International Journal of Disaster Risk Reduction 51 (2020) 101760

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: http://www.elsevier.com/locate/ijdrr

Leveraging multimodal social media data for rapid disaster damage assessment

Haiyan Hao [a], Yan Wang, PhD, Assistant Professor [b,*]

[a] Department of Urban and Regional Planning and Florida Institute for Built Environment Resilience, College of Design, Construction and Planning, University of Florida, 1480 Inner Road, Gainesville, FL, 32601, USA

[b] Department of Urban and Regional Planning and Florida Institute for Built Environment Resilience, University of Florida, P.O. Box 115706, Gainesville, FL, 32611, USA

ARTICLE INFO

Keywords:
Computer vision
Damage assessment
Disaster management
Multimodal data analysis
Social media
Text mining

ABSTRACT

During disaster response and recovery stages, stakeholders including governmental agencies collect disaster's impact information to inform disaster relief, resource allocation, and infrastructure reconstruction. The damage data collected using field surveys and satellite imagery are often not available immediately after a disaster while rapid information is crucial for time-sensitive decision makings. Some researchers turned to social media for real-time situational information of disaster damage. However, existing damage assessment research mostly focused on single data modality (i.e. text or image) and made coarse-grained predictions, which limited their practical applications in assisting city-level operations. The difficulties of retrieving useful information from vast noisy social media data have been outlined by many studies. Thus, we propose a data-driven method to locate and assess disaster damage with massive multimodal social media data. The method splits and processes two data modalities, i.e. texts and images, using two modules. The image analysis module uses five machine learning classifiers that are organized in a hierarchical structure. The text analysis module uses a keyword search-based method. They together mine various damage information including hazard types (e.g. wind and flood), hazard severities, damage types (e.g. infrastructure destruction and housing damage). The method is applied and evaluated with two recent hurricane events. In practice, the method acquires damage information throughout extreme events and supplements conventional damage assessment methods. It enables the rapid damage information access and disaster response for both first responders and the general public. The research effort contributes to achieving more transparent and effective disaster relief activities.

1. Introduction

In each year, different types of disasters including wildfires, storms, droughts, and flooding jointly cause economic losses up to hundreds of billions of dollars and claim many lives in the United States [1]. Partly due to the impact of climate change and global warming, the past decade has witnessed increased frequencies and more severe outcomes of natural disasters worldwide [1,2]. In response, human society pays tremendous effort to minimize the negative impacts of natural disasters. Disaster management thus devotes to reducing disaster risk and relieving human suffering, with four continuous phases, i.e. mitigation, preparedness, response, and recovery [3]. Disaster damage data such as the location and extent of damaged facilities is critical for disaster management operations, which is often collected in the response and recovery phases. The collected data helps official agencies convey situational information to the general public, evacuate and rescue people in affected areas, allocate resources, and plan for future repair and reconstruction.

The disaster damage data is conventionally collected with field surveys, post-disaster satellite imagery, or Unmanned Aerial Vehicle (UAV) imagery [4–6]. However, none of these authoritative sources can be accessed immediately after the disaster's occurrence due to the restricted atmospheric or environmental conditions for the deployment of labor and equipment [7]. While the timely knowledge of disaster environments and situations is crucial for emergency managers to intervene in the disaster response phase early and plan for time-sensitive operations such as rescuing affected people and optimizing shelter locations.

* Corresponting author.
E-mail addresses: hhao@ufl.edu (H. Hao), yanw@ufl.edu (Y. Wang).

https://doi.org/10.1016/j.ijdrr.2020.101760
Received 23 March 2020; Received in revised form 25 June 2020; Accepted 2 July 2020
Available online 22 July 2020
2212-4209/

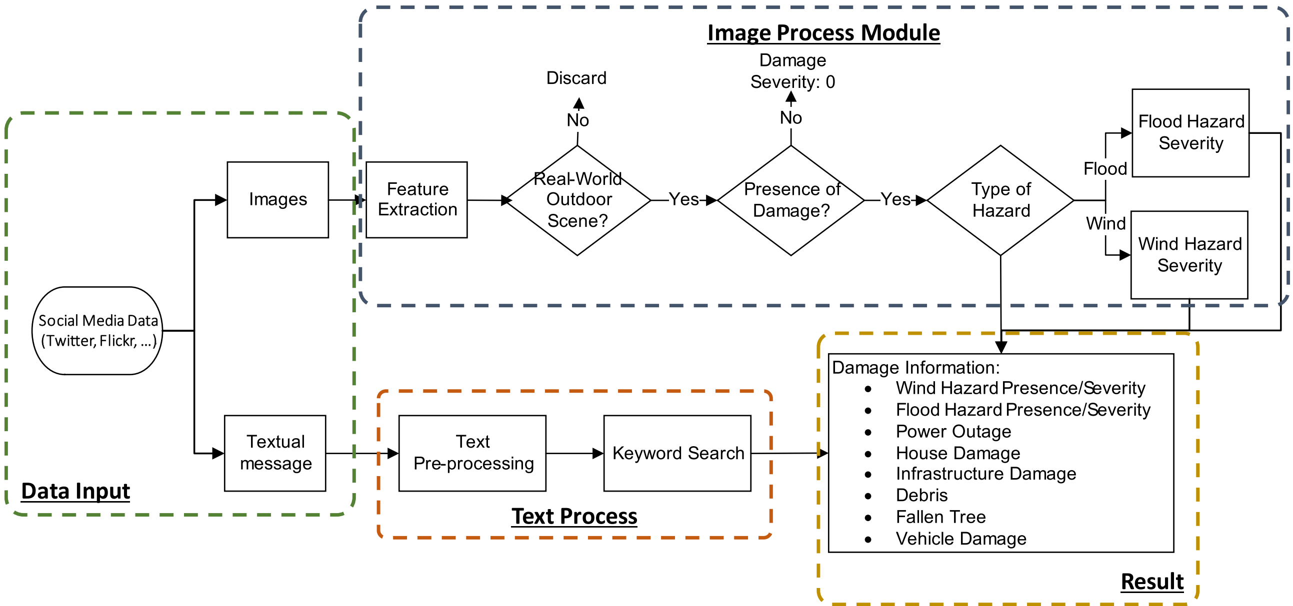

Fig. 1. The pipeline of proposed multimodal data-driven damage assessment method.

Table 1
Images used for Developing Different Classifiers.

Classification Task	Labels	# of Images
Perceived outdoor environment classifier	Total	1230
	Positive (show perceived outdoor environment)	585
	Negative (show other contents, e.g. selfies, maps.)	645
Hazard presence classifier	Total	1089
	Positive (show the evident wind or flood hazard)	577
	Negative (show normal environmental condition)	512
Hazard type classifier	Total	655
	Wind hazard	273
	Flood hazard	269
	Wind and flooding hazard	35
	None	78
Hazard severity classifier	Wind hazard:	
	Little to none	160
	Minor	121
	Severe	187
	Flood hazard:	
	Little to none	179
	Minor	139
	Severe	165

Table 2
Classifier selection and performance for each classification task.

Classification Task	Selected Classifier	Accuracy
Perceived outdoor environment classifier	SVM (Linear)	94.31%
Hazard presence classifier	LR (Binary)	88.53%
Hazard type classifier	ANN	82.01%
Wind hazard severity classifier	LR (Multinomial)	74.24%
Flood hazard severity classifier	LR (Multinomial)	83.94%

The necessity of timely disaster situation knowledge boosts research exploring user-generated data such as social media posts for disaster management applications [8,9]. Some approaches have been developed to analyze the content of social media posts with text mining and computer vision techniques. Compared to conventional data sources,

Table 3
Keyword (stemmed) search table for different damage types.

Damage Type		Word Lists
Power Outage	Object	['power', 'powerlin', 'electr', 'nopow']
	Description	['fix', 'destroy', 'broken', 'damag', 'gone', 'knock', 'lost', 'without power', 'not have power', "don't have", 'restor', 'cut', 'outag', 'no power', 'nopow', 'wait for power', 'power back', 'outta', 'lack of', 'out of', 'went off', 'flick']
Vehicle Damage	Object	['car', 'truck', 'van', 'vehicl', 'bu', 'motorcycl']
	Description	['flip', 'overturn', 'smash', 'damag', 'submerg', 'flood-damag', 'flood', 'lost', 'destroy', 'wreck', 'in the water', 'under water', 'in water']
House/Building Damage	Object	['roof', 'window', 'hous', 'home', 'wall', 'build', 'basement', 'porch', 'yard', 'door', 'backyard']
	Description	['crack', 'lose', 'lost', 'destroy', 'damag', 'destruct', 'corrupt', 'flood-damag', 'flood', 'rip', 'blow', 'reconstruct', 'clean', 'rebuild', 'reconstruct', 'pull', 'blown', 'collaps', 'submerg', 'shake', 'water in', 'water on', 'water over']
Infrastructure Destruction	Object	['street', 'road', 'dam', 'bridg', 'power cabl', 'trail', 'parkway', 'rd', 'hwi', 'highway', 'fwi', 'freeway', 'dr', 'drive', 'blvd', 'boulevard', 'ramp', 'lane', 'mainlan', 'curb', 'expressway', 'school', 'church', 'airport', 'chemic plant', 'roadway']
	Description	['destroy', 'damag', 'destruct', 'corrupt', 'flood-damag', 'flood', 'collaps', 'submerg', 'fallen', 'high water', 'under water', 'water on', 'water in', 'water over', 'underwat']
Debris	Object	['debri']
	Description	–
Tree Fall	Object	['tree', 'branch', 'limb']
	Descriptive	['fall', 'uproot', 'down', 'fallen']

social media provides real-time streaming data with various data formats. Thus, the data can be collected throughout the disaster event and the analysis can also be performed rapidly and even in real time. In the context of assessing disaster damage, social media textual contents can describe experienced or observed damage of affected people while images may represent ground-level scenes comparable to field assessors'

