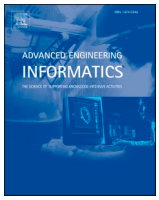
Advanced Engineering Informatics 60 (2024) 102427

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei



A Machine learning approach for Post-Disaster data curation

Sun Ho Ro, Yitong Li, Jie Gong *

Dept. of Civil and Environmental Engineering, Rutgers University, Piscataway, NJ 08854, United States

A R T I C L E I N F O

Keywords:
Natural disaster
Machine learning
Data curation
Image search

A B S T R A C T

Image data collected after natural disasters play an important role in the forensics of structure failures. However, curating and managing large amounts of post-disaster imagery data is challenging. In most cases, data users still have to spend much effort to find and sort images from the massive amounts of images archived for past decades in order to study specific types of disasters. This paper proposes a new machine learning based approach for automating the labeling and classification of large volumes of post-natural disaster image data to address this issue. More specifically, the proposed method couples pre-trained computer vision models and a natural language processing model with an ontology tailed to natural disasters to facilitate the search and query of specific types of image data. The resulting process returns each image with five primary labels and similarity scores, representing its content based on the developed word-embedding model. Validation and accuracy assessment of the proposed methodology was conducted with ground-level residential building panoramic images from Hurricane Harvey. The computed primary labels showed a minimum average difference of 13.32% when compared to manually assigned labels. This versatile and adaptable solution offers a practical and valuable solution for automating image labeling and classification tasks, with the potential to be applied to various image classifications and used in different fields and industries. The flexibility of the method means that it can be updated and improved to meet the evolving needs of various domains, making it a valuable asset for future research and development.

1. Introduction

How the post-hurricane damage data is collected, assessed, and archived builds the foundation for developing community resilience, codes, and engineering design. Precise damage assessment is paramount for all facets of disaster management [1]. Over the past decades, a notable transition has occurred from traditional manual damage assessment techniques [2–4] to the adoption of remote sensing technologies such as Unmanned Aerial Vehicles (UAV) [5–7], Light Detection and Ranging (LiDAR) [8–10], and satellite imagery [11–13] for gathering hurricane damage data. Although these technologies rapidly capture extensive affected areas, their accessibility remains limited because few researchers and governmental entities possess the capacity to deploy such broad-scope remote sensing equipment. The significant accumulation of archived disaster data, coupled with high demand, has led to the emergence of numerous disaster data repositories, project hubs, and individual websites, heralding a “Big Data” era in natural hazard research. Information-rich hurricane reconnaissance mission data is now readily available from renowned repositories such as Structural Extreme Events Reconnaissance (StEER) Data Depot [14], Automated Reconnaissance Image Organizer (ARIO) [15], United States Geological Survey, and National Oceanic and Atmospheric Administration. Although these data exist in abundance, they do not provide meaningful information before a complex processing effort, which brings Data Rich Information Poor (DRIP) situation.

The processing and extraction of useful information from complex datasets present a significant challenge for many researchers. Often, these datasets are stored in large, unorganized blocks without proper data provenance, making it difficult for researchers to locate specific data. For instance, in repositories like DesignSafe, post-hurricane UAV images are typically stored without any systematic ordering or classification. The metadata associated with these images, primarily derived from UAV data collection, offers limited informational value. This limitation creates road blocks for efficiently using these archived disaster images in various civil engineering tasks. To illustrate, when creating a fragility curve to characterize the performance of a specific type of building or building components, such as windows, during windstorm events, structural engineers often face the arduous task of searching for

* Corresponding author.
E-mail addresses: sunhoro@scarletmail.rutgers.edu (S. Ho Ro), yl2035@soe.rutgers.edu (Y. Li), jg931@soe.rutgers.edu (J. Gong).

https://doi.org/10.1016/j.aei.2024.102427
Received 10 September 2023; Received in revised form 19 January 2024; Accepted 16 February 2024
Available online 27 February 2024

images depicting these specific types of damages within a vast disaster data repository. The additional challenge of locating these images from multiple historical storm events further complicates the process. There is a need for formal methods to automate image/information search to support these engineering applications.

The traditional approach to large-scale image classification began with Content-Based Image Retrieval (CBIR) during the 90 s [16,17]. CBIR primarily targeted the extraction of visual features, like color, texture, and shape, from images. However, its emphasis on low-level visual features frequently resulted in mismatches. The advent of early semantic image search aimed to address these shortcomings by focusing on the image's context or content rather than just comparing basic elements like color histograms and texture patterns [18,19]. Still, foundational data processing beyond this stage necessitated manual labeling. Manual annotation involves marking or labeling objects within an image to serve as ground truth. For instance, in LSCOM [20], over 10,000 h of manual effort produced 33 million labels denoting the presence or absence of specific features. Web search engines grappled with similar challenges. For example, a pioneering framework was devised to autonomously augment the textual labels of an image, leveraging its initial keywords and content [21,22]. Yet, the scalability of these methods for general images remains questionable, as the quality of the augmented labels might vary depending on the image type and the presence of fitting seed keywords.

Natural disaster image classification has encountered similar challenges. Many studies have been conducted to automatically classify large volumes of natural disaster images stored in online image archives. However, past studies approached a way to classify large images by annotating a natural disaster type as a label, and no object detection was performed [23–25]. Few studies successfully generated a multi-label database that continued natural disaster type and location as the primary categories [26]. Nevertheless, the previous work can only provide a preliminary classification of images based on the incident type, lacking actual object information that each image contains and cannot be used to perform detailed image classification based on the objects or at the component level. This gap underscores the need for more advanced computational methods in image classification that not only recognize disaster types but also delve into detailed content within the image, aligning with the evolving requirements of domain-specific applications in disaster management.

Recent advancements in deep learning have revolutionized image classification. It is now possible to automatically detect and interpret higher-level features in images with precision. More specifically, deep learning-based object detection models are favored for rapid but accurate image data processing in image classification. Object detection, considered extremely challenging in the early 2000 s, became much easier with recent advancements in computer vision and AI algorithms such as You Only Look Once (YOLO) [27], regional convolutional neural network (Mask R-CNN) [28], and transformer neural network [29]. Object detection in natural disaster damage assessment aims to identify and locate critical targets, including infrastructures and buildings [13,30–32], camouflaged objects [33,34], search and rescue [35–38], change detection [39–41], crack detection [42–44], and flood damage [45–47]. Although computers can process information much faster than humans, object detection still requires initial data to train the model, which typically includes manual annotation. While this process is labor-intensive and time-consuming, it is essential for crafting high-quality datasets pivotal for training, validating, and testing computer vision models.

Despite advances in object detection algorithms, automated natural disaster damage assessment on the building component level (e.g., windows, roofs, doors, walls) [4,9] is still limited, while community-level damage (overall affected area) [12,13,48,49] and property-level damage (individual structure) [50–53] is a common practice. This limitation is not due to data scarcity but challenges in accurately classifying vast image datasets for training data. Image classification refers to assigning categories to image sets based on specific characteristics, which is crucial for understanding the damage dynamics in disaster assessment. For instance, the training data must specifically originate from wind damage events to develop a model detecting hurricane-induced roof damage on residential buildings. Using commercial building imagery from earthquake incidents would be inappropriate, as the damage mechanisms are fundamentally different. Another case may involve studying the correlation between hurricane damage to elevated buildings and their front doors. It is essential to filter images showing the front view of such buildings. Side or nadir views merely contribute as noise data. The main challenge is not just in locating post-hurricane damage data but in the labor-intensive process of filtering images to pinpoint specific events and their associated damages.

To date, no large-scale image databases exist for natural disaster classification. Despite recent advancements in AI-driven image curation and classification techniques—including automated image tagging [54], image quality assessment [55], and duplicate image detection [56]—most natural disaster repositories remain focused on raw storage without systematic curation. For researchers and practitioners leveraging these repositories for deep learning-based damage assessment, the major challenge comes from the infeasibility of manually sorting hundreds of labels into countless images and their categories. It is extremely laborious and time-consuming and cannot be completed by a small group of researchers. As a result, most studies end up with micro-level hazard analysis, failing to accomplish both micro and macro levels of image data analysis and classification.

