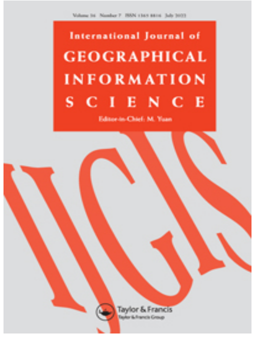

International Journal of Geographical Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Extraction and analysis of natural disaster-related VGI from social media: review, opportunities and challenges

Yu Feng, Xiao Huang & Monika Sester

To cite this article: Yu Feng, Xiao Huang & Monika Sester (2022) Extraction and analysis of natural disaster-related VGI from social media: review, opportunities and challenges, International Journal of Geographical Information Science, 36:7, 1275-1316, DOI: 10.1080/13658816.2022.2048835

To link to this article: https://doi.org/10.1080/13658816.2022.2048835

Published online: 21 Mar 2022.

Submit your article to this journal

Article views: 1336

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tgis20

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
2022, VOL. 36, NO. 7, 1275–1316
https://doi.org/10.1080/13658816.2022.2048835

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Check for updates

Extraction and analysis of natural disaster-related VGI from social media: review, opportunities and challenges

Yu Feng[a], Xiao Huang[b] and Monika Sester[a]

[a]Institute of Cartography and Geoinformatics, Leibniz University Hannover, Hannover, Germany;
[b]Department of Geosciences, University of Arkansas, Fayetteville, AR, USA

ABSTRACT

The idea of 'citizen as sensors' has gradually become a reality over the past decade. Today, Volunteered Geographic Information (VGI) from citizens is highly involved in acquiring information on natural disasters. In particular, the rapid development of deep learning techniques in computer vision and natural language processing in recent years has allowed more information related to natural disasters to be extracted from social media, such as the severity of building damage and flood water levels. Meanwhile, many recent studies have integrated information extracted from social media with that from other sources, such as remote sensing and sensor networks, to provide comprehensive and detailed information on natural disasters. Therefore, it is of great significance to review the existing work, given the rapid development of this field. In this review, we summarized eight common tasks and their solutions in social media content analysis for natural disasters. We also grouped and analyzed studies that make further use of this extracted information, either standalone or in combination with other sources. Based on the review, we identified and discussed challenges and opportunities.

ARTICLE HISTORY
Received 17 February 2021
Accepted 27 February 2022

KEYWORDS
Volunteered Geographic Information; social media; natural disaster; spatiotemporal analysis

1. Introduction

Humans have been facing the unceasing threat of natural disasters for centuries. According to the report (Wallemacq and House 2018) from United Nations Office for Disaster Risk Reduction, natural disasters between 1998 and 2017 resulted in 1.3 million deaths, 4.4 billion people injured, homeless, displaced, or in need of emergency assistance.

Rapid advances in modern technology have made it possible to obtain disaster observations through the establishment of sensor networks and remote sensing techniques. GPS networks have been used for earthquake monitoring (Crowell *et al.* 2009). River and tide gauges are used for monitoring flood and storm surge (Pegelonline 2020). Temperature and smoke sensors have been deployed for early fire detection (Molina-Pico *et al.* 2016). With satellite remote sensing and aerial imagery, floods,

CONTACT Yu Feng yu.feng@ikg.uni-hannover.de
© 2022 Informa UK Limited, trading as Taylor & Francis Group

wildfires, storm- and earthquake-induced damages can be detected (as presented in Li *et al.* 2015, Pham *et al.* 2017, Gupta *et al.* 2019). Despite that observation from sensor networks are mostly in real-time, their coverage is limited. In contrast, airborne or satellite remote sensing products have a much broader spatial coverage. Still, in many circumstances, they are hardly up-to-date given a variety of reasons that include satellite revisit time (Feng *et al.* 2015), weather conditions, and delays due to data acquisition and data processing (Ning *et al.* 2020). Although Unmanned Aerial Systems (UAS) are often used to quickly acquire imagery of natural disasters, for certain natural hazards such as floods and storms, they can usually only be deployed after the event when they are suitable for flight. All these limitations may result in the loss of first-hand information on a natural disaster event. Therefore, observations from the ground are needed as a supplement to the existing monitoring techniques.

'Citizen as sensors' is a well-known concept, where crowdsourcing is used to obtain geospatial information (Heipke 2010). Volunteered Geographic Information (VGI), first coined in Goodchild (2007), denotes crowdsourced geospatial data. The value of VGI for disaster management was first observed in the behavior of citizens during the Southern California wildfires of 2007–2009 and documented in Goodchild and Glennon (2010). Volunteers shared wildfire reports with locations on Flickr, interpreted remote sensing imagery, and established map sites presenting both VGI and official information. It was also noticed that volunteers could, in certain circumstances, provide more timely situation information than official sources.

As of today, studies on VGI for natural disaster management have been carried out for more than ten years. In this period, modern computer technology and the mobile Internet have been rapidly developed with notable changes in user habits. VGI can be distinguished into two approaches, participatory and opportunistic (Ostermann and Spinsanti 2011). A participatory approach requires conscious and active participation by the users. In contrast, the opportunistic approach acquires the information in a quasi unconscious and passive manner (Sester *et al.* 2014). Most of the early studies used participatory approaches, which rely heavily on user engagement. People usually have to install a specific App or use a dedicated web application to provide their input. In addition, users are often required to register for a new account. This is considered as inconvenient – especially if they are only infrequent users. As a result, the opportunistic VGI is more desirable and has become rapidly popular in recent years.

One of the first sources of opportunistic VGI is social media, as it can be provided at high volume and low cost in real-time. Social media is often used for information gathering, information dissemination, collaborative problem solving and response, especially in times of natural disasters, and is therefore a primary source of information for many studies. Since this data has a lot of noise, the processing and analysis of the data are particularly important. Therefore, the computational methods for social media posts with location information are the focus of this review and such information is referred to here as *social media VGI*.

Meanwhile, mobile devices, the form of social media, and the way it is handled have changed dramatically. Now, users can use their mobile devices to take high-quality photos and videos. The social media platforms also support longer texts, clearer photos, and longer videos. Great progress has been achieved in computer vision and natural language processing (NLP) after the success of deep convolutional neural networks. Information can be extracted from social media more efficiently, and new content can be obtained through the interpretation of texts and images. Therefore, we want to review the research on social media VGI for natural disasters in the last ten years and highlight the new solutions and new content that can be extracted from social media VGI.

There have been several reviews that summarized the existing studies on VGI. Yan *et al.* (2020) comprehensively reviewed the research on several application domains of VGI in the first decade of 2007–2017, including collaborative mapping, social science, disaster management, location-based services, land use and land cover, and environmental monitoring. Klonner *et al.* (2016) focused on VGI research for disasters. A total of 11 studies were reviewed, where most of them used the participatory approach for VGI acquisition. Social media data are also utilized, however, the studies around 2010–2013 (De Longueville *et al.* 2010; Middleton *et al.* 2013; Schnebele and Cervone 2013) rarely process and interpret the social media contents in an automatic and efficient manner. Wang and Ye (2018) grouped the literature concerning four dimensions (space, time, content, and network). The interpretation of social media mainly focused on text and covered only two tasks, i.e. topic and sentiment extraction. The importance of image content interpretation is noticed, however, not yet discussed. In addition, Steiger *et al.* (2015) reviewed the spatiotemporal and semantic analysis methods applied for social media data, including nine studies for disasters. These methods have also been used in recent research, meanwhile, there are other new attempts in more recent years.

Therefore, in this work, we focus on the studies of natural disasters where social media is the primary data source and processed in an automatic manner. In addition, the information provided by social media can be associated with geographic locations to provide spatial and temporal information about natural disasters. With this focus, we conducted a systematic literature search to identify the related studies.

In general, most of the research on social media VGI for natural disasters has two main components. One is information extraction, including not only the extraction of disaster-related posts but also the further details from these posts. Based on the retrieved studies and available benchmark datasets, we summarized the current research into eight typical tasks and reviewed their solutions. The other component is analysis and mapping. Existing studies are grouped on two aspects, whether social media data are analyzed standalone or combined with other sources.

This paper is organized as follows. Section 2 presents the process of the literature search. Section 3 summarizes the research for disaster-relevant information extraction, and Section 4 summarizes the research for social media VGI analysis and mapping for natural disaster analysis. Section 5 presents the identified research challenges and opportunities. Lastly, Section 6 concludes this review. The structure of this review is illustrated in Figure 1.

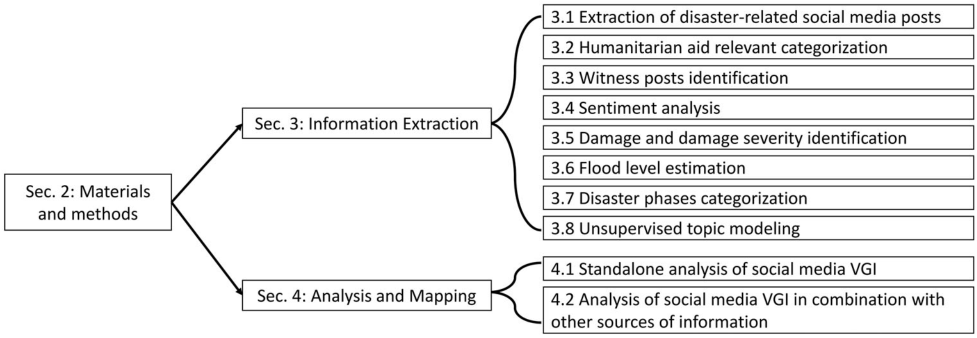

Figure 1. The structure of this review.

2. Materials and methods

As summarized in Xiao and Watson (2019), systematic literature reviews are commonly conducted with the following eight steps: (1) formulating the research problem; (2) developing and validating the review protocol; (3) searching the literature; (4) screening for inclusion; (5) assessing quality; (6) extracting data; (7) analyzing and synthesizing data; and (8) reporting the findings. This review is organized around the following two research questions:

What: What kind of information can be extracted from social media?

How: How to use the extracted information to provide spatial and temporal information about natural disasters?

We referenced the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist (Moher *et al.* 2009) as protocol, where previous review papers (Zourlidou and Sester 2019; Abed Al Ahad *et al.* 2020) applying this guideline have been used as a reference for our work. The literature search is described in Section 2.1, and the screening and eligibility check is presented in Section 2.2. The data synthesis is presented in a narrative manner and presented in Sections 3 and 4.

2.1. Search strategy and database sources

Two multidisciplinary databases: Web of Science (WoS)[1] and Scopus,[2] were used to search for relevant documents based on the article title, abstract, and keywords. Studies regarding social media and natural disasters were retrieved by applying the keyword '*social media*' and disaster-related search terms. According to the definition of the U.S. Federal Emergency Management Agency (FEMA), natural hazards include hurricanes, earthquakes, tornadoes, droughts, wildfires, winter storms, and floods.[3] Because of the differences in common hazards and vocabulary outside the United States, we expanded this list with four additional common natural disaster types: typhoons, landslides, tsunamis, and volcanoes. In total, 11 natural disaster types were considered to build the query strings as shown in Table 1. In addition, we limited the retrieved documents to be written in English and published before 2021.

Table 1. Query strings for WoS and Scopus databases (Status: 07 May 2021).

Database	Query string	Number
WoS	TS=(("social media") AND (flood OR earthquake OR wildfire OR hurricane OR tornado OR typhoon OR landslide OR volcano OR tsunami OR drought OR "winter storm")) AND LA=(English) AND PY=(2000-2020)	1076
Scopus	TITLE-ABS-KEY(("social media") AND (flood OR earthquake OR wildfire OR hurricane OR tornado OR typhoon OR landslide OR volcano OR tsunami OR drought OR "winter storm")) AND LANGUAGE (English) AND PUBYEAR < 2021	1498

2.2. Inclusion and exclusion criteria

In order to identify documents relevant to this review, inclusion and exclusion criteria are defined. According to the research question defined in Section 2, we aim to identify the studies that extract natural disaster related information automatically and the extracted information can provide spatial and temporal information about the disasters. Therefore, we try to tag each article according to the information it extracts and the methods it uses. In this process, we refer to the idea of grounded theory (Wolfswinkel *et al.* 2013), first open coding and then merging similar concepts.

There are in general five common cases for exclusion identified by screening for titles, abstracts, and keywords, where 'social media' and natural disaster-related keywords are mentioned but these documents are not related to the topic of this review.

1. Media and communication: questionnaires or interviews were used primarily to investigate the role of social media and people's attitudes during emergencies.
2. Disaster management: social media is used for information dissemination, communication, and organizing relief during emergencies.
3. Keyword ambiguity: the keywords 'flood', 'tsunami' appear in the abstract to describe the influx of information or news.
4. Other topics: natural disaster-related keywords are mentioned as examples. The main focus is on other topics, such as social events, public health, business, etc.
5. Review and short papers: the studies are review papers assigned with an incorrect document type or very short paper without experimental details.

We also notice that studies in the following six aspects used social media data posted during natural disaster events as the data source. However, because of their different focuses, they are excluded in this review based on an assessment of the full text.

1. Location inference: these studies aim to identify toponyms in social media texts to provide more posts with geolocation.
2. Fake information detection: these studies focus on identifying false information and assessing the credibility of social media posts.
3. Unsupervised event detection: users may tweet with the same words when an event occurs, such as the event name, the location or descriptive terms of the event. Events are detected using bursts of words.

4. System design: these studies focus on implementing, optimizing or evaluating the emergency management systems considering social media as the data source.
5. Network analysis: these studies analyzed the dissemination of information on social media, also including identifications of communities or key players.
6. Manual analysis: social media posts are manually analyzed and assigned to pre-defined categories or topics.

All of these aspects are necessary components of social media data analysis related to natural disasters. The first three are general topics of social media analysis that can improve the availability and reliability of social media data. Their application is not limited to research related to natural disasters. Detailed reviews of existing research on each topic can be found separately at (Ramachandran and Ramasubramanian 2018; Stock 2018; Meel and Vishwakarma 2020). The output from network analysis can reveal patterns of communication but cannot usually be used to provide a spatial and temporal overview of natural disasters. In addition, studies that entirely apply manual analysis are not considered, as information that can be extracted automatically is the focus. The process for literature search and selection is illustrated in Figure 2.

With the 550 documents retrieved from the previous steps, we conducted a narrative synthesis to summarize existing studies from two aspects, disaster-relevant information extraction in Section 3 and social media VGI analysis and mapping for natural disaster analysis in Section 4.

