Machine Learning in Disaster Management: Recent Developments in Methods and Applications

Vasileios Linardos [1], **Maria Drakaki [2,***, **Panagiotis Tzionas [3]** and **Yannis L. Karnavas [4]**

1 Archeiothiki S.A., GR-19400 Athens, Greece; vnl@archeiothiki.gr
2 Department of Science and Technology, University Center of International Programmes of Studies, International Hellenic University, 14th Km Thessaloniki-N. Moudania, GR-57001 Thermi, Greece
3 Department of Industrial Engineering and Management, P.O. Box 141, GR-57400 Thessaloniki, Greece; ptzionas@ihu.gr
4 Electrical Machines Laboratory, Department of Electrical & Computer Engineering, Democritus University of Thrace, GR-67100 Xanthi, Greece; karnavas@ee.duth.gr
* Correspondence: mdrakaki@ihu.gr

Abstract: Recent years include the world's hottest year, while they have been marked mainly, besides the COVID-19 pandemic, by climate-related disasters, based on data collected by the Emergency Events Database (EM-DAT). Besides the human losses, disasters cause significant and often catastrophic socioeconomic impacts, including economic losses. Recent developments in artificial intelligence (AI) and especially in machine learning (ML) and deep learning (DL) have been used to better cope with the severe and often catastrophic impacts of disasters. This paper aims to provide an overview of the research studies, presented since 2017, focusing on ML and DL developed methods for disaster management. In particular, focus has been given on studies in the areas of disaster and hazard prediction, risk and vulnerability assessment, disaster detection, early warning systems, disaster monitoring, damage assessment and post-disaster response as well as cases studies. Furthermore, some recently developed ML and DL applications for disaster management have been analyzed. A discussion of the findings is provided as well as directions for further research.

Keywords: disaster management; machine learning; deep learning; natural disasters; applications in disaster management

1. Introduction

Natural and man-made disasters impact the lives of millions of people worldwide each year [1]. Often it is the case that human lives are lost as a result of these events. Besides human losses, significant impacts on infrastructure and properties are caused by disasters. Disaster management operations are performed before, during and after the occurrence of a disaster, aiming at preventing human losses, protecting people and infrastructure, reducing the impacts on economy and reestablishing a state of normalcy [2]. The complexity of the disasters and the criticality and complexity of disaster operations require robust decision making, enhanced by information technology and in particular AI [2]. Effective and informed disaster management is necessary to address the scale and the impact of disasters and in recent years has been leveraged by advances in ML and DL [3].

Application fields include disasters such as hurricanes, earthquakes, floods, wildfires and landslides. The management of man-made disasters such as refugee crises, can also benefit from recent technological developments [4,5]. Yet, there is no unique definition of a disaster. According to the terminology of United Nations Office for Disaster Risk Reduction (UNISDR), a disaster is a "serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources" [6]. Based on the EM-DAT terminology, disasters can be categorized in two main

groups: natural disasters and technological disasters [7]. Using a different categorization, disasters have been categorized to natural and man-made disasters [8]. Moreover, the climate-related disasters in 2020 were mainly responsible for the 389 recorded disasters during the same year. The statistics besides the human losses placed 2020 above the average in the number of recorded disasters, with 26% more storms and 23% more floods, as well as in the annual average of economic losses [1].

Disaster management addresses disasters over time [8]. Four distinct phases have been widely adopted: mitigation, preparedness, response and recovery [2,8]. Mitigation relates to activities that will either prevent the occurrence of a disaster or reduce the impacts of a disaster. Preparedness includes the activities that prepare the communities to take actions in response to a disaster, such as emergency planning, prepositioning of supplies, training and education of the community in order to better respond when a disaster strikes or to mitigate the effects of a disaster. Response includes the operations that implement the planning made to protect the lives and property, the environment, and the socio-economic structure of the community [2]. Disaster relief and response also includes activities such as implementation of emergency plans, emergency rescue and medical care, shelter opening and management, distribution of supplies and damage assessment. It is one of the most researched phases because it is during this phase that people and infrastructure need the most immediate assistance. In this phase, time is very much of the essence, so techniques focus not only on high accuracy results but also fast performing and optimized methods. Recovery (reconstruction) includes the long-term actions that aim to bring a state of normalcy to the community. Activities in this phase include among others financial assistance and rebuilding/reconstruction (e.g., of buildings and key infrastructure). Moreover, the active involvement of local communities in disaster management can lead to community resilience [6].

The AI techniques that are used for the purpose of assisting disaster management across all of its phases consist of several ML and DL methods. ML methods include support vector machine (SVM), Naïve Bayes (NB) methods, decision trees (DT), random forest (RF), logistic regression (LR) and K-nearest neighbor (KNN) clustering algorithm. On the other hand, DL methods include different architectures of artificial neural networks (ANNs), such as convolutional neural networks (CNNs), multi-layer perceptron (MLP), recurrent neural networks (RNNs), long short-term memory neural network (LSTM), transformers architecture and generative adversarial networks (GANs) [3,9,10]. ML and DL enable the utilization of big and complex datasets in order to develop systems that can predict disasters as well as assist during the response and recovery after such events and practical decision-support tools have been developed. These techniques take advantage of the ability to manipulate different types of data from multiple sources and determine patterns that can provide intelligence which would be otherwise impossible to divulge. Big data is available from sources such as satellite imagery, unmanned aerial vehicles (UAVs), social media, crowdsourcing, geographical information systems (GIS) and wireless sensor networks.

Recent reviews in disaster management have focused on applications of AI in disaster management [3] and the use and potential of big data in natural disaster management [10]. Sun et al. [3] presented an overview of example AI applications in various application areas of disaster management such as hazard risk assessment, vulnerability assessment, early warning systems, disaster detection, event mapping, damage assessment, disaster rescue and relief and resource allocation. There is an increasing trend in using AI to analyze and process big data coming from various data sources for informed decision making in disaster management. The studies included in [3] used in total 26 AI methods for 17 application areas of disaster management. Yu et al. [10] provided a systematic review of the role of big data in disaster management. The review covered journal articles published in the period 2011 to 2018. They authors identified ML including DL as both an evolutionary technology and emerging technological topic in facilitating disaster management.

Taking into account the increasing trend in the development of ML-based methods for disaster management and the rapid advances in DL, a comprehensive review focusing

on ML and DL advances in disaster management is needed. Accordingly, in this paper, a review of literature studies utilizing ML and DL since 2017 has been conducted, covering various phases of disaster management with different types of techniques and data. The purpose of this study is to provide a comprehensive analysis of the developed ML and DL techniques for disaster management and present future trends. Moreover, recent applications developed based on ML and DL for disaster management have been included.

In the following sections, the methodology is presented next. The theoretical background in some main ML and DL methods used in disaster management follows. Recent studies presented since 2017 are analyzed next, categorized according to the subphases of disaster and hazard prediction, risk and vulnerability assessment, disaster detection, early warning systems, disaster monitoring, damage assessment and post-disaster response. Case studies for disaster management and a section presenting in more detail technological applications of ML/DL methods for disaster management follows. Discussion of the results is presented next. Finally, conclusions are drawn.

2. Methodology

For the purpose of this study, a search was conducted in Google Scholar as well as relevant journal databases, targeting articles published in the time period 2017 to (September) 2021. The keywords used in the search included the terms: "natural disaster", "disaster management", "prediction", "assessment", "mitigation", "preparedness", "response", "relief", "post-disaster analysis", "case studies", "applications", "machine learning", "deep learning". The initial search yielded 1210 articles, excluding reviews. Based on the research experience of the authors a search was further conducted on the archives of the top ranking journals of each target database. The explored databases included mainly IEEE, Elsevier, Springer, Taylor and Francis, Scopus, Web of Science and Wiley. Manual search was further performed to eliminate unrelated studies or studies that were outside the scope of this research. Accordingly, 55 papers were finally included in this review. Figure 1 shows the percentage distribution of the research papers included in this review according to the database. Table 1 shows the combination of the keywords used to collect the papers included in this study.
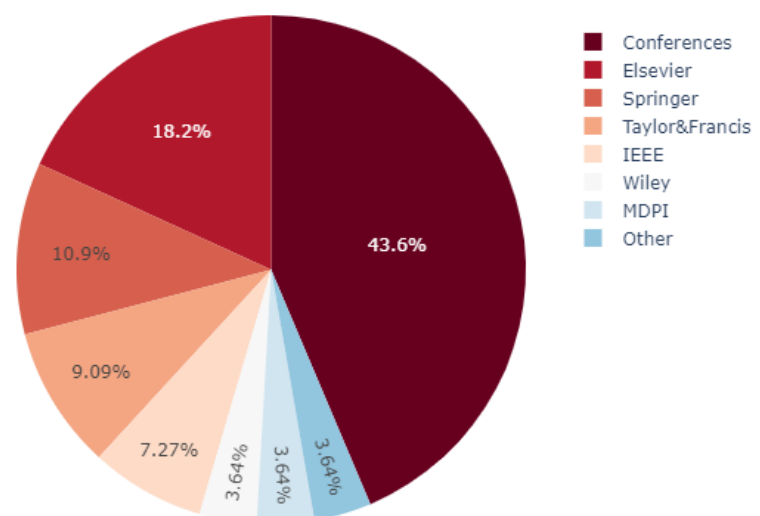

Figure 1. Percentage distribution of the research papers by database.

Table 1. Combination of the keywords used to collect the papers included in this review.

Combination No.	Keywords
1	Disaster management, machine learning, deep learning.
2	Disaster management, disaster prediction, disaster response, disaster preparedness, disaster mitigation, post-disaster analysis.
3	Disaster management case studies, disaster management applications, machine learning
4	Disaster management case studies, disaster management applications, deep learning
5	Natural disaster, deep learning, machine learning
6	Natural disaster prediction, disaster prediction, deep learning, machine learning
7	Relief, response, assessment, disaster management.