Most recently, the advancement of transformer networks and OpenAI's Contrastive Language-Image Pre-Training (CLIP) [57] has facilitated the development of rapid semantic image search engines [58]. CLIP, trained on a wide array of internet images using natural language supervision, has the capability to produce zero-shot classifiers from textual descriptions. This means that a semantic image search engine leveraging CLIP can extract text embeddings or labels from an image, like identifying objects and content (e.g., dogs and cats), without needing custom training data. However, two main challenges persist to be applied on post-natural disaster images with complex and intricate objects. First, while large pre-trained models like CLIP can identify generic objects, they might struggle with specifics; for instance, they can recognize a building but might fail to discern its structural type or assess damage. Secondly, we currently lack a reliable method to determine the accuracy of the labels assigned by these models, meaning there is no clear way to gauge how closely a label correlates with the actual content of an image.

To address these challenges, this study proposes a new approach to automatically classify large volumes of natural disaster damage images. It aims to automating disaster image search to support civil engineering tasks that include, but are not limited to, characterizing hazard exposure, damage assessment and modeling, and debris quantification. The approach couples pre-trained computer vision models and a natural language processing model with an ontology tailored to natural disasters to facilitate the search and query of specific image data types. More specifically, image labels are first extracted automatically using pre-trained models from computer vision recognition platforms, which minimizes manual image labeling work by utilizing the largest image label pre-trained model. Then, extracted labels will be incorporated into a developed word-vector query system, which quantifies the label intensity to the image, resulting in four primary labels with numerical values that can inform types of natural disasters, images, locations, and entities.

This research contributes to the field of natural disaster image data curation, addressing gaps in existing methodologies. It introduces an approach for automated curation of natural disaster images, reducing the reliance on manual labeling. This area, particularly in the context of automated data curation and efficient image classification for natural disasters, has seen limited exploration. This study introduces a method to improve data curation from large disaster repositories. The

methodological innovation of this research is the integration of advanced computer vision techniques with natural language processing. This combination is not commonly seen in current disaster management research. The proposed method enhances the curating and utilization of data in disaster management repositories, which have traditionally struggled with handling large-scale disaster data efficiently. By improving data curation processes and search capabilities, this research aids in more effective use of post-disaster data, facilitating robust post-disaster forensic studies and making extensive datasets more accessible for comprehensive disaster analysis. The final contribution of this study is to address the underutilized potential of natural disaster image data. Unlike the general object recognition models developed by companies like Google or Amazon, which rely on extensive labeled datasets, there is a lack of datasets specifically labeled for natural disaster damage assessment. This research is a step towards creating a tailored dataset for natural disaster AI research, filling a notable gap in the field and paving the way for more nuanced AI-driven disaster analysis and response strategies.

2. Proposed methodology

The overall workflow is shown in Fig. 1. The workflow begins with image uploads to cloud storage. Once uploaded, image label extraction is automatically initiated using the large pre-trained recognition models. These models may include custom labels for data gaps and can significantly reduce the need for manual labeling. This process also involves using pre-trained models for both face and text detection, with the identified areas subjected to Gaussian blurring to anonymize sensitive information like faces and license plates. All extracted labels are stored back in the cloud. These labels serve as a secondary layer in the workflow. The secondary labels undergo processing to prepare them for natural language word embedding models. This includes converting all text to lowercase, trimming unnecessary punctuation marks, and filtering out irrelevant labels (e.g., labels like *lobster* or *pilot* are not helpful in the context of natural disaster damage assessment). The labels are then evaluated against word embeddings, generating tf-idf scores for each label. The outcome of this process is a set of primary labels, each associated with a numerical value. These labels provide comprehensive insights into the nature of the natural disaster, image content, location, and entities depicted, offering a streamlined, automated approach to image analysis in disaster management scenarios.

A. Automated Label Extraction

Over the last decade, numerous advanced cloud-based computer vision recognition platforms have been developed. These platforms utilize various deep learning techniques, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer networks, to classify images and provide labels for the contents within them. Several computer vision platforms offer advanced pre-trained models based on large datasets that can be conveniently accessed and employed for research purposes. These include Google Cloud Vision API, Amazon Web Services (AWS) Rekognition, Facebook Contrastive Language-Image Pre-Training (CLIP), Microsoft Azure Computer Vision API, and IBM Watson Visual Recognition. These pre-trained models can be utilized to conduct the initial automated image label extraction pipeline. Common detection features are label, face, and text detection, making it a comprehensive solution for various image-related tasks. Example label extraction results from AWS Rekogntion and Google Cloud Vision on an aerial photo of post-hurricane Michael are shown in Fig. 2. Label Detection Results on AWS Rekognition and Google Cloud Vision

. Additionally, custom labels can be used in filling data gaps, especially when pre-trained models lack specific labels or unique identifiers.

B. Label Taxonomy.

Although many labels are extracted from image label detection, image classification and ordering sorely by these labels are inefficient. This is because each image typically produces 20 to 30 non-identical labels, and the number would increase with custom labels. With the increasing number of images, manually sorting hundreds of labels into countless images and their categories is still a labor intensive task. To reduce such human efforts, a label taxonomy is designed to facilitate image processing through natural language processing. From now on, the labels extracted from computer vision recognition are referred to as "secondary labels." This section aims to elect "primary labels" from the secondary labels in a similar group and topic, which can efficiently represent the characteristics of the image by merging the notion of the secondary labels.

To develop an ontology-based label classification system that can effectively classify post-natural disaster images, label taxonomy should

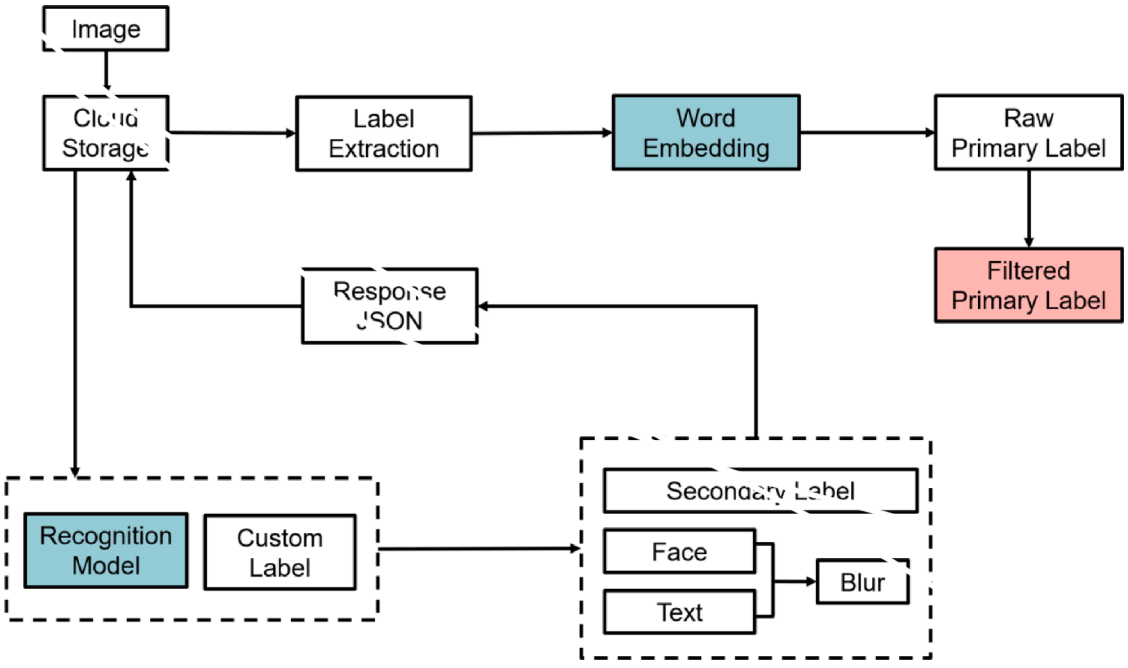


Fig. 1. The Overall Workflow of Proposed Label Extraction and Classification.