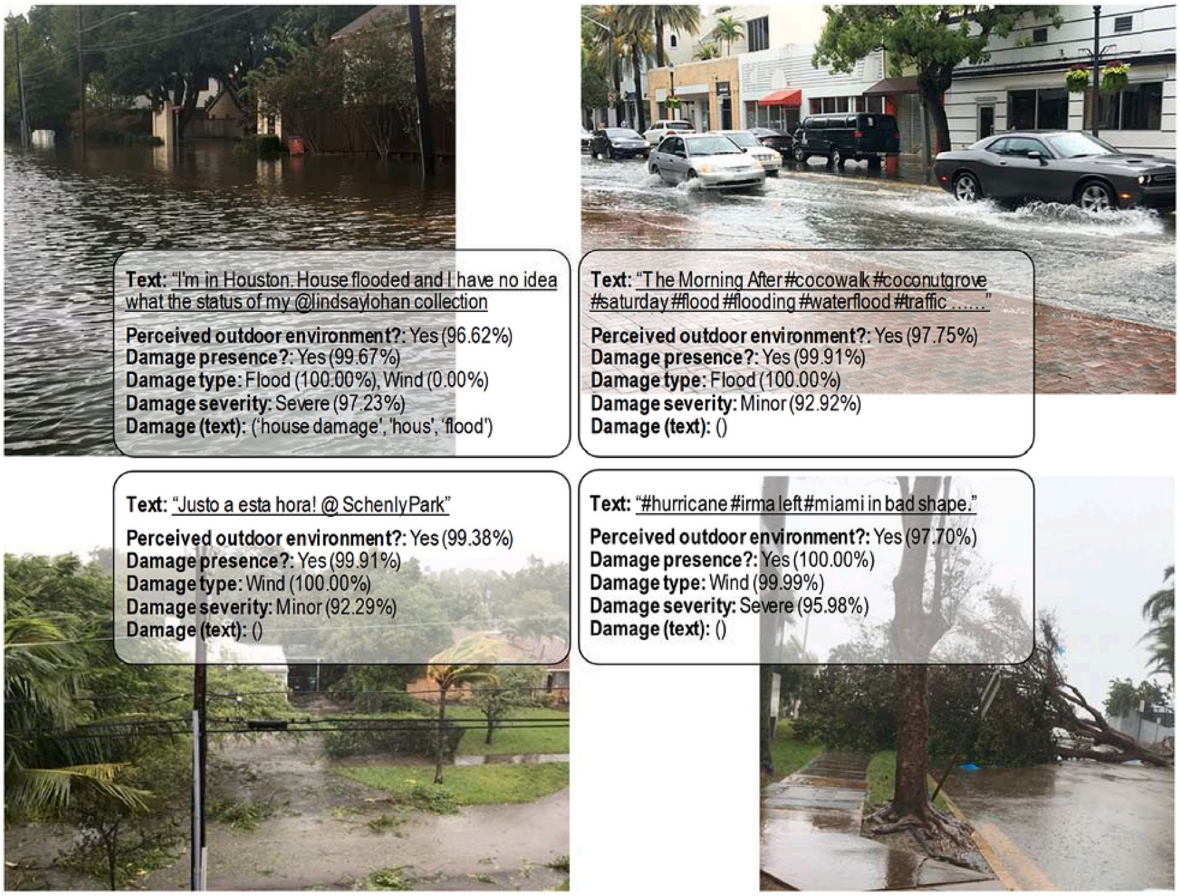

Fig. 2. Examples of damage information mining method output.

Table 4
Geotagged social media data volume.

Source	Count of Records	Temporal Span
Miami (Irma)		
Twitter	1555 images and 4006 texts	09/10/2017–09/23/2017
Flickr	94 images and associated textual descriptions	09/10/2017–09/23/2017
Houston (Harvey)		
Twitter	8642 images and 24,696 texts	08/25/2017–09/07/2017
Flickr	1011 images and associated textual descriptions	08/25/2017–09/07/2017

perceptions. Social media data are also less susceptible to adverse environmental conditions and do not require the extra deployment of field assessors or UAV pilots for data collection.

Previous studies assessing disaster damage based on social media data mostly focus on a single modality of data (e.g. textual or visual data) mine a single type of damage information, e.g. wildfire perimeters or inundation depth [7,10]. Although the massive social media data are generally considered as a source of big data, the damage assessment uses posts that are on-topic, posted in affected areas, and include location information. Researchers may still experience data shortage when searching for specific information in social media data at fine spatial- or temporal-scale. The damage-related information can reside in either texts or images. People from different groups may also prefer different social media platforms. Prior studies based on single modality and source did not maximize the use of available real-time social media data. As a result, these studies often aggregated and reported results at coarse spatial scales (e.g. state- and regional-level) [11,12], which is insufficient for practical applications such as assisting the city-level emergency operations.

Extracting useful information from vast and noisy background messages for fine-grained damage mapping is challenging. Moreover, useful information can be delivered in different formats and describe disaster damage from various perspectives. In this research, we propose a data-driven method to automatically analyze the massive raw crawled social media data and extract various damage information from social media texts and images. The method divides the overall task into steps and implements them with two modules. Each step is responsible for a single task such as filtering, classification, or keyword search. Jointly, they extract various damage information from social media posts including hazard types (i.e. wind and flood), hazard severities, and some specific damage types such as power outage, infrastructure destruction, and house/building damage. We applied the proposed method in pinpointing damage locations and assessing damages' extent in two recent hurricane damage cases, i.e. the city of Miami impacted by Hurricane Irma and the city of Houston impacted by Hurricane Harvey. The proposed method offers an additional data acquisition approach that supplements conventional damage assessment. The method identifies eyewitness reports (i.e. social media posts) of affected people that show the impact of disasters on human and community, which could be useful for humanitarian operations and emergency decision making.

2. Relevant work

An array of approaches is used to acquire disaster damage information. In practice, official agencies send human assessors to disaster sites and collect detailed damage information such as the location, number, type, and severity of damaged buildings and infrastructures [5]. The

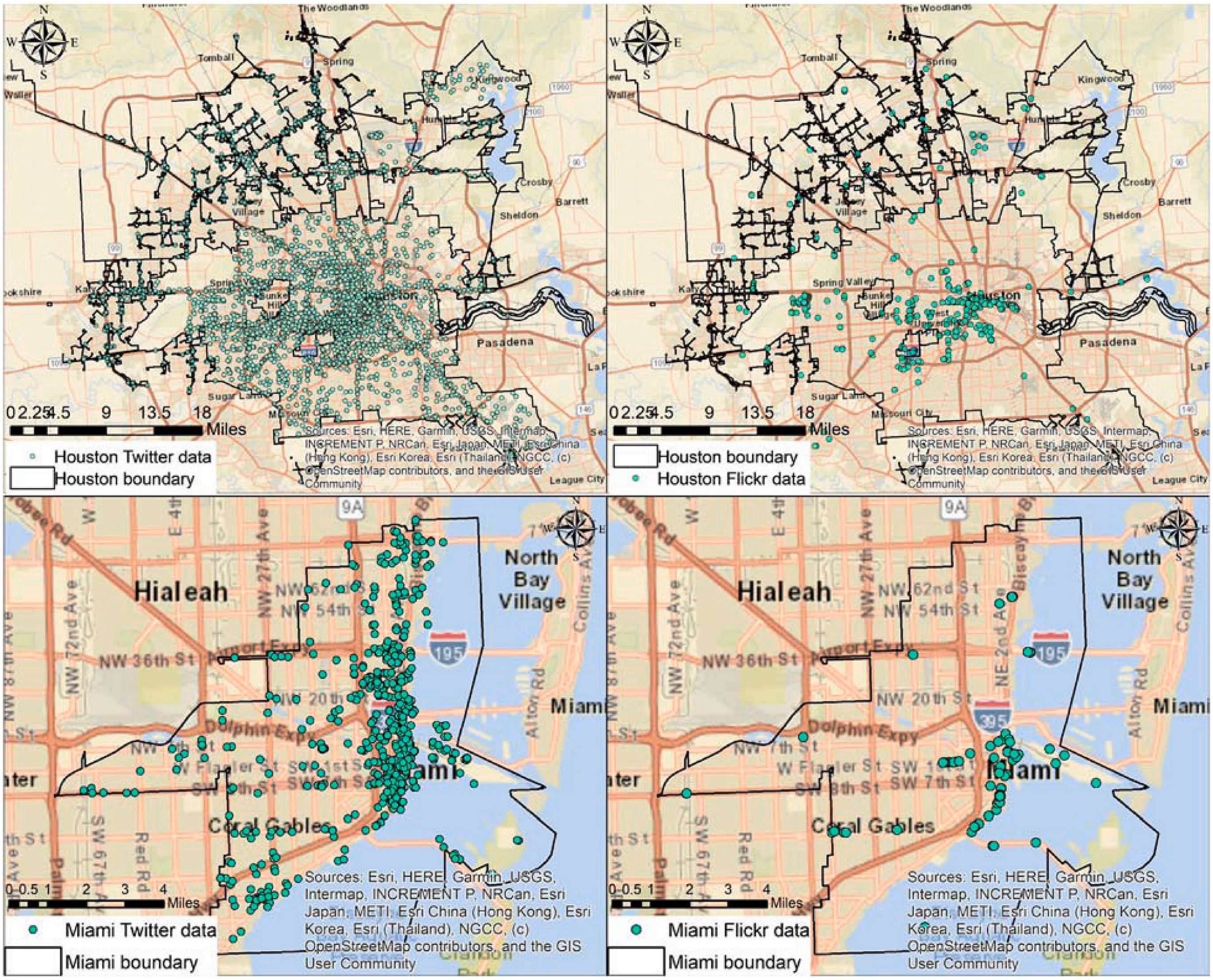

Fig. 3. Spatial Distribution of Tweets (left column) and Flickr Photos (right column) in Houston (top row) and Miami (bottom row) during Hurricanes.

Table 5
Counts of different reported damage information.