3. An Overview of ML and DL Methods

ML has effectively addressed elimination of unrelated data and provides faster processing and analysis of information on disaster events, effectively assisting in all phases of disaster management [10]. However, traditional ML methods cannot directly learn the representation of a complex system from the raw data. DL is a subclass of ML that can automatically learn the representation of a complex system for prediction, detection or classification purposes. DL uses long causal chains of neural network (NN) layers that enable higher and more abstract computational models of the real system [11,12]. DL techniques enable representations with many levels of abstraction, obtained by simple, non-linear modules that at each level transform the representation to a higher more abstract one, in order to finally learn invariant features and very complex functions [12]. DL advancements enable new approaches in the field of disaster management. CNNs are dominating in computer vision tasks, making satellite and aerial imaging systems crucial in disaster response and damage assessment [13]. ANNs are being widely used as a powerful tool for big data analysis [10,14]. On the other hand, text-based NNs, namely long short-term memory (LSTMs) and most recently transformers, utilize their architecture in order to perform natural language processing tasks [15]. These types of NNs are used in social media datasets for damage assessment studies. Two DL and one ML architectures commonly used in disaster management are presented next, namely CNN and LSTM, and SVM, respectively, in order to provide a theoretical background.

3.1. CNN

The CNN architecture is based on convolutional layers (CL). In these layers the data is propagated and applied with tabular multiplication of n*m tabular filters also called kernels, where n equals m in most cases. This process produces different representations of the input data based on the filter applied to them. In each representation various features are uncovered which are projected in feature maps that quantify the stimuli produced by each filter of the convolutional layer [11,12]. As a result of the convolution, multiple unique transformations of the input data are generated. The number and size of kernels that are used in the convolution are key for the performance of the overall network. After the convolution step, the data is passed through a pooling layer (PL) that groups the convolution results and retains only the most important ones to the network, by keeping the maximum, minimum or average of each part of the data. This entire process is repeated several times depending on the depth of the network before the data is flattened in a one-dimensional space in order to be fed in a fully connected layer (FL) which handles the classification. The softmax layer converts the data into a probability distribution. Figure 2 shows a CNN architecture. It is important to mention that each CL in the network provides a level of abstraction in the feature extraction process.
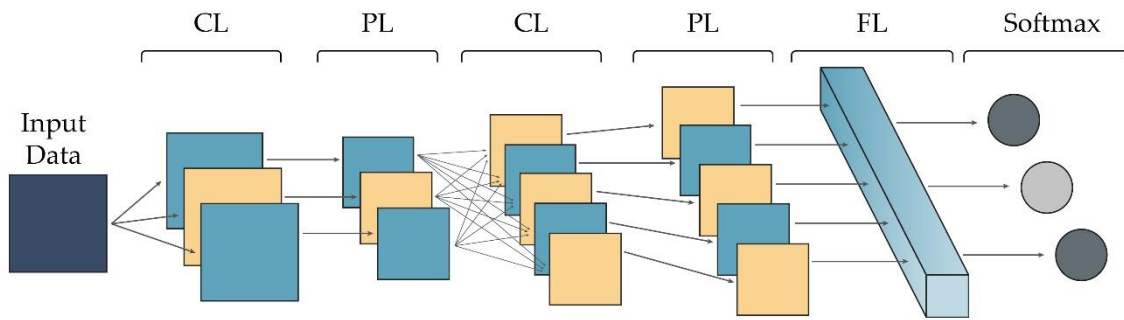
Figure 2. A CNN architecture consisting of two sets of convolution and pooling layers followed by a fully connected layer and a softmax layer.

3.2. LSTM

LSTMs are commonly used NNs for text classification in disaster management in social media datasets. LSTMs are a type of recurrent neural network (RNN). RNNs are networks that recurrently apply the same computation for every element in a set of sequential data while passing some information along to the next iteration. Consequently, in each time step, a prediction is made on the input data which in turn affects the future predictions. This operation allows the network to understand complex textual data and extract meaning based on positional data of words in sentences. LSTMs differ from regular RNNs because of their internal cell architecture. LSTMs utilize memory cells called constant error carousels (CECs), which in each time step determine how much information of the current state will be passed on the next time step along with the new input data [11]. The predictions that the LSTM can provide vary. The network output can be used with a probabilistic classifier for classification purposes, or as an element prediction of the next element in a sequence, or even to predict an entire new sequence of elements. The most important hyperparameters to fine tune in an LSTM are the learning rate and the network size. In LSTMs the hyperparameters can be tuned independently, which can save a lot of time during training and experimentation [16]. Figure 3 shows an LSTM architecture. As presented in the following section, the correct and deep analysis of textual data is very crucial in various stages of disaster management, and LSTMs have proved to work efficiently and effectively.
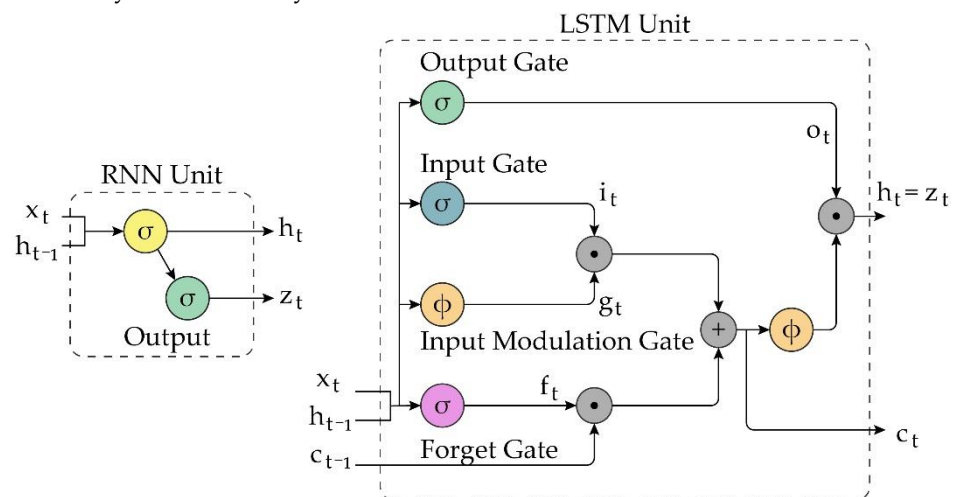

Figure 3. An LSTM cell.

3.3. SVM

The SVM is one of the simplest but most effective ML algorithms for classification and regression problems. It is a supervised learning method, which means it requires an already labeled training set, but with the support-vector clustering it can also categorize unlabeled data [16]. The idea of SVM is simple. It creates a hyperplane which separates

the data in their respective categories while trying to maximize the margin. A hyperplane in a n-dimensional Euclidean space is defined as a flat $n-1$ dimensional subset of that space which separates the space into two disconnected parts. The margin is the minimum distance between the hyperplane and the closest elements from each category. Maximizing the margin in essence makes it more clear for the model to differentiate the categories therefore making it more possible to make better predictions. While the SVM works with vectors in a linear way, it can also perform non-linear classification using kernel trick [17]. Using kernel trick, the data is augmented in a multidimensional space where the features can be represented in an easier separable way by a linear hyperplane. When reverting to the original plane, the hyperplane is projected in a nonlinear transformation. The kernel trick has made it possible for SVM to learn invariant features of complex datasets while maintaining its simple and fast processing nature. Many modern studies in disaster management use SVM as a starting point, which sometimes yields great results, without the need of more complex DL systems.

4. ML and DL Methods for Disaster Management in the Recent Literature

In an era where natural disasters are on the rise, partially due to the increasing human activities [18], ML and DL are the subfields of AI which have contributed the most in many areas of natural disaster management [3,9,19]. In the following, recent research studies presented since 2017 have been analyzed in their phases of application. Accordingly, disaster mitigation has been reviewed in terms of disaster and hazard prediction as well as risk and vulnerability assessment. Areas of disaster preparedness included in this paper are disaster prediction, disaster monitoring, disaster detection and early warning systems. The disaster response has been reviewed in terms of disaster monitoring, damage assessment and post-disaster response. Additionally, case studies and applications based on ML and DL methods developed for various areas of disaster management have been reviewed as well.

4.1. ML/DL Methods for Disaster and Hazard Prediction

In disaster and hazard prediction, data is analyzed in order to predict upcoming event or escalation of an event. It aims to minimize the unpredictability of crisis events and disasters. The identified research studies for disaster prediction, detection and risk assessment refer to the mitigation and preparedness phases of disaster management.

Yuan and Moayedi [20] proposed an optimization of the MLP classification technique for landslide prediction. For this purpose, the multilayer perceptron (MLP) NN was coupled with six evolutionary methods, namely ant colony optimization, biogeography-based optimization, evolutionary strategy, genetic algorithm (GA), probability-based incremental learning and particle swarm optimization. The MLP coupled with GA (GA-MLP) showed the highest classification accuracy for landslide prediction equal to 85%.

A flood prediction method based on temperature and rainfall intensity was proposed by Sankaranarayanan et al. [21]. The method used deep NN and achieved an accuracy of 89.71%. The developed deep NN-based model outperformed traditional ML methods, namely SVM, K-nearest neighbor (KNN) and Naïve Bayes (NB). The highest accuracy was observed just before flood occurrence.

Huang et al. [22] proposed a forecasting method (FNN–LLE) based on fuzzy NN (FNN) combined with the locally linear embedding algorithm (LLE) to predict the daily precipitation during typhoons. The methodology was applied using rainfall data from Guangxi, China. The FNN–LLE model outperformed the interpolation method of the European Centre for Medium-Range Weather Forecasts and stepwise regression approach, achieving equitable threat score (ETS) approximately equal to 1.0. and root mean square error (RMSE) value of 21.94 for tropical cyclone rainfall prediction.

Asim et al. [23] used different ML techniques, namely, pattern recognition NN, RNN, RF and linear programming boost ensemble classifier, that were separately trained with earthquake data from the Hindukush region, Pakistan. The LPBoost ensemble showed the

highest prediction accuracy of 65% with an unknown dataset, followed by RNN with an accuracy of 64%, RF with 62% and pattern recognition with 58%.

4.2. ML/DL Methods for Risk and Vulnerability Assessment

Amin et al. [24] proposed an awareness based educational system to identify the risks associated with indoors objects that can cause problems during earthquakes. For this purpose, the authors used the DL-based detection algorithm You Only Look Once (YOLO) in their system, which is deployed on their own cloud based server named as ESLS (earthquake situation learning system). A dataset was created based on the candidate problematic indoors objects to train the system. A user can use images or video-based data of indoors objects to interact with the system. Detection algorithm YOLO detects and recognizes the objects present in the images or video data, and it returns the risk tags associated with the objects. To process and communicate the results, the system just only took an average of 0.8 s. The accuracy gained by this system to detect and recognize the potential indoors harmful objects is 96%, which proves that this system can be helpful to provide awareness on a large scale to avoid harmful incidents during earthquakes.

