Harvey, a Category 4 storm with winds reaching 130 mph and a 12-foot storm surge, struck near Port Aransas, TX, on August 25, 2017, resulting in up to \$1 billion in damages in the area. The dataset for this study was derived from ground-level digital images provided by teams from Rutgers University, Princeton University, and the University of Texas at Austin (Magazine 2018). A manual damage assessment followed, which included manually geo-coding each image to its corresponding physical address for accurate damage rating. This process documented general building information and detailed damage to components such as doors, windows, walls, roofs, and garages. A total of 1,220 images, 305 images from 62 residential buildings in Bayside, TX, and 915 images from 225 residential buildings in Port Aransas, TX, were compiled. The number of images per building varied, with at least one and up to seven images per building to ensure comprehensive coverage. Bayside and Port Aransas, both affected by Hurricane Harvey, are expected to show similar damage patterns, aside from differences in storm surge impact.

During the annotation of training images, label classes were categorized into two primary damage groups: damaged and undamaged. This distinction was imperative for effectively classifying hurricane-affected building components. While initially considering a more granular approach with four damage categories (affected, minor, major, and destroyed), it became evident that this led to a significant drop in mean Average Precision during the training. This decline was attributed to the segmentation of an already limited training dataset, coupled with challenges in differentiating between minor and major damage. Consequently, the finalized classes for annotation are *Roof-Dmg*, *Roof-NoDmg*, *Wall-Dmg*, *Wall-NoDmg*, *Window-Dmg*, *Window-NoDmg*, *Door-Dmg*, *Door-NoDmg*, *Garage-Dmg*, and *Garage-NoDmg*.

4 Results discussion

4.1 Detection result

The source images, prior to processing, were grouped into four FEMA-defined damage categories to facilitate a nuanced evaluation of the model's detection capabilities (FEMA 2016). This categorization was employed to stratify the images according to the overall damage extent for comparison purposes, not as a direct part of the training dataset preprocessing. The categories are as follows: (1) Affected: Homes where damage is predominantly cosmetic; (2) Minor Damage: Homes with repairable non-structural issues; (3) Major Damage: Homes with structural impairments or other significant issues necessitating extensive repairs; and (4) Destroyed: Homes deemed a total loss. The test results depicted in Fig. 4 reflect the model's varying degrees of effectiveness in detecting component-level damage.



Fig. 4 Examples of damaged building component segmentation results

4.2 Performance metrics

First, the intersection over union (IOU) metric, a coefficient of similarity for two sets of data, was employed (Naturelles 1864). The IOU metric calculates the degree of overlap between a predicted detection (B) and the ground truth (A) and divides it by the area of their union. The performance of the object detection method is based on AP@ IoU=0.50, which refers to AP for $IoU \geq 0.5$, for each object category, shown on the below table computed from COCO Evaluator.

Table 1 shows the evaluation result for the DETR training with 50 epochs, gradient clipping value of 0.1, accumulated gradient batches of 8, logging steps of 5, and 41.5 M parameters, including 41.3 M trainable parameters and 222 K non-tradable parameters. Total training took about 1 h, and the resulting model parameter had a size of 108 MB. Based on the results obtained, an AP50 score of 0.621 indicates a fine-performing object detection model. The dimensions of annotated objects can vary significantly based on the severity and extent of damage. Generally, building components such as windows, doors, and garages can be categorized into small or medium-sized objects, whereas roofs and walls typically fall under the large object category. The average precision (AP) scores for small, medium, and large objects were found to be 0.475, 0.582,

Table 1 Evaluation of DETR training

Average precision	(AP.) @[IoU = 0.50:0.95	Area = all	maxDets = 100	0.432
Average precision	(AP.) @[IoU = 0.50	Area = all	maxDets = 100	0.621
Average precision	(AP.) @[IoU = 0.75	Area = all	maxDets = 100	0.578
Average precision	(AP.) @[IoU = 0.50:0.95	Area = small	maxDets = 100	0.475
Average precision	(AP.) @[IoU = 0.50:0.95	Area = medium	maxDets = 100	0.582
Average precision	(AP.) @[IoU = 0.50:0.95	Area = large	maxDets = 100	0.620

and 0.620, respectively. These scores reflect a model that performs moderately well, with performance improving as the size of the target object increases. Additionally, it is important to consider the quality and resolution of the images. Many of the smaller objects displayed relatively low resolution because the images were captured from a distance to prioritize safety during the reconnaissance of the damaged houses.