Damage Type	# of Reports (Miami Case)	# of Reports (Houston Case)
Wind Hazard (Image)	41	20
Flood Hazard (Image)	21	180
Power Outage (Text)	11	11
Vehicle Damage (Text)	0	10
House/Building Damage (Text)	13	37
Infrastructure (Text)	5	539
Fallen Tree (Text)	2	1
Debris (Text)	1	7
Flood Hazard (Text)	5	563
Wind Hazard (Text)	5	1

field survey yields reliable and detailed damage information but is labor-intensive and time-consuming. The field survey also inevitably exposes human assessors to dangerous environments. Recently, some researchers leveraged high-resolution satellite and UAV imagery for rapid assessment [13,14]. The broad-view imagery provides overviews of disaster-affected areas, and the high resolution enables damage assessment for individual structures. However, satellite and UAV imagery is not always available in the short aftermath of a major disaster. The deployment of UAVs should consider technical issues such as system reliability, power supply, and physical load [4]. Both data collection methods can be severely affected by adverse weather and atmospheric conditions accompanied by natural disasters such as dense clouds, heavy rains, and strong winds [4,15]. With the increasingly important role of social media and other Web 2.0 applications in disaster management, some researchers have taken advantage of the user-generated data and considered citizens as “human sensors” with five senses (i.e. touch, sight, hearing, taste, and smell) to perceive the external environment [16]. This novel conceptualization opens new avenues for disaster management and damage assessment research. We summarized damage assessment works leveraging social media data in four categories according to the analyzed data modalities, namely, activity-based, text-based, image-based, and multi-modal or fused methods.

2.1. Activity-based methods

The activity-based methods do not investigate the textual or image content of social media posts. Instead, they depict damage severity indirectly with metrics derived from tweeting frequencies. Thus the activity-based methods are computationally simple and often output

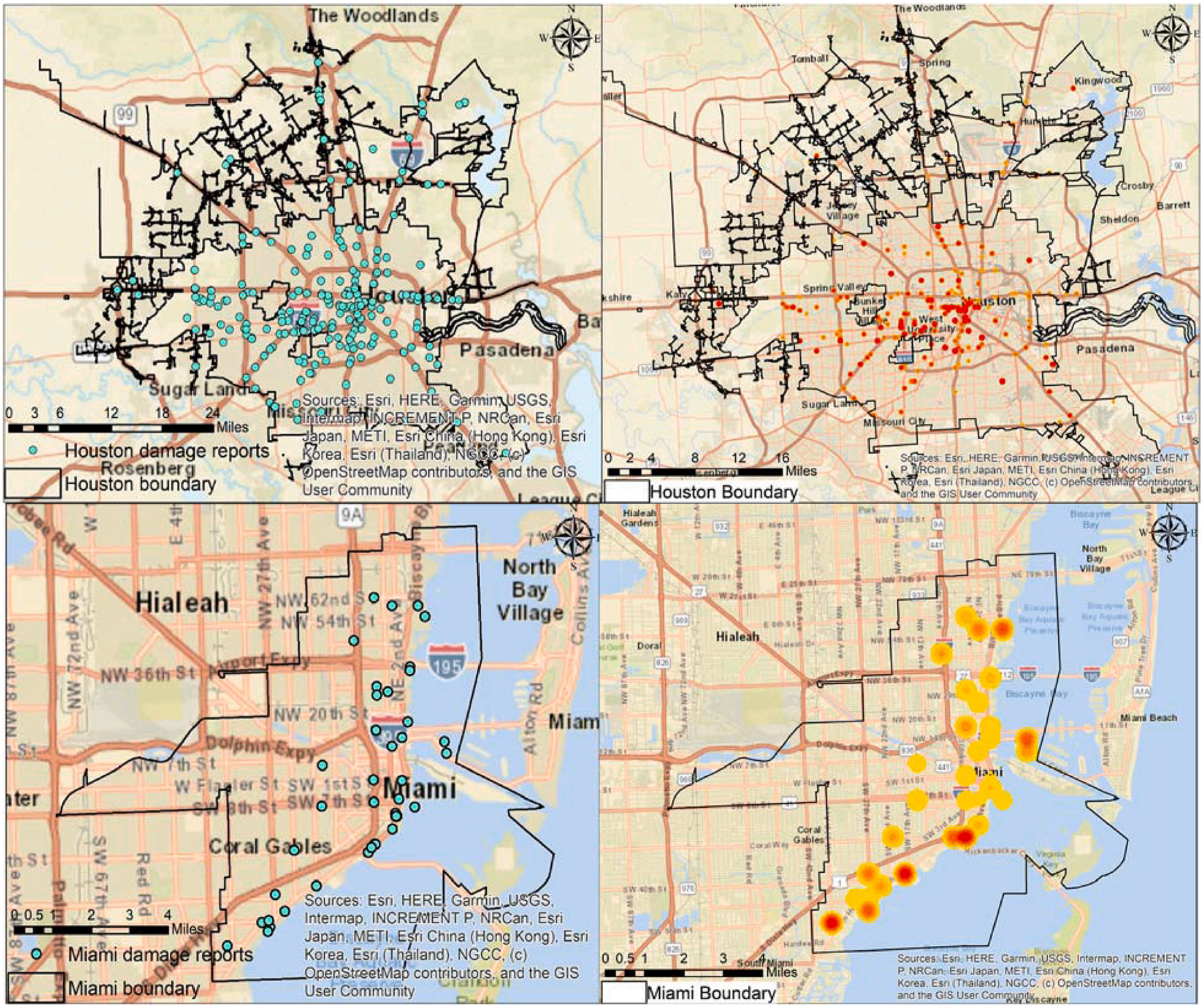

Fig. 4. Spatial distribution and density map of damage reports.

results in coarser spatial levels such as ZCTA- [17], city-, and county-level. For example, Kryvasheyeu et al. [18] identified a significant positive correlation between the number of hurricane-related tweets and the economic loss for New Jersey during Hurricane Sandy. Samuels et al. [19], considered possible loss of power or internet connection caused by the disaster, and used the variation of tweeting frequencies as the metric of damage severity. In the work of Zou et al., [20]; the ratio of disaster-related tweets and background tweets is considered as the damage severity metric. In general, the activity-based methods provide limited disaster damage information regarding both spatial and contextual details of the damage.

2.2. Text-based methods

Damage assessment methods based on social media textual posts have been studied mostly. Many have used sentiment analysis, topic modeling, and keyword search to retrieve relevant information. For instance, Wang & Taylor [12] found a significant negative correlation between average sentiment scores and the earthquake intensities in disaster-affected areas, while other studies on hurricanes found little association between the social media text sentiments and damage severities [18,20]. Some researchers used topic modeling to detect and locate trending events with geotagged social media texts [21]. This approach may not pinpoint the damage location accurately. A few recent studies included the geospatial characteristics in topic modeling to locate and track small-scale crises during disasters [22,23].

Additionally, some studies using keyword search-based methods developed pre-defined keyword lists or tables to identify useful textual posts. For example, Eilander et al. [10] used keywords such as "#(number) cm" and "#(number) m" to mine tweets containing flooding depth information and constructed situational inundation map accordingly. Smith et al. [24], considered tweets including words such as "knee-deep" and "waist-deep" as qualified reports that are used to verify the simulated flooding maps. Deng et al. [11], divided disaster damage and risk information into many subcategories such as infrastructure destruction, supply demands, and affected activities. With keyword lists developed for each subcategory, the method can identify and categorize qualified posts for more comprehensive situational knowledge. In general, keyword search-based methods often look for particular information from the social media text corpus. The colloquial nature of social media textual messages makes it expensive to enumerate all possible keywords and phrases related to a topic, although some researchers paid substantial amounts of time and effort to develop large lexicon tables for tweet collecting and mining [25].

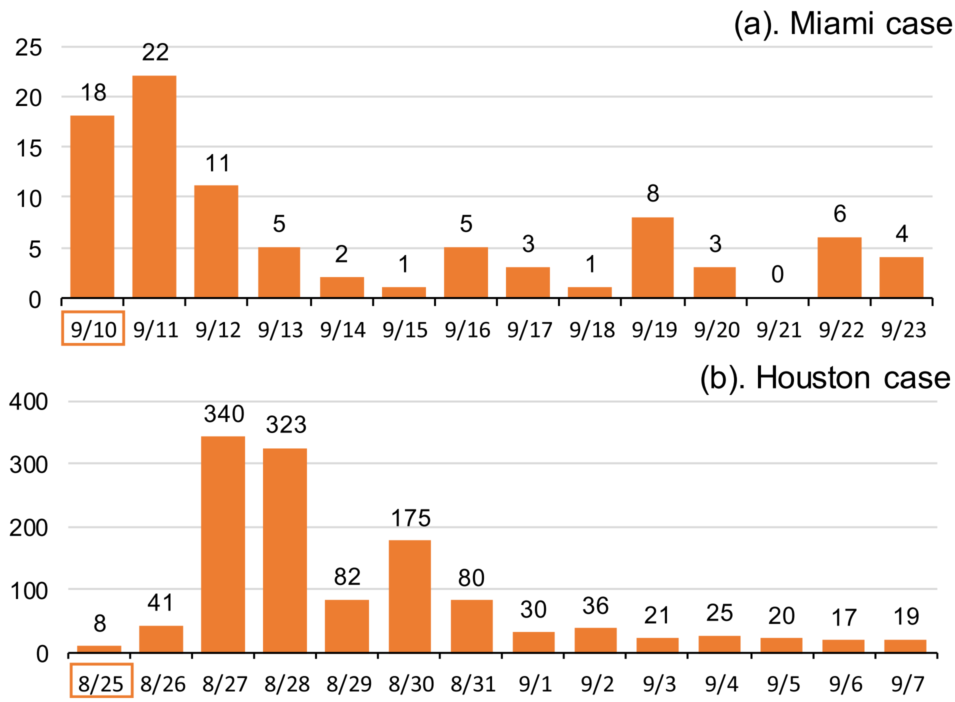

Fig. 5. Temporal distribution of damage reports.