Prasad et al. [25] proposed and implemented an ML-model-based ensemble technique in the west coast area of India to map the flood vulnerability. To address this issue, KNN, logitboost (LB), boosted regression tree (BRT), nearest shrunken centroids (NSC), and rotation forest have been used with adabag (AB) base classifier to map the vulnerabilities of the floods. The Boruta algorithm was used to select twelve such factors which were effective from the identified 210 flood areas. Different statistical measures were used to check out the validity of this approach. The authors used all the above mentioned models individually as well as with adabag (AB) as an ensemble. The results showed that the ensemble technique of AB-rotation forest models achieved the highest area under the curve (AUC) value of 94% as compared to other ensemble and individual models. According to the authors, this technique is very useful to map and study the flood areas which will result in better management and planning for these flood vulnerable areas.

Nsengiyumva and Valentino [26] addressed the problem of vulnerable landslide areas prediction using ML models. For this purpose, three ML models including Naïve Bayes tree (NBT), logistic model tree (LMT), and RF were used. This study was conducted in Rwanda to predict the vulnerable landslides in the upper Nyabarongo catchment area using map data relating to 196 landslides and using field investigations to map these slides. The authors divided the map data into training and testing data. An information gain technique was used to find out the correlation between the observed landslides and fifteen different factors associated with these landslides which were analyzed earlier. To validate the results of their study, the authors used different statistical measures including accuracy, RMSE, and precision. Moreover, area under the curve receiver operating characteristic (AUCROC) has also been used for this purpose as well. Results showed that NBT achieved the best results in terms of accuracy, precision and RMSE valued 0.799, 0.745, and 0.301, respectively. This technique also achieved the highest AUC value of 82.4%. According to the authors, this study can help to mitigate the risks associated with these vulnerable landslides and form a policy for such risks management.

Shirzadi et al. [27] studied landslide susceptibility mapping at the Bijar region, Kurdistan province (Iran) using a novel hybrid ML method. The developed method was based on NBT and random subspace (RS) ensemble, achieving an AUC value of landslide prediction equal to 0.886. The model outperformed the NBT classifier that achieved an AUC value of 0.811.

Sriram et al. [28] presented an advanced causal inference approach combined with ML to assess the vulnerability of urban infrastructure when exposed to extreme weather events. The authors proposed a deep NN-based causal approach to perform vulnerability assessment for electricity outages and roadway closures after extreme weather events. They applied the developed methodology in the context of the 2016 Hurricane Hermine at the City of Tallahassee, Florida. The results showed a 93.83% accuracy in the prediction

of power outages and a 90.54% accuracy in the prediction of roadway closures with the boosted gradient regression based forecasting method.

Wahab and Ludin [29] used the ANN technique to estimate flood vulnerability assessment. The performance was evaluated with the determination coefficient (R^2) and RMSE. The R^2 value obtained was equal to 0.996, and the RMSE was equal to 0.0035.

Mutlu et al. [30] developed a landslide mapping method using RNN to assess the landslide problem in terms of landslide susceptibility and inventory. The achieved AUC value was equal to 0.93 for the test dataset.

Pham et al. [31] developed landslide models in order to assess landslide susceptibility based on a combination of ensemble methods and the MLP classifier. The authors evaluated the performance of the various models. The MultiBoost model achieved the highest performance with an AUC value equal to 0.886.

4.3. ML/DL Methods for Disaster Detection

Gupta and Roy [32] proposed a new framework to identify the main reason of damage caused in the disaster hit areas so that effective and accurate emergency activities can be carried out. Six disasters of different types have been considered in this study. For the recognition and detection of natural disasters through this framework, a satellite images dataset of these disasters has been used here. For the feature vector construction, the framework combined two types of features which included local binary features and image wavelet scattering features. This framework had less cost in terms of the computational power that is required and also had better accuracy as compared to the deep NN models. The framework achieved better accuracy than the state of the art hand-crafted techniques and other ML techniques. Authors achieved an F1 score and accuracy of 99.40% and 99.59%, respectively, on the satellite images dataset. They argued that, using satellite images, the model can be useful to locate regions affected by any type of natural disasters from the six different types of disasters considered in the study.

A floodwater detection method based on CNN was developed by Layek et al. [33] to detect flood images from images posted on online social media. A color-based filtering followed the CNN-based detection of flood images. Real Twitter image data from a flood event were used and results showed an average accuracy of 80.07% and F1 score of 0.77.

Muhammad et al. [34] presented an early fire detection methodology using fine-tuned CNN for closed-circuit television (CCTV) surveillance cameras. An adaptive prioritization mechanism for cameras was proposed for autonomous response and a dynamic channel selection algorithm based on cognitive radio networks for reliable data transmission. An alert is sent to the disaster management system based on the adaptive prioritization mechanism for the cameras. The proposed method showed a detection accuracy of 94.39% and an F1 score of 0.89.

A wildfire detection system based on deep CNN for early wildfire detection was proposed by Lee et al. [35] using aerial images taken from UAVs. Various deep CNNs were evaluated, such as AlexNet, GoogLeNet and VGG-13. GoogLeNet achieved the highest accuracy equal to 99%.

4.4. ML/DL Methods for Early Warning Systems

Chin et al. [36] worked on improving the detection accuracy of earthquakes in early warning and detection systems in order to address the false alarm problems of these systems. These early warning systems use high-speed computer systems due to the very low available response time to transmit the earthquakes waves-related information to different centers working for this purpose. As heuristic thresholds and empirical features are used with decision algorithms in these systems to issue the warning, these systems usually issue fake warnings, causing panic to the response authorities and other tense situations. These events can also result in the disruption of services of different departments, causing heavy loss. To stop these false alarms, three ML-based algorithms have been experimented: SVM, classification tree and KNN. A criterion-based method was used to

compare the performance of these algorithms using the earthquakes related seismic data in Taiwan. In this study, ML-based algorithms significantly reduced the false alarms rate by increasing the detection accuracy.

Li et al. [37] aimed to mitigate false alerts and noise from earthquake early warning systems. The authors trained a generative adversarial network (GAN) to learn the characteristics of first-arrival earthquake P waves, using waveforms recorded in southern California and Japan. An RF classifier was trained with earthquake and noise waveforms. The classifier identified 99.2% of the earthquake P waves and 98.4% of the noise signals.

In [38], Moon et al. developed an ML-based method of an early warning system for very short-term heavy rainfall. Meteorological data were preprocessed by the selective discretization and principal component analysis. The logistic regression (LR) classifier was used for prediction purposes. The performance of the developed approach was evaluated and compared to early warning system models using other classifiers. Data from 652 locations in South Korea from 2007 to 2012 were used. The empirical results showed that the preprocessing methods improved the prediction quality and LR performed well on heavy rainfall in terms of F1 score and ETS. The proposed approach outperformed other classifier methods with an accuracy of 99.93% and an F1 score equal to 0.4601.

4.5. ML/DL Methods for Disaster Monitoring

During a disaster, the organization and live monitoring of information is crucial, and many actions aim to optimize disaster monitoring.

Gopal et al. [39] used online news data for disaster monitoring. The authors used a data scraping approach to crawl the online news data from different websites and online sources which are based on various hazard emergencies. As crawlers collect all the information and cannot differentiate between the useful and non-useful information and data, ML-based approaches have also been incorporated in this study to target only useful data by filtering out the irrelevant data. Text classification has been carried out by using supervised ML-based approaches which classify news data which is important and collected from different news articles and stories. The developed method can be applied to monitor online news related to disasters in order to facilitate disaster preparedness and response.

Domala et al. [40] worked towards better crisis management using news data by incorporating natural language processing and ML models. A scraping technique was used by the authors to scrape the news related to disasters from several news outlets, and then natural language processing and ML techniques were applied on these news data to identify the relevant news data, which was then shared and shown on the crisis management websites. For scraping news data from different English news websites, a spider-scraper tool was used. This is a totally automated system which classified the news data by using ML models into two categories which were disaster-relevant news data and disaster irrelevant news data that was then published on websites related to crisis management.

Fan et al. [15] worked on discovering disaster related events from social media posts across various locations using hybrid ML methods for an effective disaster response. In earlier studies conducted on this issue with the use of geotagged locations and detection of coarse grained events, it was found that the posts' content does not always contain the entirety of the information which may hinder the situation's awareness and enshroud important information. So, the credibility of situation awareness social media data can be improved using important event information and accurate location detection. To address the above-mentioned limitations of earlier studies, the authors proposed a hybrid ML system which considers all the tweets related to the disaster across different locations to uncover all the disaster related events. To identify the credible information, graph-based clustering was used while the posts classification task was handled by a BERT transformer model. Moreover, for the detection of locations mentioned in the posts, named entity recognition (NER) was used. For denoising the data and extracting the location coordinates, the authors utilized a location fusion approach, which is a version of fusion approach.

For this study, the authors used the 2017 Hurricane Harvey related Twitter posts data in Houston and this application gained good accuracy by successfully mapping the events mentioned in the posts based on different time and space. The authors suggested that the application can be helpful to carry out emergency operations across different locations in a timely manner and it can be used for awareness purposes as well. Furthermore, the application can be used for risk mitigation as well as post-disaster response.

4.6. ML/DL Methods for Damage Assessment

Damage assessment includes the techniques that can put to scale the damage done as well as calculate the resources that are needed for disaster response.

To assess the disaster situation, Wang et al. [41] proposed a novel framework which used DL-based multimodal approach. This framework included VGG-19, CNN and LSTM models. This framework significantly improved the performance of the model as it used an automatic method of loss weighting instead of manual weight tuning process. A main advantage of this framework was that it could also capture the correlation between different kinds of data and concepts. A large-scale Twitter-based dataset was used for the experiment in order to identify the level of damage caused by the disaster. The ability of this model to learn the multi tasks with weight loss outperformed all other single task-based models by achieving an F1 score of 0.857.

Resch et al. [42] addressed the shortcomings of limited spatial and temporal resolution and high temporal lags. The authors combined the ML techniques, namely, Latent Dirichlet Allocation (LDA), which helps to identify the hotspot areas by extracting the semantic information and carrying out a temporal and spatial analysis on the social media posts to assess the damage in the result of natural disasters. Authors successfully and accurately identified the earthquakes and other natural disasters in advance with different spatial and temporal properties. Moreover, they successfully generated a loss map caused by these natural disasters, which is validated by using the HAZUS loss model [43]. Furthermore, identification and prediction of earthquakes was also validated by using the official earthquake footprint provided by the US Geological Survey.