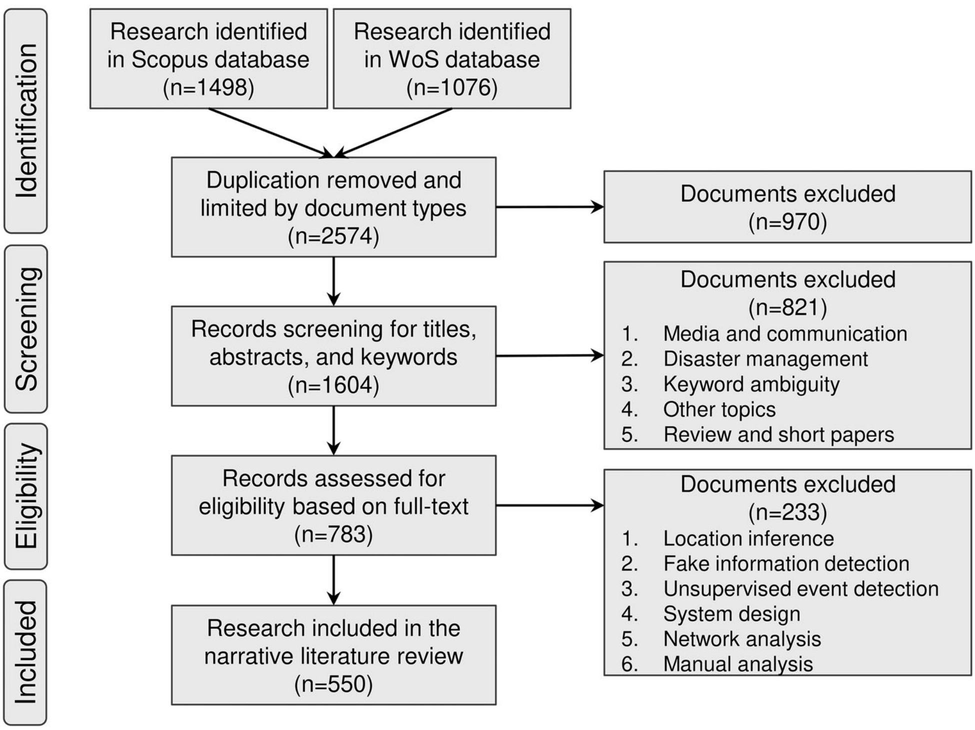

Figure 2. The flow diagram illustrating the literature search process (Status: 07 May 2021).

3. Information extraction

Social media encompass all aspects of human life and contain various types of information. Therefore, the extraction of natural disaster-relevant information is an essential step before using it. This step appears in most of the relevant studies. The accuracy of the information extraction directly affects the quality of the subsequent analysis of social media VGI. First and foremost, it is important to know what information can be extracted from social media.

The review of the information extraction component is based primarily on existing publicly available datasets. These datasets focus on research tasks that are representative in the research field and often reused for different studies. Ofli *et al.* (2020) has summarized the existing publicly available datasets for extracting information related to natural disasters from social media data. In addition, we have expanded several datasets that are not included as shown in Table 2. Six types of tasks can be outlined based on these datasets, which are

- (1) Extraction of disaster-related social media posts
- (2) Humanitarian aid relevant categorization
- (3) Witness posts identification
- (4) Sentiment analysis
- (5) Damage and damage severity identification
- (6) Flood level estimation

The connections between these tasks and datasets are presented in Table 2. Most of the datasets focus on tasks (1) and (2). In addition to those, we identified another two types of common tasks from the rest of the retrieved documents.

Table 2. Datasets for natural disaster-related social media data analytic.

Name	Host	Source	Task
CrisisLexT6	CrisisLex[a]	Olteanu *et al.* (2014)	1
CrisisLexT26	CrisisLex	Olteanu *et al.* (2015)	1,2,3
SWDM2013	CrisisNLP[b]	Imran *et al.* (2013a)	1,2
ISCRAM13	CrisisNLP	Imran *et al.* (2013b)	1,2
CrisisNLP	CrisisNLP	Imran *et al.* (2016)	1,2
Disaster Response Data (DRD)	CrowdFlower	Udacity course[c]	1,2
Disasters on Social Media (DSM)	CrowdFlower	data.world[d]	1
CrisisMMD	CrisisNLP	Alam *et al.* (2018c)	1,2,5
ISCRAM18	CrisisNLP	Alam *et al.* (2018a)	1,2,4
ASONAM17 Damage Images	CrisisNLP	Nguyen *et al.* (2017b)	5
ACL ICWSM 2018	CrisisNLP	Alam *et al.* (2018b)	1
Eyewitness Tweets	CrisisNLP	Zahra *et al.* (2020)	3
ASONAM20	CrisisNLP	Alam *et al.* (2020b)	1,2,5
CrisisBench	CrisisNLP	Alam *et al.* (2020c)	1,2
HumAID	CrisisNLP	Alam *et al.* (2021)	2
MediaEval'17 MMS Task	MediaEval[e]	Bischke *et al.* (2017b)	1
MediaEval'18 MMS Task	MediaEval	Bischke *et al.* (2018)	1,6
MediaEval'19 MMS Task	MediaEval	Bischke *et al.* (2019)	6

[a]CrisisLex: Crisis-Related Social Media Data and Tools. https://crisislex.org/.
[b]CrisisNLP. https://crisisnlp.qcri.org/.
[c]Multilingual Disaster Response Messages. https://www.appen.com.cn/datasets/combined-disaster-response-data/.
[d]Disasters on Social Media – dataset by crowdflower. https://data.world/crowdflower/disasters-on-social-media.
[e]MediaEval Benchmarking Initiative. http://www.multimediaeval.org/ [Accessed 07 May 2021].

- (7) Disaster phases categorization
- (8) Unsupervised topic modeling

Disaster phases identification has appeared in several studies, e.g. Iyengar *et al.* (2011), Chowdhury *et al.* (2013), and Huang and Xiao (2015). However, public datasets are not yet available. *Unsupervised topic modeling* aims to assign text documents into topics. Usually only the number of topics is defined and the topics are presented directly to the users, so there is hardly any annotated data.

Our review is based on the aforementioned eight tasks and details the representative solutions in the following subsections. Each task corresponds to a category of information that can be extracted from social media to benefit natural disaster research. Of course, we acknowledge the existence of other tasks that are beyond what we defined in this review.

3.1. Extraction of disaster-related social media posts

This is the most basic task, which can be applied for a broad concept of natural disasters or for a specific type of natural disaster. Since there exist a large number of studies related to this task, existing studies are summarized with respect to their modality of input, i.e. text, image, fused text and image.

3.1.1. Text

Keyword filtering is the most straightforward solution that has been used in many early studies to extract disaster-relevant posts. Predefined keywords are filtered to retrieve social media posts related to flood (Fuchs *et al.* 2013; Murthy and Longwell 2013; Fohringer *et al.* 2015; Li *et al.* 2018), earthquake (Earle *et al.* 2011; MacEachren *et al.* 2011), hurricane (Mandel *et al.* 2012; Kogan *et al.* 2015), tornado (Ukkusuri *et al.* 2014), fire (De Longueville *et al.* 2009), etc. For multilingual regions, keywords in multiple languages are defined, e.g. German and English keywords were used to filter Tweets in Germany (Fohringer *et al.* 2015). For applications with a global focus, keywords related to nine types of natural disaster (including *earthquake*, *blizzard*, *tornado*, *drought/heatwave*, *cyclonic storm*, *hail/thunder*, *flood*, *tsunami*, *volcanic eruption*) have been filtered in 43 languages in Dittrich and Lucas (2014). However, the selected keywords are often ambiguous in meaning. For example, the keyword *flood* also has other possible meanings under different contexts, such as in *flood light*, the term *flooded by people*, thus leading to a limited performance in information retrieval.

In comparison, text classification in NLP provides better solutions for extracting disaster-related information. Social media documents are classified into binary or multiple categories with supervised classification based on the manually annotated corpus datasets, often prepared in advance. Different types of features are summarized from social media posts. Statistical features, such as word n-grams (e.g. unigrams, bigrams), text length, the number of hashtags, user mentions, URLs, whether it is a retweet, whether it is a reply to another Tweet, POS (Part-of-Speech) tags, to list a few, are commonly used, e.g. in Sakaki *et al.* (2010), Yin *et al.* (2012), Karimi *et al.* (2013), and Cresci *et al.* (2015). *tf-idf* (term frequency-inverse document frequency) (Salton and

Buckley 1988) is a special case of word n-grams, where the word frequency is normalized by word document frequency. It represents the importance of the word to a document based on the whole corpus, which has been frequently used for text classification tasks, e.g. in Xiao *et al.* (2018) and Khare *et al.* (2018). These statistical features can be used to train a binary or multi-class classifier using supervised machine learning (ML) methods, such as Support Vector Machine (SVM), random forest, logistic regression, naive Bayes.

Word embedding is the technique that aims to represent words or phrases as vectors of real numbers. This strategy has gained more attention after the rise of artificial neural networks. With a shallow neural network, vector representations of words can be learned based on a large corpus in an unsupervised manner. Word2vec (Mikolov *et al.* 2013a, 2013b), the most well-known model, was widely applied for the text classification of disaster-related social media texts. Since the vector representations are learned for individual words, sentences can be classified in multiple ways. Sentences can be represented by the averaged word vectors (Tkachenko *et al.* 2017; Bischke *et al.* 2017a) and classified using classic machine learning methods. Stowe *et al.* (2016) demonstrated that such features outperformed many combinations of statistical features and showed the highest importance in the ablation study for both the binary and multi-class text classification tasks. There are also further developed deep learning solutions to summarize sentence representations for supervised short text classification, such as TextCNN (Kim 2014) (used in Feng and Sester 2018; Feng *et al.* 2018; Huang *et al.* 2020), and LSTM/Bi-LSTM (Liu *et al.* 2016; Zhou *et al.* 2016) (used in Lopez-Fuentes *et al.* 2017; Sit *et al.* 2019).

Meanwhile, word embedding with better performance has been developed. GloVe (Pennington *et al.* 2014), an improvement from word2vec, captured both global statistics and local statistics of a corpus. In addition, fastText (Joulin *et al.* 2017) represents each word as a bag of character n-grams. Character n-grams represent words with a continuous sequence of n characters, e.g. the word *flood* is represented as *fl, flo, loo, ood, od* for character 3-grams. The word representation is the sum of the character n-grams (Bojanowski *et al.* 2017). A model similar to word2vec is trained to learn the vector representation of a word considering the subword information. These two improved word embedding models have also been used in disaster-related text classification, e.g. in Lopez-Fuentes *et al.* (2017) and Feng *et al.* (2018).

In recent years, contextualized word embedding has received wide attention, as the same word may have different semantic meanings under different contexts. Instead of using a fixed vector to represent a word, ELMo (Peters *et al.* 2018) provided a vector representation of words considering the context. BERT (Devlin *et al.* 2019) is one successful extension of the idea of ELMo, where the Bi-LSTM has been replaced by a more powerful encoder using attention mechanism – Transformer (Vaswani *et al.* 2017). It has also been used in some of the very recent works for disaster-relevant social media content analysis. Recently proposed pre-trained models of contextualized word embedding such as BERT and ELMo were compared with word2vec and GloVe features for disaster-related text classification in Jain *et al.* (2019) and Alam *et al.* (2020c). It is noticed that for a binary classification task, the contextualized word embedding does not necessarily outperform the TextCNN architecture or using Word2vec embedding.

Table 3 summarizes the aforementioned studies based on the types of natural disaster, classification categories, and applied methods. It can be observed that studies from 2017 mainly adopted word embedding techniques instead of statistical features. In addition, we also notice that systems designed to retrieve disaster-related social media posts are mostly suitable for situations where a single language (mainly English) is used; cross-language situations and data acquisition on a global scale are less often considered.

3.1.2. Images

Many early studies rely on human interpretation to extract visual observations of disaster events, e.g. in Kutija *et al.* (2014), Fohringer *et al.* (2015), and Le Coz *et al.* (2016). Research on automatic extraction of disaster-related posts has emerged in recent

Table 3. Extraction of disaster-related social media posts based on texts.

Input	Topic	Category	Method[a]	Example papers
Text	Flood	Binary	keyword filtering	Fuchs *et al.* (2013), Fohringer *et al.* (2015) Murthy and Longwell (2013), Li *et al.* (2018)
			tf-idf features + classic ML	Hanif *et al.* (2017), Xiao *et al.* (2018)
			word2vec + classic ML	Tkachenko *et al.* (2017), Bischke *et al.* (2017a)
			word2vec/fastText + TextCNN	Feng and Sester (2018), Feng *et al.* (2018) Huang *et al.* (2020)
			word2vec/GloVe + LSTM/Bi-LSTM	Lopez-Fuentes *et al.* (2017), Sit *et al.* (2019)
			tf/tf-idf/word2vec + classic ML	Moumtzidou *et al.* (2018)
	Earthquake	Binary	keyword filtering — MacEachren, Earle statistical features þ classic ML — Sakaki, Yin, Cresci	MacEachren *et al.* (2011), Earle *et al.* (2011) Sakaki *et al.* (2010), Yin *et al.* (2012) Cresci *et al.* (2015)
	Hurricane /Tornado	Binary	keyword filtering	Mandel *et al.* (2012), Ukkusuri *et al.* (2014) Kogan *et al.* (2015)
	Fire	Binary	keyword filtering	Craglia *et al.* (2012), Wang *et al.* (2016b) De Longueville *et al.* (2009)
			statistical features + SVM	Power *et al.* (2013), Power *et al.* (2015)
	Multiple disasters	9 Disaster types binary	keyword filtering statistical features + classic ML	Dittrich and Lucas (2014) Khare *et al.* (2018)
		7 Disaster types binary	statistical features + classic ML statistical features/ word2vec + classic ML	Karimi *et al.* (2013) Stowe *et al.* (2016)
		Informative or not	word2vec/GloVe/ELMo/BERT TextCNN/BERT	Jain *et al.* (2019) Alam *et al.* (2020c)

[a]Methods are represented in the form of *features + classification methods*, e.g. statistical features + classic ML. *classic ML* includes common machine learning methods such as Naive Bayes, logistic regression, SVM, random forest, etc.

years, all thanks to the rapid development of computer vision, and in particular, the success of deep convolutional neural networks (DCNN).

Before using DCNN, engineered visual features, such as SIFT, SURF, and their derivatives, have been used for detecting floods (Jing *et al.* 2016a, 2016b) and fire (Daly and Thom 2016) from social media images around 2016. *Multimedia Satellite (MMSat) Task* in the *MediaEval'17* benchmarking initiative (Bischke *et al.* 2017b) is a well-known task in the community to retrieve flood-relevant Flickr posts based on visual and textual features. The organizers prepared engineered visual features, such as Auto color correlogram (Huang *et al.* 1997), Color and Edge Directivity Descriptors (Chatzichristofis and Boutalis 2008), etc., which have been used to retrieve flood relevant images in Tkachenko *et al.* (2017), Zhao and Larson (2017), and Hanif *et al.* (2017).

The use of DCNN for disaster-related image classification started around 2016. The image classification algorithms using deep learning are rarely trained from scratch. Instead, transfer learning techniques are commonly adopted to fine-tune models trained on much larger datasets, such as ImageNet (Deng *et al.* 2009), or Places365 (Zhou *et al.* 2018). Lagerstrom *et al.* (2016) fine-tuned an early DCNN model named OverFeat (Sermanet *et al.* 2013), which is pre-trained on the ImageNet dataset. More deep learning models for image classification have appeared since then, such as VGG (Simonyan and Zisserman 2014), InceptionV3 (Szegedy *et al.* 2016), ResNet (He *et al.* 2016), etc. Many of them have also been used for image classification tasks related to natural disasters.

As for transfer learning, two strategies are commonly used. One replaces the output layer of a pre-trained network (e.g. 1000 categories for ImageNet) with a softmax layer corresponding to the desired categories, such as a certain disaster-relevant and irrelevant. This process is considered as fine-tuning of pre-trained networks. This strategy has been used in many studies, such as (Lopez-Fuentes *et al.* 2017; Nogueira *et al.* 2017; Huang *et al.* 2020). Another is to view DCNN as a feature generator and further apply classic machine learning methods, such as SVM, to the extracted features. This strategy has been explored in Ahmad *et al.* (2017b) and Avgerinakis *et al.* (2017). In addition, machine learning classifiers can be trained on the deep features from different pre-trained DCNNs, i.e. pre-trained on both ImageNet and Places365 datasets. The final output is the fusion of prediction scores of multiple classifiers, e.g. the ensemble of SVM classifiers in Ahmad *et al.* (2017a). This fusion strategy was further investigated in Ahmad *et al.* (2018), where better results were achieved by introducing scene-level information, i.e. a pre-trained model on the Places365 dataset. There are also experiments with an ensemble of multiple DCNNs at feature level as in Said *et al.* (2018) and Feng *et al.* (2018).