Mask R-CNN (He et al. 2017), equipped with a ResNet-101-FPN backbone, was trained using the identical dataset to facilitate a comprehensive comparison. Mask R-CNN enhances the Faster R-CNN framework by introducing a segmentation mask prediction branch for each Region of Interest (RoI) alongside the conventional classification and bounding box regression branches. The ResNet-101-FPN backbone, notable for its complexity, provides a detailed multi-scale feature representation, optimizing the model for detecting objects of varying sizes. The performance comparison result is shown in Table 2.

When evaluating Average Precision (AP) across IoU thresholds from 0.5 to 0.95, DETR fine-tuning with the COCO dataset attained an AP of 0.432, surpassing Mask R-CNN's performance, which stood at 0.389. This discrepancy suggests DETR's superior capability in consistently detecting objects across a spectrum of IoU thresholds, possibly due to its enhanced ability to understand the overall context of damaged building components. At the specific IoU threshold of 0.5 (AP@0.5), DETR and Mask R-CNN demonstrated nearly equivalent performance, with DETR marginally leading (0.621 vs. 0.617). This parity indicates that both frameworks effectively identify damaged building components under less rigorous overlap criteria. These findings underscore DETR's efficacy in object detection tasks within post-disaster damage assessment. DETR's edge in performance across various IoU thresholds can be ascribed to its innovative detection approach, combining fine-tuning and a distinct model structure, which renders it more adept at handling complex detection scenarios. Conversely, Mask R-CNN's slightly diminished performance, especially at higher IoU thresholds, could stem from multiple challenges, including the difficulty of segmenting extensively damaged components, the diversity in types of damage, and the intricacy of disaster-impacted scenes.

Table 2 Performance comparison between DETR and Mask-RCNN

Framework	Backbone	Dataset	AP _[0.5,0.95]	AP _{0.5}
DETR	ResNet-50	Hurricane Harvey affected buildings (1,220 images)	0.432	0.621
Mask R-CNN	ResNet-101-FPN		0.389	0.617

Table 3 Undamaged component precision-recall analysis

	Roof	Wall	Window	Door	Garage
TP	412	626	920	302	157
FP	51	42	334	62	52
FN	69	72	317	58	78
Precision	0.86	0.90	0.74	0.84	0.67
Recall	0.89	0.94	0.73	0.83	0.75

Table 4 Damaged component precision-recall analysis

	Roof	Wall	Window	Door	Garage
TP	227	347	633	152	67
FP	87	48	385	45	35
FN	196	31	216	73	84
Precision	0.54	0.92	0.75	0.68	0.44
Recall	0.72	0.88	0.62	0.77	0.66

A comprehensive evaluation using traditional performance metrics, precision and recall was used to better understand the automated component-level damage assessment performance. Tables 3 and 4 below show the precision-recall analysis for undamaged and damaged component detection. A notable difference in the performance metrics between undamaged and damaged roofs exists. Damaged roofs registered a precision of 0.54, significantly lower by 32% compared to the undamaged counterpart of 0.86. The ground-level perspective from which many images were captured restricted visibility to only parts of the roof, further complicating the differentiation between roofs and walls, as evinced by the elevated False Negative (FN) values. In contrast, walls were consistently well-represented in images due to the ground-level perspective. This facilitated the quality and quantity wall annotation and the model's performance, with a precision of 0.90 and a recall of 0.94, while the damaged ones hold similarly impressive values of 0.92 and 0.88, respectively. This data indicates that the model robustly detects and differentiates walls irrespective of their damage status. For windows, doors, and garages, moderate precision values were noted. However, a considerable count of False Positives (FP) and FNs were recorded, attributed largely to their shared rectangular morphology. Particularly, windows and doors exhibited elevated FP and FN rates. Enhanced feature discrimination or a diversified training dataset could mitigate this ambiguity. Lastly, damaged garages demonstrated suboptimal precision and recall values of 0.44 and 0.66, respectively. A primary contributor to this underperformance was the prevalence of broken or absent garage doors, leading to pronounced indoor shadows. Such shadows have been consistently recognized as detrimental factors in object detection algorithms. Addressing shadow effects or integrating illumination normalization might bolster detection accuracy in such scenarios.