Assessment of the damage caused by the natural disasters is important for the timely response and relief activities. For this purpose, Presa-Reyes and Chen [13] proposed a CNN architecture that was two-streamed network. The proposed architecture overcomes the shortcomings of pre-existing CNN applications, which classified well the structures between destroyed and intact, yet these applications do not perform well if there are more than two damage levels. The proposed architecture overcomes this problem very effectively and could differentiate up to four damage levels with better accuracy. For this purpose, pre- and post-hurricane disaster aerial images have been taken and then evaluated using the proposed architecture. Concatenated features from the pre- and post-damage aerial images played important role for the prediction as these features were used to train the model. To evaluate this architecture an open source dataset that was fully labelled was used.

Akshya and Priyadarsini [14] proposed a hybrid-in-nature approach and incorporated ML techniques to classify the areas which are flood affected. In this study, the authors used the drone to capture aerial images of the areas from various heights as drones can capture high resolution images with various features. For the classification purpose, K-means clustering and SVM have been used in combinations. Aerial images were input to the system which then classified them according to whether the area in the image was flood affected or not. The authors also used different kernel functions which were used in SVM to check the performance of the system. It was also shown that when quadratic SVM was used, then the training and prediction time of the system decreased. Overall, the proposed system achieved an accuracy of 92% by classifying the flood affected areas accurately.

Yang and Cervone [44] presented a method that combined DL and ML in order to assess damage based on information extracted from aerial images. A pre-trained DL CNN model initially identified critical infrastructure from images. Various ML techniques such as SVM, RF, DT, LR and KNN, were then used to capture the features associated with the

damaged areas. An ensemble max-voting classifier was then constructed from the trained ML classifiers. The proposed methodology was applied to assess the damage of flooded areas in the state of Texas using aerial images collected in 2015 and showed an accuracy of 85.6% accuracy and an F1 score of 89.09%.

Nguyen et al. [45] used deep CNN to perform damage assessment caused by natural disasters. The domain-specific fine-tuned deep CNNs outperformed other techniques such as Bag-of-Visual-Words (BoVW). The VGG-16-fine-tuned achieved an accuracy of 0.84 and an F1 score of 0.82 for the Nepal earthquake data.

Rizk et al. [46] proposed a two-stage multi-modal damage classification scheme. During the first stage, classifiers were trained on visual and semantic features. Visual features included color, shape and texture features and semantic features were based on BoW. The outcome of the classification of the first stage was used in the second stage to create a multi-modal feature vector that was used to train a new classifier. A dataset was created using Twitter data. An accuracy of 92.43% was achieved with a linear SVM classifier.

Dotel et al. [47] proposed a DL based approach to assess the impact of water related disasters in urban and rural areas using satellite image data. CNNs were employed to assess the disaster impact in urban areas by segmenting topographical features such as roads in images obtained before and after a disaster and identifying regions of maximal change. Regarding rural areas, a bitemporal image classification method was developed to assess the disaster impact directly comparing pre- and post-disaster area images. The method developed for urban areas was tested based on an image provided by DigitalGlobe with labeled data showing the impacts of Hurricane Harvey. The method developed for rural areas was tested using images from the South Asian Monsoon Flooding of 2017.

Li et al. [48] developed a method based on CNNs and class activation maps to identify and assess damage in disaster-hit areas. The components of the method included: a CNN that classified images into the classes damage or no damage; class activation mapping; damage severity score. The proposed method was evaluated based on image data from different disaster events. A dataset consisted of social media image data from the Nepal earthquake, Hurricane Matthew, Ecuador earthquake and typhoon. An accuracy of 90.1% was achieved using image data from the Ecuador earthquake.

4.7. ML/DL Methods for Post-Disaster Response

In the post-disaster response, the aftermath of an event is evaluated in order to respond to immediate needs, perform search and rescue operations, quantify the damage done and the impact of the disaster in the community. In this phase, data is also collected that will be later used to prevent future similar disasters or cope with similar disasters.

In disaster-hit areas, supply of relief aid is one of the main activities conducted for disaster response. For this purpose, the traditional way of estimation of relief supply depends upon the census data which may cause imbalance in demand and supply due to the bias during the data capture operations, known as participation bias. For this purpose, a dynamic calculation of the population distribution is spatio-temporal which can be helpful to meet the supply demand. To overcome such issues, Lin et al. [49] proposed a dynamic model for the demand of relief supplies which was driven by big data, as the emergence of big data has changed the scenario due to various sources of data such as crowdsourcing, web mapping and social media data. In this paper, the authors worked on the urban flood disaster using Baidu big data and incorporating the MLP to improve the accuracy of relief supplies across the affected population. The MLP NN was trained on the Baidu big data as well as historical data of floods for the effective estimation of relief supplies. Due to its accuracy, the authors argued that the model can be an ideal choice for the governments to effectively carry out the relief activities in pre- and post-disaster areas.

O'Neal et al. [50] addressed the problem of noisy social media data related to natural disasters by creating an ML-based supervised model on the qualitative data. These data were collected from the 2017 Hurricane Harvey disaster from the on-field rescue workers and those rescued; data were captured through interviews and their social media posts, in

terms of images, videos and text from their social media accounts. Data was also collected from Twitter, Facebook and Instagram posts. Google vision API was used here to detect the attributes from these posts which have been proven more accurate as compared to the human attributes-detection, and eight various ML-based supervised classification algorithms were used. Between signal (high-quality) data and noise, authors achieved an accuracy on classifying data equal to 99%, and the same accuracy was achieved on respondent type classification. This ML-based automatic human role classifier and pattern recognition application had an edge over other applications as it pursued signal data instead of noisy social media data.

Post disaster, various activities are carried out, and medical rescue is one of them. Li et al. [51] carried out qualitative research for medical rescue so that better disaster management, in terms of medical rescue, can be achieved. By accurate classification of the disaster, effective and correct decision making aids the medical response during the event. A decision table based on medical rescue was constructed by incorporating the medical features of different kinds of disaster types. To classify the disasters, medical features were then analyzed on various bases using a genetic algorithm. Based on this classification and the common features with the disaster personality features, recommendations were issued, which in turn assisted a rescue management system based on medical emergencies to plan the rescue work.

Ehara et al. [52] proposed a system using UAVs for the recognition of individual people and their critical status to recommend them to the rescue teams for immediate aid. This system used supervised ML techniques to classify the status of the individuals whether they were standing, sitting or lying down on the ground. UAVs have proved their advantages due to their wide area coverage through videos and photographs in a very short span of time, so authors capitalized this advancement in UAVs to obtain the data related to the status of the people from the disaster struck areas. The system successfully classified all the three statuses of the individuals with an accuracy of 95.6%. It also proved to be effective in real-time scenarios in a disaster struck area where it successfully recognized the status of the peoples. This is a prime example of how a combination of new cutting-edge technologies can play a vital role and save human lives in real scenarios.

In [53], Reynard and Shirgaokar used ML algorithms to characterize geolocated tweets about Hurricane Irma, in Florida. The research aim was to identify whether Twitter data can be used in planning disaster response operations during and after the event. The authors used sentiment analysis to categorize tweets containing damage or transportation related information. They trained ML techniques to identify negative, neutral, or positive sentiment.

Chaudhuri and Bose [54] presented a method of effective classification of images from earthquake-hit smart urban settlements. The authors applied a DL method to identify survivors in debris, based on CNN variants, namely AlexNet, Inception-V3, and ResNet-50. ML algorithms were also used, namely, ANN and SVM. Performance evaluation results showed that DL methods outperformed ML methods for image classification, with ResNet-50 showing the best performance, achieving a positive predictive value (PPV) score of 90.81% and F1 score of 0.9205.

Li et al. [55] presented a methodology to automatically analyze tweets in order to assist in post-disaster response. The proposed methodology was based on domain adaptation classifiers, using both source labelled data and unlabeled target data to train classifiers. The method combined NB with an iterative self-training method (NB-ST). A dataset of tweets called CrisisLexT6 was used to implement the experiments. The NB-ST method outperformed traditional supervised ML classifiers such as NB, RF, SVM and LR, in identifying tweets relevant to a disaster showing an accuracy of 86.91%.

Bejiga et al. [56] presented a DL-based methodology in order to facilitate avalanche search and rescue operations with UAVs equipped with vision cameras. A pre-trained CNN was used to extract features from images of an avalanche debris. A trained linear SVM was used to detect objects. A post-processing method based on a Hidden Markov

Model was further employed to enhance the prediction accuracy of the classifier which achieved a value of 96.93%.

Robertson et al. [57] compared the DL and ML techniques on the social media image data posted by the people during Hurricane Harvey to check the communication with the emergency responders during disaster. For this purpose, the authors used a portion of a dataset composed of the images posted by the people on social media. A total 1128 images were used, which were randomly selected out of 17,483 images. During disaster situations, emergency services can be overloaded. It is often the case when people post images on social media platforms to better reach out to the emergency responders. The framework proposed by the authors used MLP to classify the images posted by the users while to extract the important features from the images, VGG-16 CNN had been used. According to the findings, ML methods cannot always be useful to capture the human experiences during disaster so these methods can work better when used together targeted at the important requests and content on the social media.

Using DL models, Huang et al. [58] also used the social media posts data to overcome the problem of retrieving a specific topic related to social media posts with textual and visual information during a disaster. The proposed approach was an automated approach which used combined features of visual and textual information of social media posts to label the related topic of social media posts. To extract the visual and textual features from the posts of social media, the authors used embedded CNN and Inception-V3 CNN. Extracted features were trained on the train data and then combined together to be passed forward for the classification purpose of social media posts. Relevant or on-topic social media posts were classified automatically based on the pictures and text data which helped to formulate a timely document of the events. By combining these posts with geotagging features, social media can play a vital role to formulate an approach to mitigate the disaster. During the fused classification phase, selected ML algorithms including LR, DT, RF, SVM-Linear, SVM-RBF, and SVM-Sigmoid were employed. The methodology was tested with a case study of the 2017 Houston flood using Twitter posts. The classification performance results for LR showed the highest classification accuracy of 95.2% and an AUC value equal to 0.945.

Kundu et al. [59] presented an LSTM-based architecture to classify tweets into various classes relevant to post-disaster activities such as resource need and availability and activities of non-governmental organizations (NGOs). The dataset used in the study was obtained from Forum for Information Retrieval Evaluation 2016 (FIRE2016) and 2017 (FIRE2017) and included data on the Nepal earthquake. The proposed method performed better than BoW and term frequency-inverse document frequency (tf-idf)-related approaches, achieving a precision of 0.9234 and an F1 score of 0.9159.