Besides floods, other natural disasters have been studied much less individually. For example, fire images were classified in Daly and Thom (2016) and Lagerstrom *et al.* (2016). Damages caused by earthquakes and hurricanes were classified in Nia and Mori (2017) and Li *et al.* (2019), which are further detailed in Section 3.5. More compact models have been developed in recent years, where social media images can be categorized by the type of natural disaster (Alam *et al.* 2020b), covering the most frequent natural disasters such as earthquakes, fires, floods, hurricanes, landslides, etc. EfficientNet (Tan and Le 2019) outperforms the current commonly used models (e.g. ResNet, DenseNet) by around 1–2% on weighted averaged F_1-scores. Weber *et al.*

Table 4. Extraction of disaster-related social media posts based on images.

Input	Topic	Category	Method[a]	Example papers
Image	Flood	Binary	SIFT-like features + classic ML	Jing *et al.* (2016a, 2016b)
			Multiple engineered visual features + classic ML	Hanif *et al.* (2017) Zhao and Larson (2017) Tkachenko *et al.* (2017)
			Pre-trained DCNN + classic ML	Bischke *et al.* (2017a) Avgerinakis *et al.* (2017) Feng and Sester (2018)
			Fine-tune pre-trained DCNN	Lopez-Fuentes *et al.* (2017) Nogueira *et al.* (2017) Huang *et al.* (2020) Ning *et al.* (2020)
			Ensemble of SVMs trained on features from multiple pre- trained DCNNs	Ahmad *et al.* (2017a, 2018)
			Concatenation of features from multiple pre-trained DCNNs + softmax/SVM	Feng *et al.* (2018) Said *et al.* (2018)
	Fire	Binary	SIFT-like features + SVM	Daly and Thom (2016)
			Pre-trained DCNN + random forest	Lagerstrom *et al.* (2016)
	Multiple disasters	6 Topics about earthquake, oil spill, hurricane	Engineered visual features	Yang *et al.* (2011)
		Informative or not / 7 disaster type categories	Fine-tune pre-trained DCNN	Alam *et al.* (2020b)
		43 Disaster and incidents categories	Multi-task DCNN with class-negative loss	Weber *et al.* (2020)

[a]Methods are represented in the form of *features + classification methods*, e.g. statistical features + classic ML. *classic ML* includes common machine learning methods such as Naive Bayes, logistic regression, SVM, random forest, etc.

(2020) prepared a large-scale image dataset for disaster and incidents, which was trained using a multi-task DCNN architecture to predict 43 kinds of incidents (e.g. flood, collapse, car accidents, fire, etc.) at 49 location types (e.g. bridge, forest, highway, etc.).

Table 4 summarizes the aforementioned studies that target images as input. It can be observed that the application of DCNN disaster-related image classification started to emerge around 2017, while fine-tuning and ensemble of pre-trained DCNN models are dominant approaches till now.

3.1.3. Fused text and image

Social media text and image modalities are often fused to extract disaster-relevant posts, as both visual and textual information may contain important information during a natural disaster event. In general, information from different modalities can be fused at the raw data level, feature level, and decision level (Roheda *et al.* 2018). Due to the vast difference of texts and images, the fusion of textual and visual information is mainly conducted at the feature- and decision-level. Feature fusion concatenates features from text and images and performs a classification, which has been used in Lopez-Fuentes *et al.* (2017), Hanif *et al.* (2017), and Bischke *et al.* (2017a) for flood. Decision fusion, also known as late fusion, combines the outputs of the text and image classifiers into a final decision, which has been used in Dao and Pham (2017), Ahmad *et al.* (2017a, 2017b), Avgerinakis *et al.* (2017), Dao *et al.* (2018), and Feng and Sester (2018) for flood-relevant social media retrieval.

As for the classification of social media posts for multiple natural disasters, feature fusion was used in Yang *et al.* (2011) to concatenate engineered features from both text (word frequency) and image, i.e. color features in HSV space (Sural *et al.* 2002) and location features using SPCPE (Chen *et al.* 1999), for a classification. Mouzannar *et al.* (2018) compared feature fusion and decision fusion based on visual features from pre-trained DCNN and textual features from word2vec. A similar performance was achieved. The decision fusion outperforms only 0.02% over feature fusion for a multi-modal classification task with six categories.

Table 5 summarizes the aforementioned studies. The study of fused text and image information has been explored, and many of the studies proved that such a fusion is beneficial for filtering social media data, given the close relationship between visual and textual features, especially during a natural disaster event (Bischke *et al.* 2017a; Mouzannar *et al.* 2018).

Table 5. Extraction of disaster-related social media posts based on fused texts and images.

Input	Topic	Category	Method	Example papers
Text + image	Flood	Binary	Feature fusion	Lopez-Fuentes *et al.* (2017), Hanif *et al.* (2017) Bischke *et al.* (2017a), Huang *et al.* (2020)
			Decision fusion	Dao and Pham (2017), Ahmad *et al.* (2017a, 2017b) Avgerinakis *et al.* (2017), Dao *et al.* (2018) Feng and Sester (2018)
	Multiple disasters	6 Topics about earthquake, oil spill, hurricane	Feature fusion of engineered visual and text statistical features	Yang *et al.* (2011)
		6 Disaster categories	Compared feature fusion and decision fusion	Mouzannar *et al.* (2018)

3.2. Humanitarian aid relevant categorization

The previous task focused on identifying disaster-related posts and the general categorization of posts. However, there is a lack of efforts to retrieve useful information from social media data that benefits for emergency response. More fine-grained information from user-generated texts is needed. Emergency relief needs information from a variety of aspects of a natural disaster event, e.g. casualties, injury, or damage. Vieweg (2012) summarized a list of 32 types of disaster-related information that can be identified from social media and are useful for enhancing situational awareness. This finding is further grouped as the taxonomy of humanitarian aid relevant categories. According to this taxonomy, Imran *et al.* (2013b) applied supervised classification to assign social media texts into four categories: *Caution and advice*, *Information source*, *Donations*, and *Casualties and damage*. In the following studies, this taxonomy was refined, e.g. the category of *Casualties and damage* was separated into *Infrastructure and utilities damage*, *Injured or dead people* in Imran *et al.* (2016). The taxonomy has also been extended with additional aspects, like *Affected individual*, *Missing and found people*, *Displaced and evacuations*, *Requests or needs*, *Response efforts*, *Sympathy and support*. Only recently, Alam *et al.* (2020c) consolidated the existing datasets for humanitarian types classification into a new taxonomy with 11 categories.

The methods used are similar to that of the task in Section 3.1. At the early stages, statistical features and Bag-of-Words features were used for text classification, e.g. Imran *et al.* (2013b). With the advent of word2vec, more approaches were developed to tackle this multi-label classification task using word embedding (Yu *et al.* 2019). The contextual word embedding methods have also been applied and compared with the method using word2vec embedding and TextCNN, e.g. ELMo in Jain *et al.* (2019), BERT in Alam *et al.* (2020c). Alam *et al.* (2020c) demonstrated that BERT outperforms TextCNN by 3.5% on weighted averaged F_1-score and achieves better single-class evaluation performance overall.

The aforementioned methods are text classification based on supervised learning. Since the taxonomy is more complex than a simple binary annotation, crowdsourcing workers were hired, and annotations were confirmed by multiple workers to generate datasets. As reported in Imran *et al.* (2013b), 4406 Tweets annotated by 50–100 workers cost less than $200 in 2012.

In addition to supervised text classification that requires a very time-consuming annotation process, weakly supervised learning, requiring only the annotation of the clusters of text documents, has also been applied (Yao *et al.* 2020). Social media posts were classified into nine fine-grained categories similar to what Imran *et al.* (2016) defined, including *Preventative measure*, *Help and rescue*, *Casualty*, *Housing*, *Utilities and supplies*, *Transportation*, *Flood control infrastructures*, *Business*, *Built-environment hazards*. It can achieve similar performance to that of supervised classification with the annotation effort of 1–2 person-hours. This task has also been performed on images via fine-tuning pre-trained DCNN models (Alam *et al.* 2020b), where the complex taxonomy is simplified as four categories, *Affected, injured, or dead people*, *Infrastructure and utility damage*, *Not humanitarian*, *Rescue volunteering or donation effort*.

Table 6 summarizes the aforementioned studies for this fine-grained classification task. Even though similar methods as in Section 3.1 have mostly been applied, it is

Table 6. Categorisation of social media posts according to the topics of humanitarian aids.

Input	Topic	Category	Method[a]	Example papers
Text	Humanitarian topics	4 Categories	Statistical/BoW features + classic ML	Imran *et al.* (2013b)
		5–10 Categories	BoW features + classic ML	Imran *et al.* (2016) Alam *et al.* (2020a)
			word2vec + TextCNN	Nguyen *et al.* (2017a) Yu *et al.* (2019)
		5 Categories	word2vec/GloVe/ELMo/BERT	Jain *et al.* (2019)
		11 Categories	TextCNN/BERT	Alam *et al.* (2020c)
		9 Categories	Weakly supervised text classification	Yao *et al.* (2020)
Image	Humanitarian topics	4 Categories	Fine-tune pre-trained DCNN	Alam *et al.* (2020b)

[a]Methods are represented in the form of *features + classification methods*, e.g. statistical features + classic ML. *classic ML* includes common machine learning methods such as Naive Bayes, logistic regression, SVM, random forest, etc.

worth highlighting new advances in the use of contextual word embedding and weakly supervised training strategies to improve classification accuracy and efficiency.

3.3. Witness posts identification

As observed in previous studies, reports of casualties and needs for help appear earlier on social media than on traditional media such as news and television during natural disaster events (Zahra *et al.* 2020). However, information from eyewitnesses and bystanders is mixed in social media. The former, which is considered to be more credible (Truelove *et al.* 2015), is generally preferred over the information from the latter. Thus, the identification of eyewitness reports can potentially provide firsthand and more credible information for disaster relief. Distinguishing real observers in social media from bystanders is often considered as a supervised classification task.

In Zahra *et al.* (2020), social media posts were annotated as *eyewitness*, *non-eyewitness*, and *unknown*. Manual analysis was conducted to further refine the eyewitness class with fine-grained categories: *direct eyewitnesses*, *indirect eyewitnesses*, and *vulnerable eyewitnesses*. With this, the characteristics of different types of eyewitness reports were analyzed, e.g. a direct eyewitness often uses first-person pronouns like *I* and *we*; an indirect eyewitness often mentions people related to the author, like *mom* or *dad*; and a vulnerable eyewitness may mention warnings and alerts. These observations provide the basis for extracting additional features from user-generated texts.

As for the content analysis based on texts, statistical features were mostly used, e.g. in Imran *et al.* (2013b) and Zahra *et al.* (2020). In addition, linguistic features (e.g. *first-person pronouns*, *exclamative or emotive punctuation*, *lexical exclamations and expletives*, etc.) in Doggett and Cantarero (2016), meta features (e.g. *length of Tweet*, *mentions or hashtags*, etc.) in Fang *et al.* (2016) and domain-expert features (e.g. *words indicating perceptual senses*, *personalized location markers*, *expletives*, etc.) in Zahra *et al.* (2020) were explored and proved to be beneficial for this task.

The aforementioned studies are summarized in Table 7. As for this task, more task-specific features have been summarized and used to identify eyewitness social media posts. At the same time we also note that for this task, studies tend to focus mainly on the identification step. Further validation of such information, whether it helps to provide analysis and mapping results of better quality is often not considered.

Table 7. Identification of social media eyewitness posts.

Input	Topic	Category	Method[a]	Example papers
Text	Witness posts	Binary	Statistical/BoW features + classic ML	Imran *et al.* (2013b)
			Rule-based filtering	Doggett and Cantarero (2016)
			Statistical/domain expert/word2vec features + SVM	Fang *et al.* (2016)
		Eyewitness/non-eyewitness/unknown	Domain-expert/BoW features + random forest	Zahra *et al.* (2020)

[a]Methods are represented in the form of *features + classification methods*, e.g. statistical features + classic ML. *classic ML* includes common machine learning methods such as Naive Bayes, logistic regression, SVM, random forest, etc.

3.4. Sentiment analysis

The sentiment of social media users during natural disasters is also a common research topic. Sentiment analysis is a relatively mature technique that is now widely used in areas such as e-commerce and other domains of social media analytics.

In the early years, there were studies that trained sentiment analysis models for disaster-related social media data on their own. For example, Caragea *et al.* (2014) trained a sentiment classifier using statistical features and lexicon-based features – SentiStrength.[4] With Naive Bayes or SVM, texts are classified into three categories, *negative*, *neutral*, and *positive*. More studies using statistical and lexicon-based features before 2016 can be found in the review by Beigi *et al.* (2016). Li and Fox (2019) applied a pre-trained RNN model to predict the sentiment scores regarding fear, sadness, and surprise for data collected during earthquake and hurricane events.

In recent years, fewer studies focused on the training of natural disaster sentiment analysis models. Scholars started to focus on analyzing changes in sentiment over time using existing sentiment analysis tools. Alam *et al.* (2020a) directly applied the Stanford sentiment analysis toolkit (Socher *et al.* 2013) for text classification with the same three categories as (Caragea *et al.* 2014). The sentiment dynamics of Twitter posts during three hurricane events were analyzed. It is often observed that negative sentiments dominate during the natural disaster period.

Visual sentiment analysis has also been performed to evaluate user sentiments during disasters based on social media images (Hassan *et al.* 2019). Fine-grained categories have been defined, including *destruction*, *happiness*, *hope*, *neutral*, *pain*, *rescue*, and *shock*. Since the same picture may belong to several categories, a multi-label classifier was trained with DCNN models pre-trained on ImageNet. Similar performance was achieved by these models with an accuracy of around 80%. The above-mentioned studies are summarized in Table 8.

3.5. Damage and damage severity identification

As natural disasters often lead to severe damage, obtaining rapid damage information greatly benefits emergency response. The occurrence of damage and the severity of the damage are also contents that can be interpreted from social media data. Both text and image can be used to extract such information. Based on texts, binary categories, i.e. damage-relevant or damage-irrelevant, have been defined to perform

Table 8. Sentiment analysis of disaster-related social media posts.

Input	Topic	Category	Method[a]	Example papers
Text	Hurricane	3 (Negative, neutral, positive)	BoW/SentiStrength/statistical features + classic ML	Caragea *et al.* (2014)
			Stanford sentiment analysis toolkit	Alam *et al.* (2020a)
	Earthquake/Hurricane	3 (Fear, sadness, surprise)	Pre-trained RNN	Li and Fox (2019)
Image	Multiple disasters	7 Visual sentiment categories	pre-trained DCNN + sigmoid for each class (multi-label)	Hassan *et al.* (2019)

*Methods are represented in the form of *features + classification methods*, e.g. statistical features + classic ML. *classic ML* includes common machine learning methods such as Naive Bayes, logistic regression, SVM, random forest, etc.

supervised text classification (Madichetty and Sridevi 2019). Statistical features, similar to the tasks as summarized in Section 3.1 and Section 3.2, were used. Damage has also been included as one of the humanitarian aid relevant categories as in Imran *et al.* (2016) and Alam *et al.* (2020c) or mixed with the *Casualties* category as in Imran *et al.* (2013b) and Yu *et al.* (2019). Details can be found in Section 3.2.