4.3 Challenges in automated component-level damage assessment

Building damage patterns can be intricate and multifaceted. Beyond these complexities, several other factors play a crucial role in ensuring a precise and comprehensive automated

component-level damage assessment. The following delve into some of these pivotal considerations.

- **Post-disaster alteration:** A building in this state retains noticeable remnants of its original form, making it optimal for data acquisition and processing. However, recovery efforts typically commence within 72 h following a natural disaster, altering the initial damage. Consequently, subsequent analyses may not directly correlate with the disaster's cause, diminishing the effectiveness of understanding the damage mechanism. Often, debris piles (Fig. 5a) emerge as a result of accumulating damaged materials. These should not serve as primary sources for damage assessment unless the objective is to identify and quantify the overall debris.
- **Data bias:** Given that the vast majority of the training data originates from residential buildings, the detection process struggles with structures of a distinct type. For instance, Fig. 5b displays a damaged boat rack, which fundamentally differs from the training data. Consequently, damage detection on such datasets warrants re-evaluation. While the training data predominantly features exposed wooden structural components, a boat rack mainly consists of steel columns and beams.
- **Obscured damage:** Post-disaster structures are often shielded with blue tarpaulins (Fig. 5c) to prevent further damage. Such measures obscure the extent of building damage, rendering them unsuitable for detailed component-level damage assessment.
- **Classification blind spot:** Some structures display tilted or misaligned columns (Fig. 5d) due to the lateral forces exerted by hurricane events, greatly compromising their integrity and stability. However, some of these buildings may be mistakenly classified as undamaged, given the absence of labeled classes for tilted columns. Separate training

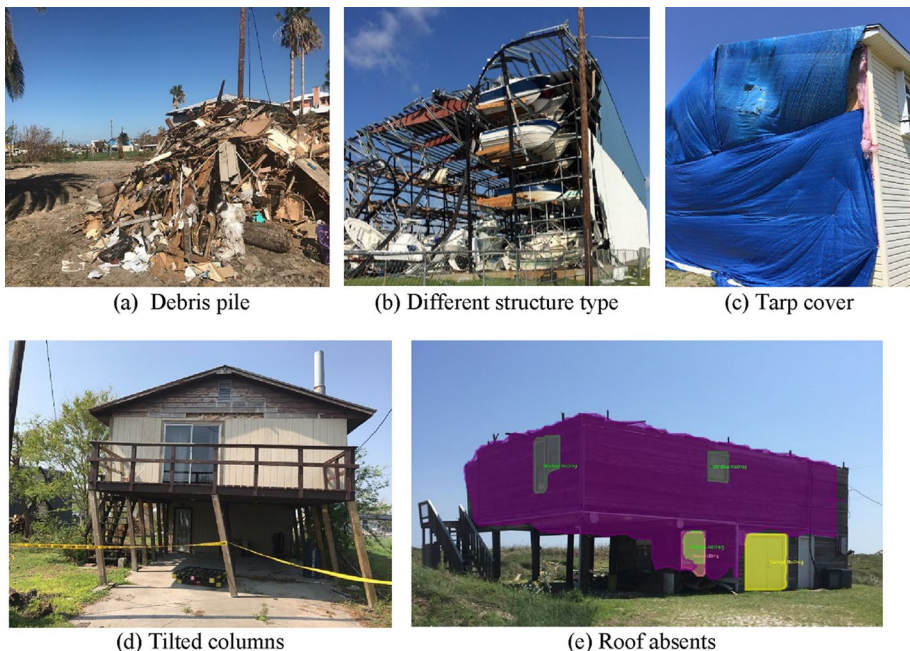


Fig. 5 Key challenges in automated damage assessment

data should be curated to account for such damages and enable accurate reclassification.

- **Component absence challenge:** The algorithm might identify buildings with walls and windows that seem structurally intact. Yet, a missing roof (Fig. 5e), evident from top-down analysis, signals considerable damage, emphasizing the necessity of a holistic assessment method. The complete absence of a component poses significant challenges in damage evaluation. This issue is among the most daunting in automated damage assessment, as specialized models are needed to distinguish missing elements—easily spotted by human observers but potentially overlooked by detection algorithms.