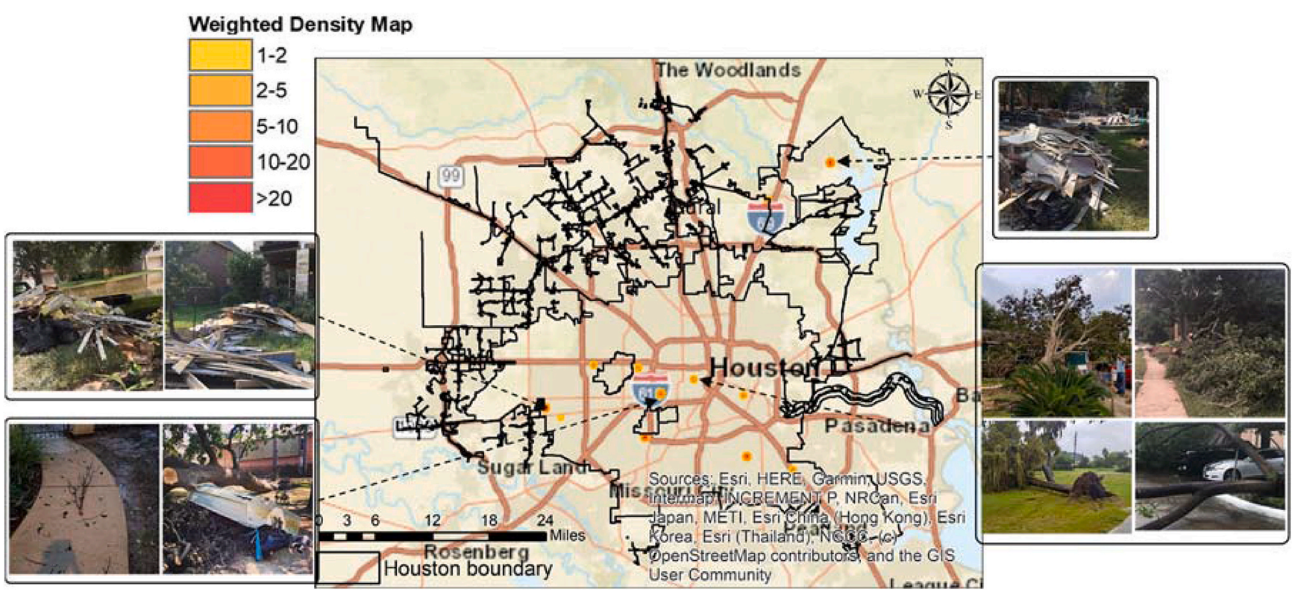

Fig. 6. Density map of wind hazard reports in Houston during Hurricane Harvey.

2.3. Image-based methods

Compared to social media texts, images convey more objective and more useful information [26], but much fewer studies explored the use of social media images and computer vision (CV) for damage assessment. The limited quantity of studies relied on transfer learning and convolutional neural networks (CNNs) for classifying damage severity levels and locating damage contents [27,28]. Transfer learning is an approach that leverages pre-trained deep learning models to solve new problems. The pre-training is often conducted on large datasets such as the well-known ImageNet dataset, which contains more than ten million images in twenty thousand categories. The pre-trained model is then re-trained on annotated datasets for the new problem. Both Alam et al. [29], and Nguyen et al. [28], annotated more than one hundred thousand social media images for the model retraining. ImageNet is a dataset for object classification (e.g. pizza, bird, and soccer), however, some researchers argued that social media images include more "scenes" (e.g. highway, bedroom, and parks) than "objects". Thus they experimented with the scene-level features that were obtained with models pre-trained for scene-related CV tasks such as scene recognition or scene parsing [30]. The scene-level features are also used in some built environment studies [31].

2.4. Multi-modal and fused methods

A few prior works harnessed multi-modal social media data for

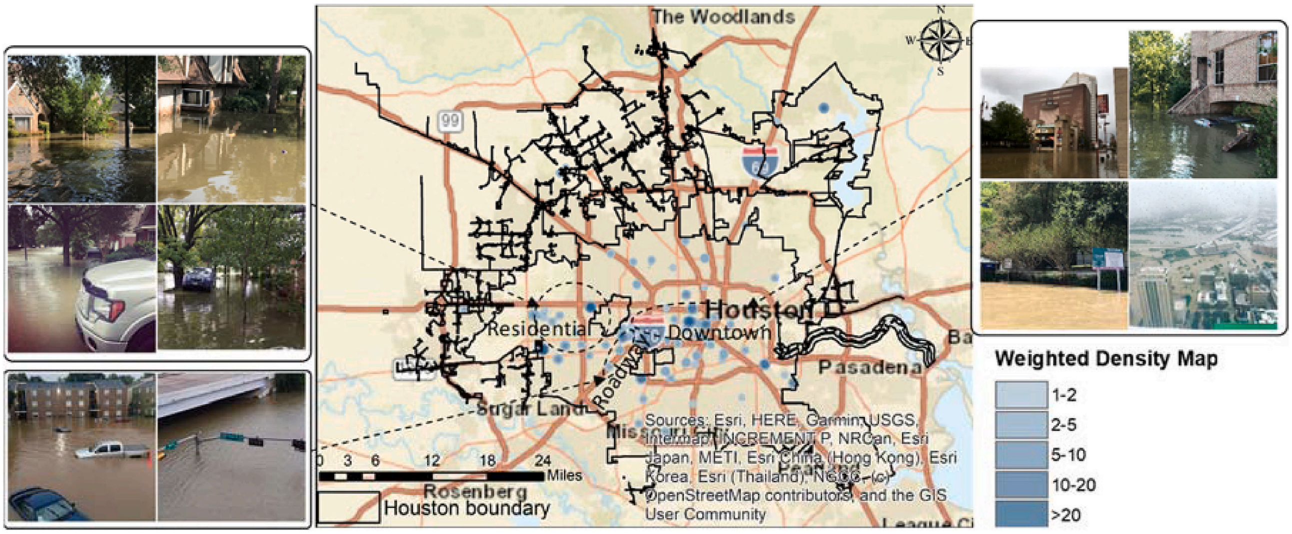

Fig. 7. Density map of flood hazard reports in Houston during Hurricane Harvey.

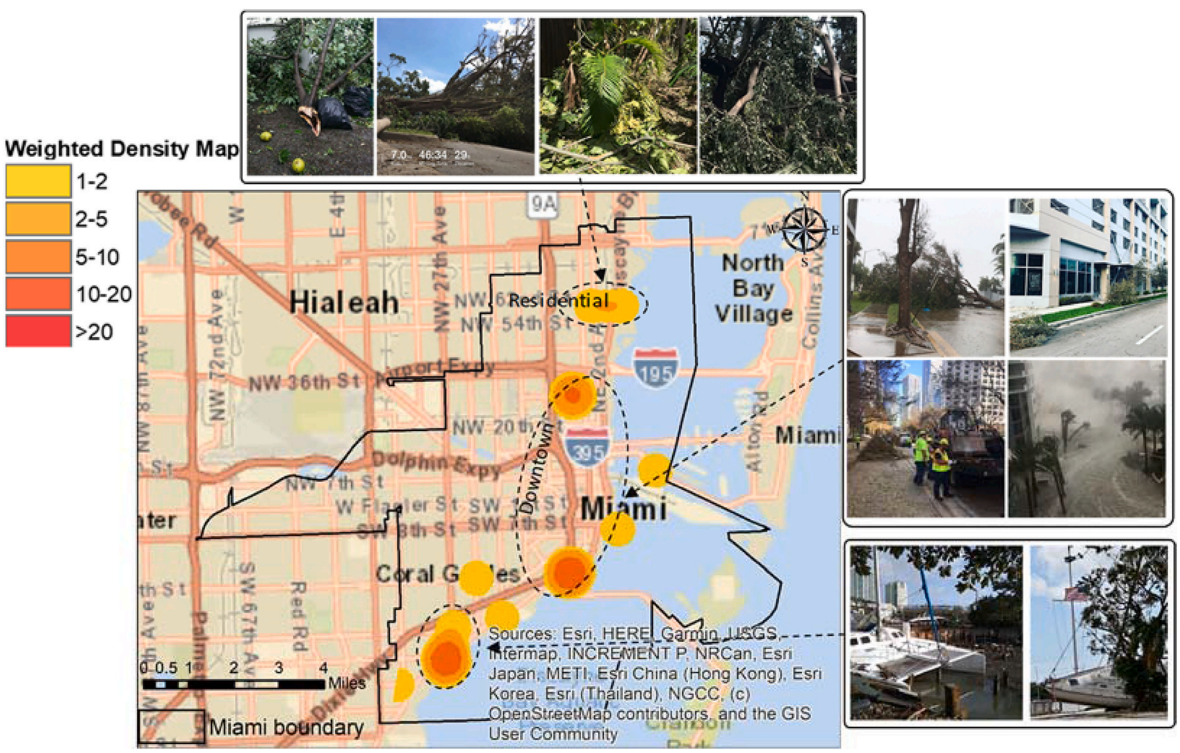

Fig. 8. Density map of wind hazard reports in Miami during Hurricane Irma.

damage type classification [32] and flood detection [33,34]. These works considered pre-trained CNNs as feature generators and used two CNNs to extract features from social media images and texts separately, in which the extracted features are concatenated for the classification task. This concatenation did not consider the correlation between different modalities. Pouyanfar et al. [35], accounted for this by fusing the visual and audio features into a ranking matrix with Multiple Correspondence Analysis. The final decision was trained with the fused matrix. Similarly, Lazaridou et al. [36] fused the visual and textual in a cross-modal mapping matrix for a multimodal skip-gram model. The model is used for image labeling and retrieval. Some research learned the common semantic information of different modalities by forcing the models to learn similar representations for different modalities [37,38] which were implemented by adjusting the objective function.