More research has been performed to extract damage information from social media images. Natural disasters like hurricanes and earthquakes often cause structural damages, e.g. destroyed or shattered bridges, buildings, and roads. Nguyen *et al.* (2017b) and Alam *et al.* (2017, 2018d) grouped images into three categories, namely severe damage, mild damage, and little-to-no damage, by fine-tuning pre-trained DCNN models, i.e. VGG16 (Simonyan and Zisserman 2014). Subsequently, this approach has been used to analyze social media images during three hurricane events in the U.S. in Alam *et al.* (2020a). In a recent work, Alam *et al.* (2020b) reported a weighted averaged F_1-score of 0.758. However, as Alam *et al.* (2018d) noticed, the category of *mild* damage is the most challenging class because this category usually has fewer images to train on (Alam *et al.* 2018d), and the judgment between *mild* and *severe* is very subjective (Li *et al.* 2019). Further experiments using domain adaptation were performed for a simpler setting of binary damage image classification (Li *et al.* 2019).

On the contrary, Nia and Mori (2017) defined even finer categories (including *no damage*, *slight damage*, *moderate damage*, *heavy damage*, *total destruction*) and proposed an innovative approach. They used not only the deep features from the entire image but also ensembled features from inside and outside of a segment mask generated by a segmentation network, which detects foreground from the background. With such design, the network focuses more on the objects that are considered more relevant to damages. We summarized the aforementioned studies in Table 9.

3.6. Flood level estimation

Flood level estimation is an emerging task that has received much attention in recent years. The in-time estimation of flood extent and depth improves situational awareness and is beneficial for hydrological studies. A few studies tried to derive water levels from social media text via manual interpretation (Li *et al.* 2018) or template matching. Combinations of numbers and length units (e.g. *m*, *cm*, *in*) are searched in

Table 9. Identification of disaster damage and damage severity from social media posts.

Input	Topic	Category	Method	Example papers
Text	Earthquake	Binary	BoW/tf-idf features + classic ML	Madichetty and Sridevi (2019)
Image	Hurricane/Earthquake	3 (Severe/mild/no)	Fine-tune pre-trained DCNN	Nguyen *et al.* (2017b) Alam *et al.* (2017, 2018d, 2020a, 2020b)
		5 Severity categories	Ensemble of features from entire image, inside and outside of a segment mask	Nia and Mori (2017)
		Binary	Domain adaptation	Li *et al.* (2019)

the user-generated texts (Eilander *et al.* 2016). Pre-defined keywords like 'knee-deep' have also been used as water level indicators (Smith *et al.* 2017). Despite the success of the above efforts, social media users who mentioned flood depth in texts during flood events are rare.

Visual information from social media contributes more to the water level estimation. In many early studies, water levels were manually extracted from social media images that contained objects of known size submerged in water (Assumpção *et al.* 2018). The most commonly used indicators for such a manual analysis are standing people and wheels of cars in water (Kutija *et al.* 2014).

With the development of DCNNs, the efficiency and accuracy of image classification and object recognition have been greatly improved. Pereira *et al.* (2019) classified social media images into three water level categories (i.e. no flood, below 1 m, and above 1 m). Deep features, extracted from the entire image by DenseNet (Huang *et al.* 2017) and EfficientNet (Tan and Le 2019), were used. Other studies explored the application of object detection to assess water levels. Partially submerged objects in the water received more attention. Chaudhary *et al.* (2019, 2020) detected person, car, bus, bicycle, and house by a Mask R-CNN model (He *et al.* 2017) as water level indicators. With the local deep features around these detections, objects are classified into 11 water levels, which correspond to the water height intervals in real numbers, e.g. 0 cm, 1 cm, 10 cm, 21 cm, until 170 cm.

With the improved performance of human keypoint detection, e.g. OpenPose,[5] scholars started to investigate the possibility of using human keypoints to estimate water levels. Quan *et al.* (2020) made use of detected human pose and well-designed rules to compare the relation between body keypoints and person segments. Multiple empirical thresholds were applied on ratios between different body parts to represent such a water level situation. Two categories (i.e. above the knee and below the knee) were assigned for the images containing people in the flood scenarios. Based on the results of body keypoints, object detection, and image segmentation, Feng *et al.* (2020) developed a distance feature, i.e. distances between body keypoints and the detected water level, to represent the relationship between human and water. People in the scenarios can be classified into five categories with respect to submerged body parts. These five categories include: ankle, knee, hip, chest, and none. Vehicles have also been used as water level indicators. Huang *et al.* (2020) detect vehicle tires with

Table 10. Flood water level estimation from social media posts.

Input	Topic	Prediction	Method	Example papers
Text	Water level	Real numbers	Text template matching and manual extraction	Eilander *et al.* (2016) Smith *et al.* (2017), Li *et al.* (2018)
Image	Water level	3 (Above 1 m/below 1 m/no)	Fine-tune pre-trained DCNN	Pereira *et al.* (2019)
		11 Water levels	Adapted Mask R-CNN using local deep features	Chaudhary *et al.* (2019, 2020)
		2 (Above/below the knee)	Apply pre-designed rules on Mask R-CNN and human keypoints outputs	Quan *et al.* (2020)
		5 (Chest/hip/knee/ankle/none)	Classify distance features generated from Mask R-CNN, human keypoints, semantic segmentation	Feng *et al.* (2020)
		Real numbers	Car tires detection with Mask R-CNN	Huang *et al.* (2020)

Mask R-CNN. The water levels of the real numbers can be estimated based on the submerged tire segment.

The above-mentioned studies are summarized in Table 10. Flood level estimation is a rapidly developing field of research, especially in recent years. It is worth noting that classification models are limited by the granularity of the pre-defined categories and are mostly used as indicators of flood severity to enhance situational awareness. The frequently used water level indicator – humans, poses certain challenges, given the discrepancies in individuals' height. For a more precise water level estimation, cars are regarded as an indicator with much less variation, where an averaged height difference of 6–8 cm is reported in the very recent regression-based approach (Park *et al.* 2021). However, only a small proportion of photos on social media contain cars or humans in flood. Therefore, a reasonable combination of predictions based on different targets with different precision levels needs to be addressed in future research. In addition, LiDAR-equipped mobile phones are becoming popular and have the opportunity to provide accurate 3 D measurements.

3.7. Disaster phases categorization

In disaster management, natural disasters are regarded as recurring events, typically divided into four phases (also known as the disaster management cycle), namely *preparedness*, *response*, *recovery*, and *mitigation* (Neal 1997). Such division facilitates our understanding of disaster events. However, adjacent phases may overlap, and the severity of the disaster affects the length of individual phases. Social media can therefore serve as an information source that benefits the identification of such temporal boundaries among disaster phases.

In early studies, social media posts are categorized into simply pre-defined disaster phases like *before*, *during*, and *after* a disaster, e.g. in Iyengar *et al.* (2011) and Chowdhury *et al.* (2013). Other research is based on the disaster management cycle. Huang and Xiao (2015) adapted the ontology but excluded the *Mitigation* phase

Table 11. Identification of the disaster phases of social media posts.

Input	Topic	Category	Method[a]	Example papers
Text	Hurricane	3 (Before, during, after)	BoW/verb tense/statistical features + SVM	Iyengar *et al.* (2011)
	Hurricane/earthquake	3 (Pre-, during-, post-incident)	BoW/verb tense/keyword occurrence + classic ML	Chowdhury *et al.* (2013)
	Hurricane/flood	22 Categories w.r.t 4 phases (preparedness, response, impact, recovery)	BoW + logistic regression	Huang and Xiao (2015)
Image	Hurricane/flood	4 (Preparedness, response, impact, recovery)	Fine-tune pre-trained DCNN	Wang *et al.* (2020)

[a]Methods are represented in the form of *features + classification methods*, e.g. statistical features + classic ML. *classic ML* includes common machine learning methods such as Naive Bayes, logistic regression, SVM, random forest, etc.

because it is a relatively long-term activity. An additional category, i.e. *Impact*, is included because it contains crucial information for disaster response (Huang and Xiao 2015) (Table 11).

Despite the differences in the category definitions, the solution to this supervised text classification task is, in general, similar to the multi-class text classification presented in Section 3.1 and Section 3.2. It is worth noting that special attention is given to the tense of the verbs, e.g. in Iyengar *et al.* (2011) and Chowdhury *et al.* (2013), as this task focuses on the temporal dimension of the contents in the text. Since some of the words can be directly used to identify disaster phases, the occurrence of keywords, such as *alert*, *aftermath*, were also included as additional features in Chowdhury *et al.* (2013). Huang and Xiao (2015) reported a weighted average F_1-score of 0.664 over 22 categories. Besides texts, social media images have also been used for this task. With the same ontology, social media images are classified into these four same disaster phases (Wang *et al.* 2020). Due to the imbalance of the training dataset, only the class *Impact* achieved a satisfying F_1-score of 0.88, while the worst class *Recovery* is 0.02. The aforementioned studies are summarized in Table 6.

3.8. Unsupervised topic modeling

In addition to tasks using pre-defined categories and manually annotated training datasets, topic modeling was used to extract disaster-related information in an unsupervised manner. LDA – Latent Dirichlet allocation (Blei *et al.* 2003) that divides text documents into topics is one of the best-known methods. The number of topics is a hyper-parameter, which needs to be defined in advance.

LDA was embedded in visual analytics (Chae *et al.* 2012), where users of the system can apply a varying number of topics and explore topics with different granularity. Others applied LDA to extract social media posts related to a particular disaster topic. Wang *et al.* (2016a) applied LDA on Chinese social media posts during a rainstorm. Usually, the best number of topics was selected after repeated attempts. Since the derived topics are difficult to interpret, these topics were manually cleaned and grouped into five categories (*traffic*, *weather*, *disaster information*, *loss and influence*, *rescue information*). An SVM classifier was further trained on this topic-level annotated

Table 12. Unsupervised topic modelling on social media posts.

Input	Topic	Category	Method	Example papers
Text	Earthquake	Varies by user	Visual analytics for LDA topics	Chae *et al.* (2012)
		25 Topics	Select disaster-related LDA topics by keyword	Resch *et al.* (2018)
	Flood	5 Categories	Annotate LDA topics to train supervised models	Wang *et al.* (2016a)

data to cope with new incoming data. Resch *et al.* (2018) applied LDA to extract earthquake-related Tweets. A topic with the keyword *earthquake* is selected, and a cascading LDA was applied to this topic to detect subtopics. However, the subtopics still need human interpretation.

The above-mentioned studies are summarized in Table 12. The general challenge of using unsupervised topic modeling to extract disaster-related information is the proper selection of the number of topics and the interpretation of the detected topics.

4. Analysis and mapping

After the extraction of natural disaster-relevant information from social media, the next step is to analyze their distribution and generate mapping results. In general, the analysis and mapping of disaster-related social media data can be divided into two categories. Social media data can be analyzed standalone to detect spatial, temporal, and spatiotemporal patterns. They can also be combined with other data sources to improve their validity. In the following, these two aspects are presented in detail.

4.1. Standalone analysis of social media VGI

Almost all of the social media data are available with timestamps, however, only around 2% of the Twitter posts have location information (Dittrich and Lucas 2014). In the following, the research of social media VGI analysis of disasters is introduced from three aspects: temporal analysis, spatial analysis, and spatiotemporal analysis.

4.1.1. Temporal analysis

Temporal analysis is one of the most basic components and has been widely used in research on disaster-related social media posts. Time series of the number of social media posts over time are compared with significant events reported by local newspapers, e.g. during a wildfire event (De Longueville *et al.* 2009), aiming to investigate the full course of the natural disaster event and the behavior of users on social media platforms. In addition, such temporal analysis can also be carried out over a long time frame. Daly and Thom (2016) analyzed social media photos associated with fires over a seven-year period and linked peaks in the temporal analysis to 13 major fire events. For natural disaster events, an exponential distribution on the time series of the disaster-related posts is often observed after a disaster, such as the observed pattern during an earthquake and a typhoon in Sakaki *et al.* (2010).

With the time series of the number of social media posts, anomaly detection can be performed to detect events, e.g. using Seasonal Trend Decomposition (STL) as in Chae *et al.* (2012). In addition, the change of sentiment over time has been observed

Table 13. Standalone analysis of social media VGI for disasters.

Type	Information	Method and visualization tools	Example papers
Temporal	Count of posts	Number of posts as time series	De Longueville *et al.* (2009) Sakaki *et al.* (2010), Daly and Thom (2016)
	Anomaly/event	STL on time series of LDA topics	Chae *et al.* (2012), Chae *et al.* (2014)
	Sentiment score	Sentiment score as time series	Alam *et al.* (2020a)
Spatial	Distribution	Location points or markers on static or online maps	Alam *et al.* (2020a), Feng *et al.* (2020)
	Heatmap	KDE generated raster heatmap	Wang *et al.* (2016b), Cervone *et al.* (2016)
	Heat regions	Aggregation counts to grids	MacEachren *et al.* (2011), Stefanidis *et al.* (2013)
		Aggregation counts to administrative polygons	Crooks *et al.* (2013), Cresci *et al.* (2015)
		Aggregation counts to Voronoi polygons	Cerutti *et al.* (2016), Wang *et al.* (2016a)
	Hotspot	Hotspot detection on grids with Getis-Ord Gi*	Resch *et al.* (2018)
		Hotspot detection on administrative polygons with Getis-Ord Gi*	Feng and Sester (2018)
		Hotspot detection with Getis-Ord Gi* and interpolated with KIB	Panteras and Cervone (2018)
	Clusters	DBSCAN OPTICS DBSCAN>Daly and Thom OPTICS>Wang	Daly and Thom (2016) Wang *et al.* (2016a)
Spatio-temporal	Typhoon path	Kalman filter and particle filter	Sakaki *et al.* (2010)
		Mean center of typhoon related VGI	Zhao *et al.* (2019)
	Heatmaps over time	Heatmaps generated separately by KDE for multiple time slots	Chae *et al.* (2014), Zhu *et al.* (2018)
		ST-KDE and presented in space-time cube	Kersten and Klan (2020)
	Hot spots over time	Get-Ord Gi* applied on different dates	Kersten and Klan (2020)
	ST-clusters	OPTICS and visualized in time-space cube	Fuchs *et al.* (2013), Cerutti *et al.* (2016)
		ST-DBSCAN	Huang *et al.* (2018c) Kersten and Klan (2020)

in Alam *et al.* (2020a) to visualize the changes in negative sentiment score in response to a disaster event (Table 13).

4.1.2. Spatial analysis

Spatial analysis is mainly used to visualize the spatial distribution of social media posts containing location information. Earlier studies presented locations of natural disaster-relevant posts simply as point symbols on static maps or markers with pop-up windows on online maps.

Kernel Density Estimation (KDE) is often used to generate heatmaps from location points, where the concentration of location points can be represented as a raster image, e.g. flood in Cervone *et al.* (2016), and wildfire in Wang *et al.* (2016b). In a KDE, the bandwidth or the radius of the kernel is a hyperparameter which is mostly chosen empirically. Since social media locations have a strong bias due to the uneven distribution of population and social media users, population data has been used to normalize the KDE results, such as (Wang *et al.* 2016b).