A methodology was proposed by Basu et al. [60] for informed post-disaster response management regarding resource needs and resource availabilities based on information retrieved from tweets. The authors proposed two unsupervised neural retrieval models that combined word-level and character-level embeddings. They investigated various supervised classification methods, unsupervised pattern matching and unsupervised information retrieval methods. The dataset consisted of tweets posted during the 2015 Nepal earthquake and parts of India and the earthquake in central Italy, in August 2016. For the Nepal earthquake data, the proposed unsupervised information retrieval method outperformed other methodologies. The proposed word-level and character-level attention-based embeddings method achieved a classification accuracy of 0.57 and F1 score of 0.191 for the Nepal earthquake data.

Neppalli et al. [61] developed and evaluated pre-trained CNN and RNN models in identifying informative Twitter data which contain critical information about the situation of a disaster event. They engineered a set that combined BoW features with features extracted from tweet content, user details and polarity clues and trained NB classifiers with the feature set as well as with each feature independently. The CNN and RNN models

performed better than the NB models in identifying informative Twitter data regarding disasters and generalized better across different disasters using the CrisisLexT26 dataset.

Paul et al. [62] extracted information from Twitter data about power and communication outages relevant to seven major hurricanes that hit the USA between 2012 and 2018. Various ML models such as SVM and LR were used to filter out tweets related to outages. They also applied transfer learning models such as BERT to detect the various types of outages.

Kabir and Madria [63] proposed a method to address the challenges of effective rescue scheduling during disasters. The developed method combined an attention based Bi-directional LSTM (bi-LSTM) and CNN to classify Twitter data and used feature engineering to increase the model accuracy. They also developed a hybrid scheduling method for rescue operations. The proposed method was evaluated based on Twitter data from Hurricanes Harvey and Irma, as well as a merged dataset of different disasters from CrisisNLP and CrisisLex.

Peng et al. [64] addressed urban flood mapping at high resolution by using a residual patch similarity convolutional neural network (ResPSNet). They also used data augmentation to remove the impact of varying illuminations due to different data acquisition conditions. Experiments were performed based on data from the 2017 Hurricane Harvey flood in Houston. An accuracy of 94.97% was achieved.

4.8. ML/DL Methods in Case Studies for Disaster Management

Nagendra et al. [65] addressed the problem of relief operations management that can be disrupted due to the destruction or damage of communication infrastructure as the result of disasters. If information and communication technology (ICT) infrastructure is damaged during a disaster, it can be difficult to carry out relief operations. This disruption can cause serious problems in communications and the identification process of the areas where immediate relief operations are needed. The authors proposed and applied a methodology to identify those areas where relief operations are needed based on priority using big data analytics techniques. They used the census, geospatial and satellite images data and deployed this system using the cloud platform of AWS (Amazon Web Services). For validation purposes, authors deployed this system in the state of Kerala in the 2018 flood and involved the rescue teams to carry out the targeted relief operations. This system proved to be helpful as it improved the timely and targeted relief operations with improved communication between the rescue teams.

Critical communication and knowledge are very important to carry out the rescue operations, which can be difficult in disaster-hit areas. To address this problem, Laverdiere et al. [66] proposed a framework using remote sensing and DL techniques. For this purpose, to support the rescue operations carried out by rescue teams in 2018 in the wake of a lava flow incident in Hawaii Island, the authors used the satellite high-resolution images. In this framework, CNN models were used which showed the ability to generalize well in the training phase on the available data so that mapping of the structure can be performed very quickly on the pre- and post-incident satellite images. Rescue agencies were provided the important and timely information to carry out the assessment of damage and loss due to this disaster. In this case study, the future direction was provided of how this study can be beneficial on a large scale in such incidents.

Sit et al. [67] worked on a case study of the Hurricane Irma by identifying and analyzing the tweets related to this disaster using natural language processing, DL and ML techniques. The purpose of this study was to identify the services which are disrupted, the affected people and the damaged infrastructure. Moreover, another purpose of this study was to differentiate the various impacted areas, as well as the time period and other related information. For the disaster-related binary classification, authors used DL and ML techniques. LSTM was also used for the classification purpose as it was, according to the authors, one of the best networks which outperformed the other methods. In this case, LSTM takes the whole text structure for semantic and feature dependencies analysis. For

affected people and disrupted services, the authors also used unsupervised learning for the multi-label classification of the tweets. The authors used 500 million tweets based on location data and keywords which were published before, during and after the disaster. The results were promising as the framework successfully identified the areas with highly affected people and damaged infrastructure.

Zhou et al. [68] studied landslide susceptibility using a case study at Longju, in the Three Gorges Reservoir area in China. Two types of landslide, namely, the colluvial and rockfall, were considered in the landslide susceptibility modeling. The SVM model outperformed the ANN and LR models, achieving an overall AUC value of landslide prediction 0.881 compared to 0.836 for ANN and 0.697 for LR.

4.9. ML/DL Methods in Developed Applications for Disaster Management

With the introduction of 5G technology, many innovations are taking place in the various industries due to its low latency. By combining distributed edge computing with the 5G technology, the automation process in many sectors is gaining pace. One of such examples, presented in Ardiansyah et al. [69], is 5G-DIVE. In this paper, authors used this 5G-DIVE with the autonomous drones for navigation, surveillance and real-time emergency situation detection. The authors also used the edge computing ML models for these real-time emergency situations detection. This system was named as EagleEYE which was basically an aerial system for disaster relief operations. By reusing existing datasets and making a mechanism of object fusion, this system reduced the time required for training. The system also worked parallel both for the detection and response tasks. A new algorithm, merged object detection (MOD), was proposed by the authors, while CNN and YOLO V3 have been used as well. For evaluation purposes, the authors tested this system on two datasets including COCO and Google Open Images dataset, as well as in real-time, and gained an accuracy of 87% for detection purpose with 90% reduced inference latency.

Zhang et al. [70] worked on an application for the phase of damage assessment. Many applications have been proposed earlier by various researchers which incorporate AI- and DL-based deep NN approaches. To assess the damage and the severity of the impacts in the disaster-hit areas, post-disaster image data was used in these applications. However, the black box nature of AI algorithms makes it difficult to achieve better accuracy. For this purpose, research focus has been shifting and adopting new ways to address this problem. In this paper, the authors also incorporated the crowd sourced information gained through social media data with AI-based algorithms. Prior to the use of the machine intelligence, crowd source information was used to tune, improve and troubleshoot the black box nature of AI algorithms. Then, this crowd source information worked closely with the machine intelligence in a cooperative way within the system. A damage assessment application based on DL models was designed which is termed as CrowdLearn: crowd-AI hybrid application. CrowdLearn was based on the crowdsourcing platform and exploited the combination of crowd intelligence and machine intelligence to perform damage assessment. This system was evaluated in real-time and proved that it can provide timely and accurate assessment of the disasters. It outperformed existing AI algorithms such as VGG-16, achieving a classification accuracy of 0.877 and F1 score of 0.894.

Alam et al. [71] worked on the solution of problems faced during relief operations by using image processing techniques on images. The system was called Image4Act. To assist the humanitarian organizations carry out the relief operations, the system used image data of social media posts. This system collected the data, denoised it and then classified it during natural disasters. Perceptual hashing and deep NN techniques have been used to remove the noisy data. To assess the damage of infrastructure caused due to the disaster, a case of real-time cyclone natural disaster was used. Evaluation of the system on pre-existed dataset of disasters as well as in real time proved that this system was very effective and reliable to use in real-time natural disasters.

Song et al. [72] developed a disaster management system, called DeepMob, which predicted and simulated people's evacuation behavior and mobility as well as evacuation routes following various types of earthquakes. DeepMob used heterogeneous big data sources such as GPS records, transportation network data and Japan earthquake data. The system achieved an accuracy of 87.8% in predicting people's evacuation behavior.

5. Discussion

Appendix A Table A1 summarizes the technical analysis of the reviewed papers, based on the disaster phase and subphase, the ML and DL techniques used, the data sources used for evaluation purposes, performance metrics and the disaster type. The proposed studies aimed to provide solutions in various subphases/areas of disaster management: disaster and hazard prediction, risk and vulnerability assessment, disaster detection, early warning systems, disaster monitoring, damage assessment and post-disaster response. Based on the results shown in Table A1 and Figure 4, research studies focused on response operations outnumbered those focused on other disaster management phases. Figure 4 shows the number of research studies in terms of the disaster subphases in percentage form. The largest percentage of studies focused on post-disaster response (38.2%), followed by damage assessment (20%), risk and vulnerability assessment (14.5%), disaster and hazard prediction (9.09%), early warning systems (5.45%), disaster detection (5.45%) and disaster monitoring (5.45%). Accordingly, more than 50% of the research studies focused on disaster response, followed by disaster mitigation and then by disaster preparedness.
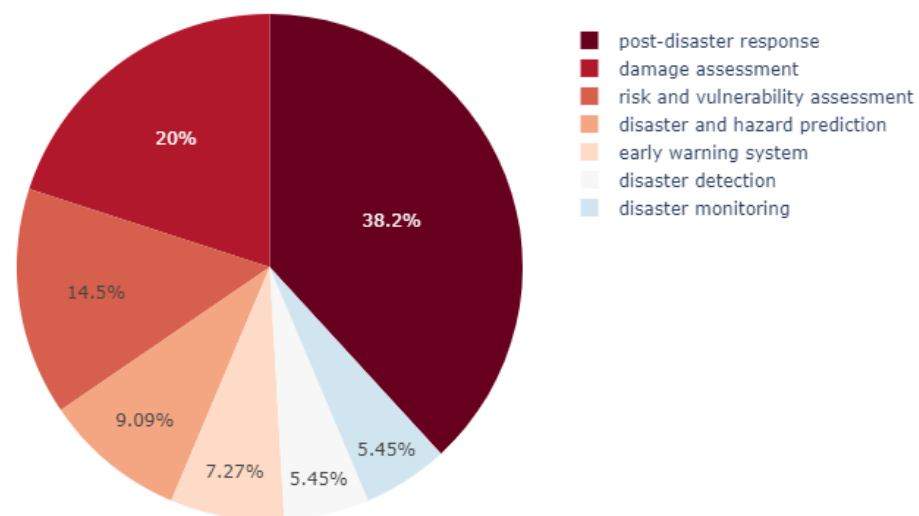

Figure 4. Research publications by disaster subphase in percentage form.