5 Integration of segmented component and manual damage assessment data

Many esteemed natural disaster data platforms, including repositories of the National Oceanic and Atmospheric Administration (NOAA) and the United States Geological Survey (USGS), primarily serve as repositories for raw image data without extensive data curation. While they offer significant storage capacities, their data structures have a notable absence of uniformity and compatibility. A few sophisticated natural disaster repositories (Gurram et al. 2017; Park et al. 2019) employ deep learning techniques for data curation, like object detection and visualization. However, none currently bridges the gap between extensive manual damage assessment data and the corresponding post-disaster building images archived for decades. This study proposes a Hurricane Image Analysis Viewer (HIAV) to overcome the limitations of current natural disaster repositories. This prototype seamlessly integrates segmented building component images with manual damage assessment outcomes. Developed using HTML, JavaScript, and PHP for backend support, HIAV's central feature is its digital association between segmented building elements and damage assessment data. Within HIAV, the primary interface comprises a data filter (Fig. 6). Sections A (General building information) and B (Hurricane building damage) facilitate the extraction of manual damage assessment findings. Users can engage with each category through text or a selection mechanism. Upon selecting the desired data, hitting the search button refines the image list to match the criteria.

After selecting the necessary filters, the user can use the “Search” button to update the list of images based on the selected filters. A “Reset” button also clears all filters and resets

A. General Building Information

Address <input type="text"/>	Number Of Stories 2	Attached Garage No	Building Use Single Family
Roof Shape select	Roof Cover Format select	Roof Cover Material select	Roof Framing select
Structural wall select	Wall covercladding select	Structure Elevated No	Shuttersstorm panels Yes
Hurricane clips Yes	Damage cause Surge,Wind		

B. Hurricane Building Damage

Total Damage Rating Moderate	Structure Usable Yes	Roof cover dmg select	Roof sheathing dmg select
Roof framing dmg select	Wall cover dmg select	Wall sheathing dmg select	Wall framingstructure dmg select
Windows dmg 0.25	Garage door dmg Yes	Patio doors dmg Yes	Entry doors dmg Yes

Fig. 6 Data filter tool by sections

the image list to its original state. The image list displays each image along with its name, and each image has a checkbox next to it. The user can select one or multiple images by checking the corresponding checkboxes. Once the desired images are selected, the user can click the “Download” button to download the selected images. A “Select All” button allows the user to select all images in the list with a single click. This can be useful if the user wants to download all images simultaneously. Upon selecting the desired filters, users can click the “Search” button to refine the image list accordingly. A “Reset” button is available to clear all filters, reverting the list to its default state. Images are presented with accompanying names and checkboxes for selection. Users can choose individual or multiple images by checking the associated checkboxes. To download chosen images, hit the “Download” button. For convenience, a “Select All” button is provided, allowing users to select and download the entire image list in one go. Following the image selection, HIAV transitions to the image visualization phase. This section showcases the original image and its version with component segmentation results (Figs. 7 and 8).

6 Conclusions, limitations, and future work

This research delved into refining the assessment of hurricane-caused building damage, introducing an advanced workflow that bridges the gap between segmented images of damaged components and their corresponding manual damage assessments. An in-depth performance evaluation was conducted by implementing the transformer-network-based fine-tuning object detection, trained meticulously to understand the intricacies of post-hurricane damages. The results highlighted the model’s capabilities, revealing that it excelled in detecting larger components like walls while encountering challenges with smaller or more ambiguous components like windows and doors. External challenges, post-disaster alternation, data bias, obscured damage, classification blind spot, and component absence challenges notably influenced the model’s efficacy. HIAV was proposed, offering a comprehensive platform to integrate segmented images seamlessly with manual damage assessment outcomes. It is evident that while the proposed methodology has made significant strides in automating building component-level damage assessments post-hurricanes, there remains scope for further refinement. This study, however, undoubtedly lays the foundation for future endeavors to enhance our understanding and response to the aftermath of hurricanes.