In addition, the location points can be aggregated to a variety of spatial levels, such as grids (MacEachren *et al.* 2011; Stefanidis *et al.* 2013), administrative polygons (Crooks *et al.* 2013), or Voronoi polygons (Cerutti *et al.* 2016; Wang *et al.* 2016a), by counting the number of posts. The distribution of these aggregated points can be visualized as choropleth maps that present the concentration of location points with respect to these spatial units. In general cases, population data from agencies (Cresci *et al.* 2015) or the averaged historical social media posts amounts (Feng and Sester 2018) are used to normalize these aggregation results.

With the information aggregated to spatial units, hot spot analysis can be further performed, e.g. using Getis-Ord-Gi* (Ord and Getis 2010). The output z scores and p scores represent the statistical significance of spatial clustering, based on the values in the spatial units. High or low values clusters (i.e. hot spots or cold spots) can be identified spatially. In terms of disaster-related VGI, Getis-Ord-Gi* has been applied to hotspot analysis of floods (Feng and Sester 2018; Panteras and Cervone 2018) and earthquakes (Resch *et al.* 2018). In addition, KIB (Kernel Interpolation with barriers) has been applied on the result of Getis-Ord-Gi* to provide a smoother visualization (Panteras and Cervone 2018). Even though such a mechanism of aggregating location points into areal units has been widely applied, this strategy may suffer from Modifiable Area Unit Problem – MAUP (Ratcliffe 2004) (i.e. the identified spatial patterns can vary with a changing spatial unit).

Clustering is an approach that does not need predefined spatial units. Since the number of clusters is normally unknown in advance, density-based clustering is often applied, e.g. DBSCAN (Ester *et al.* 1996) for wildfire events detection in Daly and Thom (2016), OPTICS (Ankerst *et al.* 1999) for urban flood event in Wang *et al.* (2016a). Wang *et al.* (2016a) clustered the Weibo posts (similar to Twitter) in China during the flood event in Beijing in June 2012.

4.1.3. Spatiotemporal (ST) analysis

Spatiotemporal (ST) analysis is often used to present the changes of disaster-reported locations over time. Kalman filter and particle filter were used to estimate the path of a typhoon event based on the locations of social media posts (Sakaki *et al.* 2010), which has been further compared with the actual typhoon path. In addition, KDE has been applied to generate heatmaps at multiple temporal periods separately to discover the changes of spatial distribution patterns over time (Chae *et al.* 2014; Zhu *et al.* 2018). Instead of partitioning the time axis into intervals, ST-KDE takes time as an additional dimension for three-dimensional density estimation and has been visualized in the space-time cube (Kersten and Klan 2020).

In addition, ST-clusters are detected based on the VGI points with timestamps. OPTICS has been applied considering the temporal dimension, and the ST-clusters were visualized in the space-time cubes (Fuchs *et al.* 2013). The detected spatiotemporal clusters were manually validated with the evidence on the Internet, which confirmed the potential of social media data to be used as a distributed sensor for flooding. ST-DBSCAN has also been utilized in Huang *et al.* (2018c) and Kersten and Klan (2020) to detect flood-related ST-clusters as events.

4.2. Analysis of social media VGI in combination with other sources of information

Due to the sparseness and uncertainty of social media locations, standalone analysis of social media VGI can hardly provide information with full coverage and high-level details. Therefore, another branch of research focuses on the integration of social media VGI with other information sources for natural disaster mapping.

4.2.1. Digital terrain models (DTM)

Digital terrain models (DTM) provide the basic relief information of an area. The terrain itself has bulges and depressions, which indicate where there is a high chance of flooding. A straightforward way is to estimate a flood surface with the water level information. There are a series of early studies utilizing social media for the analysis of Queensland floods in 2011 (McDougall 2011a, 2011b; McDougall and Temple-Watts 2012). Texts, photos and videos from Flickr and Facebook were manually interpreted. In order to obtain precise water levels and exact locations for these user observations, field surveys were conducted with Real-time kinematic (RTK) GPS and conventional survey methods. In this way, 23 selected sites with photographs of the 2011 flooding in Brisbane, Australia, were verified. A flood surface was estimated with these measures and the flood extent was generated by subtracting the DTM by this flood surface (McDougall and Temple-Watts 2012).

For the fluvial flood in 2013 in Dresden, Germany, Tweets were filtered by flood-related keywords. Experts or voluntary annotators were asked to estimate the relevance regarding inundation mapping and the water level from social media photos on a web-based platform. Five inundation depth estimates were used to estimate a flood surface with DTM via bilinear spline interpolation (Fohringer *et al.* 2015). Instead of estimating one global flood surface, Li *et al.* (2018) estimated a simple flood plane based on each water level estimate from either social media or river gauges. However, each estimate has an impact only on the area around it and decreases with distance (i.e. Inverse Distance Weighting – IDW). The flood probability of all estimates is summed and then normalized to the 0–100 range. The results show a high agreement with the USGS flood mapping results. However, methods that use only terrain information ignore the hydrological and hydraulic aspects of flood events.

4.2.2. Disaster simulation

Disaster simulation is another data source that is often compared with social media VGI. For flood events, hydrodynamic models provide the estimation of flood-prone and inundation areas based on a DTM. Flood-related social media posts can be used as evidence to evaluate the flood modelling results (e.g. Aulov *et al.* 2014; Kutija *et al.* 2014; Smith *et al.* 2017). Aulov *et al.* (2014) validated the surge model forecasts from NOAA with social media data. Smith *et al.* (2017) applied hydrodynamic modelling for the 2012 flood events in Newcastle upon Tyne, UK with a 2D hydraulic model. Modelling results were compared with the locations of flood-relevant social media posts. Eilander *et al.* (2016) applied flood mapping based on 888 water level mentions from social media texts during three days in 2015 in Jakarta, Indonesia. Combined with DTM and hydraulic models, flood extent and water depth maps were generated.

4.2.3. Remote sensing

Remote sensing has also received a lot of attention. Contemporary technology can identify floods, fires, and building damages (Dong and Shan 2013) from remote sensing products. Flood is the most studied natural disaster, allowing information on the extent of flooding to be obtained for disaster management and emergency response. Flood extent can be extracted from remote sensing imagery to generate a flood probability map, e.g. using NDWI – Normalized Difference Water Index (e.g. Huang *et al.* 2018b), Modified NDWI (e.g. Rosser *et al.* 2017) or machine learning models (e.g. Sarker *et al.* 2019). However, for densely built-up urban areas, the performance of flood detection from remote sensing products is often compromised. In addition, occlusion due to observation angles and shadows of buildings and trees may also lead to misses of flood detections. In contrast, VGI data appear more frequently in cities, as more users live there. Thus, flood-related social media posts with geolocation can be used as an ideal local complement to remote sensing flood detection. VGI locations have been used to generate flood probability maps by applying kernel smoothing, e.g. with a quadratic kernel in Schnebele *et al.* (2014b), Gaussian kernel in Cervone *et al.* (2016). They are merged with remote sensing detection or other data sources (e.g. flood hazard map based on DTM and river gauge data) with a weighted sum overlay. Due to their uncertainty, social media data were only given low weight (Schnebele *et al.* 2014b; Cervone *et al.* 2016). Still, it has a considerable effect, even if only using a small amount of VGI data, as demonstrated in Schnebele and Cervone (2013).

Based on a Digital Terrain Model, Huang *et al.* (2018a) queried the height of each flood-related VGI location and marked areas below that height as having a higher probability of flooding. In order to limit the impact range of individual VGI location, this probability decreases with increasing distance, which is similar to IDW (Inverse distance weighting). VGI points have been assigned with weights based on the NDWI wetness values around each point. By applying a weighted sum, the flood probability map was generated. In another research for the same event, Huang *et al.* (2018b) created a basic flood probability map on DTM and gauge observations, which was integrated with a flood probability map generated using quadratic kernel smoothing on NDWI. A local morphological dilation was applied to increase the flood probability for the area with VGI data points. This study showed that even though flood-relevant information takes up only a very small proportion of the social media data streams, the geotagged flood-relevant posts can still contribute to flood monitoring and extent mapping. In further, Wang *et al.* (2018a) introduced a theoretical and algorithmic framework for heterogeneous data fusion of remote sensing data and social media data based on the maximum entropy and the least effort principle.

As for fire, Boulton *et al.* (2016) examined the spatial and temporal distribution between the remote sensing fire detection and locations of fire-related social media posts. Bischke *et al.* (2016) enriched the visualization of remote sensing fire detection with ground observations from social media. Rashid *et al.* (2020) detected fire reports from social media and dispatched drones for more reliable sensing.

4.2.4. Participatory VGI

Participatory VGI can also be combined with social media VGI to increase the quantity or refine the quality of natural disaster reports from citizens. The most well-known form of Participatory VGI is collaborative mapping. When natural disasters occur, OpenStreetMap contributors are quickly mobilized to digitize map data of buildings and roads for the affected areas, such as Humanitarian OpenStreetMap Team (HOG),[6] MapAction.[7] This data can directly serve disaster relief and is rarely integrated with social media VGI. Furthermore, there are many efforts where voluntary citizens can directly provide or interpret information related to natural disasters.

Agencies and researchers provide websites, online maps, and mobile apps to collect disaster reports from volunteers, e.g. Ushahidi.[8] They are often treated similarly to social media VGI overlaid for visualization (McDougall and Temple-Watts 2012; Wang *et al.* 2018b) or combined with remote sensing detections via kernel smoothing. Such crowdsourced information is often considered to have a higher location accuracy than social media VGI (Wang *et al.* 2018b). In addition to reporting location as point coordinates, the U-Flood project (CBS News 2017) asked volunteers to mark flooded streets in Houston during Hurricane Harvey. Similarly, sketch maps of roads have been used

Table 14. Analysis of social media VGI in combination with other sources of information.

Source	Disaster	Method and purpose	Example papers
DTM	Flood	Estimation of a flood surface using interpolation of water levels from post-event survey on VGI locations	McDougall (2011a, 2011b) McDougall and Temple-Watts (2012)
		Estimation of a flood surface using interpolation of water levels from social media images	Fohringer *et al.* (2015)
		Estimation and integration of local flood surfaces based on water level estimates from gauges and VGI	Li *et al.* (2018)
Disaster simulation	Flood	Validate flood hydraulics simulation	Aulov *et al.* (2014) Eilander *et al.* (2016) Smith *et al.* (2017)
Remote sensing	Flood	Kernel smoothing generate VGI layer merged with RS and others using a weighted sum overlay	Schnebele and Cervone (2013) Schnebele *et al.* (2014b) Cervone *et al.* (2016)
		Integrate Tweets with NDWI flood detection by applying Gaussian kernel	Huang *et al.* (2018a)
		Integrate gauge data and Tweets with NDWI flood detection by applying kernel-smoothing and local morphological dilation	Huang *et al.* (2018b)
		Integrate Tweets with RS flood detection using maximum entropy and the least effort principle	Wang *et al.* (2018a)
	Fire	Correlation analysis with RS fire detection	Boulton *et al.* (2016)
		Enrich visualization of RS fire detection with ground observations	Bischke *et al.* (2016)
		Social-media-driven drone sensing	Rashid *et al.* (2020)
Participatory VGI	Flood	Volunteer reports with geo-location available via websites, online maps, and mobile apps	McDougall and Temple-Watts (2012) Wang *et al.* (2018b)
		Interpretation of damage and location by volunteers, followed by fusion using the kernel method	Schnebele *et al.* (2014a, 2014b) Cervone *et al.* (2016)

for collecting flood risk perceptions of residents in less-developed countries (Klonner *et al.* 2018; Brandt *et al.* 2020). Such line segments are not often integrated into flood maps, but they can be similarly integrated as presented in Schnebele *et al.* (2014b), where road closure information collected from newspapers was integrated with flood maps via kernel methods (Table 14).

In addition to providing disaster reports, the participatory approach can also be used to interpret disaster information or refine the location. Collaborative damage mapping based on aerial and satellite images has been presented in many studies (Kerle and Hoffman 2013), where volunteers manually process aerial photos to identify damaged buildings and infrastructures. In addition, volunteers who are familiar with the area can provide locations for non-geotagged images and videos from social media and news stories. For example, the GISCorps team[9] provided the PhotoMappers platform for participatory interpretation of social media images in terms of location and damage. Such information has been considered for fusion together with geotagged social media posts as presented in Schnebele *et al.* (2014a, 2014b) and Cervone *et al.* (2016). The geotagged aerial photos from Civil Air Patrol (Schnebele *et al.* 2014a; Cervone *et al.* 2016), traffic camera images (Schnebele *et al.* 2014b), Youtube videos, Twitter posts were combined to provide the spatial and temporal maps of a flood. In the fusion step, these data sources are applied with the kernel smoothing and combined with the weights.

In addition, there are studies that lead to further engagement of social media users, inviting the authors of disaster-related posts to provide details of the location and severity of the disaster, e.g. PetaJakarta (Ogie *et al.* 2019).

5. Opportunities and challenges

Although social media data VGI has been widely considered in many natural disaster studies, there are still many opportunities and challenges in need of further investigation. From the review above, we could identify the following five main aspects.

5.1. VGI quality on time, location, and content

The quality issue for VGI is mentioned in the vast majority of previous reviews or research (Klonner *et al.* 2016; Yan *et al.* 2020), and it remains a challenge today. Even though deep learning-based algorithms extract the textual and visual information far more accurately and efficiently than before, it is still difficult to know whether a certain post is what the user is observing at the stated time and the stated location. Social media posts with precise GPS locations are very rare, as users are more likely to provide abstract location names than their actual coordinates.

There is a branch of studies that utilized Name-Entity Recognition to identify locations mentioned in the social media texts. Users will usually provide location descriptions at the administrative unit level in their posts. In emergency situations, they tend to provide very detailed location descriptions, such as exact house numbers and road intersections (Hu and Wang 2020). However, it is still a great challenge for current geoparsing services to correctly identify and locate place names, especially with different languages, dialects, and cultural differences.

Efforts have also been made to take advantage of visual information, aiming to infer geographic coordinates, e.g. in Muller-Budack *et al.* (2018). Unfortunately, the accuracy of geolocation estimation approaches, e.g. based on global-scale Im2GPS[10] dataset, is still far from sufficient for the purpose of analyzing and mapping natural hazards. However, for urban scenarios with street-level geo-tagged images, e.g. San Francisco Dataset (Chen *et al.* 2011), the 2 D image retrieval-based method has achieved 10 m-level localization error and 5 m-level after combining with a local Structure from Motion (SfM) reconstructions (Sattler *et al.* 2017). With the availability of detailed 3 D LiDAR data, sub-meter level camera pose estimation has been achieved (Cattaneo *et al.* 2019, 2020). As more geotagged images and LiDAR data accumulate, and as computing power increases, vision-based localization has great potential to provide precise location estimation also for social media images.

Furthermore, the current interpretation of natural disaster-related information focused primarily on posts in English. However, natural disaster events are described differently for users who speak different languages, dialects, and have cultural differences. Currently, only little research has been done in this area.

5.2. Triage of rescue requests

When natural disasters occur, emergency services such as 911 are often overloaded. Many people seek for help on social media. Humanitarian aid relevant categories as summarized in Section 3.2 can identify posts on *Casualties and damage*, and *Requests or needs*. The work summarized in Section 3.5 can identify severely damaged structures, and the work in Section 3.6 identifies people suffering from a severe flood. However, this does not yet allow for identifying those victims who are most in need of rescue. Therefore triage of help-seekers is very important. Unfortunately, the literature on automated social media analysis rarely touches on this aspect.