(a) Input aerial image

AWS Rekognition		Google Cloud Vision	
Label	Confidence	Label	Confidence
Nature	99.6%	Building	95%
Landscape	99.6%	Property	94%
Outdoors	99.6%	Window	88%
Scenery	99.4%	Urban Design	86%
Aerial View	93%	House	84%
Roof	92.1%	Tree	84%
Neighborhood	83.4%	Neighborhood	82%
Building	83.4%	Residential Area	81%
Urban	83.4%	Landscape	80%
Earthquake	61.8%	Roof	78%
Soil	60.1%	Facade	77%
Land	57.8%	Bird's-eye View	72%
Housing	56.5%	Event	66%
Ground	55%	Aerial Photo	63%

(b) Label detection outcome

Fig. 2. Label Detection Results on AWS Rekognition and Google Cloud Vision.

be considered first. Initially, images with metadata already include time, GPS, and camera settings. Therefore, this information does not need to be labeled and is excluded from the taxonomy. Overall, five main aspects are considered for modeling label taxonomy:

(1) Entity: Representation of objectively existing units of information. Examples are, humans, structures, vehicles, trees, roads, traffic signs, lights, and debris.

(2) Hazard: Possible natural hazards related to the state of damage. Examples are, wind, flood, tornado, rain, current, and storm surge.

(3) Image type: Describes the type of image. Examples are, ground, aerial oblique, nadir, satellite, and panorama.

(4) Land Use: Describe the type of the affected area. Examples are, residential, commercial, vegetation, agriculture, coastal, and infrastructure.

(5) Component: Possible structural and non-structural building components present. Examples are, roof, window, door, wall, garage, column, foundation, ventilation, HVAC, and sheathing.

Table 1 presents the final list of primary for each label type, which will be the main information to classify, group, and sort out a large volume of images effectively. This is because multiple primary labels are needed that can be stacked to describe the image in the way of hazard language. Frequent labels from AWS label extraction results were compacted into the ten (E1, E2, E3, H1, H2, I1, I2, L1, L2, and L3) primary labels in four categories. A set of custom labels was introduced for the "Building Component" category to streamline the classification process. While a wide array of structural and non-structural components are easily recognizable to the human eye, it is not feasible to identify, label, and train a model for every single one. This is especially true for smaller or less visible components (like gutters, vents, railings, and cladding elements) and those not typically exposed (such as columns, beams, foundations) or severely damaged components that no longer retain their original form. To facilitate a more manageable and practical classification, five key custom labels were defined: door, window, roof, wall, and garage door. A total of 1,220 images from post-Hurricane Harvey were manually labeled, encompassing a mix of damaged and undamaged buildings. These images were sourced from ground-level digital photographs taken by the reconnaissance teams of Rutgers University, Princeton University, and the University of Texas at Austin [59].

C. Natural Language Model.

To develop a word-vector query system to quantify the similarity between primary and secondary labels (i.e., to extract word vector

Table 1
Word Embeddings for the Query System.

Label Type	Primary label	Word Embeddings
Entity 1 (E1)	People	people, human, men, women, boy, girl, child
Entity 2 (E2)	Building	building, house, villa, condo, duplex, apartment, structure, property, porch, balcony, roof, door, wall, window, garage, lawn, driveway, mailbox, garden
Entity 3 (E3)	Vehicle	vehicle, car, truck, wheel, motor, automobile
Hazard 1 (H1)	Hurricane	hurricane, typhoon, cyclone, storm, wind, tornado, destruction, emergency, damage, landfall, weather, surge, flood, disaster, debris
Hazard 2 (H2)	Flood	flooding, flood, water, river, storm, mud, submerge, surge, plunge, tide, landslide, rain, torrential, downpours, tsunami, disaster, debris
Image Type 1 (I1)	Ground	ground, street, road, lane, route, roadway, panorama
Image Type 2 (I2)	Aerial	aerial, air, drone, UAV, scenic, drone, flying, satellite, vertical, nadir
Land Use 1 (L1)	Residential	residential, resident, property, urban, suburban, neighborhood, house, home, apartment, condo, townhouse, duplex, villa, garden
Land Use 2 (L2)	Commercial	commercial, business, office, retail, warehouse, hotel, industrial, urban, street, parking, parkinglot, mall, shopping
Land Use 3 (L3)	Vegetation	vegetation, landscape, greenery, nature, forest, tree, habitat, plant, green, park, flora, woodland, shrubs
BuildingComponent (C1)	Roof	roof
BuildingComponent (C2)	Wall	wall
BuildingComponent (C3)	Window	window
BuildingComponent (C4)	Door	door
BuildingComponent (C5)	Garage	garage

representation and categorize image labels into four classes described above), word embedding models from natural language processing can be implemented. They can learn vector representations of words to capture the semantic meaning and relationships with other words. Popular word embedding models trained on large amounts of text data include Global Vectors for Word Representation (GloVe) [60] and Word2Vec [61]. Although state-of-the-art transformer network-based natural language models like GPT [62] and BERT [63] have advanced self-attention mechanisms that enable them to process entire sentences or paragraphs, this research employs conventional natural language models because the current technology for automated label extraction can only return single or multi-letter labels rather than descriptions of the contents within images.

D. Primary Labels and Word Embeddings.

The world embeddings for each primary label are provided in Table 1Table 1. Word embeddings, a learned representation for text with the same or similar representation, were determined by analyzing the context in which words appear within large text corpora, thereby capturing their semantic meaning and relationships to one another (example shown in Fig. 5). For instance, when the "People" primary label is used, related terms within that semantic field—such as *human, men, women, boy, girl,* and *child*—are selected as word embeddings. These terms are chosen based on their contextual relevance and similarity to the primary label, creating a cluster of words with shared meanings. The initial clusters of closely associated words form the preliminary word embeddings. These preliminary embeddings are further refined through manual intervention to ensure relevance and accuracy. This process involves reviewing and adjusting the associations to better reflect the nuanced relationships between words, particularly in the context of disaster response and damage assessment.

Unlike other primary labels, the world embeddings for Building Components C1, C2, C3, C4, and C5 are unique in that they only encompass the exact terms corresponding to each component. This design is intentional and serves the specific purpose of detecting the presence of these particular building components without the inclusion of broader or related terms.

The generation of primary labels relies on pre-trained recognition models, but it is important to understand the nuances of this process. For example, identifying the specific type of natural disaster based solely on the "Hazard" primary label and visible damage in the images can be quite challenging. This is because the process is not a supervised detection method and largely depends on recognizing common patterns of damage. For instance, flood patterns are indicative of hurricanes, while severe structural collapses might suggest earthquakes. However, these damage patterns can often overlap or be similar across different disaster types. Therefore, relying exclusively on the "Hazard" primary label might not always yield accurate results. It is advisable to use this label in conjunction with other primary labels to get a more comprehensive understanding of the disaster.

The term frequency-inverse document frequency (tf-idf) is used to associate the secondary labels with the primary labels. Tf-idf score is a numerical statistic intended to reflect how important a word is to a document in a collection. It is a famous method to rank the content of data, focusing on term frequency rather than simple keyword counts of observing a word occurrence in specific topics. For example, in this study, if a secondary label, *roof*, extracted from the framework, is inputted to the query system, the output is a score on how similar the *roof* is to the word embeddings.

$$TFIDF_{t,d} = TF_{t,d}\log(\frac{N}{DF_t})$$

For the text t in a document d from the document set N is calculated as the multiplication of the term frequency of a word in a document $TF_{t,d}$ and the inverse document frequency of the word across a set of documents $\log(\frac{N}{DF_t})$ where DFt is the number of documents containing the input text [64]. For example, if the word *truck* occurs five times in a 1,000-word document, *TF* (term frequency) is 5/1000 = 0.005. Considering the GloVe model, which utilizes a comprehensive corpus size of 804 billion, the IDF (inverse document frequency) would be calculated as log(804 billion / N). In detail, this extensive corpus was employed over world embeddings of each primary label and computing the term frequency scores for the secondary labels inputted. At the end of the process, images are classified into four classes (i.e., entity, hazard, image type, and location) based on the ontology model word vector result.