Research efforts are also found to integrate heterogeneous data sources for damage estimation. For example, Smith et al. [24], integrate the rainfall intensity data and social media posts for rapid flood mapping. The method iteratively simulated flooding maps with rainfall intensity data. Valid social media posts were used for verification purposes. The simulation result, which mostly conforms to social media posts, was adopted as the final output. In a method proposed by Huang et al., [39]; satellite imagery, high-resolution elevation data, and crowdsourced reports were integrated for flood mapping. The method calculated flooding probability layers for each validated crowdsourced report based on the elevation data and satellite imagery. The output flooding map was the weighted combination of different flooding

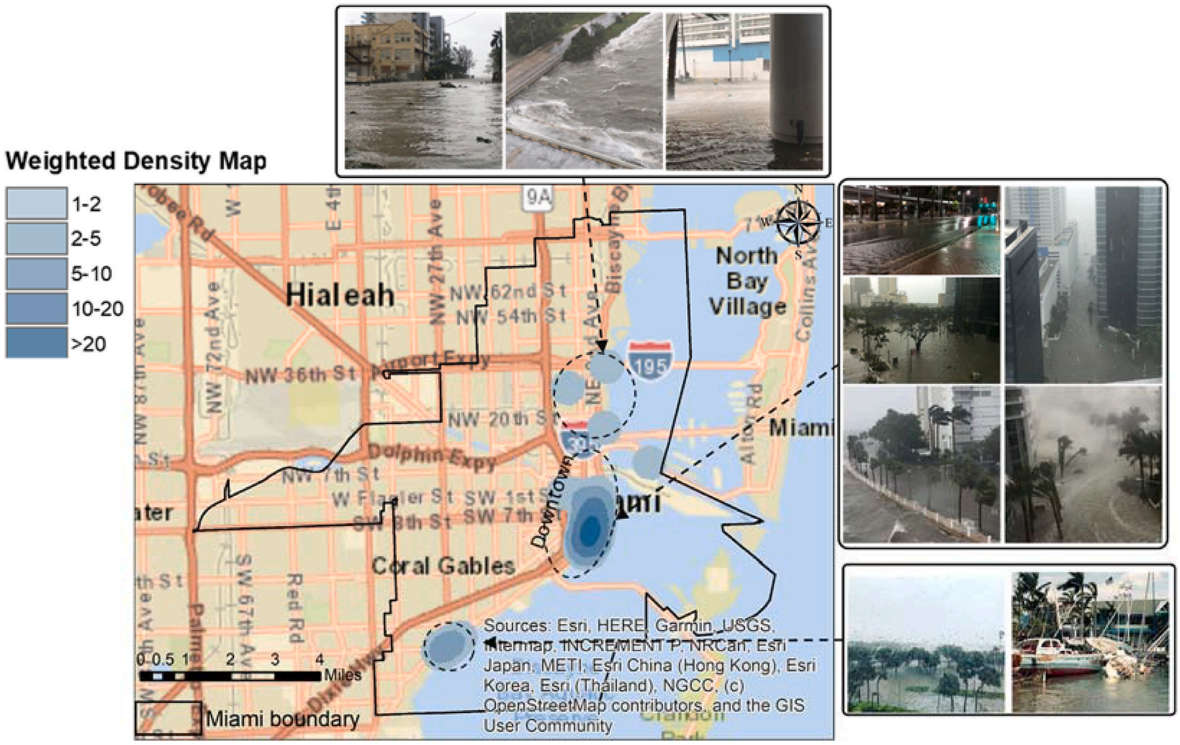

Fig. 9. Density map of flood hazard reports in Miami during Hurricane Irma.

probability layers.

In summary, we identified a few research gaps in existing social media data-based damage assessment works. Many methods relied on single data modality and text-based methods have been mostly studied. Social media textual contents are limited by short text length, high subjectivity, low information quality, and colloquial expression [40]. The damage information mined from textual posts could be inadequate and unreliable. Therefore, many activity- and text-based methods aggregate and report results at state-, county-, and ZCTA-level, with averaging or summation to alleviate individual estimation errors. Second, estimating the disaster damage with social media images is still challenging due to the loosely-defined forms of disaster damage, poor signal-to-noise ratio of raw crawled social media images, and the subjectivity of the damage severity level [28]. Although some researchers fused two data modalities in a single model for prediction, not many social media posts contain both data modalities. In addition, the semantic contents of social media texts and associated images are weakly correlated in many cases [41]. The fusion of textual and visual features may not yield promising results in these cases and possibly overlook valuable damage information that only resides in one modality. Moreover, these image-based and multimodal models developed for analyzing social media images are limitedly tested with lab-developed datasets, which comprised human-sorted images with relatively balanced sample distribution. However, the raw crawled social media images were extremely unbalanced and only included a small portion of on-topic images for damage assessment [42]. Models that perform well in lab-developed datasets still need to be validated over raw crawled social media images.

Therefore, we proposed a method that takes raw crawled multimodal social media data (i.e. both textual messages and images) and outputs various damage information. Instead of fusing different modalities in one model, we split and analyzed them in different modules acknowledging that the textual contents and images often convey different levels of information. The textual messages are subjective in describing damage situations while images do not tell abstractive information such as power outage. We used a pre-trained Resnet18 CNN to extract scene-level features from Twitter and Flickr images. For textual messages, we adapted a keyword search-based method considering that the disaster damage only occupied a small portion of the raw crawled texts. Other methods either infer the damage severities with indirect metrics (e.g. activity-based methods and sentiment analysis) or group textual messages for analyses (e.g. topic modeling), which may be manipulated by other dominating disaster-related topics rather than disaster damage.

3. Multimodal data-driven damage assessment method

The proposed method aims to automatically locate and summarize damage information from visual and textual contents of the massive social media posts. It consists of four modules (Fig. 1). The Data Input module is responsible for data collection. The collected images and texts are separated for analyses considering that their contents are often weakly correlated and describe damages from different perspectives. In the context of disaster events, social media images can convey situational information like ambient environmental conditions and hazard severities while texts often describe damaged objects and consequent impacts. Images then go through five image classifiers in the Image Process module and information including hazard types (i.e. wind and flood) and associated severity levels is extracted. The Text Process module uses a pre-defined keyword search table to examine whether the textual message contains any of six pre-defined damage types that are mostly shared in the textual part of social media posts, including power outage, vehicle damage, house/building damage, infrastructure destruction, fallen trees, and debris. The Result module integrates the mined damage information from the Image Process module and Text Process module.

3.1. Data input module

We mainly use two sources of social media data: Twitter and Flickr in this study. Twitter is a popular social media platform with more than 500 million tweets posted daily around the world. Twitter allows users to post a maximum of 140-character (expanded to 280-character in

Table 6
Examples of mined damage information from textual reports.

Damage type	Example of textual damage reports
Power Outage	*True positive:* e1: "Almost 10 days after #hurricaneharvey hit #Houston our building is still closed without power…" *False positive:* e2: "I know many are without power, if you know someone that may need help being evacuated please…" e3: "Trying to enjoy the electricity before it goes out @ [user] Houston, Houston, Texa" *Indeterminate:* e4: "Some of you can't live without power!!!Always have a good book on hand during storms!"
Vehicle Damage	*True positive:* e5: "Not pictured … half of my car sitting in water. @ [user] -…" *False positive:* e6: "This is the new must-have vehicle in Houston! Rule the flood! #hurricaneharvey #houstonstrong." e7: "People have lost everything, due to Hurricane Harvey's damage. Homes, lives, memories, vehicle…"
House Damage	*True positive:* e8: "The work doesn't stop! Carpet cleaning because of roof leaks. Support the long term recovery…" *False positive:* e9: "Flooded Home??Here are some helpful tips on what to do with your the wet or damaged materials…" e10: "This is Baby Susie, my new best friend. Her and her dad, Dennis lost their home in Meyerland" *Indeterminate:* e11: "I went out in the back yard in my wellies last night to inspect the flood and came back with…"
Infrastructure Damage	*True positive:* e12 "Closed due to flooding. in #Baytown on I-10 Baytown E Fwy Inbound between Crosby Lynchburg and Magnolia…" *False positive:* e13 "Day 4. More rain. Luckily the water runs off to the street. If the street floods, we're screwed" *Indeterminate:* e14 "Flooding, what flooding? @ Briargrove Drive Townhousescondominium Association"
Fallen Tree	*True positive:* e15 "Oak Tree down in Memorial Northwest subdivision @ Memorial Northwest, Spring, Texas" *False positive:* e16 "Water has receded down to the trees. Like the Stars Spangled Banne"
Debris	*True positive:* e17 "Harvey bags and more debris on front yards in the greater Meyerland borough of Houston a week…" *False positive:* e18 "Full day of unloading relief supplies & 18-wheeler trailers, cleaning flood debris at a partner" e19 " #debris removal guidelines #FEMA #harvey #texas @ Houston, Texas"

Table 7
Counts and percentages of true and false positive predictions of the Text Process Module.

Damage Type	True positive	False positive	Indeterminate	Total
Power Outage	18 (80.8%)	3 (13.6%)	1 (4.5%)	22
Vehicle Damage	6 (60.0%)	2 (20.0%)	2 (20.0%)	10
House Damage	22 (44.0%)	21 (42.0%)	7 (14%)	50
Infrastructure Damage	536 (98.5%)	5 (0.9%)	3 (0.6%)	544
Fallen Tree	3 (100.0%)	0 (0.0%)	0 (0.0%)	3
Debris	7 (87.5%)	1 (12.5%)	0 (0.0%)	8

November 2017) textual messages with links to images and videos. Around one percent of the messages are geotagged. During disaster events, many affected people tweet to report observations, express urgent needs, and seek helps, thus making Twitter an ideal data source for disaster management related research [43,44]. Flickr is also a popular social media platform, allowing users to share geotagged photos with optional textual descriptions. Flickr photos are mostly high-resolution images of the natural and built environment, thus have been used in some tourism studies [45]. Both data can be accessed via their Application Program Interface (API). As the Twitter streaming API only returns around one percent of tweets due to the rate limit [46], we restrict the crawled tweets to be geotagged. For each crawled tweet, we script to check whether it includes a link referring to images and download the image if it does. Flickr API can return archived geotagged photos. We set the time and location windows on Flickr API to access photos posted during disaster-affected periods and geotagged in affected areas.