In the healthcare system, machine learning methods to analyze vital signs, symptom descriptions, and active medical history as presented in Berlyand *et al.* (2018) demonstrated the potential to automatically triage patients. In some cases, it even outperforms humans (Levin *et al.* 2018). However, individuals on social media can hardly provide such expert information. Chaudhry and Yuksel (2019) presented a Public Safety Framework design that proposed to use AI robots to communicate with people who request for help on social media per phone call and determine the level of emergency. Victim triage often requires detailed questioning. However, it is worth noting that during natural disasters, many victims may provide their detailed information to enhance credibility. Therefore, victim triage may also become a reality through a more intelligent analysis of the text and images provided by the victims and the establishment of appropriate communication mechanisms.

5.3. Integration of social media VGI with other sources of information

As presented in Section 4.2, information from various sources can be integrated to provide a comprehensive and detailed overview to assist disaster responses (Schnebele and Cervone 2013; Schnebele *et al.* 2014b; Cervone *et al.* 2016; Huang *et al.* 2018a, 2018b). Existing studies so far have proposed solutions to integrate social media VGI with other information sources through heuristic weighting. However, the proper determination of weights remains a challenge, given its subjective nature. Thus, guidelines need to be provided to properly choose the weights based on the characteristic of different information sources. Individual VGI record also owns varying degrees of confidence and time-location quality that are worth considering individually in future research.

5.4. Fake and synthetic user-generated contents

Fake news and content on Twitter are often retweeted by many more users and spread far more rapidly (Vosoughi *et al.* 2018). Fake news regarding natural disasters may lead to misallocation of resources, and in extreme cases, endangering people's lives (Johnson 2020). Detection of fake news (Zhou and Zafarani 2020) has been studied mainly based on content interpretation. However, this requires not only the understanding of current social media content provided by users but also the comparison and verification with information from other sources and other users.

Current developments in computer vision, such as Generative Adversarial Networks (GANs), can generate photo-realistic human faces (Karras *et al.* 2019) and videos (e.g. text-driven video synthesis in Thies *et al.* 2020). It is believed that in the near future, even novices can also be proficient in using these techniques to generate fake pictures or videos. Therefore, it is crucial to develop methods that can efficiently and automatically detect such fake and synthetic content.

5.5. Video and LiDAR, potential social media content for disaster-related VGI analysis

In recent years, short videos are receiving increasing attention. Conventional social media platforms, such as Facebook, Twitter, Instagram, Weibo, all support users to share videos. Tiktok,[11] as an emerging platform for short videos, is very popular among young people. Disaster-related videos have been considered in early-stage studies via manual interpretation (McDougall 2011a, 2011b; McDougall and Temple-Watts 2012). However, few studies explored the possibility of adopting automated procedures for analyzing social media videos. Meanwhile, user-generated videos are often shaking or blurry, posing challenges for traditional video analytics.

The dynamic nature of videos facilitates the retrieval of other essential information. For flood, scholars started to use social media videos to estimate water flow velocity (Le Boursicaud *et al.* 2016). However, such an approach needs to survey ground reference points. The growing popularity of LiDAR on mobile devices has made it much easier for users to obtain 3D measurements. With this, social media users can potentially provide a rough estimation of the water level or flow velocity easily in the future.

6. Conclusions

In this paper, we provided an comprehensive review of studies that extract and analyze natural disaster-related VGI from social media. We identified eight popular tasks from which disaster-related information can be extracted, including (1) Extraction of disaster-related social media posts, (2) Humanitarian aid relevant categorization, (3) Witness posts identification, (4) Sentiment analysis, (5) Damage and damage severity identification, (6) Flood level estimation, (7) Disaster phases categorization, and (8) Unsupervised topic modeling. In addition, two research branches that make further use of the extracted information were reviewed, namely the standalone analysis of social media VGI and the analysis of social media VGI in combination with other sources of information. For each task and research branch, we grouped the studies with similar components and summarized the existing solutions that were selected based on an extensive literature search. Based on the results of the review, we identified challenges from five aspects, in which new research opportunities arise. In addition to the most frequently discussed challenge on VGI quality and better integration of social media VGI with other sources of information, we also noticed the potential to triage rescue requests and the new possibilities offered by emerging social media platforms and devices, such as the new information that can be extracted from videos, or LiDAR on smartphones. At the same time, with the rapid spread of computer vision technologies and applications, fake as well as synthetic information will become a major challenge for the future use of social media data in the emergency management domain.

Due to the large number of relevant articles, this paper can only compile the mainstream tasks and solutions through a narrative review after a literature search. Therefore, there are certainly other tasks and solutions that receive less attention and need to be considered in the next step. Nevertheless, our work provides a valuable reference for the development and design of social media VGI analysis frameworks, so that it is possible to quickly identify content that can be considered for information extraction and tools for spatiotemporal analysis.

Notes

1. Web of Science. https://www.webofknowledge.com [Accessed 07 May 2021].
2. Scopus. https://www.scopus.com/ [Accessed 07 May 2021].
3. Natural Disasters – Homeland Security. https://www.dhs.gov/natural-disasters [Accessed 07 May 2021].
4. SentiStrength. http://sentistrength.wlv.ac.uk/ [Accessed 07 May 2021].
5. OpenPose. https://github.com/CMU-Perceptual-Computing-Lab/openpose [Accessed 07 May 2021].
6. Humanitarian OpenStreetMap Team. https://www.hotosm.org/ [Accessed 07 May 2021].
7. MapAction – The humanitarian mapping charity. https://mapaction.org/ [Accessed 07 May 2021].
8. Ushahidi. https://www.ushahidi.com/ [Accessed 07 May 2021].

9. NAPSG Foundation Disaster Response Partnership – Crowdsourcing Projects. https://www.giscorps.org/napsg-crowdsource-photos/ [Accessed 07 May 2021].
10. Im2GPS Dataset: http://graphics.cs.cmu.edu/projects/im2gps/ [Accessed 07 May 2021].
11. Tiktok. https://www.tiktok.com/ [Accessed 07 May 2021].

Acknowledgements

The authors would like to thank associate editor Dr. Urska Demsar and anonymous reviewers for their insightful comments and suggestions, which greatly improved this article.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data and codes availability statement

The data and codes that support this study are available in GitHub with the link https://github.com/yuzzfeng/SocialMediaVGI4disasters.

Funding

This work was supported by the German Federal Ministry of Education and Research (BMBF) under grant number [033W105A].

Notes on contributors

Yu Feng is currently a postdoctoral researcher at the Institute of Cartography and Geoinformatics (ikg), Leibniz University Hannover. He received his M.Sc. and PhD in Geodesy and Geoinformatics from the same institution in 2015 and 2021. His research interests include social media VGI analysis, LiDAR data processing, map generalization, and deep learning.

Xiao Huang received his BS degree from Wuhan University in 2015. He obtained his Master's degree from Georgia Institution of Technology in Geographic Information Science and Technology in 2016 and PhD in Geography from the University of South Carolina in 2020. He is currently an assistant professor in the Department of Geosciences at the University of Arkansas. His research interests cover GeoAI, deep learning, and human-environmental interactions.

Monika Sester is a surveying engineer by training (University Karlsruhe) and earned her PhD on a topic of Machine Learning at the University of Stuttgart in 1995, and her habilitation in 2000 on the automatic generation of multiple representations of geodata. Since 2000 she has been professor and head of the Institute of Cartography and Geoinformatics (ikg) at Leibniz University Hannover. She and her group work on the automation of spatial data processing with methods from computational geometry, optimization and AI.

ORCID

Yu Feng http://orcid.org/0000-0001-5110-5564
Xiao Huang http://orcid.org/0000-0002-4323-382X
Monika Sester http://orcid.org/0000-0002-6656-8809

References

Abed Al Ahad, M., *et al.*, 2020. The effect of air-pollution and weather exposure on mortality and hospital admission and implications for further research: a systematic scoping review. *PLoS One*, 15 (10), e0241415.

Ahmad, K., *et al.*, 2017a. CNN and GAN based satellite and social media data fusion for disaster detection. *In*: *Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Ahmad, K., *et al.*, 2018. A comparative study of global and deep features for the analysis of user-generated natural disaster related images. *In*: *2018 IEEE 13th image, video, and multidimensional signal processing workshop (IVMSP)*. New York: IEEE, 1–5.

Ahmad, S., *et al.*, 2017b. Convolutional neural networks for disaster images retrieval. *In*: *Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Alam, F., *et al.*, 2018a. A Twitter tale of three hurricanes: Harvey, Irma, and Maria. *In*: *Proceedings of the international ISCRAM conference*, Rochester, UK, 20–23 May 2018, 553–572.

Alam, F., *et al.*, 2020b. Deep learning benchmarks and datasets for social media image classification for disaster response. *arXiv preprint arXiv:2011.08916*.

Alam, F., *et al.*, 2020c. Standardizing and benchmarking crisis-related social media datasets for humanitarian information processing. *arXiv preprint arXiv:2004.06774*.

Alam, F., *et al.*, 2021. Humaid: human-annotated disaster incidents data from twitter. *In*: *15th International conference on web and social media (ICWSM)*.

Alam, F., Imran, M., and Ofli, F., 2017. Image4act: Online social media image processing for disaster response. *In*: *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, Sydney, Australia, 31 July – 3 August 2017, 601–604.

Alam, F., Joty, S., and Imran, M., 2018b. Graph based semi-supervised learning with convolution neural networks to classify crisis related tweets. *In*: *Proceedings of the international AAAI conference on web and social media*, Palo Alto, CA, 25–28 June 2018, vol. 12, 556–559.

Alam, F., Ofli, F., and Imran, M., 2018c. Crisismmd: multimodal Twitter datasets from natural disasters. *In*: *Proceedings of the international AAAI conference on web and social media*, Palo Alto, CA, 25–28 June 2018, vol. 12, 465–473.

Alam, F., Ofli, F., and Imran, M., 2018d. Processing social media images by combining human and machine computing during crises. *International Journal of Human–Computer Interaction*, 34 (4), 311–327.

Alam, F., Ofli, F., and Imran, M., 2020a. Descriptive and visual summaries of disaster events using artificial intelligence techniques: case studies of hurricanes Harvey, Irma, and Maria. *Behaviour & Information Technology*, 39 (3), 288–318.

Ankerst, M., *et al.*, 1999. Optics: ordering points to identify the clustering structure. *ACM SIGMOD Record*, 28 (2), 49–60.

Assumpção, T.H., *et al.*, 2018. Citizen observations contributing to flood modelling: opportunities and challenges. *Hydrology and Earth System Sciences*, 22 (2), 1473–1489.

Aulov, O., Price, A., and Halem, M., 2014. Asonmaps: a platform for aggregation visualization and analysis of disaster related human sensor network observations. *In: ISCRAM* 2014 Conference Proceedings, University Park, PA, 1 May 2014, 802–806.

Avgerinakis, K., *et al.*, 2017. Visual and textual analysis of social media and satellite images for flood detection@ multimedia satellite task mediaeval 2017. *In*: *Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Beigi, G., *et al.*, 2016. An overview of sentiment analysis in social media and its applications in disaster relief. *In*: W. Pedrycz and S.-M. Chen, eds. *Sentiment analysis and ontology engineering*. Cham: Springer, 313–340.

Berlyand, Y., *et al.*, 2018. How artificial intelligence could transform emergency department operations. *The American Journal of Emergency Medicine*, 36 (8), 1515–1517.

Bischke, B., *et al.*, 2016. Contextual enrichment of remote-sensed events with social media streams. *In: Proceedings of the 24th ACM international conference on Multimedia*, Amsterdam, Netherlands, 15–19 October 2016, 1077–1081.

Bischke, B., *et al.*, 2017a. Detection of flooding events in social multimedia and satellite imagery using deep neural networks. *In: Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Bischke, B., *et al.*, 2017b. The multimedia satellite task at mediaeval 2017: emergency response for flooding events. *In: Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Bischke, B., *et al.*, 2018. The multimedia satellite task at mediaeval 2018 emergency response for flooding events. *In*: *Working notes proceedings of the MediaEval 2018 workshop*, Sophia Antipolis, France, 29–31 October 2018.

Bischke, B., *et al.*, 2019. The multimedia satellite task at mediaeval 2019. *In*: *Proceedings of the MediaEval 2019 workshop*, Sophia Antipolis, France, 27–30 October 2019.

Blei, D.M., Ng, A.Y., and Jordan, M.I., 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3 (Jan), 993–1022.

Bojanowski, P., *et al.*, 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146.

Boulton, C., Shotton, H., and Williams, H., 2016. Using social media to detect and locate wildfires. *In*: *Proceedings of the international AAAI conference on web and social media*, Cologne, Germany, 17–20 May 2016, 178–186.

Brandt, K., *et al.*, 2020. Integrating sketch mapping and hot spot analysis to enhance capacity for community-level flood and disaster risk management. *The Geographical Journal*, 186 (2), 198–212.

Caragea, C., *et al.*, 2014. Mapping moods: geo-mapped sentiment analysis during hurricane sandy. *In*: *ISCRAM 2014 Conference Proceedings*, University Park, PA, 1 May 2014, 642–651.

Cattaneo, D., *et al.*, 2019. Cmrnet: camera to LiDAR-map registration. *In*: *2019 IEEE intelligent transportation systems conference (ITSC)*. New York: IEEE, 1283–1289.

Cattaneo, D., Sorrenti, D.G., and Valada, A., 2020. Cmrnet++: map and camera agnostic monocular visual localization in LiDAR maps. *arXiv preprint arXiv:2004.13795*.

CBS News. 2017. Interactive map shows where Harvey flooding is worst. https://www.cbsnews.com/news/houston-harvey-u-flood-maps/?ftag=CNM-00-10aab6a&linkId=41670697 [Accessed 07 May 2021].

Cerutti, V., *et al.*, 2016. *Identification of disaster-affected areas using exploratory visual analysis of georeferenced tweets: application to a flood event*. Helsinki, Finland: Association of Geographic Information Laboratories in Europe, 5.

Cervone, G., *et al.*, 2016. Using twitter for tasking remote-sensing data collection and damage assessment: 2013 boulder flood case study. *International Journal of Remote Sensing*, 37 (1), 100–124.

Chae, J., *et al.*, 2012. Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition. In: *2012 IEEE conference on visual analytics science and technology (VAST)*. New York: IEEE, 143–152.

Chae, J., *et al.*, 2014. Public behavior response analysis in disaster events utilizing visual analytics of microblog data. *Computers & Graphics*, 38, 51–60.

Chatzichristofis, S.A., and Boutalis, Y.S., 2008. Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. *In: International conference on computer vision systems*. Santorini, Greece, 12–15 May 2008, Cham: Springer, 312–322.

Chaudhary, P., *et al.*, 2019. Flood-water level estimation from social media images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W5 (2/W5), 5–12.

Chaudhary, P., *et al.*, 2020. Water level prediction from social media images with a multi-task ranking approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 252–262.

Chaudhry, S., and Yuksel, M., 2019. Engaging communities in public safety via social media. *In: IEEE INFOCOM 2019-IEEE conference on computer communications workshops (INFOCOM WKSHPS)*. New York: IEEE, 560–565.

Chen, D.M., *et al.*, 2011. City-scale landmark identification on mobile devices. *In: CVPR 2011*. New York: IEEE, 737–744.