Now, if the primary labels are utilized as designed, the desired result should be as shown in Table 2Table 2. Suppose we want to use the three images below that are found in online image archives and lack metadata (stock.adobe.com). The first image shows an aerial view of a flooded residential area. We can easily assume that the image was taken after a flood event with a drone or helicopter. Therefore, with the label classification, the desired result would show a high tf-idf score for aerial image type, flood for hazard type, residential for land use, and building for entity type. The second image shows a road and a connected bridge near a water body. The sky darkens ominously, and the sea surges, indicating a severe hurricane scene. From this image, the desired result would show high tf-idf scores of ground for image type and hurricane for hazard type. The last image shows damaged buildings and debris in the residential area with some vegetation. From a human perspective, we can see that the roads are clear, which means the initial recovery effort has been completed to get the road access. Therefore, the desired primary label detection result for the last image would have high tf-idf scores of ground for the image type, residential and vegetation for land use, and building and vehicle for the entity type.

The type of data shown in Table 2 can be pivotal for post-disaster decision-making, primarily due to its role in enhancing AI and deep learning algorithms through the efficient categorization of disaster-related images. Creating a structured dataset from various images, regardless of their origin or metadata, significantly improves AI's accuracy in object detection and automation of disaster responses. This adaptability ensures a comprehensive understanding of diverse disaster scenarios, which is crucial for effective damage assessment and resource allocation. Furthermore, the labeling process not only facilitates machine learning advancements but also supports historical analysis and trend prediction. This is essential in forecasting future disaster patterns, thereby aiding in the development of more informed and efficient response strategies. Addressing the gap in disaster-specific AI, this research equips decision-makers with nuanced, AI-driven analysis tools, paving the way for more accurate and rapid disaster response and preparedness measures.

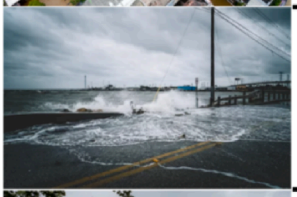
3. Hurricane Harvey use case

As previously discussed, the foremost objective in creating the primary label set is to develop a novel workflow designed for categorizing and classifying a vast array of natural disaster images—a task that has not been feasible until now. The introduction of primary labels to these images paves the way for their use as foundational training data in research on natural disaster hazards. Two main datasets were utilized for the comparison of data characteristics: reconnaissance images from Hurricane Harvey, enriched with metadata, and a collection of search results featuring buildings impacted by the hurricane, which lack metadata.

A. Data Sets.

Major data sources come from images of post-Hurricane Harvey, 2017, at Port Aransas and Rockport, Texas. Streel-level images were taken from a mobile system equipped with a Ladybug spherical digital camera. The camera's multi-sensor captures six views with 1024 × 768 pixels: one vertical view and five circularly configured horizontal views. Roughly 80 pixels were overlapped between the views to be composited into a pharaonic image.

Table 2
Desired Primary Label Result.

	Type	Hazard	Location	Entity
	Aerial	Hurricane Flood	Residential	Building
	Ground	Hurricane Flood	-	-
	Ground	Hurricane Earthquake	Residential Vegetation	Building Vehicle Debris

The second data comes from a collection of web-searchable hurricane-affected building images from web search engines. Web crawling or scraping often refers to a subset of the web scraping tool for creating a large web dataset, such as URLs, images, video, text, etc. In recent years, image crawling tools have been designed specifically for large search engines (e.g., Google, Bing, Naver) to discover automatically, scan, and download targeted web data. In order to automatically download hurricane images from online archives and create a large image database with image crawling, CygnusX1(Google crawling) [65] and AutoCrwaler (Naver crawling) [66] have been implemented. During the crawling process, images that are too dark or in poor resolution have been neglected from the selection. Lastly, a simple Laplacian of the Gaussian algorithm [67] was used to filter out blurred images. In the end, 3,721 images and 782 images each from Google and Naver, a total of 4,053 images, have been collected with keywords of *'hurricane'* + *'building'* as a sample online archive dataset (careful selection of keywords can result in much difference in the tendency of crawled images as shown in Fig. 3).

B. Methodology implementation

A cloud-based computer vision platform, Amazon Web Services (AWS) Rekognition, was implemented to construct the initial automated image label extraction due to its accurate recognition model and convenient connection with cloud storage and detection event trigger. Recent studies have demonstrated the high accuracy of AWS Rekognition's pre-trained model, further validating its effectiveness for image labeling and classification. [68,69]. AWS S3, a web service interface for object storage, is mainly used for image upload and response storage. AWS Lambda is linked between Rekognition and S3 to manage and connect event triggers. AWS Lambda is an event-driven computing service that runs in response to other AWS events via AWS software development kits (SDKs).

The four major steps for label extraction are: (1) add five custom labels, including roof, wall, window, door, and garage; (2) upload image: input images are uploaded to a designated folder in an S3 bucket. During this initial process, image EXIF data are stored as Amazon S3 Object Metadata (e.g., camera information, location, time); (3) detect request: Lambda function allows automatic analysis when images are uploaded to S3 Bucket. When the Lambda function is triggered, it calls StartLabelDetection to start the label detection of uploaded images; (4) response data: once StartLabelDetection is finished, the second Lambda function, GetLabelDetection, is called to store the analysis result back in S3 Bucket. Both Start and Get functions are repeated for

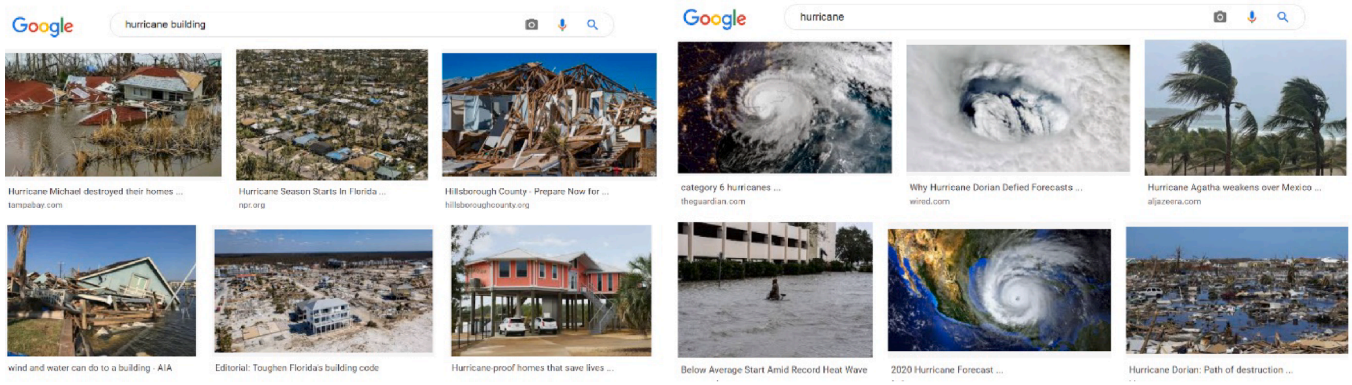


Fig. 3. Image search results with keywords *'hurricane'* + *'building'* on the left and *'hurricane'* on the right. A combination of keywords can easily lead to images of hurricane-damaged buildings, while using a single 'hurricane' often shows satellite view or general hurricane images (google.com).

TextDetectionand FaceDetection.

In this study, the choice to employ a minimum confidence threshold of 50 % for detection accuracy, as opposed to the typical 70–80 %, is informed by the specific objectives of label extraction rather than precise object detection. The confidence score is a measure typically used to determine the accuracy of detection results, often associated with creating bounding boxes around recognized objects in an image. However, our focus is on extracting simple text labels associated with the overall image, which is a less stringent requirement than identifying and delineating objects. By setting the confidence threshold at a lower level, we allow for a greater number of potential labels to be included. The response of the label detection API is saved in the JSON structure, which includes the detected object, bounding box location, label, and confidence. Bounding box locations from text and face detection are later used for personal information blurring, as shown in Fig. 4.

The extracted labels are returned in JSON format for each image. While 2,580 image labels are offered from the pre-trained Rekognition model, not all labels are particularly useful for building damage assessment. Small object labels, such as pencil and shrimp, or irrelevant labels, such as surgeon and elephant, can cause possible interruptions to the later analysis. Therefore, the selected 654 labels related to building and natural disaster damage were filtered to be used for post-hurricane building damage image label classification. Lastly, an additional step to combine multi-letter words with single-letter words was performed to minimize vector weight confusion for the label classification.