The percentage distribution of research studies by disaster type is shown in Figure 5. As can be seen from Table A1 and Figure 5, research efforts have been put on developing ML/DL methods applicable for various types of disasters. Floods have been mostly studied (20.3%), followed by earthquakes and hurricanes (18.8% each) followed by general type (any disaster type) (15.9%) and landslides (10.1%). Other types of disasters that have been studied include (heavy) rainfall, typhoon, volcano, wildfire, avalanche and tsunami.
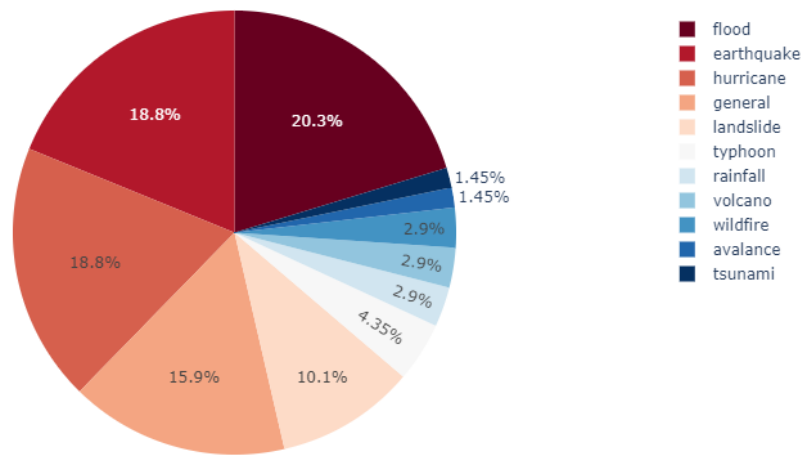
Publications by Natural Disaster

Figure 5. Research publications by type of disaster (the percentage distribution of research studies by disaster type).

Figure 6 shows the percentage distribution of studies developed using ML and DL by disaster phase. In terms of the mitigation phase, the research studies focused on ML have largely outnumbered those focused on DL. Accordingly, 69.2% of the studies have been developed based on ML techniques and 30.8% of the studies have been developed based on DL techniques. For the disaster preparedness phase, equal number of studies have focused on ML and DL techniques (50%). In terms of disaster response, the studies developed based on DL techniques have largely outnumbered those based on ML techniques. Accordingly, 69.2% of the studies have focused on DL and 30.8% on ML.

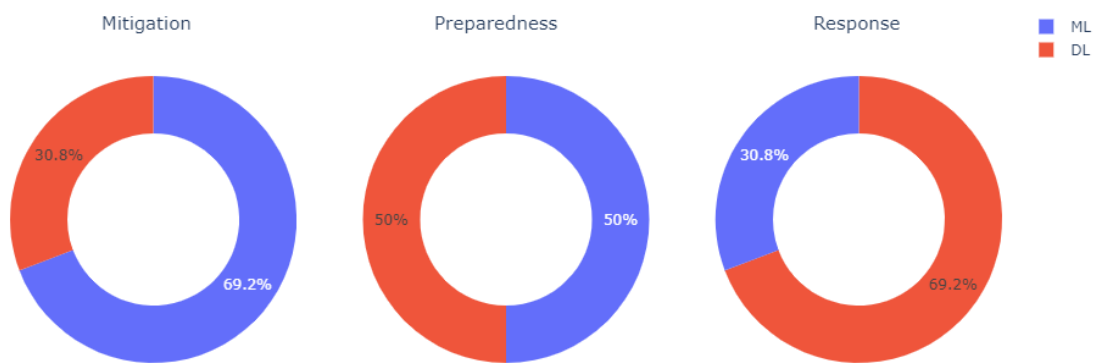Method distribution by disaster phase

Figure 6. Distribution of the research studies in percentage form by disaster phase.

Figure 7 shows the percentage distribution of ML and DL in the presented research studies in terms of disaster subphases. Accordingly, DL-based methods have outnumbered ML-based methods in damage assessment and post-disaster response with 83.3% and 63%, respectively, and disaster detection with 75%. ML-based methods have outnumbered DL-based methods in disaster and hazard prediction with 60%, early warning systems with 75%, risk and vulnerability assessment 66.7% and disaster monitoring with 66.7%.
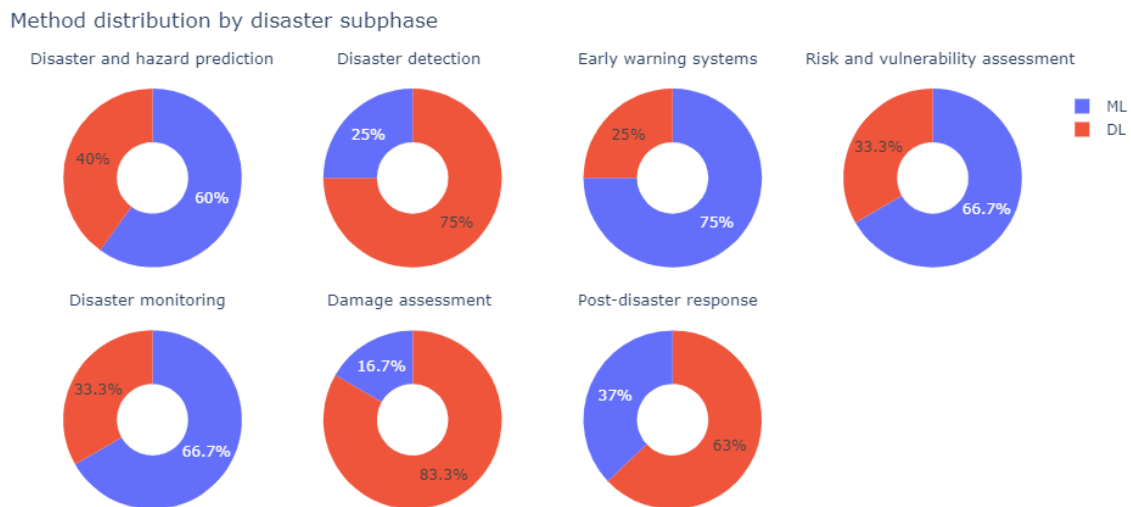
Method distribution by disaster subphase

Figure 7. Distribution of ML and DL in the research studies in terms of disaster subphases in percentage form.

Online text, image and video data posted on social media platforms has gained importance in crisis and disaster management. Effective decision making by emergency responders and other decision makers can be facilitated by the information posted online on social media platforms such as Twitter. In disaster response, research has mostly focused on DL using social media and in particular Twitter data. The emphasis on post-disaster response studies has been placed on processing and analyzing Twitter data [49,50,53,55,57,58,60–63,67]. Furthermore, besides post-disaster response, social media data and mainly Twitter data, has been used in studies focused on damage assessment, disaster monitoring and disaster detection. Other data sources used in different subphases include the following: satellite imagery for disaster risk and vulnerability assessment and disaster detection; inventory maps for disaster risk and vulnerability assessment; sensor data for disaster and hazard prediction and early warning systems; video for disaster detection; online news for disaster monitoring; aerial imagery for damage assessment and post-disaster response; other image data such as from the AIDR platform and Google for damage assessment. Furthermore, the crowdsourcing platforms have been used to improve the performance of DL-based methods for damage assessment in terms of noise (false data labeling) reduction and the black box nature [70].

Technologies such as UAVs and drones have been used in obtaining aerial imagery. In [14], SVM was employed to assess the damage of flood affected areas using aerial image data obtained from drones. A method based on CNN VGG-16 to assess post-disaster information for disaster response operations using UAV aerial image data was proposed in [52]. Wildfire detection based on various CNN variants used UAV aerial images in [35]. Drone aerial imagery and 5G technologies were used in DL-based methods [69] for disaster detection and post-disaster response.

CNN and its variants have been employed for the development of methods processing and analyzing social media image data and especially Twitter data [41,45,70,71] for damage assessment. Traditionally, disaster damage assessment was performed by domain experts. The drawbacks in this approach include the inability of domain experts to handle large amounts of data and the high labor cost. The challenges of analyzing large volumes of image data include not clearly defined images such as destroyed roads, poor signal-to-noise ratio, subjective assessment of the damage severity by human annotators, and difficulty of acquiring sufficient number of labeled data at the onset of a disaster. A pre-trained CNN was developed in [45] to overcome the difficulty of obtaining large datasets to train CNNs. A combination of CNN and bi-LSTM leveraged information from both text and image Twitter data [41].

Studies on post-disaster response aimed to leverage information mostly provided by social media and especially Twitter in order to achieve the following: improve communication between affected population and emergency responders [57,66]; estimate demand of relief supplies [49]; search and rescue operations [52,56]; evaluate Twitter performance in planning and assisting disaster response operations [53,55]; identify resource needs and resource availabilities based on Twitter data [59,60]; enhance the robustness of developed methods in the context of noise reduction [50]; identify affected population and disrupted services [67]; manage relief due to disruption or damage of communication infrastructure [65]; retrieve information for general purposes [22,54,61,69,71]; extract information about power and communication outages [62]; rescue scheduling [63].

Based on the review results, accuracy, F1 score, precision, recall and AUC were mainly used as performance metrics to assess the ML- and DL-based methods for disaster management. Figure 8 shows the performance of the developed ML- and DL-based methods in terms of accuracy by disaster subphase and disaster type. Accordingly, highest accuracies have been achieved for post-disaster response regarding hurricanes, disaster detection regarding floods and wildfires and early warning systems regarding rainfalls and tsunamis.
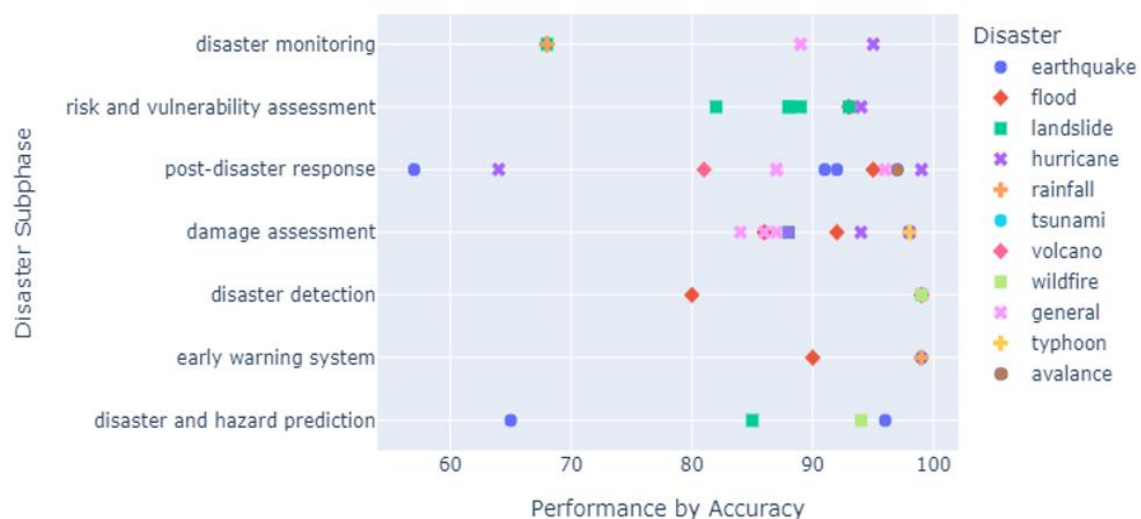

Figure 8. Performance of the developed methods in terms of accuracy by disaster subphase and disaster type.