Chen, S., *et al.*, 1999. Indexing and searching structure for multimedia database systems. *In:* M.M. Yeung, B.-L. Yeo, and C.A. Bouman, eds. *Storage and retrieval for media databases 2000*. Bellingham, WA: International Society for Optics and Photonics, vol. 3972, 262–270.

Chowdhury, S.R., *et al.*, 2013. Tweet4act: using incident-specific profiles for classifying crisis-related messages. *In: ISCRAM 2013 Conference Proceedings*. Baden-Baden, Germany, 12–15 May 2013, 834–839.

Craglia, M., Ostermann, F., and Spinsanti, L., 2012. Digital earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, 5 (5), 398–416.

Cresci, S., *et al.*, 2015. Crisis mapping during natural disasters via text analysis of social media messages. *In: International conference on web information systems engineering*. Miami, FL, 1–3 November 2015, Cham: Springer, 250–258.

Crooks, A., *et al.*, 2013. # earthquake: Twitter as a distributed sensor system. *Transactions in GIS*, 17 (1), 124–147.

Crowell, B.W., Bock, Y., and Squibb, M.B., 2009. Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks. *Seismological Research Letters*, 80 (5), 772–782.

Daly, S., and Thom, J.A., 2016. Mining and classifying image posts on social media to analyse fires. *In: Proceedings of the ISCRAM 2016 Conference*, Rio de Janeiro, Brazil, 1–14 May 2016.

Dao, M. S., and Pham, Q., 2017. A domain-based late-fusion for disaster image retrieval from social media. *In: CEUR workshop proceedings*. New York, NY: ACM Press, 1–3.

Dao, M.S., *et al.*, 2018. A context-aware late-fusion approach for disaster image retrieval from social media. *In: Proceedings of the 2018 ACM on international conference on multimedia retrieval*. New York, NY: ACM, 266–273.

De Longueville, B., *et al.*, 2010. Digital earth's nervous system for crisis events: real-time sensor web enablement of volunteered geographic information. *International Journal of Digital Earth*, 3 (3), 242–259.

De Longueville, B., Smith, R.S., and Luraschi, G., 2009. "OMG, from here, I can see the flames!" a use case of mining location based social networks to acquire spatio-temporal data on forest fires. *In: Proceedings of the 2009 international workshop on location based social networks*. New York, NY: ACM, 73–80.

Deng, J., *et al.*, 2009. Imagenet: a large-scale hierarchical image database. *In: 2009 IEEE conference on computer vision and pattern recognition*. New York: IEEE, 248–255.

Devlin, J., *et al.*, 2019. Bert: pre-training of deep bidirectional transformers for language understanding. *In*: *NAACL HLT 2019 - 2019 conference of the North American chapter of the Association for Computational Linguistics: Human Language Technologies – Proceedings of the Conference*, Minneapolis, MN, 2–7 June 2019, vol. 1, 4171–4186.

Dittrich, A., and Lucas, C., 2014. Is this Twitter event a disaster? *In: 17th AGILE conference on geographic information science. Connecting a digital Europe through location and place, proceedings of the AGILE'2014 international conference on geographic information science*, Castellón, E., 3–6 June, 2014. Ed.: J. Huerta.

Doggett, E., and Cantarero, A., 2016. Identifying eyewitness news-worthy events on twitter. *In*: *Proceedings of the fourth international workshop on natural language processing for social media*, Austin, TX, 1 November 2016, 7–13.

Dong, L., and Shan, J., 2013. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 84, 85–99.

Earle, P.S., Bowden, D.C., and Guy, M., 2011. Twitter earthquake detection: Earthquake monitoring in a social world. *Annals of Geophysics*, 54 (6), 708–715.

Eilander, D., *et al.*, 2016. Harvesting social media for generation of near real-time flood maps. *Procedia Engineering*, 154, 176–183.

Ester, M., *et al.*, 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *In: Proceedings of the second international conference on knowledge discovery and data mining*, KDD'96. Palo Alto, CA: AAAI Press, 226–231.

Fang, R., *et al.*, 2016. Witness identification in twitter. *In: Proceedings of the fourth international workshop on natural language processing for social media*, Austin, TX, 1 November 2016, 65–73.

Feng, Q., Liu, J., and Gong, J., 2015. Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier–a case of Yuyao, China. *Water*, 7 (12), 1437–1455.

Feng, Y., Brenner, C., and Sester, M., 2020. Flood severity mapping from volunteered geographic information by interpreting water level from images containing people: a case study of Hurricane Harvey. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, 301–319.

Feng, Y., *et al.*, 2018. Ensembled convolutional neural network models for retrieving flood relevant tweets. *In: Working notes proceedings of the MediaEval 2018 workshop, Sophia Antipolis, France, 29–31 October 2018*.

Feng, Y., and Sester, M., 2018. Extraction of pluvial flood relevant volunteered geographic information (VGI) by deep learning from user generated texts and photos. *ISPRS International Journal of Geo-Information*, 7 (2), 39.

Fohringer, J., *et al.*, 2015. Social media as an information source for rapid flood inundation mapping. *Natural Hazards and Earth System Sciences*, 15 (12), 2725–2738.

Fuchs, G., *et al.*, 2013. Tracing the German centennial flood in the stream of tweets: first lessons learned. *In: Proceedings of the second ACM SIGSPATIAL international workshop on crowdsourced and volunteered geographic information*. New York, NY: ACM, 31–38.

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69 (4), 211–221.

Goodchild, M.F., and Glennon, J.A., 2010. Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, 3 (3), 231–241.

Gupta, R., *et al.*, 2019. Creating xBD: a dataset for assessing building damage from satellite imagery. *In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, Long Beach, CA, 16–20 June 2019, 10–17.

Hanif, M., *et al.*, 2017. Flood detection using social media data and spectral regression based kernel discriminant analysis. *In: Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Hassan, S.Z., *et al.*, 2019. Sentiment analysis from images of natural disasters. *In: International conference on image analysis and processing*. Trento, Italy, 9–13 September 2019, Cham: Springer, 104–113.

He, K., *et al.*, 2016. Deep residual learning for image recognition. *In*: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, 26 June 2016 – 1 July 2016, 770–778.

He, K., *et al.*, 2017. Mask R-CNN. *In: Proceedings of the IEEE international conference on computer vision*, Venice, Italy, 22–29 October 2017, 2961–2969.

Heipke, C., 2010. Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65 (6), 550–557.

Hu, Y., and Wang, J., 2020. How do people describe locations during a natural disaster: an analysis of tweets from Hurricane Harvey. *In*: *11th international conference on geographic information science (GIScience 2021)-Part I*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Huang, G., *et al.*, 2017. Densely connected convolutional networks. *In: Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, 21–26 July 2017, 4700–4708.

Huang, J., *et al.*, 1997. Image indexing using color correlograms. *In: Proceedings of IEEE computer society conference on computer vision and pattern recognition*. New York: IEEE, 762–768.

Huang, J., *et al.*, 2020. A novel approach to measuring urban waterlogging depth from images based on mask region-based convolutional neural network. *Sustainability*, 12 (5), 2149.

Huang, Q., and Xiao, Y., 2015. Geographic situational awareness: mining tweets for disaster preparedness, emergency response, impact, and recovery. *ISPRS International Journal of Geo-Information*, 4 (3), 1549–1568.

Huang, X., *et al.*, 2020. Identifying disaster related social media for rapid response: a visual-textual fused cnn architecture. *International Journal of Digital Earth*, 13 (9), 1017–1023.

Huang, X., Wang, C., and Li, Z., 2018a. A near real-time flood-mapping approach by integrating social media and post-event satellite imagery. *Annals of GIS*, 24 (2), 113–123.
Huang, X., Wang, C., and Li, Z., 2018b. Reconstructing flood inundation probability by enhancing near real-time imagery with real-time gauges and tweets. *IEEE Transactions on Geoscience and Remote Sensing*, 56 (8), 4691–4701.
Huang, Y., Li, Y., and Shan, J., 2018c. Spatial-temporal event detection from geo-tagged tweets. *ISPRS International Journal of Geo-Information*, 7 (4), 150.
Imran, M., *et al.*, 2013a. Practical extraction of disaster-relevant information from social media. *In: Proceedings of the 22nd international conference on World Wide Web*, Rio de Janeiro, Brazil, 13–17 May 2013, 1021–1024.
Imran, M., *et al.*, 2013b. Extracting information nuggets from disaster-related messages in social media. *In: ISCRAM 2013 Conference Proceedings*, Baden-Baden, Germany, 12–15 May 2013, 791–801.
Imran, M., Mitra, P., and Castillo, C., 2016. Twitter as a lifeline: Human-annotated twitter corpora for nlp of crisis-related messages. *In: 10th international conference on language resources and evaluation, LREC 2016*. Paris: European Language Resources Association (ELRA), 1638–1643.
Iyengar, A., Finin, T., and Joshi, A., 2011. Content-based prediction of temporal boundaries for events in twitter. *In: 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. New York: IEEE, 186–191.
Jain, P., Ross, R., and Schoen-Phelan, B., 2019. Estimating distributed representation performance in disaster-related social media classification. *In: 2019 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. New York: IEEE, 723–727.
Jing, M., *et al.*, 2016a. The application of social media image analysis to an emergency management system. *In: 2016 11th international conference on availability, reliability and security (ARES)*. New York: IEEE, 805–810.
Jing, M., *et al.*, 2016b. Integration of text and image analysis for flood event image recognition. *In: 2016 27th Irish signals and systems conference (ISSC)*. New York: IEEE, 1–6.
Johnson, B., 2020. Fake news during disasters putting people's lives at risk, Warns Intel Bulletin. https://www.hstoday.us/subject-matter-areas/emergency-preparedness/fake-news-during-disasters-putting-peoples-lives-at-risk-warns-intel-bulletin/ [Accessed 07 May 2021].
Joulin, A., *et al.*, 2017. Bag of tricks for efficient text classification. *In: Proceedings of the 15th conference of the European chapter of the association for computational linguistics: volume 2, short papers*, April. Morristown, NJ: Association for Computational Linguistics, 427–431.
Karimi, S., Yin, J., and Paris, C., 2013. Classifying microblogs for disasters. *In: Proceedings of the 18th Australasian document computing symposium*, Brisbane, Australia, 5–6 December 2013, 26–33.
Karras, T., Laine, S., and Aila, T., 2019. A style-based generator architecture for generative adversarial networks. *In: Proceedings of the IEEE conference on computer vision and pattern recognition*, Long Beach, CA, 16–20 June 2019, 4401–4410.
Kerle, N., and Hoffman, R.R., 2013. Collaborative damage mapping for emergency response: the role of cognitive systems engineering. *Natural Hazards and Earth System Sciences*, 13 (1), 97–113.
Kersten, J., and Klan, F., 2020. What happens where during disasters? A workflow for the multifaceted characterization of crisis events based on twitter data. *Journal of Contingencies and Crisis Management*, 28 (3), 262–280.
Khare, P., Burel, G., and Alani, H., 2018. Classifying crises-information relevancy with semantics. *In*: *European semantic web conference*. New York: Springer, 367–383.
Kim, Y., 2014. Convolutional neural networks for sentence classification. *In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, Doha, Qatar, 25–29 October 2014, 1746–1751.
Klonner, C., *et al.*, 2016. Volunteered geographic information in natural hazard analysis: a systematic literature review of current approaches with a focus on preparedness and mitigation. *ISPRS International Journal of Geo-Information*, 5 (7), 103.

Klonner, C., *et al.*, 2018. Capturing flood risk perception via sketch maps. *ISPRS International Journal of Geo-Information*, 7 (9), 359.

Kogan, M., Palen, L., and Anderson, K.M., 2015. Think local, retweet global: Retweeting by the geographically-vulnerable during hurricane sandy. *In: Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, Vancouver, Canada, 14–18 March 2015, 981–993.

Kutija, V., *et al.*, 2014. Model validation using crowd-sourced data from a large pluvial flood. *In: 11th international conference on hydroinformatics*, New York, USA, 17–21 August 2014. New York: CUNY Academic Works.

Lagerstrom, R., *et al.*, 2016. Image classification to support emergency situation awareness. *Frontiers in Robotics and AI*, 3, 54.

Le Boursicaud, R., *et al.*, 2016. Gauging extreme floods on YouTube: application of LSPIV to home movies for the post-event determination of stream discharges. *Hydrological Processes*, 30 (1), 90–105.

Le Coz, J., *et al.*, 2016. Crowdsourced data for flood hydrology: feedback from recent citizen science projects in Argentina, France and New Zealand. *Journal of Hydrology*, 541, 766–777.

Levin, S., *et al.*, 2018. Machine-learning-based electronic triage more accurately differentiates patients with respect to Clinical Outcomes Compared With the Emergency Severity Index. *Annals of Emergency Medicine*, 71 (5), 565–574.

Li, L., and Fox, E.A., 2019. Understanding patterns and mood changes through tweets about disasters. *In: Proceedings of the 16th international conference on information systems for crisis response and management,* Valencia, Spain, 19–22 May 2019, 756–767.

Li, L., *et al.*, 2015. Sub-pixel flood inundation mapping from multispectral remotely sensed images based on discrete particle swarm optimization. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101, 10–21.

Li, X., *et al.*, 2019. Identifying disaster damage images using a domain adaptation approach. *In: Proceedings of the international ISCRAM conference,* Valencia, Spain, 19–22 May 2019, 633–645.

Li, Z., *et al.*, 2018. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science*, 45 (2), 97–110.

Liu, P., Qiu, X., and Huang, X., 2016. Recurrent neural network for text classification with multi-task learning. *arXiv preprint arXiv:1605.05101*.

Lopez-Fuentes, L., *et al.*, 2017. Multi-modal deep learning approach for flood detection. *In: Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

MacEachren, A.M., *et al.*, 2011. Senseplace2: Geotwitter analytics support for situational awareness. *In: 2011 IEEE conference on visual analytics science and technology (VAST)*. New York: IEEE, 181–190.

Madichetty, S., and Sridevi, M., 2019. Disaster damage assessment from the tweets using the combination of statistical features and informative words. *Social Network Analysis and Mining*, 9 (1), 42.

Mandel, B., *et al.*, 2012. A demographic analysis of online sentiment during Hurricane Irene. *In: Proceedings of the second workshop on language in social media*, Montreal, Canada, 7 June 2012, 27–36.

McDougall, K., 2011a. Understanding the impact of volunteered geographic information during the Queensland floods. *In: Proceedings of the 7th international symposium on digital earth (ISDE 7)*. Midland, WA: Western Australian Land Information System, 66–74.

McDougall, K., 2011b. Using volunteered information to map the Queensland floods. *In: Proceedings of the 2011 surveying and spatial sciences conference: innovation in action: working smarter (SSSC 2011)*. South Brisbane: Surveying and Spatial Sciences Institute, 13–23.

McDougall, K., and Temple-Watts, P., 2012. The use of LiDAR and volunteered geographic information to map flood extents and inundation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-4, 251–256.

Meel, P., and Vishwakarma, D.K., 2020. Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Systems with Applications*, 153, 112986.

Middleton, S.E., *et al.*, 2013. Multi-disciplinary approaches to intelligently sharing large-volumes of real-time sensor data during natural disasters. *Data Science Journal*, 12, WDS109–WDS113.

Mikolov, T., *et al.*, 2013a. Efficient estimation of word representations in vector space. *In*: *1st International conference on learning representations, ICLR 2013 – workshop track proceedings*. Scottsdale, AZ, 2–4 May 2013.