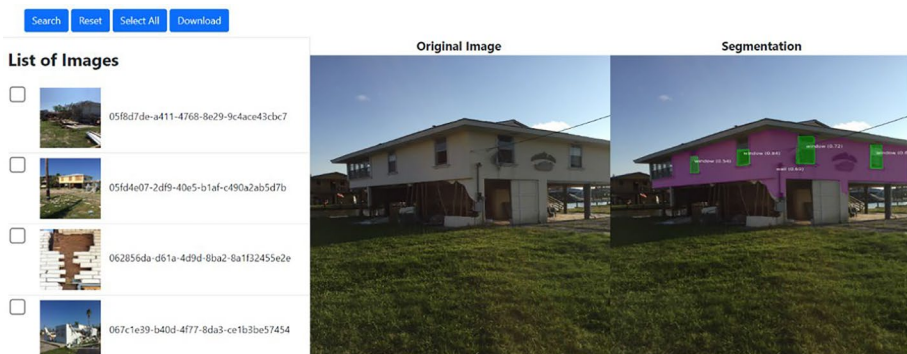


Fig. 7 Image list and visualization

Building Info	Building Damage
Latitude: 28.16525308	Total damage rating: 2=Moderate
Longitude: -97.01632904	Structure usable: no
Gps altitude: 3.653815926	Roof cover dmg: 0.25
Date of survey: 9/16/2017	Roof sheathing dmg: 0.25
Street address full: 199 Sailfish Dr Rockport Aransas TX 78382 US	Roof framing dmg: n/a
Number of stories: 1	Wall cover dmg: n/a
Age yrs: n/a	Wall sheathing dmg: n/a
Attached garage: no	Wall framingstructure dmg: n/a
Building use: Single Family	Windows dmg: 0.25
Roof shape: Hip	Garage door dmg: n/a
Roof cover format: Shingles	Patio doors dmg: no
Roof cover material: Asphalt	Entry doors dmg: no
Roof framing: Wood	

Fig. 8 Example of damage data of the selected image

One potential limitation in accurately assessing hurricane-induced building damage could be data bias. Despite the considerable size of the training dataset, it might not fully encompass the diverse range of damage types and building characteristics, highlighting the need for additional data to improve the model's comprehensiveness and precision. This necessity becomes even more pronounced when considering the variability in damage sources characteristic of different hurricanes. For example, the devastation from Hurricane Sandy in 2012 was largely due to storm surge, whereas wind was the primary factor for Hurricane Harvey in 2017, and Hurricane Michael in 2018 showcased the combined forces of wind and storm surge. To ensure the model's effectiveness across various scenarios, the dataset should include a broad spectrum of damage instances. Moreover, the timing of image acquisition is critical; preferably securing images within 72 h post-hurricane is ideal for capturing the initial damages before any recovery efforts alter the scene. This approach guarantees that the training data accurately reflects the direct consequences of hurricanes, which is essential for a precise assessment of their impact.

Reflecting on the scope for further advancement and acknowledging the current constraints of the proposed approach, the subsequent key areas are identified for future work:

- Incorporate a more diverse and comprehensive dataset that includes a wider variety of building types and damage conditions to reduce bias and improve the model's robustness.
- Move beyond binary classification by introducing a multi-tiered damage severity scale, which could provide a more detailed and accurate damage assessment.

- Developing an application for real-time damage assessment, to help emergency responders efficiently allocate resources and respond swiftly in the aftermath of a disaster.
- Enhancing existing post-disaster data repositories by integrating our workflow and data viewer, enabling more effective data analysis and interpretation to aid recovery and planning.
- Building a larger, more inclusive database that captures a wide array of natural disasters, aiming to improve the development of models capable of assessing damage across different disaster types for a global response initiative.