3.2. Image Process Module

The processing of Twitter and Flickr images starts with converting images into numeric feature vectors and then uses five classifiers to extract the disaster damage information. The five classifiers are responsible for different tasks with a defined semantic hierarchy. Specifically: 1) one classifier filters out images showing a perceived built environment; 2) one classifier identifies images showing hazards; 3) one classifier classifies the damage type; and 4) two classifiers assign severity levels to identified wind and flood hazards respectively. The five classifiers are organized in a hierarchical structure (Fig. 1). This design helps locate the limited images showing exposed hazards from the sheer amount of posted images. The hierarchical processing can remove irrelevant images in early steps and keep remaining images similar in contents, i.e. displayed scenes and objects. As the succeeding classifiers work on the remaining images with less noise, the classifiers can better focus on learning the difference between positive and negative samples with defined semantic labels (Table 1). Therefore, the classifiers can achieve satisfactory performance even with relatively small training data. In fact, we use 1795 images for developing the five classifiers. These images were collected from social media images posted in affected areas during historical hurricane events and two existing databases: YFCC100 M and CrisisMMD [29,47]. Two researchers worked together to annotate the images. Table 1 summarizes the counts and labels of annotated images for different classification tasks. Some images are repeatedly used in developing different classifiers. Note that we also include the false positive images, predicted by preceding classifiers, in the training set of its succeeding classifier during the training. In this way, the succeeding classifier can gain a certain ability to remove false positive predictions and mitigate the Type I error with the proposed hierarchical structure [48]).

Feature extraction represents 2-D images with 1-D numeric features that are used for the following image classifications. In this study, we adopt a ResNet18 CNN pre-trained on the Places365 dataset as the feature generator, which yields the scene-level features. The ResNet18 CNN has a relatively compact size and performs well in many CV competitions [49] while the Places365 is a dataset consisted of more than 10 million images for scene recognition [50]. We made this selection after comparing it with two other feature generators, namely, an Inception-v3 CNN trained on ILSVRC dataset that extracts object-level features [51] and an AutoEncoder trained on the ADE20K dataset which also returns scene-level features [52]. The ResNet18 CNN performs equally or better than the other two feature generators for the five classification tasks and is hence selected for this study. Features are extracted as the output of the penultimate layer of the CNN. We adapted the ResNet18 CNN model with the PyTorch library in Python [53]. The features only need to be extracted once for each crawled image and are repeated for use by the following image classifiers.

The raw-crawled crowdsourced images can include screenshots of texts, maps, selfies, posters, cartoons, advertisements, and so on [42]. These images occupied a large portion of social media images but provide little information on disaster situations. So our first step is to sort out "informative" images that show perceived real-world environmental conditions. We further restrict the perceived environments to be outdoor environments. Images taken inside a building may reveal damage conditions for that individual building, however, they vary too much in terms of exhibited objects and backgrounds from the outdoor environment. Also, we found very limited images to train a separate classifier for it. We use 585 positive samples showing perceived outdoor environments and 645 negative samples for classifier development (Table 1). The samples are divided into 80% training set and 20% testing set, which is stratified according to labels (the same setting applies to the development of other image classifiers). We experiment with different machine learning classifiers including logistic regression (LR) models,

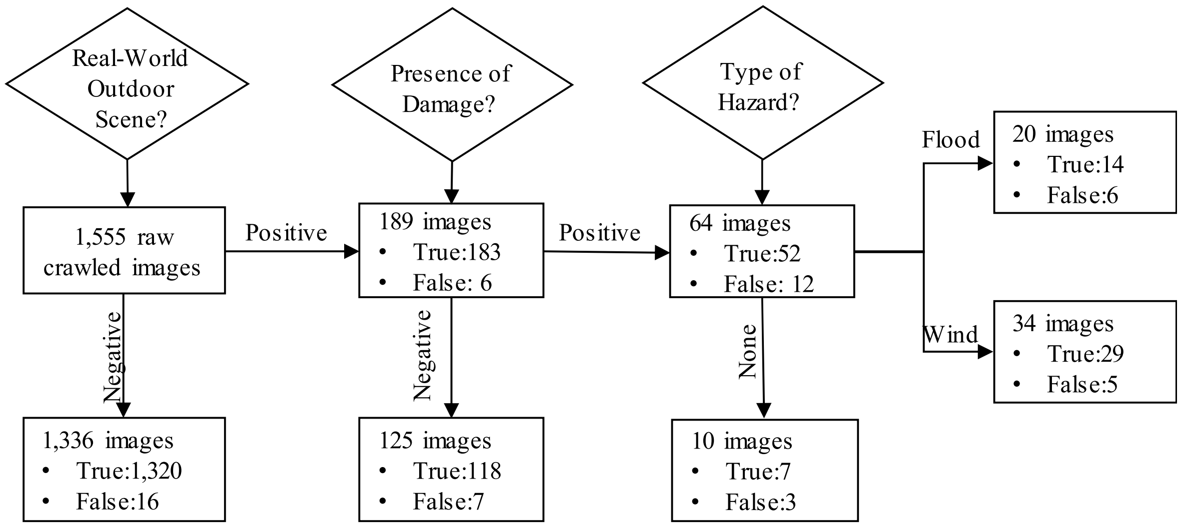

Fig. 10. The filtering and classification of Twitter images collected for the Miami case.

Table 8
Performance metrics of the hierarchical image classifiers on identifying hazards.

Hazard type	Precision	Recall	Specificity	Accuracy
Wind hazard	85.3%	78.4%	99.7%	99.2%
Flood hazard	70.0%	87.5%	99.6%	99.5%

decision trees, and support vector machines (SVMs). We used the Scikit-learn package in Python for the machine learning classifiers [54]. The SVM with a linear kernel achieves the highest classification accuracy of 94.31% (Table 2).

The next step distinguishes images showing evident natural hazards (positive) from images showing normal outdoor environment views (negative). Examples of common hazard content in social media images include inundated roads or uprooted trees. We select the binary LR model with 88.53% accuracy for this step (Table 2).

The third step details the hazard types, which serves as a prerequired procedure for classifying hazard severity levels. Some previous work modeled the disaster damage severity with social media images directly without considering distinct hazard types. However, images showing different hazard types are often highly disparate in terms of contents. For example, an image showing flood hazards usually contains water body while an image showing wind hazards is generally represented with uprooted trees or roofs with missing tiles. The classification of hazard type is a typical multi-label task as an image can include either wind or flood hazard or both. We use the artificial neural network (ANN) model for this classification task as ANN considers the possible correlation between different labels. The ANN model can accurately predict both labels for 82.01% of testing images (Table 2).

The final steps use two classifiers to determine the severity levels of wind/flood hazards. We assign each image showing hazard contents with one of the following three severity levels:

- Little to None: images show no damage, minor adverse weather conditions, or little damage that does not cause any economic loss or impact human activities (e.g. transportation);
- Minor: images show damage that requires money for repair or recovery, and partially affect human activities; and
- Severe: images show severe damage that suggests extreme environmental conditions, associated with significant economic loss, or severely impact human activities.

We select the multinomial logistic regression models for the severity level classification, which achieves 83.94% accuracy for flood hazard severity classification and 74.24% accuracy for wind hazard severity classification (Table 2).

Table 2 summarizes the selected classifiers and their associated accuracies for each classification task. Note that these accuracies are based on test images.

3.3. Text Process module

Social media users can describe their observations with different wording, phrases, and word sequences in textual posts. The colloquial expressions impede conventional keyword search-based methods with limited keywords and phrases from identifying much information, while developing a large keyword table takes enormous efforts and time. Therefore, we adopt a two-list search method in the Text Process module (Fig. 1) aimed to detect on-topic texts maximally with fewer efforts spent on enumerating possible word/phrases combinations for search. The two-list keyword search method identifies pre-defined damage types with one list collecting physical damaged objectives (e.g. roadway) and the other list collecting descriptive words of the damages (e.g. submerged). A textual post is considered to describe a type of disaster damage when it concurrently has words/phrases in both lists of the same damage type. This is performed after we remove the punctuations, URLs, emoji, numbers, and stopwords in texts as well as stemming each word to its root form. We defined six types of damage information that are mostly discussed on social media platforms including power outage, vehicle damage, house damage, infrastructure destruction, fallen trees, and debris. The fallen trees and debris are not damages themselves, however, they indicate the presence of strong winds or flood water in reported locations and take money for removal. The fallen trees may also damage properties like houses and vehicles.

The qualified textual damage reports should be the ones posted by affected people and discussing their own experiences or observations. These texts often express damages with colloquial phrases and details, such as the detailed damaged objects (e.g. “roofs” vs. “houses”), locations (e.g. “curbs” vs. roads), and damage forms (e.g. “blow”, vs. “destroy”). We particularly include these words in the keyword search table. Besides, we also refer to sources like EMTerms collection [25] for keywords collection. Table 3 presents the two-list keyword search table. Note some descriptive words such as “submerge” and “blown” indicate the type of hazards that cause the damage. We also collect this information from the textual reports.

3.4. Result module

The Result module integrates different types of hazards and damage information mined with the Image Process and Text Process modules. The outputs are individual-level estimations. Fig. 2 shows some example outputs. The percentages in brackets are estimated probabilities.

4. Empirical case study

4.1. Case descriptions and data collections

This section presented two case studies that applied the proposed method to assess damage situations in 1) the city of Miami, Florida as affected by Hurricane Irma; and 2) the city of Houston, Texas as affected by Hurricane Harvey. Both cities were impacted severely during hurricane events. Miami experienced sustained winds of 45–55 kt during Hurricane Irma [55]. The storm tide and urban runoff caused a 3–5 ft. inundation along Biscayne Bay shoreline and in downtown Miami. There were also widespread tree and power pole damage reported in the metro area [55]. Houston was less affected by the wind hazard during Hurricane Harvey compared to Miami. However, the exceptional rainfall and storm tides caused massive flooding that inundated nearly one-third of the city, disabled major roads, and cut power connections to households [56]. The severe impacts and versatile damage types make these two cases ideal for evaluating the proposed method.