Figure 9 shows the ML/DL-based method performance in terms of accuracy distribution for the different disaster phases. Most studies have resulted in performance accuracy between 80% and 98%. According to Figure 9, disaster response distribution by accuracy peaks at 91% accuracy in 5.3% of the publications following normalization of the distribution. Disaster mitigation shows highest representation of 5.1% of publications also on 91%. On the contrary, different data is shown in disaster preparedness distribution after normalization. Publications about disaster preparedness are more evenly distributed peaking at 95% accuracy at 3.9% of the publications. This figure shows that most ML and DL models used in publications on disaster management, in spite of the various different phases are performing very well, yielding results mostly around 88.8% accuracy. Another useful metric is the correlation between the accuracy and the type of data used by the models. More specifically, models that used image data display slightly better results in terms of performance in accuracy. Image-based models average 88.82% accuracy. while text-based models average 88.17% accuracy. Finally, models that used data in a structured form also display average accuracy of 88.65%. These results are very promising for future research as

they clearly show that regardless of the type of data used, ML and DL techniques can yield accurate results. Each type of data can be used exclusively or in combination with other types of data.
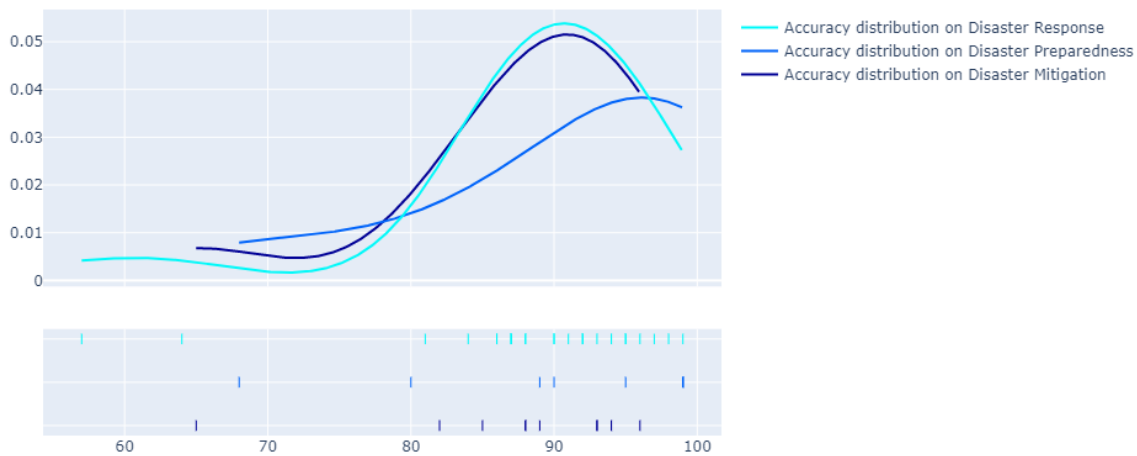

Figure 9. Performance of the developed methods in terms of accuracy distribution for the different disaster phases.

By examining in more detail the methods used in the reviewed sources and their accuracy, some methods stand out in terms of performance accuracy and yield better results in their respecting disaster phases. In disaster preparedness, 99.93% accuracy was achieved using the ML method of logistic regression (LR) and 99.2% was achieved with DL using GANs. In disaster mitigation, the highest accuracy of 96% was observed with DL using CNN, while the best ML method was KNN with accuracy of 92.74%. Finally, in the phase of disaster response, the highest accuracy was 99% and it was achieved with the ML method of SVM. The best DL method in disaster response was CNN with 98% accuracy. It is important to mention that across all the reviewed studies, the highest accuracy results of both ML and DL methods were achieved on structured data. More specifically, 99.93% was achieved with ML methods on meteorological data from locations in South Korea in a structured form. Additionally, 99.2% was achieved with DL method using GANs upon recorded waveforms in California and Japan in a structured form. This is an indication that both ML and DL methods perform much better on structured data. Although this analysis points out which methods of ML/DL perform better in terms of accuracy in each disaster phase, there are many factors that contribute to the overall efficacy of the methods used. For instance, unstructured data tend to be harder to analyze, preprocess and model, therefore leading to worse performance results and possible overfitting when fed to a ML/DL based system. A system that is trained in low quality datasets could potentially prove confusing when applied in a real-world scenario, due to its biased results that are a byproduct of its poor quality training data. A case specific comparison as presented in some of the studied literature is appropriate to determine which method is better in each scenario [15,23,25,26,44,50,54,62,67,68]. Additionally, the performance of the developed methods can be enhanced by integrating big data coming from different sources [13]. This conclusion highlights the need for the development and the expansion of additional data capture techniques in the domain of disaster management to produce clean data and big datasets of structured data that will support and fuel further research on the field and yield better results. Moreover, data quality can be enhanced by employing ML to automate the integration of multiple big data sources [10].

5.1. Limitations

According to the results in this paper, research studies focused on ML/DL approaches for (long-term) disaster recovery have not been identified. Although the keyword "disaster recovery" was not included in the search, the keyword "disaster management" could have

yielded relevant results for disaster recovery. Disaster recovery covers a broad range of activities aiming to bring communities back in normalcy, build community resilience and design of efficient policies, yet it has been understudied in the literature. Therefore, future research should focus on long-term disaster recovery, facilitated or driven by advances in both ML and DL.

Due to the manual filtering that followed the initial keyword search, some research articles may have not been included in this review.

5.2. Future Research Trends and Challenges

Bottlenecks in DL that need to be addressed in future studies, in order to enhance the robustness of DL-based methods for effective disaster management, include the limited amount of available labeled training data and the human labeling of the datasets [44,45]. DL overcomes the manual feature engineering process by automatically learning complex structures, yet with the expense of requiring very large amount of labeled data manually annotated in order to automatically learn the features. Moreover, social media data contain high levels of noise [50], therefore, more methods should be developed to effectively differentiate signal from noise in the above data.

According to the results of this paper, the prevalence and importance of Twitter has been shown as a tool to extract both qualitative and quantitative information valuable to emergency responders and other decision makers in all phases of disaster management. Considering that Twitter contains geospatial information, in contrast to other social media platforms, the extracted information can provide answers not only to what happened but also to where it happened. However, due to the fact that Twitter users may not be representative of all the population affected, cross-validation with information extracted by other means such as 911 calls could enhance the performance of the developed methods [53].

Crowdsourcing platforms such as Amazon Mechanical Turk have been shown to assist in the development of novel methods that improve the performance of AI algorithms for disaster management tasks [70]. Crowd intelligence combined with machine intelligence can increase the accuracy of the ML/DL developed methods and reduce labor costs of using domain experts for data labeling. More research is needed in different areas of disaster management to leverage the advantages of crowdsourcing when combined with ML/DL-based methods.

Online news data is shown to be a reliable source of disaster related information compared to social media data [39]. The process of extracting the right information in the right time from wireless sensor networks and other relevant technologies deployed in the disaster affected regions can be enhanced by using of web crawling or web scraping for effective preparedness and response operations.

Furthermore, focus on all phases and areas of disaster management is needed, as already identified in previous reviews [3]. In particular, disaster recovery management using ML/DL-based methods is under-researched. Long-term disaster recovery includes sustainable development efforts ultimately leading to community resilience building. The resilience of key infrastructure influences the effectiveness and progress of disaster recovery efforts [28]. More research is needed on the development of novel ML- and DL-based methods to assess the resilience of key infrastructure. Vulnerability assessment of an urban electricity and transportation infrastructure was studied in the context of resilience, using a hybrid graphical causal method [28]. Additionally, the accurate damage and loss assessment is necessary for efficient management of financial aid allocation during the recovery phase. Novel methods based on CNNs and satellite imagery data can enhance the performance of damage assessment of residential buildings [13]. Furthermore, the integration of multiple remote sensing data can improve the overall performance.

Moreover, ML- and DL-based methods can assist in the evaluation of the performance of disaster management operations. For instance, monitoring can follow the progress of disaster response and recovery operations. Hybrid ML-based methods have been used for disaster monitoring [15,40]. The performance of disaster management operations should

also be measured in terms of the distress of the affected population [4]. ML- and DL-based methods can be developed to measure the distress and feelings of affected population based on the analysis of social media data during disaster response and recovery [53].

Hybrid ML- and DL-based methods have been developed for different areas of disaster management to better cope with the complexity of operations. The criticality and complexity of disaster operations highlights the need for robust, validated and trustworthy AI solutions. Robust AI models should also explain both the outcome and the process which led to it to the human experts [73]. Accordingly, in the context of medicine, verification and explainability has been identified as a key frontier research area [73]. Yet, it is important, besides in the area of medicine, to include explainability in research focused on all application areas that affect human life, including disaster management.