Funding This research project was sponsored by the US National Science Foundation under award 2103754 and by FEMA under HMGP DR4488. The funding and support from the US National Science Foundation and FEMA are appreciated.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alzughaihi AA (2018) Post-disaster structural health assessment system using personal mobile-phones. In: UC Irvine electronic theses and dissertations
- Asad MH, Asim MM, Awan MN, Yousaf MH (2023) Natural disaster damage assessment using semantic segmentation of UAV imagery. In: 2023 International conference on robotics and automation in industry (ICRAI). IEEE, pp 1–7
- Berke P, Smith G, Lyles W (2012) Planning for resiliency: evaluation of state hazard mitigation plans under the disaster mitigation act. *Nat Hazard Rev* 13:139–149
- Bhuyan K, Van Westen C, Wang J, Meena SR (2023) Mapping and characterising buildings for flood exposure analysis using open-source data and artificial intelligence. *Nat Hazards* 119(2):805–835
- Bloice MD, Roth PM, Holzinger A (2019) Biomedical image augmentation using Augmentor. *Bioinformatics* 35:4522–4524
- Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A (2020) Language models are few-shot learners. *Adv Neural Inf Process Syst* 33:1877–1901
- Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S (2020) End-to-end object detection with transformers. In: Computer vision—ECCV 2020: 16th European conference, Glasgow, UK, August 23–28, 2020, proceedings, Part I 16. Springer, pp 213–229
- Chen C, Seff A, Kornhauser A, Xiao J (2015) Deepdriving: Learning affordance for direct perception in autonomous driving. In: Proceedings of the IEEE international conference on computer vision, pp 2722–2730
- Da Y, Ji Z, Zhou Y (2022) Building damage assessment based on Siamese hierarchical transformer framework. *Mathematics* 10:1898
- Daud SM, Mohd S, Yusof MYPM, Heo CC, Khoo LS, Singh MKC, Mahmood MS, Nawawi H (2022) Applications of drone in disaster management: a scoping review. *Sci Justice* 62:30–42

- Elharrouss O, Almaadeed N, Al-Maadeed S, Akbari Y (2020) Image in painting: a review. *Neural Process Lett* 51:2007–2028
- FEMA (2016) Damage assessment operations manual: a guide to assessing damage and impact, United States of America
- FEMA (2019) Post-disaster building safety evaluation guidance. FEMA P-2055, November 2019, United States of America
- Friedland CJ (2009) Residential building damage from hurricane storm surge: proposed methodologies to describe, assess and model building damage (Louisiana State University and Agricultural and Mechanical College)
- Gao Z, Ding M (2022) Application of convolutional neural network fused with machine learning modeling framework for geospatial comparative analysis of landslide susceptibility. *Nat Hazards* 113:833–858
- Gong J, Maher A (2014) Use of mobile lidar data to assess hurricane damage and visualize community vulnerability. *Transp Res Rec* 2459:119–126
- Gupta R, Goodman B, Patel N, Hosfelt R, Sajeev S, Heim E, Doshi J, Lucas K, Choset H, Gaston M (2019) Creating xBD: a dataset for assessing building damage from satellite imagery. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp 10–17
- Gupta R, Goodman B, Patel N, Hosfelt R, Sajeev S, Heim E, Doshi J, Lucas K, Choset H, Gaston M (2019) xbd: A dataset for assessing building damage from satellite imagery. *arXiv preprint. arXiv: 1911.09296*
- Gupta R, Mubarak S (2021) Rescuenet: joint building segmentation and damage assessment from satellite imagery. In: *2020 25th international conference on pattern recognition (ICPR)*. IEEE, pp 4405–4411
- Gurram H, Subramanian C, Pinelli JP, Basu R (2017) Processing data from hurricane matthew in DesignSafe. ci. In: *Proceedings of the 2017 Americas conference on wind engineering*, pp 21–24
- Hatzikyriakou A, Lin N, Gong J, Xian S, Xuan Hu, Kennedy A (2016) Component-based vulnerability analysis for residential structures subjected to storm surge impact from Hurricane Sandy. *Nat Hazard Rev* 17:05015005
- Hayes MP, Gough PT (2009) Synthetic aperture sonar: a review of current status. *IEEE J Oceanic Eng* 34:207–224
- He K, Gkioxari G, Dollár P, Girshick R (2017) Mask r-cnn. In: *Proceedings of the IEEE international conference on computer vision*, pp 2961–2969
- Huang X, Cheng X, Geng Q, Cao B, Zhou D, Wang P, Lin Y, Yang R (2018) The apolloscape dataset for autonomous driving. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp 954–960.
- Inc, Matterport (2018) Mask_RCNN, GitHub
- JST, CREST (2015) Change detection from a street image pair using cnn features and superpixel segmentation
- Kakooei M, Baleghi Y (2017) Fusion of satellite, aircraft, and UAV data for automatic disaster damage assessment. *Int J Remote Sens* 38:2511–2534
- Kanistras K, Martins G, Rutherford MJ, Valavanis KP (2013) A survey of unmanned aerial vehicles (UAVs) for traffic monitoring. In: *2013 international conference on unmanned aircraft systems (ICUAS)*. IEEE, pp 221–234
- Kaur N, Lee CC, Mostafavi A, Mahdavi-Amiri A (2023) Large-scale building damage assessment using a novel hierarchical transformer architecture on satellite images, *Computer-Aided Civil and Infrastructure Engineering*
- Khan SH, He X, Porikli F, Bennamoun M (2017) Forest change detection in incomplete satellite images with deep neural networks. *IEEE Trans Geosci Remote Sens* 55:5407–5423
- Kikuchi T, Sakita K, Nishiyama S, Takahashi K (2023) Landslide susceptibility mapping using automatically constructed CNN architectures with pre-slide topographic DEM of deep-seated catastrophic landslides caused by Typhoon Talas. *Nat Hazards* 117:339–364
- Kim J, Bae H, Kang H, Lee SG (2021) CNN algorithm for roof detection and material classification in satellite images. *Electronics* 10:1592
- Kwon Y, Won J-H, Kim BJ, Paik MC (2020) Uncertainty quantification using Bayesian neural networks in classification: application to biomedical image segmentation. *Comput Stat Data Anal* 142:106816
- Lee CH, Liu Z, Wu L, Luo P (2020) Maskgan: towards diverse and interactive facial image manipulation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 5549–5558
- Lenjani A, Yeum CM, Dyke S, Bilonis I (2020) Automated building image extraction from 360 panoramas for post-disaster evaluation. *Computer-Aided Civil Infrastruct Eng* 35:241–257

- Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL (2014) Microsoft coco: common objects in context. In: Computer vision—ECCV 2014: 13th European conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13. Springer, pp 740–755
- Lindell MK, Prater CS (2003) Assessing community impacts of natural disasters. *Nat Hazard Rev* 4:176–185
- Lv Z, Huang H, Li X, Zhao M, Benediktsson JA, Sun W, Falco N (2022) Land cover change detection with heterogeneous remote sensing images: review, progress, and perspective. In: Proceedings of the IEEE Magazine, Rutgers (2018) Damage control. <https://ucmweb.rutgers.edu/magazine/1419archive/insights/damage-control.html>
- Massarra CC (2012) Hurricane damage assessment process for residential buildings. In: LSU digital commons
- Min B, Ross H, Sulem E, Veyseh AP, Nguyen TH, Sainz O, Agirre E, Heintz I, Roth D (2021) Recent advances in natural language processing via large pre-trained language models: a survey. *ACM Comput Surv* 56:1–40
- Naturelles, Société Vaudoise des Sciences (1864) Bulletin de la Société vaudoise des sciences naturelles (Librairie F. Rouge and Cie)
- Neumann B, Vafeidis AT, Zimmermann J, Nicholls RJ (2015) Future coastal population growth and exposure to sea-level rise and coastal flooding—a global assessment. *PLoS ONE* 10:e0118571
- Nex F, Duarte D, Tonolo FG, Kerle N (2019) Structural building damage detection with deep learning: assessment of a state-of-the-art CNN in operational conditions. *Remote Sensing* 11:2765
- Ngo EB (2001) When disasters and age collide: reviewing vulnerability of the elderly. *Nat Hazard Rev* 2:80–89
- Nicholls RJ, Small C (2002) Improved estimates of coastal population and exposure to hazards released. *EOS Trans Am Geophys Union* 83:301–305
- Oludare V, Kezebou L, Panetta K, Agaian S (2021) Semi-supervised learning for improved post-disaster damage assessment from satellite imagery. In: Multimodal image exploitation and learning 2021. SPIE, pp 172–182
- OpenAI. clip-vit-base-patch32. <https://huggingface.co/openai/clip-vit-base-patch32>
- Ou HN, Ro SH, Gong J, Zhu Z (2021) Building an annotated damage image database to support AI-assisted hurricane impact analysis. In: 2021 IEEE international conference on imaging systems and techniques (IST). IEEE, pp 1–6
- Pant S (2019) Climate change impact on US hurricane risk to residential buildings
- Park J, Yeum CM, Choi J, Liu X (2019) Automated Image classification for post-earthquake reconnaissance images. *J Comput Vision Imaging Syst* 5:1–1
- Purser A, Marcon Y, Dreutter S, Hoge U, Sablotny B, Hehemann L, Lemburg J, Dorschel B, Biebow H, Boetius A (2018) Ocean floor observation and bathymetry system (OFOBS): a new towed camera/sonar system for deep-sea habitat surveys. *IEEE J Oceanic Eng* 44:87–99
- Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
- Resarch, Facebook. DETR COCO evaluator. <https://github.com/NielsRogge/coco-eval/tree/main>
- Schaefer M, Teeuw R, Day S, Zekkos D, Weber P, Meredith T, Van Westen CJ (2020) Low-cost UAV surveys of hurricane damage in Dominica: automated processing with co-registration of pre-hurricane imagery for change analysis. *Nat Hazards* 101:755–784
- Skalski P (2023) How to use the segment anything model (SAM)
- Song X, Wang P, Zhou D, Zhu R, Guan C, Dai Y, Su H, Li H, Yang R (2019) Apollo3d: A large 3d car instance understanding benchmark for autonomous driving. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5452–5462
- Tounsi A, Temimi M (2023) A systematic review of natural language processing applications for hydro-meteorological hazards assessment. *Nat Hazards* 116:2819–2870
- Valentijn T, Margutti J, van den Homberg M, Laaksonen J (2020) Multi-hazard and spatial transferability of a cnn for automated building damage assessment. *Remote Sens* 12:2839
- Van Ackere S, Verbeurgt J, De Sloover L, De Wulf A, Van de Weghe N, De Maeyer P (2019) Extracting dimensions and localisations of doors, windows, and door thresholds out of mobile Lidar data using object detection to estimate the impact of floods. In: Gi4DM 2019: geoinformation for disaster management. International Society for Photogrammetry and Remote Sensing (ISPRS), pp 429–436
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. *Adv Neural Inf Process Syst*, 30
- Wada K. 2021. 'labelme', GitHub. <https://github.com/labelmeai/labelme>