We collected social media posts that were posted within two weeks after hurricanes' landfall from Twitter and Flickr. Table 4 shows the data sources, volumes, and temporal spans of collected data, and in Fig. 3 plotted the spatial distribution of collected data for the two cases. Retweets were not included in the analyses.

4.2. Identified damage reports

For each social media post, the method examined whether it contained any of the ten defined hazard/damage information. 89 (2.17%) and 793 (3.08%) posts were identified to include at least one type of damage/hazard information (as reported by either text or image or both) for the Miami and Houston case respectively. We presented the counts of identified damage reports in Table 5. It was found that Miami suffered from both wind and flood hazards during Hurricane Irma. However, Houston was mainly affected by the flooding with most identified damage reports, i.e. 180 images and 563 texts, indicating flood hazards. In Fig. 4, we showed the distribution of identified damage reports in points and densities. We used the bandwidth of 500 m for the density mapping. The density is weighted by the ratio of the damage report counts and the raw crowdsourced posts counts. Fig. 4 shows that the identified damage reports clustered in the central downtown area and distributed along roadways for the Houston case. The damage reports were generally located along the shoreline for the Miami case, which indicated that these areas experienced noticeable damage during hurricanes.

We plotted the counts of identified damage reports by days in Fig. 5. Irma hit Florida on September 10, 2017 and Harvey struck Texas on the night of August 25, 2017. We found that most damage reports were posted in the first three to four days following the hurricanes' landfall. These reports can be mined immediately once they were posted.

4.3. Mining hazard types and severity information with Image Process module

More widespread flood hazard reports (180) were found in Houston during Hurricane Harvey compared to wind hazard reports (20). In comparison, Miami received slightly more wind hazard reports (41) than flood hazard reports (21) during Hurricane Irma (Table 5). Figs. 6–9 show the density maps of identified wind and flood hazards as well as some representative images for the two cases. The densities were weighted according to estimated severity levels. Images showing wind hazards for the Houston case were located very sparsely across the city (Fig. 6). In general, wind hazards were represented as fallen trees and wrecked boats (Figs. 6 and 8). Some images showing debris piles were falsely identified as wind hazard reports (Fig. 6). The proposed method can identify flooding in different environments including residential, downtown, and streets (see Figs. 7 and 9).

4.4. Mining damage types with Text Process module

We mined six types of damage information from textual messages (Table 1). Most textual reports identified with the proposed two-list search approach were related to disaster damage. Some of them did not reflect situations that happened at the geotagging locations and some may not refer to the damage experienced or observed by users who posted the textual message. Table 6 shows some representative examples of truly and falsely identified textual damage reports. The falsely identified damage reports included general concerns and comments on the disaster events expressed by affected people (e.g., e2 and e7) or official accounts (e.g., e9), situations about other affected people (e.g., e10), negation (e.g., e3), assumption (e.g., 13), and advertisements (e.g., e6). Moreover, some texts are too vague to determine whether the contents are about users' own experiences or observations (e.g. e4, e11, and e14).

We then checked the textual posts that were identified to convey damage information. Table 7 records the number of identified true and false positive cases. The indeterminate ones are the textual posts that we cannot tell whether the user observes a damage based on the present text (e.g. e4, e11, and e14 in Table 6).

The negative cases are not counted as the false negative cases can always be reduced with more keywords considered. It can be seen from Table 7 that the method works for most predefined damage. Especially for the "Infrastructure Damage". However, the prediction of "House Damage" is associated with low precision. We found that most of the false positive predictions are due to negation.

4.5. Performance of image classifiers on identifying wind and flood hazards

To evaluate how the hierarchical image classifiers performed with raw crawled data, we annotated the 1555 Twitter images collected for the Miami case and tracked how these images flow through the Image Process module. The process was demonstrated in Fig. 10.

When working as an integral, the image classifiers correctly identified 29 and 14 images showing wind and flood hazards (true positive) from the 1555 raw crawled Twitter images, and only missed very few images that present one of the two hazard types (false negative) and only 5 and 6 images were falsely predicted. Overall, the image classifiers are effective in dealing with the noisy semantic contents of social media images. We summarized the performance metrics of the hierarchical image classifiers on identifying the two hazards in Table 8.

5. Discussion

A data-driven method is proposed to assess damage information with multimodal social media data. The method leverages a set of machine learning approaches to process textual messages and visual images separately. While many previous studies relied on single-modal data and aggregated results in coarser spatial levels, our method outputs individual estimations that can assist city-level emergency operations. A few studies (e.g. Ref. [32,33]) concatenated image and textual features for disaster damage assessment. They were not efficient in analyzing posts with missing modalities or which the two modalities were weakly correlated. The separation of two data modalities also allows the method to include more data sources with visual and textual data formats.

Social media textual messages are limited by its short text length, subjectivity, low information quality, colloquial expressions, and so forth [40]. We have addressed some of these limitations with the proposed two-list keyword search method that uses two keyword lists for each pre-defined damage type. We adopt this method intending to detect damage efficiently from texts with few efforts spent on enumerating word/phrases combinations. The results show that the textual damage reports identified with this approach are generally on-topic, though some do not talk about the users' own experience or observations. Previous related works [28] also found it difficult to assess the damage with social media images due to the poor signal-to-noise ratio and loosely-defined damage forms. We address these challenges by devising five image classifiers with a defined semantic hierarchy to find out images providing information of interest, i.e. wind and flood hazards. The hierarchical structure also serves as a robust filter that removes irrelevant and less informative images in early steps, which is effective for the noisy and unbalanced social media images. Our method shows satisfactory performance when tested with raw crawled social media data.

The proposed method is possibly limited in a few aspects and is open for further improvements. First, we used around 1800 training and testing images for the development of five classifiers. Though the classifiers show comparable performance to other related works analyzing social media data (e.g. Ref. [27,28,42]), the performance can always be improved with more annotated images for classifier development. Second, both studied cases found few (2–3%) posts containing damage information and yielded few qualified reports as a consequence. A reason is that we downloaded images about two years after the event occurrence and many image links were invalid. Also, we restricted crawled posts to be geotagged. Some recent works using approaches such as geoparsing and reverse geocoding can extract location information from textual messages [57]. The method can yield more estimations once these approaches are considered. Future research in this direction should also keep paying attention to other emerging data courses for disaster management operations. Third, the data-driven damage assessment approach is built based on the assumption that the crawled eyewitness reports accurately reflect the ground truth scenarios at the time when the post is created for the location where the report is geotagged. This assumption may not always be true as affected population sometimes do not immediately report their observed damages. They may report damages when they moved to a safer location [10]. However, the result of this study showed that most damage reports were posted within two to three days after the hurricanes' landfall (Fig. 5), which suggested that most affected people reported their observations without much time lapse. Figs. 6–9 also showed that the built environment scenarios (e.g. residential, commercial, natural) identified from the images correspond to their associated locations. For the textual damage reports, the infrastructure destruction reports were found located on major roadways and bridges. These demonstrate the overall validity of the assumption for a course-grained time period (e.g. a day) and spatial scale (e.g. a neighborhood). Future studies can inspect the temporal and locational consistency between the reported damage and the ground truth for individual posts once data is available. Moreover, social media data can suffer from spatial bias [58]. Regions with no damage report identified do not translate to no damage presence [7]. Potential users of similar data-driven methods should also be informed by this limitation when using analysis results of social media data. Last, though the method can process social media data rapidly, a real-time setting could facilitate the method in practical applications.

6. Conclusion

A data-driven method is proposed to mine rapid, fine-grained, and comprehensive damage information from multimodal social media data, and its effectiveness is tested with two recent hurricane cases. Emergency managers and first responders frequently engage in time-sensitive decision-making and operations throughout the disaster events. The early access to fine-grained disaster damage information can largely relieve uncertainties in these processes, which in turn improves the rapidity and efficiency of disaster response and reduce consequent losses. Our method offers a supplementary resource for the acquisition of timely disaster damage information, which could be useful in the absence of authoritative data acquisition approaches. The proposed method also provides experience for future research in this direction on countering the noisy, unstructured, and unbalanced social media data, especially for research looking for specific information that only resides in few social media posts. In addition to our presented case study, the general research framework also applies to other disasters and extreme events with data of similar modalities from other sources.

Declaration of competing interest

None.

Acknowledgments

This material is based upon work supported by the early-career faculty start-up fund and graduate research assistantships at the University of Florida. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the University of Florida.