GloVe's pre-trained model was implemented to develop a word-vector query system for label classification. GloVe is modeled for dimensionality reduction on the co-occurrence count matrix to find the lower-dimensional representations. It allows one to take a corpus of text and intuitively transform each word into a high-dimensional space. It can capture the context of a word in a document, semantic and syntactic similarity, and relation with other words. In other words, pre-trained word weight will count the vector relationship between labels, providing a direction of similarity between the labels. Fig. 5Err or! **Reference source not found.** shows a similarity visualization among common secondary labels extracted from AWS using the Hurricane Harvey reconnaissance mission. GloVe pre-trained word vector with 840 billion token corps, 2.2 million vocabularies, and 300 dimensions was used and visualized with the t-distributed stochastic neighbor embedding (t-SNE) method. As a result, words with similar topics are placed closer together. For example, words clustered in a red circle are related to natural disaster terms: *storm, wind, flooding, hurricane,* and *cyclone*. The yellow circle also shows a cluster of buildings and their components. The green circle also shows a cluster of representations of humans, such as *women, men, boy,* and *girl*.

4. Results discussion

Determining what constitutes a high or low tf-idf score can be inherently challenging due to its subjective nature. This subjectivity stems from the fact that tf-idf is a relative measure; its significance is contextually bound to the specific dataset or corpus to which it is applied. In different collections of documents, the same numeric tf-idf score might represent different levels of importance. However, this ambiguity can be mitigated by applying simple statistical measures. Analyzing the distribution (mean, median, and standard deviation) of tf-idf scores across a corpus makes it feasible to establish a data-driven threshold. It allows for a more systematic approach in differentiating between terms of varying significance, thereby providing a more grounded and justifiable means of classifying high and low tf-idf scores within the unique context of the given dataset. The value of 15 was identified as the threshold for categorizing a score as high, which was determined to be higher than both the mean and the median of the tf-idf scores across the dataset.

Table 3 shows the tf-idf result of four common labels detected from a sample Hurricane Harvey reconnaissance image (Fig. 6) containing multiple buildings affected by wind and flood damage and recovery efforts. Of the 22 secondary labels detected with a confidence of 50 % or more, the four labels with the highest confidence level were selected for comparison. The first secondary label identified was *neighborhood*, which showed high-frequency scores of 17.81 for L1 Residential and 16.03 for L2 Commercial, indicating a high probability of the image being located in a residential or commercial area. The second secondary label, *road*, showed high-frequency scores of 16.31 for E3 Vehicle, 17.27 for I1 Ground, 15.20 for L1 Commercial, and 15.40 for L3 Vegetation. These scores suggest that the image was taken at ground level, with the possibility of a vehicle present near a commercial or vegetated area. The third secondary label, *car*, showed a high-frequency score of 29.41 for E3 Vehicle. The exceptionally high score for E3 indicates a very high probability of a vehicle being present. The fourth secondary label, the *debris* label, showed high-frequency scores of 16.37 for H2 Flooding and 19.26 for C1 Roof. Notably, the label returned a high score for flooding, indicating potential debris and damages caused by the flood event. The last two secondary labels, *window* and *roof*, showed overall exceptional high-frequency scores for all building component primary labels. This data suggests that both windows and roofs, being prevalent features in various structures, significantly indicate the presence of buildings and their other components, particularly in typical residential or commercial areas.

All labels displayed low scores for I2 Aerial, suggesting that the image originated from ground level, aligning with expectations. Significantly, the high-term frequency scores derived from the input data were compiled and stored in a distinct JSON file. These files serve as an additional dimension for metadata analysis, offering a nuanced

Fig. 4. Text Detection and Blur.

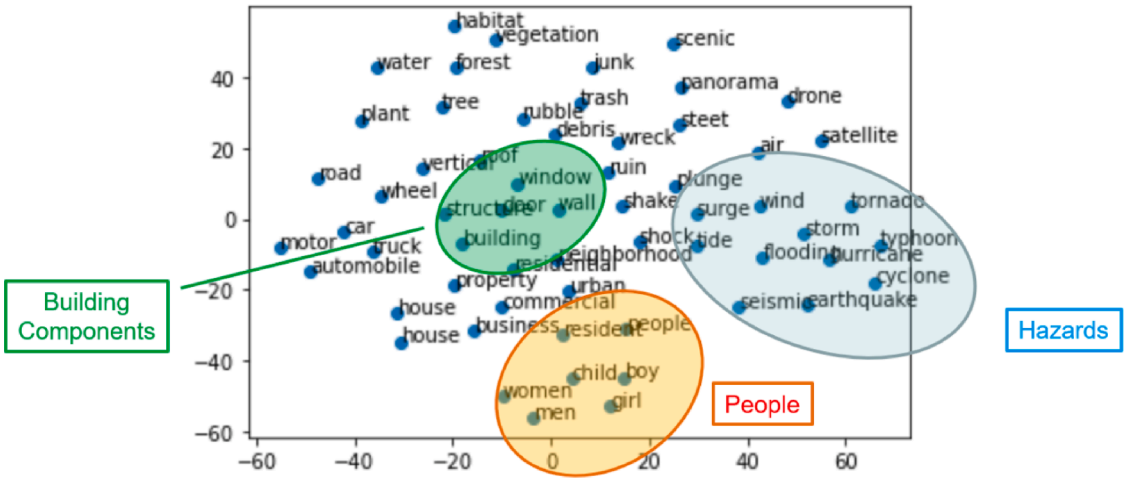


Fig. 5. Visualization of Similarly Among Common Secondary Labels.

Table 3
Tf-idf Score Result of Fig. 6.

Primary Label	Secondary Labels							
	Neighborhood	Road	Car	Debris	Window	Roof		
Entity	E1	People	12.24	11.76	14.29	7.52	8.24	7.56
	E2	Building	11.03	12.04	13.24	10.84	**16.42**	7.65
	E3	Vehicle	9.59	**16.31**	**29.41**	11.44	14.35	11.43
Hazard	H1	Hurricane	4.10	5.76	5.02	8.45	6.72	6.60
	H2	Flooding	6.09	9.63	8.56	**16.37**	5.94	11.48
Image								
Type	I1	Ground	9.24	**17.27**	10.43	7.98	8.26	9.16
	I2	Aerial	3.39	7.11	8.15	7.05	5.69	7.17
Land Use	L1	Residential	**17.81**	12.75	13.27	6.91	**15.41**	9.77
	L2	Commercial	**16.03**	**15.20**	14.90	7.64	**16.84**	9.26
	L3	Vegetation	9.76	**15.40**	9.85	13.38	4.54	13.36
Building								
Component	C1	Roof	11.24	13.11	16.22	**19.26**	**23.45**	**36.32**
	C2	Wall	14.20	14.78	15.41	14.24	**21.87**	**21.14**
	C3	Window	13.16	11.67	17.97	13.78	**30.73**	**23.45**
	C4	Door	14.41	12.99	21.18	11.34	**24.55**	**19.82**
	C5	Garage	15.91	13.02	21.78	10.75	**16.93**	**17.42**

Fig. 6. Sample Reconnaissance Image for tf-idf Score Computation.

means of comparing and categorizing images based on similar input parameters. Its utility is particularly evident in segregating specific subsets of images from extensive image datasets. For instance, applying a high-score threshold for E2, I1, and L1 can effectively isolate ground-level images predominantly featuring residential structures. Additionally, incorporating specific Building Component primary labels enhances the precision in identifying buildings with specific components.

5. Validation of results

A manual labeling process was implemented to establish a baseline for comparison and validate the accuracy of the computed primary labels. Images are divided into four categories, with 40 images each: (1) affected, (2) minor damage, (3) major damage, and (4) destroyed to illustrate post-hurricane damage on residential buildings. These 160 images were manually assigned labels that should describe the overall character of the image from a human perspective and serve as the ground truth. The approach and list of recommended manual label assignments are shown in the Table 4. Eight people with sufficient experience and knowledge in hurricane damage assessment completed manual labeling on each dataset. Building component labels (door, window, roof, wall, and garage) were excluded from manual label assigning because they are detected from a custom label separated from the pre-trained Rekognition model. The results were validated by comparing primary labels assigned manually and secondary labels extracted from Rekognition.