Mikolov, T., *et al.*, 2013b. Distributed representations of words and phrases and their compositionality. *In: Advances in neural information processing systems*, Lake Tahoe, 5–10 December 2013, 3111–3119.

Moher, D., *et al.*, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, 6 (7), e1000097.

Molina-Pico, A., *et al.*, 2016. Forest monitoring and wildland early fire detection by a hierarchical wireless sensor network. *Journal of Sensors*, 2016, 1–8.,

Moumtzidou, A., *et al.*, 2018. Flood relevance estimation from visual and textual content in social media streams. *In: Companion proceedings of the web conference 2018*, Lyon, France, 23–27 April 2018, 1621–1627.

Mouzannar, H., Rizk, Y., and Awad, M., 2018. Damage identification in social media posts using multimodal deep learning. *In: Proceedings of the International ISCRAM* Conference. Rochester, NY, 20–23 May 2018, 529–543.

Muller-Budack, E., Pustu-Iren, K., and Ewerth, R., 2018. Geolocation estimation of photos using a hierarchical model and scene classification. *In: Proceedings of the European conference on computer vision (ECCV)*, Munich, Germany, 8–14 September 2018, 563–579.

Murthy, D., and Longwell, S.A., 2013. Twitter and disasters: the uses of twitter during the 2010 Pakistan floods. *Information, Communication & Society*, 16 (6), 837–855.

Neal, D.M., 1997. Reconsidering the phases of disasters. *International Journal of Mass Emergencies and Disasters*, 15 (2), 239–264.

Nguyen, D., *et al.*, 2017a. Robust classification of crisis-related data on social networks using convolutional neural networks. *In: Proceedings of the international AAAI conference on web and social media,* Montreal, Canada, 15–18 May 2017, 632–635.

Nguyen, D.T., *et al.*, 2017b. 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*, Sydney, Australia, 31 July – 3 August 2017, 569–576.

Nia, K.R., and Mori, G., 2017. Building damage assessment using deep learning and ground-level image data. *In: 2017 14th conference on computer and robot vision (CRV)*. New York: IEEE, 95–102.

Ning, H., *et al.*, 2020. Prototyping a social media flooding photo screening system based on deep learning. *ISPRS International Journal of Geo-Information*, 9 (2), 104.

Nogueira, K., *et al.*, 2017. Data-driven flood detection using neural networks. *In: Working notes proceedings of the MediaEval 2017 Workshop*, Dublin, Ireland, 13–15 September 2017.

Ofli, F., Imran, M., and Alam, F., 2020. Using artificial intelligence and social media for disaster response and management: an overview. *In*: T.V. Vijay Kumar and K. Sud, eds. *AI and robotics in disaster studies*. Singapore: Springer Singapore, 63–81.

Ogie, R.I., *et al.*, 2019. Crowdsourced social media data for disaster management: lessons from the petajakarta.org project. *Computers, Environment and Urban Systems*, 73, 108–117.

Olteanu, A., *et al.*, 2014. Crisislex: a lexicon for collecting and filtering microblogged communications in crises. *In: Proceedings of the international AAAI conference on web and social media*, Ann Arbor, MI, 1–4 June 2014, 376–385.

Olteanu, A., Vieweg, S., and Castillo, C., 2015. 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*, Vancouver, Canada, 14–18 March 2015, 994–1009.

Ord, J.K., and Getis, A., 2010. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27 (4), 286–306.

Ostermann, F.O., and Spinsanti, L., 2011. A conceptual workflow for automatically assessing the quality of volunteered geographic information for crisis management. *In: Proceedings of AGILE*, Utrecht, Netherlands, 1–6.

Panteras, G., and Cervone, G., 2018. Enhancing the temporal resolution of satellite-based flood extent generation using crowdsourced data for disaster monitoring. *International Journal of Remote Sensing*, 39 (5), 1459–1474.

Park, S., *et al.*, 2021. Computer vision–based estimation of flood depth in flooded-vehicle images. *Journal of Computing in Civil Engineering*, 35 (2), 04020072.

Pegelonline. 2020. PEGELONLINE. https://www.pegelonline.wsv.de/gast/start [Accessed 07 May 2021].

Pennington, J., Socher, R., and Manning, C.D., 2014. Glove: global vectors for word representation. *In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, Doha, Qatar, 25–29 October 2014, 1532–1543.

Pereira, J., *et al.*, 2019. Assessing flood severity from georeferenced photos. *In: Proceedings of the 13th workshop on geographic information retrieval*, Lyon, France, 28–29 November 2019, 1–10.

Peters, M.E., *et al.*, 2018. Deep contextualized word representations. *In: Proceedings of NAACL-HLT*. New Orleans, LA, 1–6 June 2018, 2227–2237.

Pham, H.X., *et al.*, 2017. A distributed control framework for a team of unmanned aerial vehicles for dynamic wildfire tracking. *In: 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. New York: IEEE, 6648–6653.

Power, R., *et al.*, 2015. A case study for monitoring fires with twitter. *In: ISCRAM 2015 Conference Proceedings*. Kristiansand, Norway, 24–27 May 2015.

Power, R., Robinson, B., and Ratcliffe, D., 2013. Finding fires with twitter. *In: Proceedings of the Australasian Language Technology Association workshop 2013 (ALTA 2013)*, Brisbane, Australia, 4–6 December 2013, 80–89.

Quan, K.A.C., *et al.*, 2020. Flood level prediction via human pose estimation from social media images. In: *Proceedings of the 2020 international conference on multimedia retrieval*, Dublin, Ireland, 8–11 June 2020, 479–485.

Ramachandran, D., and Ramasubramanian, P., 2018. Event detection from twitter–a survey. *International Journal of Web Information Systems*, 14 (3), 262–280.

Rashid, M.T., *et al.*, 2020. Compdrone: towards integrated computational model and social drone based wildfire monitoring. *In: 2020 16th international conference on distributed computing in sensor systems (DCOSS)*. New York: IEEE, 43–50.

Ratcliffe, J.H., 2004. The hotspot matrix: a framework for the spatio-temporal targeting of crime reduction. *Police Practice and Research*, 5 (1), 5–23.

Resch, B., Usländer, F., and Havas, C., 2018. Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and Geographic Information Science*, 45 (4), 362–376.

Roheda, S., *et al.*, 2018. Decision level fusion: an event driven approach. *In: 2018 26th European signal processing conference (EUSIPCO)*. Rome, Italy, 3–7 September 2018, New York: IEEE, 2598–2602.

Rosser, J.F., Leibovici, D., and Jackson, M., 2017. Rapid flood inundation mapping using social media, remote sensing and topographic data. *Natural Hazards*, 87 (1), 103–120.

Said, N., *et al.*, 2018. Deep learning approaches for flood classification and flood aftermath detection. *In: Working notes proceedings of the MediaEval 2018* workshop, Sophia Antipolis, France, 29–31 October 2018.

Sakaki, T., Okazaki, M., and Matsuo, Y., 2010. Earthquake shakes twitter users: real-time event detection by social sensors. *In: Proceedings of the 19th international conference on World Wide Web*. Raleigh, NC, 26–30 April 2010, 851–860.

Salton, G., and Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24 (5), 513–523.

Sarker, C., *et al.*, 2019. Flood mapping with convolutional neural networks using spatio-contextual pixel information. *Remote Sensing*, 11 (19), 2331.

Sattler, T., *et al.*, 2017. Are large-scale 3d models really necessary for accurate visual localization? *In: Proceedings of the IEEE conference on computer vision and pattern recognition*, 1637–1646.

Schnebele, E., and Cervone, G., 2013. Improving remote sensing flood assessment using volunteered geographical data. *Natural Hazards and Earth System Sciences*, 13 (3), 669–677.

Schnebele, E., Cervone, G., and Waters, N., 2014a. Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Sciences*, 14 (4), 1007–1015.

Schnebele, E., *et al.*, 2014b. Real time estimation of the Calgary floods using limited remote sensing data. *Water*, 6 (2), 381–398.

Sermanet, P., *et al.*, 2013. Overfeat: integrated recognition, localization and detection using convolutional networks. *In*: *2nd International conference on learning representations, ICLR 2014 – conference track proceedings*. Banff, Canada, 14–16 April 2014.

Sester, M., *et al.*, 2014. Integrating and generalising volunteered geographic information. *In: Abstracting geographic information in a data rich world*. Cham: Springer, 119–155.

Simonyan, K., and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Sit, M.A., Koylu, C., and Demir, I., 2019. Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: a case study of Hurricane Irma. *International Journal of Digital Earth*, 12 (11), 1205–1229.

Smith, L., *et al.*, 2017. Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *Journal of Flood Risk Management*, 10 (3), 370–380.

Socher, R., *et al.*, 2013. Recursive deep models for semantic compositionality over a sentiment treebank. *In: Proceedings of the 2013 conference on empirical methods in natural language processing*, Seattle, WA, 18–21 October 2013, 1631–1642.

Stefanidis, A., Crooks, A., and Radzikowski, J., 2013. Harvesting ambient geospatial information from social media feeds. *GeoJournal*, 78 (2), 319–338.

Steiger, E., De Albuquerque, J.P., and Zipf, A., 2015. An advanced systematic literature review on spatiotemporal analyses of twitter data. *Transactions in GIS*, 19 (6), 809–834.

Stock, K., 2018. Mining location from social media: a systematic review. *Computers, Environment and Urban Systems*, 71, 209–240.

Stowe, K., *et al.*, 2016. Identifying and categorizing disaster-related tweets. *In: Proceedings of The fourth international workshop on natural language processing for social media*, Austin, TX, 1 November 2016, 1–6.

Sural, S., Qian, G., and Pramanik, S., 2002. Segmentation and histogram generation using the HSV color space for image retrieval. *In: Proceedings. International conference on image processing*. New York: IEEE, vol. 2, II.

Szegedy, C., *et al.*, 2016. Rethinking the inception architecture for computer vision. *In: Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, 26 June – 1 July 2016, 2818–2826.

Tan, M., and Le, Q., 2019. Efficientnet: rethinking model scaling for convolutional neural networks. *In: International conference on machine learning*, Long Beach, CA, 9–15 June 2019, 6105–6114.

Thies, J., *et al.*, 2020. Neural voice puppetry: audio-driven facial reenactment. *In: European conference on computer vision*. Glasgow, UK, 23–28 August 2020. Cham: Springer, 716–731.

Tkachenko, N., Zubiaga, A., and Procter, R., 2017. Wisc at mediaeval 2017: multimedia satellite task. *In: Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Truelove, M., Vasardani, M., and Winter, S., 2015. Towards credibility of micro-blogs: characterising witness accounts. *GeoJournal*, 80 (3), 339–359.

Ukkusuri, S.V., *et al.*, 2014. Use of social media data to explore crisis informatics: study of 2013 Oklahoma Tornado. *Transportation Research Record: Journal of the Transportation Research Board*, 2459 (1), 110–118.

Vaswani, A., *et al.*, 2017. Attention is all you need. *In: Advances in neural information processing systems*, Long Beach, CA, 4–9 December 2017, 5998–6008.

Vieweg, S.E., 2012. *Situational awareness in mass emergency: a behavioral and linguistic analysis of microblogged communications*. Thesis (PhD). Citeseer.

Vosoughi, S., Roy, D., and Aral, S., 2018. The spread of true and false news online. *Science*, 359 (6380), 1146–1151.

Wallemacq, P., and House, R., 2018. *Economic losses, poverty and disasters 1998-2017*. Geneva, Switzerland: The United Nations Office for Disaster Risk Reduction, 33.

Wang, H., *et al.*, 2018a. Fusing heterogeneous data: a case for remote sensing and social media. *IEEE Transactions on Geoscience and Remote Sensing*, 56 (12), 6956–6968.

Wang, R.Q., *et al.*, 2018b. Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. *Computers & Geosciences*, 111, 139–147.

Wang, R.Q., *et al.*, 2020. Tracking flooding phase transitions and establishing a passive hotline with AI-enabled social media data. *IEEE Access*, 8, 103395–103404.

Wang, Y., *et al.*, 2016a. Using social media for emergency response and urban sustainability: a case study of the 2012 Beijing rainstorm. *Sustainability*, 8 (1), 25.

Wang, Z., and Ye, X., 2018. Social media analytics for natural disaster management. *International Journal of Geographical Information Science*, 32 (1), 49–72.

Wang, Z., Ye, X., and Tsou, M.H., 2016b. Spatial, temporal, and content analysis of twitter for wildfire hazards. *Natural Hazards*, 83 (1), 523–540.

Weber, E., *et al.*, 2020. Detecting natural disasters, damage, and incidents in the wild. *In: European conference on computer vision*. Glasgow, UK, 23–28 August 2020. Cham: Springer, 331–350.

Wolfswinkel, J.F., Furtmueller, E., and Wilderom, C.P., 2013. Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22 (1), 45–55.

Xiao, Y., Li, B., and Gong, Z., 2018. Real-time identification of urban rainstorm waterlogging disasters based on Weibo big data. *Natural Hazards*, 94 (2), 833–842.

Xiao, Y., and Watson, M., 2019. Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39 (1), 93–112.

Yan, Y., *et al.*, 2020. Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience. *International Journal of Geographical Information Science*, 0 (0), 1–27.

Yang, Y., *et al.*, 2011. Hierarchical disaster image classification for situation report enhancement. *In: 2011 IEEE international conference on information reuse & integration*. New York: IEEE, 181–186.

Yao, W., *et al.*, 2020. Weakly-supervised fine-grained event recognition on social media texts for disaster management. *In: Proceedings of the AAAI conference on artificial intelligence*, New York, 7–12 February 2020, 532–539.

Yin, J., *et al.*, 2012. Using social media to enhance emergency situation awareness. *IEEE Intelligent Systems*, 27 (6), 52–59.

Yu, M., *et al.*, 2019. Deep learning for real-time social media text classification for situation awareness–using hurricanes Sandy, Harvey, and Irma as case studies. *International Journal of Digital Earth*, 12 (11), 1230–1247.

Zahra, K., Imran, M., and Ostermann, F.O., 2020. Automatic identification of eyewitness messages on twitter during disasters. *Information Processing & Management*, 57 (1), 102107.

Zhao, Q., *et al.*, 2019. Extracting and classifying typhoon disaster information based on volunteered geographic information from Chinese Sina Microblog. *Concurrency and Computation: Practice and Experience*, 31 (9), e4910.

Zhao, Z., and Larson, M., 2017. Retrieving social flooding images based on multimodal information. *In: Working notes proceedings of the MediaEval 2017 workshop*, Dublin, Ireland, 13–15 September 2017.

Zhou, B., *et al.*, 2018. Places: a 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40 (6), 1452–1464.

Zhou, P., *et al.*, 2016. Text classification improved by integrating bidirectional lstm with two-dimensional max pooling. *In: Proceedings of COLING 2016, the 26th international conference on computational linguistics: technical papers*, Osaka, Japan, 11–16 December 2016, 3485–3495.

Zhou, X., and Zafarani, R., 2020. A survey of fake news: fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53 (5), 1–40.

Zhu, R., *et al.*, 2018. Geo-tagged social media data-based analytical approach for perceiving impacts of social events. *ISPRS International Journal of Geo-Information*, 8 (1), 15.

Zourlidou, S., and Sester, M., 2019. Traffic regulator detection and identification from crowd-sourced data—a systematic literature review. *ISPRS International Journal of Geo-Information*, 8 (11), 491.