References

[1] A. Smith, 2018's billion dollar disasters in context | NOAA Climate.gov. Climate. Gov, Retrieved from, https://www.climate.gov/, 2019.
[2] D.B. McWethy, T. Schoennagel, P.E. Higuera, M. Krawchuk, B.J. Harvey, E. C. Metcalf, C. Schultz, C. Miller, A.L. Metcalf, B. Buma, A. Virapongse, J.C. Kulig, R. C. Stedman, Z. Ratajczak, C.R. Nelson, C. Kolden, Rethinking resilience to wildfire, Nat. Sustain. 2 (9) (2019) 797–804.
[3] James A. Gordon, Comprehensive Emergency Management for Local Governments:: Demystifying Emergency Planning, Rothstein Associates Inc, Brookfield, CT, 2015.
[4] M. Erdelj, M. Król, E. Natalizio, Wireless Sensor Networks and Multi-UAV systems for natural disaster management, Comput. Network. 124 (2017) 72–86.
[5] FEMA, Damage assessment operations manual, Retrieved from, https://www.fema.gov/, 2016.
[6] M. Yu, C. Yang, Y. Li, Big data in natural disaster management: a review, Geosciences 8 (5) (2018) 165.
[7] X. Zhong, M. Duckham, D. Chong, K. Tolhurst, Real-time estimation of wildfire perimeters from curated crowdsourcing, Sci. Rep. 6 (2016) 24206.
[8] S. Anson, H. Watson, K. Wadhwa, K. Metz, Analysing social media data for disaster preparedness: understanding the opportunities and barriers faced by humanitarian actors, Int. J. Disaster Risk Reduct. 21 (2017) 131–139.
[9] C. Granell, F.O. Ostermann, Beyond data collection: objectives and methods of research using VGI and geo-social media for disaster management, Comput. Environ. Urban Syst. 59 (2016) 231–243.
[10] D. Eilander, P. Trambauer, J. Wagemaker, A. Van Loenen, Harvesting social media for generation of near real-time flood maps, Procedia Eng. 154 (2016) 176–183.
[11] Q. Deng, Y. Liu, H. Zhang, X. Deng, Y. Ma, A new crowdsourcing model to assess disaster using microblog data in typhoon Haiyan, Nat. Hazards 84 (2) (2016) 1241–1256.
[12] Y. Wang, J.E. Taylor, Coupling sentiment and human mobility in natural disasters: a Twitter-based study of the 2014 South Napa Earthquake, Nat. Hazards 92 (2) (2018) 907–925.
[13] B.R. Jordan, A bird's-eye view of geology: the use of micro drones/UAVs in geologic fieldwork and education, GSA Today (Geol. Soc. Am.) 50–52 (2015).
[14] G. Novikov, A. Trekin, G. Potapov, V. Ignatiev, E. Burnaev, Satellite imagery analysis for operational damage assessment in emergency situations, Lecture Notes Bus. Inf. Process. 320 (2018) 347–358.
[15] T.R. Robinson, N. Rosser, R.J. Walters, The spatial and temporal influence of cloud cover on satellite-based emergency mapping of earthquake disasters, Sci. Rep. 9 (1) (2019) 1–9.
[16] M.F. Goodchild, Citizens as sensors: the world of volunteered geography, Geojournal 69 (4) (2007) 211–221.
[17] U.S. Census Bureau, (n.d.). ZIP Code Tabulation Areas (ZCTAs). Retrieved from https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html.
[18] Y. Kryvasheyeu, H. Chen, N. Obradovich, E. Moro, P. Van Hentenryck, J. Fowler, M. Cebrian, Rapid assessment of disaster damage using social media activity, Sci. Adv. 2 (3) (2016) 1–11.
[19] R. Samuels, J. Taylor, N. Mohammadi, The sound of silence: exploring how decreases in tweets contribute to local crisis identification, in: 15th International Conference on Information System for Crisis Response and Management (ISCRAM), Rochester, NY, May, 2018.

[20] L. Zou, N.S.N. Lam, S. Shams, H. Cai, M.A. Meyer, S. Yang, K. Lee, S.J. Park, M. A. Reams, Social and geographical disparities in twitter use during hurricane Harvey, Int. J. Digital Earth 12 (11) (2019) 1300–1318.
[21] B. Resch, F. Usländer, C. Havas, Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment, Cartogr. Geogr. Inf. Sci. 45 (4) (2018) 362–376.
[22] Y. Wang, J.E. Taylor, DUET: data-driven approach based on latent Dirichlet allocation topic modeling, J. Comput. Civ. Eng. 33 (3) (2019), 04019023.
[23] F. Yao, Y. Wang, Tracking urban geo-topics based on dynamic topic model, Comput. Environ. Urban Syst. 79 (2019) 101419.
[24] L. Smith, Q. Liang, P. James, W. Lin, Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework, J. Flood Risk Manag. 10 (3) (2017) 370–380.
[25] I. Temnikova, C. Castillo, S. Vieweg, EMTerms 1.0: a terminological resource for crisis tweets, in: ISCRAM 2015 Conference Proceedings - 12th International Conference on Information Systems for Crisis Response and Management, 2015-January, 2015, pp. 147–157.
[26] M. Bica, L. Palen, C. Bopp, Visual representations of disaster, in: Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, 2017, pp. 1262–1276.
[27] F. Alam, F. Ofli, M. Imran, Processing social media images by combining human and machine computing during crises, Int. J. Hum. Comput. Interact. 34 (4) (2018) 311–327.
[28] D.T. Nguyen, F. Ofli, M. Imran, P. Mitra, 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, ASONAM 2017, 2017, pp. 569–576.
[29] F. Alam, F. Ofli, M. Imran, CrisisMMD: multimodal twitter datasets from natural disasters, in: Proceedings of International AAAI Conference on Web and Social Media (ICWSM), Stanford, California, USA, 2018, pp. 465–473.
[30] K. Ahmad, K. Pogorelov, M. Riegler, O. Ostroukhova, P. Halvorsen, N. Conci, R. Dahyot, Automatic detection of passable roads after floods in remote sensed and social media data, Signal Process. Image Commun. 74 (2019) 110–118.
[31] L. Liu, E.A. Silva, C. Wu, H. Wang, A machine learning-based method for the large-scale evaluation of the qualities of the urban environment, Comput. Environ. Urban Syst. 65 (2017) 113–125.
[32] H. Mouzannar, Y. Rizk, M. Awad, Damage identification in social media posts using multimodal deep learning, in: 15th International Conference on Information System for Crisis Response and Management (ISCRAM), Rochester, NY, May, 2018, pp. 529–543.
[33] L. Lopez-Fuentes, J. Van De Weijer, M. Bolaños, H. Skinnemoen, Multi-modal deep learning approach for flood detection, in: Proceeding of the MediaEval 2017 Workshop. 1–3, 2017.
[34] X. Huang, C. Wang, Z. Li, H. Ning, A visual–textual fused approach to automated tagging of flood-related tweets during a flood event, Int. J. Digital Earth 12 (11) (2019) 1248–1264.
[35] S. Pouyanfar, Y. Tao, H. Tian, S.C. Chen, M.L. Shyu, Multimodal deep learning based on multiple correspondence analysis for disaster management, World Wide Web 22 (5) (2019) 1893–1911.
[36] A. Lazaridou, N.T. Pham, M. Baroni, Combining language and vision with a multimodal skip-gram model, in: 2015 Conference of the North American Chapter of the Association for Computational Linguistics, Human Language Technologies, 2015, 153–163.
[37] Q. You, J. Luo, H. Jin, J. Yang, Cross-modality consistent regression for joint visual-textual sentiment analysis of social multimedia, in: Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, 2016, pp. 13–22.
[38] F. Feng, X. Wang, R. Li, November). Cross-modal retrieval with correspondence autoencoder, in: Proceedings of the 22nd ACM International Conference on Multimedia, 2014, pp. 7–16.
[39] X. Huang, C. Wang, Z. Li, A near real-time flood-mapping approach by integrating social media and post-event satellite imagery, Spatial Sci. 24 (2) (2018) 113–123.
[40] N. Agarwal, Y. Yiliyasi, Information quality challenges in social media, in: Proceedings of the 2010 International Conference on Information Quality (ICIQ). 12-14 November 2010 Little Rock, Arkansas, SA, 2010.
[41] L. Vadicamo, F. Carrara, A. Cimino, S. Cresci, F. Dell'Orletta, F. Falchi, M. Tesconi, Cross-media learning for image sentiment analysis in the wild, in: Proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017, 2018-January, 2017, pp. 308–317.
[42] Ning, Hodgson Li, Wang, Prototyping a social media flooding photo screening system based on deep learning, ISPRS Int. J. Geo-Inf. 9 (2) (2020) 1–18.
[43] A. Olteanu, S. Vieweg, C. Castillo, What to expect when the unexpected happens: social media communications across crises, in: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, 2015, pp. 994–1009.
[44] H. Tanev, V. Zavarella, J. Steinberger, Monitoring disaster impact: detecting micro-events and eyewitness reports in mainstream and social media, in: 14th International Conference on Information System for Crisis Response and Management (ISCRAM), 2017.
[45] Y. Hu, S. Gao, K. Janowicz, B. Yu, W. Li, S. Prasad, Extracting and understanding urban areas of interest using geotagged photos, Comput. Environ. Urban Syst. 54 (2015) 240–254.
[46] Y. Wang, Q. Wang, J.E. Taylor, Aggregated responses of human mobility to severe winter storms: an empirical study, PloS One 12 (12) (2017), e0188734.
[47] B. Thomee, D.A. Shamma, G. Friedland, B. Elizalde, K. Ni, D. Poland, D. Borth, L. J. Li, YFCC100M: the new data in multimedia research, Commun. ACM 59 (2) (2016) 64–73.
[48] Haiyan Hao, Yan Wang, Hurricane Damage Assessment with Multi-, Crowd-Sourced Image Data: A Case Study of Hurricane Irma in the City of Miami, in: Amanda Hughes, Fiona McNeill, Christopher W. Zobel (Eds.), ISCRAM 2020 Conference Proceedings – 17th International Conference on Information Systems for Crisis Response and Management, Virginia Tech, Blacksburg, VA (USA), 2020, pp. 825–837.
[49] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[50] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, A. Torralba, Places: a 10 million image database for scene recognition, IEEE Trans. Pattern Anal. Mach. Intell. 40 (6) (2018) 1452–1464.
[51] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December, 2016, pp. 2818–2826.
[52] B. Zhou, H. Zhao, X. Puig, T. Xiao, S. Fidler, A. Barriuso, A. Torralba, Semantic understanding of scenes through the ADE20K dataset, Int. J. Comput. Vis. 127 (3) (2019) 302–321.
[53] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, A. Desmaison, Pytorch: an imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems, 2019, pp. 8026–8037.
[54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, J. Vanderplas, Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[55] J.P. Cangialosi, A.S. Latto, R. Berg, Hurricane irma, Retrieved from, https://www.nhc.noaa.gov/, 2018.
[56] E.S. Blake, D.A. Zelinsky, Hurricane harvey, Retrieved from, https://www.nhc.noaa.gov/, 2018.
[57] S.E. Middleton, G. Kordopatis-Zilos, S. Papadopoulos, Y. Kompatsiaris, Location extraction from social media: geoparsing, location disambiguation, and geotagging, ACM Trans. Inf. Syst. 36 (4) (2018) 1–27.
[58] G. Zhang, A.X. Zhu, The representativeness and spatial bias of volunteered geographic information: a review, Spatial Sci. 24 (3) (2018) 151–162.