Images are divided into four categories, with 40 images each: (1) affected, (2) minor damage, (3) major damage, and (4) destroyed to illustrate post-hurricane damage on residential buildings. These 160 images were manually assigned labels that should describe the overall character of the image from a human perspective and serve as the ground truth. The approach and list of recommended manual label assignments are shown in the table below. Eight people with sufficient experience and knowledge in hurricane damage assessment completed manual labeling on each dataset. Building component labels (door, window, roof, wall, and garage) were excluded from manual label assigning because they are detected from a custom label separated from the pre-trained Rekognition model. The results were validated by comparing primary labels assigned manually and secondary labels extracted from Rekognition.

The validation aimed to assess the reliability and consistency of the manual labeling process and determine the effectiveness of the automated labeling approach. Images from each damage category and their secondary labels and primary labels computed from Rekognition and manual labels are compared in Table 5. The analysis revealed that the number of secondary labels detected by Rekognition exceeded those obtained through manual labeling, with a wide range of categories identified. However, it is important to note that not all of these labels were accurate. For instance, false positives were detected as irrelevant labels for the first affected image, such as indoors and yacht. In contrast, the manual labeling process was found to be intuitive and efficient, avoiding unnecessary or unrelated ones. Additionally, the manual process enabled the identification of information-rich labels, such as *tarp,* from recovery efforts and the ability to determine whether a building is elevated, which would be difficult to detect through an automated approach without sufficient training datasets.

Table 4
Protocol for Manual Assignment of Image Labels.

Entity	• Is there a building present? What type of building is it? What other objects exist, vehicles, trees, and roads?
Hazard	• Is there a natural hazard damage present? What is the extent of building damage, intact, damaged, or destroyed?
Image Type	• Is the image taken from ground-level or aerial? Is the image panoramic?
Land Use	• What is the building use type, residential or commercial? What environment is the building located in?

Table 6 displays the comparison between automatically computed primary labels and manually assigned labels across a dataset that encompasses ten primary labels and four distinct damage categories: affected, minor damage, major damage, and destroyed. The comparison is quantified through a ratio, which is calculated by dividing the tf-idf scores of the automatically generated labels by those of the manual labels. Overall, the primary labels obtained through Rekognition demonstrated consistency with the manual primary labels, with a 13.32 % difference in the tf-idf score.

A higher tf-idf score indicates a richer set of attributes related to the primary label. However, the primary labels for Hurricane (H1) and Flood (H2) showed significant improvement when manually labeled, indicating that Rekognition is ineffective at identifying the relationship between damaged buildings and natural disasters. Specifically, the difference ratio for the affected category remained at 1.25, while the minor, major, and destroyed categories exhibited more than double the scores. The manual labeling approach is more accurate since it is easy and natural for humans to identify damaged buildings, their debris, and the extent of damage and relate them to natural disasters. Primary labels for Ground (I1) and Aerial (I2) also demonstrated higher accuracy with manual labeling. Similar to the hazard category, it is more natural for humans to list labels directly related to the image type (ground), while primary labels from Rekognition act as a recomputed similarity score based on various labels located in similar vector space. Hence, the manual primary labels performed better than the Rekognition labels, with an average difference of only 13.32 %. This means that the pre-trained Rekognition model is a suitable and efficient tool for extracting labels from a large number of post-hurricane images and classifying them effectively.

The study also involved a detailed comparison between automatically computed secondary labels and manually assigned labels for five specific building component custom labels: roof, window, wall, door, and garage. The objective was not to evaluate the final tf-idf scores of primary labels but to assess the precision with which automated custom labels could identify varying degrees of building damage, using manually assigned labels as a ground truth for accuracy. As illustrated in Table 7, a minor discrepancy between automated and manual labels in categories of affected and minor damage is observable. However, this disparity grows more pronounced in categories of major damage, with a difference of −16.88 %, and in the destroyed building category, with a difference of −48.47 %. Despite the 1,120 training data and advanced computer vision technology used in automated label generation systems, they are still outstripped by the human eye's ability to discern building damage, particularly in complex scenarios.

6. Conclusions and future work

This study introduces a novel workflow capable of automatically extracting images from large volumes of data and classifying them. By leveraging pre-trained computer vision and natural language processing models in conjunction with a developed word-embedding model, the process assigns primary labels and similarity scores to each processed image. These labels and scores can be used to efficiently search and query specific types of image data.

The validation process and the accuracy assessment of AWS Rekognition covered in many studies [68,70] enhance the credibility of the data set and ensure that the proposed pipeline can be used effectively and efficiently to classify images. AWS Rekognition's pre-trained model label detection result was modified to return 654 labels related to post-hurricane images. Then, the GloVe-based word-vector query system was developed to generate primary labels to classify images based on the similarity score computed in four hazard categories. Overall, the Rekognition primary labels demonstrated upright scores compared to

Table 5
Primary Label Comparison With Unobstructed Building View.

Image	Rekognition Secondary Label	Manual Secondary Label		Primary Label (Automated)	Primary Label (Manual)	
fx2	truck	condo	building	E1	8.96	11.67
Affected	vehicle	villa	house	E2	9.75	12.95
	architecture	porch	residential	E3	12.2	13.12
	building	grass	car	H1	3.86	4.9
	housing	plant	vehicle	H2	6.44	8.13
	car	pickup	truck	I1	7.12	8.9
	machine	garage	tree	I2	5.13	6.45
	wheel	indoors	intact	L1	10.17	13.27
	house	balcony	railing	L2	10.38	12.3
	city	handrail	ground	L3	9.84	10.06
	mobile home	yacht	driveway			
fx3	garage	chair	building	E1	8.45	10.69
Minor Damage	indoors	furniture	house	E2	10.91	12.3
	architecture	grass	residential	E3	8.87	11.03
	building	plant	tree	H1	3.29	7.63
	house	city	debris	H2	5.95	11.65
	housing	condo	railing	I1	6.89	8.48
	porch	balcony	damaged	I2	3.94	6.8
	villa	cottage	ground	L1	10.96	12.46
	water	door	hurricane	L2	10.2	11.35
	waterfront			L3	10.67	10.1
fx4	architecture	car	building	E1	8.75	10.69
Major Damage	building	vehicle	house	E2	10.35	12.3
	house	nature	elevated	E3	10.25	11.03
	housing	staircase	residential	H1	3.52	7.63
	porch	villa	tree	H2	6.55	11.65
	gate	deck	tarp	I1	7.4	8.84
	grass	cottage	debris	I2	4.89	8.48
	plant	siding	damaged	L1	10.44	12.46
	outdoors	neighbor	ground	L2	10.33	11.35
	indoors	shelter	hurricane	L3	10.87	10.1
	garage		driveway			
			railing			
fx5	architecture	porch	building	E1	10.21	10.66
Destroyed	building	play area	house	E2	9.73	12.13
	house	wood	elevated	E3	7.46	10.87
	housing	city	residential	H1	4.04	7.76
	staircase	country	debris	H2	7.08	12.09
	outdoors	hut	destroyed	I1	7.28	6.48
	nature	rural	ground	I2	4.79	6.98
	handrail	shelter	hurricane	L1	11.23	13.01
				L2	10.53	11.68
				L3	11.75	11.14

Table 6
The tf-idf Score Ratio of 160 Topic Data Primary Labels, Rekognition Label vs. Manual Label.

	Affected	Minor Damage	Major Damage	Destroyed
E1	1.36	1.30	1.21	1.03
E2	1.31	1.21	1.18	1.28
E3	1.07	1.24	1.08	1.41
H1	1.25	2.30	2.21	2.07
H2	1.27	1.94	1.75	1.85
I1	1.29	1.26	1.18	0.87
I2	1.30	1.75	1.74	1.40
L1	1.34	1.15	1.27	1.17
L2	1.21	1.09	1.05	1.10
L3	1.00	0.92	0.98	0.90

Table 7
Average Building Component Labels Detected on 160 Topic Data Topic Data, Rekognition/Custom Label vs. Manual Label.