- Weber E, Kané H (2020) Building disaster damage assessment in satellite imagery with multi-temporal fusion. arXiv preprint. [arXiv:2004.05525](https://arxiv.org/abs/2004.05525)
- Wengrowski S (2019) Comprehensive damage assessment and analysis of damage mechanisms from hurricane harvey. Rutgers, the State University of New Jersey
- Wilson R, Wood N, Kong L, Shulters M, Richards K, Dunbar P, Tamura G, Young E (2015) A protocol for coordinating post-tsunami field reconnaissance efforts in the USA. *Nat Hazard* 75:2153–2165
- Xie Y, Zhang J, Shen C, Xia Y (2021) Cotr: efficiently bridging cnn and transformer for 3d medical image segmentation. In: Medical image computing and computer assisted intervention–MICCAI 2021: 24th international conference, Strasbourg, France, September 27–October 1, 2021, proceedings, Part III 24. Springer, pp 171–180
- Yang J, Shi R, Wei D, Liu Z, Zhao L, Ke B, Pfister H, Ni B (2023) MedMNIST v2-A large-scale lightweight benchmark for 2D and 3D biomedical image classification. *Scientific Data* 10:41
- Yeom J, Han Y, Chang A, Jung J (2019) Hurricane building damage assessment using post-disaster UAV data. In: *IGARSS 2019–2019 IEEE international geoscience and remote sensing symposium*. IEEE, pp 9867–9870
- Zhao Y, Wang G, Tang C, Luo C, Zeng W, Zha ZJ (2021) A battle of network structures: an empirical study of cnn, transformer, and mlp. arXiv preprint. [arXiv:2108.13002](https://arxiv.org/abs/2108.13002)
- Zhong Y, Xin Hu, Luo C, Wang X, Zhao Ji, Zhang L (2020) WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sens Environ* 250:112012
- Zhou Z, Gong J (2018) Automated analysis of mobile LiDAR data for component-level damage assessment of building structures during large coastal storm events. *Computer-Aided Civil Infrastruct Eng* 33:373–392
- Zhu J, Shen Y, Zhao D, Zhou B (2020) In-domain gan inversion for real image editing. In: European conference on computer vision. Springer, pp 592–608
- Zhu X, Liang J, Hauptmann A (2021) Msnet: a multilevel instance segmentation network for natural disaster damage assessment in aerial videos. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp 2023–2032

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.