	Automatic Label	Manual Label	Difference
Affected	3.875	4.425	−12.43 %
Minor damage	3.725	4.15	−10.24 %
Major damage	3.325	4	−16.88 %
Destroyed	2.1	4.075	−48.47 %

the manual primary labels, with a minimum average difference of 13.32 %. This indicates that the pre-trained Rekognition model is a good fit and can be used effectively and efficiently to extract labels from a large volume of post-hurricane images for classification. However, Rekognition's labels are inefficient in identifying the relationship between natural disasters and damaged buildings, and custom data supplementation is necessary.

Yet, the use of custom labels showed clear limitations on damage-building component detection. This limitation is deeply rooted in the intricate and nuanced process of damage assessment, where visual indicators can be subtle, complex, and heavily reliant on context. Automated detection struggles to grasp the full extent of these contextual subtleties, a feat that the human eye can manage more adeptly. This is particularly true in cases of severe destruction or when buildings are reduced to rubble. Identifying completely damaged or collapsed structures using computer vision remains a formidable challenge, as these systems often fail to recognize the chaotic and unpredictable nature of such devastation. Consequently, while automated tools are indispensable for processing data on a large scale, the discerning judgment of human evaluators is essential for the accurate detection and evaluation of severely damaged or collapsed building components.

The framework also showed low capital and quick processing time, as it takes about four minutes and costs $2.80 to process 1,000 images. In the long term, a stand-alone pre-trained weight should be created for post-hurricane image classification using custom-labeled data and other computer vision pre-trained models, which can be used in various hurricane research and applications. To develop this weight, images should be retrained with primary labels. Although the framework relies on AWS systems due to the impracticality of manually training over 600 labels, an independent training weight can be established with enough primary and secondary labels. As a result, the AWS system or other computer vision pre-trained models can be eliminated from the framework. Ultimately, this independent and original framework can be utilized for various offline projects, including reconnaissance missions and image classification simultaneously from the site.

The proposed method offers an advanced system for automatically labeling and classifying large volumes of post-natural disaster image data, including images with and without metadata that have been archived for decades. The reconnaissance images with metadata can be curated to a higher level to enhance their management and organization in natural disaster data repositories and infrastructures. Even images without metadata from online archives can be classified using the proposed method to filter specific sets of images and create valuable training datasets for deep learning-based object detection applications. By automating this process, the proposed method saves time and effort that would otherwise be spent manually labeling and sorting images while ensuring that the resulting training datasets are accurate and relevant. This not only improves the efficiency of the training process but also helps to improve the overall performance of deep learning-based object detection systems. In particular, the methodology excels in reducing the typically daunting task of extracting suitable data for specific research needs, which often involves laboriously sifting through an extensive pool of images. For instance, it simplifies locating images depicting key aftermath scenarios, such as debris identification [71,72], multilevel segmentation with aerial footage [50], flood damage modeling [73], and wildfire damage pattern analysis [74]. This targeted data extraction capability is highly sought after by researchers and practitioners in the field of disaster management.

This study did not involve any innovation of the object detection or image classification algorithm because this is not in the sense of model improvement but more about rearranging the return and utilizing it in a new perspective to extract and classify labels of post-natural disaster images automatically. It is the first attempt to curate abundant natural disaster images to add supplementary but forefront information to organize the hazard database. The key contribution of this study is significant in that the proposed method can be applied to a wide range of image classifications. This method is versatile and adaptable, provided that the word embedding can be updated to meet the standard of the domain being targeted. Making the method applicable to various image classifications can be utilized in diverse fields and industries, offering a practical and valuable solution for image labeling and classification tasks. Moreover, the inherent flexibility of this method ensures that it can evolve alongside the changing needs of various domains. These evolutionary potential positions the method not just as a solution for current challenges but also as a foundational asset for future research and development.

In considering the potential for future development and the recognition of existing limitations in the proposed methodology, the following points emerge as key areas:

- Mitigating data quality dependence: The method's effectiveness heavily relies on the quality and volume of available data. In areas with scarce or poor-quality data, performance may falter. Future enhancements could involve merging multiple data sources and algorithms to overcome this limitation, ensuring consistent performance regardless of data quality. This approach is particularly pertinent in the context of detecting damaged building components, ensuring more reliable and consistent performance across diverse data conditions.
- Enhanced object detection and damage assessment: Current limitations in detecting and classifying complex objects, particularly in varied disaster types and damaged building components, highlight the need for improved models. Future work could develop deep learning models that not only identify damage but also assess its severity, incorporating extensive custom data to cover various damage patterns and building components.
- Real-time disaster monitoring and response: While suited for post-disaster analysis, the method's application in real-time monitoring is limited. Future advancements could focus on developing real-time image analysis capabilities that align with ongoing data acquisition at disaster sites, enhancing immediate response strategies.
- Predictive analysis applications: Utilizing classified image data for predictive analytics presents a significant opportunity for future research. This could involve developing models to predict the impact of imminent disasters using historical image data, which would be instrumental in improving disaster preparedness and mitigation strategies.
- Application to different disaster contexts: adapting the current workflow for use in various natural disasters like earthquakes and fires. This may include categorizing building damages and components specific to earthquakes and fires, such as structural faults from earthquakes or burn path patterns from fires.

CRediT authorship contribution statement

Sun Ho Ro: . **Yitong Li:** Formal analysis, Validation, Writing – review & editing. **Jie Gong:** Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

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

References

[1] Z. Zhou, *Computer Vision-based Assessment of Coastal Building Structures during Hurricane Events.* Rutgers, the State University of New Jersey, 2017.
[2] G.L. Chiu, S.J. Wadia-Fascetti, Assessment and quantification of hurricane induced damage to houses, Wind and Structures 2 (3) (1999) 133–150.
[3] J.H. Crandell, V. Kochkin, *Scientific damage assessment methodology and practical applications*. in *Structures Congress 2005*, Metropolis and Beyond, 2005.
[4] A. Hatzikyriakou, et al., Component-based vulnerability analysis for residential structures subjected to storm surge impact from Hurricane Sandy, Natural Hazards Review 17 (1) (2016) 05015005.
[5] S.A.H. Mohsan, et al., Towards the unmanned aerial vehicles (UAVs): A comprehensive review, Drones 6 (6) (2022) 147.
[6] S.H. Alsamhi, et al., UAV computing-assisted search and rescue mission framework for disaster and harsh environment mitigation, Drones 6 (7) (2022) 154.
[7] M. Rahnemoonfar, T. Chowdhury, R. Murphy, RescueNet: A High Resolution Post Disaster UAV Dataset for Semantic Segmentation, UMBC Student Collection (2021).
[8] Van Ackere, S., et al. *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*. 2019. International Society for Photogrammetry and Remote Sensing (ISPRS).
[9] Z. Zhou, J. Gong, Automated analysis of mobile LiDAR data for component-level damage assessment of building structures during large coastal storm events, Computer-Aided Civil and Infrastructure Engineering 33 (5) (2018) 373–392.
[10] J. Gong, A. Maher, Use of mobile lidar data to assess hurricane damage and visualize community vulnerability, Transportation Research Record 2459 (1) (2014) 119–126.
[11] Oludare, V., et al. *Semi-supervised learning for improved post-disaster damage assessment from satellite imagery*. in *Multimodal Image Exploitation and Learning 2021*. 2021. SPIE.
[12] Gupta, R. and M. Shah. Rescuenet: Joint building segmentation and damage assessment from satellite imagery. in 2020 25th International Conference on Pattern Recognition (ICPR). 2021. IEEE.
[13] Weber, E. and H. Kané, *Building disaster damage assessment in satellite imagery with multi-temporal fusion.* arXiv preprint arXiv:2004.05525, 2020.
[14] Kijewski-Correa, T., et al., *StEER-Hurricane Michael.* 2020.
[15] S. Dyke, X. Liu, Learning from Earthquakes Using the Automatic Reconnaissance Image Organizer (ARIO). In *Proceeding of the 17th World Conference on Earthquake Engineering*, 2021.
[16] V.N. Gudivada, V.V. Raghavan, Content based image retrieval systems, Computer 28 (9) (1995) 18–22.
[17] Y. Rui, et al., Relevance feedback: a power tool for interactive content-based image retrieval, IEEE Transactions on Circuits and Systems for Video Technology 8 (5) (1998) 644–655.
[18] J.Z. Wang, et al., Content-based image indexing and searching using Daubechies' wavelets, International Journal on Digital Libraries 1 (1998) 311–328.
[19] Y. Kiyoki, T. Kitagawa, T. Hayama, A metadatabase system for semantic image search by a mathematical model of meaning, ACM Sigmod Record 23 (4) (1994) 34–41.
[20] Kennedy, L. and A. Hauptmann, *LSCOM lexicon definitions and annotations (version 1.0)*. 2006.
[21] Wang, X.-J., et al. *Annosearch: Image auto-annotation by search.* in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. 2006. IEEE.
[22] Fergus, R., et al. *Learning object categories from google's image search*. in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*. 2005. IEEE.
[23] D.T. Nguyen, et al., Damage assessment from social media imagery data during disasters. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 2017.
[24] A.K. Gautam, et al., Multimodal Analysis of Disaster Tweets, IEEE, 2019.
[25] E. Weber, et al., Detecting natural disasters, damage, and incidents in the wild. European Conference on Computer Vision, Springer, 2020.
[26] Weber, E., et al., *Incidents1M: a large-scale dataset of images with natural disasters, damage, and incidents.* arXiv preprint arXiv:2201.04236, 2022.
[27] J. Redmon, et al., You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
[28] K. He, et al., Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision*, 2017.
[29] A. Vaswani, et al., Attention is all you need, Advances in Neural Information Processing Systems 30 (2017).
[30] Y. Pi, N.D. Nath, A.H. Behzadan, Convolutional neural networks for object detection in aerial imagery for disaster response and recovery, Advanced Engineering Informatics 43 (2020) 101009.
[31] Lam, D., et al., *xview: Objects in context in overhead imagery*. arXiv preprint arXiv: 1802.07856, 2018.
[32] Z. Zhou, J. Gong, Automated residential building detection from airborne LiDAR data with deep neural networks, Advanced Engineering Informatics 36 (2018) 229–241.
[33] T.-N. Le, et al., Anabranch network for camouflaged object segmentation, Computer Vision and Image Understanding 184 (2019) 45–56.
[34] Fan, D.-P., et al. *Camouflaged object detection*. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.
[35] B. Mishra, et al., Drone-surveillance for search and rescue in natural disaster, Computer Communications 156 (2020) 1–10.
[36] M.B. Bejiga, et al., A convolutional neural network approach for assisting avalanche search and rescue operations with UAV imagery, Remote Sensing 9 (2) (2017) 100.
[37] D. Hu, et al., Detecting, locating, and characterizing voids in disaster rubble for search and rescue, Advanced Engineering Informatics 42 (2019) 100974.
[38] T. Bloch, R. Sacks, O. Rabinovitch, Interior models of earthquake damaged buildings for search and rescue, Advanced Engineering Informatics 30 (1) (2016) 65–76.
[39] Z. Zheng, et al., Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters, Remote Sensing of Environment 265 (2021) 112636.
[40] T. Liu, L. Yang, D. Lunga, Change detection using deep learning approach with object-based image analysis, Remote Sensing of Environment 256 (2021) 112308.
[41] Amit, S.N.K.B. and Y. Aoki. *Disaster detection from aerial imagery with convolutional neural network*. in *2017 international electronics symposium on knowledge creation and intelligent computing (IES-KCIC)*. 2017. IEEE.
[42] J. Chen, Y.K. Cho, CrackEmbed: Point feature embedding for crack segmentation from disaster site point clouds with anomaly detection, Advanced Engineering Informatics 52 (2022) 101550.
[43] F. Ni, et al., A generative adversarial learning strategy for enhanced lightweight crack delineation networks, Advanced Engineering Informatics 52 (2022) 101575.
[44] C. Koch, et al., A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure, Advanced Engineering Informatics 29 (2) (2015) 196–210.
[45] M. Rahnemoonfar, et al., Floodnet: A high resolution aerial imagery dataset for post flood scene understanding, IEEE Access 9 (2021) 89644–89654.
[46] Munawar, H.S., et al. *After the flood: A novel application of image processing and machine learning for post-flood disaster management*. in *Proceedings of the 2nd International Conference on Sustainable Development in Civil Engineering (ICSDC 2019), Jamshoro, Pakistan*. 2019.
[47] J. Chen, et al., Augmenting a deep-learning algorithm with canal inspection knowledge for reliable water leak detection from multispectral satellite images, Advanced Engineering Informatics 46 (2020) 101161.
[48] Gupta, R., et al., *xbd: A dataset for assessing building damage from satellite imagery*. arXiv preprint arXiv:1911.09296, 2019.
[49] Gupta, R., et al. *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*. 2019.
[50] Zhu, X., J. Liang, and A. Hauptmann. *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*. 2021.
[51] Yeom, J., et al. *Hurricane building damage assessment using post-disaster UAV data*. in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. 2019. IEEE.
[52] S.M.S.M. Daud, et al., Applications of drone in disaster management: A scoping review, Science & Justice 62 (1) (2022) 30–42.
[53] M.K. Lindell, C.S. Prater, Assessing community impacts of natural disasters, Natural Hazards Review 4 (4) (2003) 176–185.
[54] Wu, J., et al. *Deep multiple instance learning for image classification and auto-annotation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
[55] S. Bosse, et al., Deep neural networks for no-reference and full-reference image quality assessment, IEEE Transactions on Image Processing 27 (1) (2017) 206–219.
[56] K. Thyagharajan, G. Kalaiarasi, A review on near-duplicate detection of images using computer vision techniques, Archives of Computational Methods in Engineering 28 (2021) 897–916.
[57] OpenAI, *clip-vit-base-patch32. https://openai.com/research/clip*.
[58] F. Chiusano, Two minutes NLP — Semantic search of images with CLIP and Unsplash, Meidum (2022).
[59] R. Magazine, Damage, Control. (2018).
[60] Pennington, J., R. Socher, and C.D. Manning. *Glove: Global vectors for word representation*. in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 2014.
[61] Mikolov, T., et al., *Efficient estimation of word representations in vector space.* arXiv preprint arXiv:1301.3781, 2013.
[62] T. Brown, et al., Language models are few-shot learners, Advances in Neural Information Processing Systems 33 (2020) 1877–1901.
[63] Devlin, J., et al., *Bert: Pre-training of deep bidirectional transformers for language understanding.* arXiv preprint arXiv:1810.04805, 2018.
[64] W.T. Chuang, et al., A fast algorithm for hierarchical text classification. Data Warehousing and Knowledge Discovery: Second International Conference, DaWaK 2000 London, UK, September 4–6, 2000 Proceedings 2, Springer, 2000.
[65] datnnt1997, *CygnusX1*. 2021. https://github.com/datnnt1997/CygnusX1.
[66] Kim, Y., *AutoCrawler*. 2018. https://github.com/YoongiKim/AutoCrawler.
[67] H. Kong, H.C. Akakin, S.E. Sarma, A generalized Laplacian of Gaussian filter for blob detection and its applications, IEEE Transactions on Cybernetics 43 (6) (2013) 1719–1733.
[68] Spichkova, M., et al., *Easy mobile meter reading for non-smart meters: Comparison of aws rekognition and google cloud vision approaches.* arXiv preprint arXiv:1910.12617, 2019.
[69] Al-Omair, O.M. and S. Huang. *A comparative study on detection accuracy of cloud-based emotion recognition services*. in *Proceedings of the 2018 International Conference on Signal Processing and Machine Learning*. 2018.
[70] J. Schütz, Kontinuierliche versus diskrete Modelle der Rekognition und des Quellengedächtnisses, Logos Verlag Berlin GmbH, 2011.
[71] H. Jalloul, et al., A systematic approach to identify, characterize, and prioritize the data needs for quantitative sustainable disaster debris management, Resources, Conservation and Recycling 180 (2022) 106174.
[72] Jalloul, H., et al. *Toward Sustainable Management of Disaster Debris: Three-Phase Post-Disaster Data Collection Planning*. in *Construction Research Congress 2022*. 2022.
[73] V. Kumar, et al., Comprehensive overview of flood modeling approaches: A review of recent advances, Hydrology 10 (7) (2023) 141.
[74] J. Ma, et al., Real-time detection of wildfire risk caused by powerline vegetation faults using advanced machine learning techniques, Advanced Engineering Informatics 44 (2020) 101070.