6. Conclusions

Natural disasters are one of the major causes of human lives loss and damage to infrastructure and property. Advances in ML and DL have been increasingly used to manage with the complexity of disasters. In this paper, a review study has been carried out to investigate how ML and DL techniques have been used in various areas of disaster management to assist disaster management operations and improve their performance. For this purpose, those papers which have been published since 2017 have been targeted, dividing them in categories which include the different phases and subphases of a disaster event. In these categories, various ML and DL techniques have been used for different types of disasters including floods, lava flow, earthquakes, typhoons, hurricanes, landslides among others. Reviewed studies focused on the areas of disaster and hazard prediction, risk and vulnerability assessment, disaster detection, early warning systems, disaster monitoring, damage assessment and post-disaster response as well as cases studies and applications. Challenges and future research directions have been discussed. Future research should be directed towards leveraging ML and DL for improving the performance of disaster recovery operations. Disaster recovery operations should be sustainable; therefore, research should focus on using ML and DL to enhance mitigation efforts, reduce vulnerabilities, and assess resilience including of key infrastructure.

The criticality and complexity of disaster operations requires robust and validated ML and DL solutions. Disaster operations affect human life; therefore, the developed models should also be explainable to be understood by domain experts and decision makers [73]. Moreover, research should focus on improving the quality of the data and developing novel data capture techniques as well as using crowdsourcing to improve the performance of ML/DL-based methods for disaster management operations.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Technical analysis of the ML/DL based research studies.

Ref. No.	Year	Disaster Phase(s)/ Subphase(s)	Disaster Type	Techniques	Dataset	Performance Metrics
[24]	2021	Mitigation/ Disaster and hazard prediction	Earthquake	YOLO (CNN)	Own indoor image dataset "TUHO" consisting of 19,200 images	Accuracy = 96%

Table A1. *Cont.*

Ref. No.	Year	Disaster Phase(s)/ Subphase(s)	Disaster Type	Techniques	Dataset	Performance Metrics
[25]	2021	Mitigation/ Risk and vulnerability assessment	Flood	KNN, LB, BRT, NSC, and rotation forest used individually and with adabag as an ensemble	Flood inventory map with 210 flood localities	AUC = 92.74%,
[26]	2020	Mitigation/ Risk and vulnerability assessment	Landslide	NB tree, Logistic model tree, RF	Inventory map of 196 passed landslides data	Accuracy = 79.9%, Precision = 74.5%, AUC = 82.4% RMSE = 0.301
[68]	2018	Mitigation/ Risk and vulnerability assessment Case study	Landslide	SVM, ANN, LR	High-resolution remote sensing imagery data of Pleiades-1 (9/22/2014) and GF-1 (3/30/2015) and historical landslide data, Longju in the Three Gorges Reservoir area in China.	For SVM: AUC = 0.881
[20]	2019	Mitigation/ Disaster and hazard prediction	Landslide	GA-MLP, traditional ML classification techniques		For GA-MLP: Accuracy = 85%
[27]	2017	Mitigation/ Risk and vulnerability assessment	Landslide	Hybrid ML method based on NB trees (NBT) and random subspace (RS) ensemble	Aerial photographs and satellite imagery data of Bijar, Kurdistan, obtained from the Forests, Rangelands and Watershed Management Organization (Iran),	AUC = 0.886
[28]	2019	Mitigation/ Risk and vulnerability assessment	Extreme weather events including hurricane	Deep NN-based causal approach	Data from Hurricane Hermine 2016, and January Storm 2017 on the Tallahassee, Florida	For the boosted gradient regression based forecasting: Accuracy = 93.83%
[22]	2018	Mitigation/ Disaster and hazard prediction	Tropical cyclone, Typhoon	Fuzzy NN combined with the locally linear embedding algorithm	Rainfall data, Guangxi, China Guangxi Meteorological Service, China	RMSE = 21.94
[29]	2018	Mitigation/ Risk and vulnerability assessment	Flood	ANN-MLP	Satellite and sensor data from Muar region, Johor, Malaysia	RMSE = 0.0035
[23]	2017	Mitigation/ Disaster and hazard prediction	Earthquake	Pattern Flood vulnerability assessment intensity map recognition NN, RNN, NN, RF, linear programming boost ensemble	Earthquake data from the Hindukush region, Pakistan.	For LPBoost ensemble: Accuracy = 65%
[30]	2019	Mitigation/ Risk and vulnerability assessment	Landslide	RNN	Landslide inventory map of the Buyukkoy catchment area, Turkey	AUC = 0.93
[21]	2020	Preparedness/Disaster prediction, early warning system	Flood	Deep NN	Data from flood affected districts from Bihar and Orissa, India	Accuracy = 89.71%
[31]	2017	Mitigation/ Risk and vulnerability assessment	Landside	Combination of ensemble methods, MLP	Landslide inventory map, Google Earth images data from landslide inventory map, part of Himalaya, India	For MultiBoost ensemble: AUC = 0.886
[36]	2019	Preparedness/Early warning system	Earthquake	SVM, KNN and classification tree	453 waveforms recorded data, Time series data, Central Weather Bureau (CWB) data	——
[37]	2018	Preparedness/Early warning systems	Earthquake	GAN RF	Waveforms recorded in southern California and Japan	Accuracy = 99.2%
[38]	2018	Preparedness/Early warning systems	Heavy rainfall	LR	Meteorological data from locations in South Korea	Accuracy = 99.93% F1 score = 0.4601
[32]	2020	Preparedness/ Disaster detection	Earthquake, Flood, Tsunami, Volcanic eruption, and Hurricane	Wavelet feature vector, LBP feature vector and Bagged decision tree classifier	Satellite imagery data	Accuracy = 99.59%
[33]	2019	Preparedness/Disaster detection	Flood	CNN	Twitter image data from a flood event	Accuracy = 80.07%, F1 score = 77%

Table A1. *Cont.*

Ref. No.	Year	Disaster Phase(s)/ Subphase(s)	Disaster Type	Techniques	Dataset	Performance Metrics
[34]	2018	Preparedness/Disaster prediction	Fire	CNN	Images and videos from various fire datasets	Accuracy = 94.39%, F1 score = 89%
[15]	2020	Preparedness, Response/ Disaster Monitoring	Hurricane Harvey	NER, BERT Transformer	Twitter data	Accuracy = 95.55%
[39]	2020	Preparedness, Response/ Disaster Monitoring	Landslide, Flood and Rain	SVM, K-NN, Naïve Bayes, MLP, LR	Online News Data	Accuracy = 68% F1 Score = 68%
[40]	2020	Preparedness Response/ Disaster Monitoring	General	LR, SVM with BoW	Online English News Data	Accuracy = 89% F1 Score = 88%
[13]	2020	Response/ Damage Assessment	Hurricane	CNN and ResNet50	Own aerial image data	F1 Score = 94%
[14]	2019	Response/ Damage assessment	Flood	SVM and K-means clustering	UAV aerial image data	Accuracy = 92%
[41]	2020	Response/ Damage assessment	General	VGG-19, CNN and LSTM	CrisisMMD Twitter Disaster data which includes 16,097 tweets and 18,126 images	F1 Score = 85.7%
[42]	2018	Response/ Damage assessment	Earthquake	ML Technique LDA and HAZUS loss model	Twitter Data	Accuracy = 86.45%
[44]	2019	Response/Damage assessment	Flood	CNN, SVM, RF, DT, LR and K-NN	Aerial image data from flooded areas in the state of Texas	Accuracy = 85.6%, F1 score = 89.09%
[45]	2017	Response/ Damage assessment	General	Deep CNN	Twitter Image data from Artificial Intelligence for Disaster Response (AIDR) platform Google image data	VGG-16: Accuracy = 84%, F1 score = 82%
[70]	2019	Response/ Damage assessment Disaster application	General, Earthquake	Crowd and machine intelligence DL such as VGG-16 Ensemble, Hybrid active learning	960 images from Twitter and Instagram about the Ecuador Earthquake in 2016 Amazon Mechanic Turk	Accuracy = 87.7%, Precision = 0.904, Recall = 0.885, and F1 score = 89.4%
[71]	2017	Response/ Damage assessment Disaster application	Earthquake, Typhoon, Hurricane	Deep NN, perceptual hashing, CNN, VGG-16, Transfer Learning	Social Media data, Twitter image data of Nepal earthquake, Ecuador earthquake, Ruby Typhoon and Hurricane Matthew	AUC 0.98
[46]	2019	Response/ Damage assessment	Earthquake, Flood, General	Linear SVM, kernel SVM, ANN, ensemble learning	Own dataset with text and image Twitter data, SUN dataset with image data	For linear SVM: Accuracy = 92.43%
[47]	2020	Response/ Damage assessment	Flood, Hurricane	CNN	Satellite image from DigitalGlobe about Hurricane Harvey, data about South Asian Monsoon Flooding of 2017	Accuracy = 94% (for the South Asian Monsoon Flooding of 2017)
[48]	2018	Response/ Damage assessment	Nepal earthquake, Ecuador earthquake, Typhoon Ruby, and Hurricane Matthew	CNN, VGG-19	Google image data, social media image data about Nepal earthquake, Ecuador earthquake, Typhoon Ruby, and Hurricane Matthew	Accuracy = 90.1% (for the Ecuador Earthquake)
[57]	2019	Response/Post-disaster response	Hurricane	VGG-16 CNN and MLP	Twitter image data	Accuracy = 64%
[58]	2020	Response/ Post-disaster response	Flood	Inception-V3 CNN and Embedded CNN	Social Media Twitter Data	Accuracy = 95.4% AUC = 0.945
[49]	2020	Response/ Post-disaster response	Flood	MLP	Baidu Big Data	Correlation coefficient = 0.923
[50]	2018	Response/ Post-disaster response	Hurricane Harvey	NB, KNN, SVM, and MLP	Social Media image data including Facebook, Instagram and Twitter data	Accuracy = 99%
[51]	2020	Response/ Post-disaster response	25 different types	Genetic algorithm	——	——
[52]	2020	Response/ Post-disaster response	General	CNN	UAV dataset of 420 images	Accuracy = 95.6%
[53]	2019	Response/ Post-disaster response	Hurricane	Multinomial logit model Sentiment analysis	Twitter data about Hurricane Irma, in Florida	

Table A1. *Cont.*

Ref. No.	Year	Disaster Phase(s)/ Subphase(s)	Disaster Type	Techniques	Dataset	Performance Metrics
[54]	2020	Response/ Post-disaster response	Earthquake	CNN variants: AlexNet, Inception-V3, and ResNet-50	Image data from the earthquake-hit regions, Tohoku, Japan (2011) and Bologna, Italy (2012).	PPV = 90.81%, F1 score = 92.05%
[55]	2018	Response/ Post-disaster response	General	NB-ST	Twitter data, CrisisLexT6	Accuracy = 86.91%
[56]	2017	Response/ Post-disaster response	Avalanche	CNN, SVM	Image data of an avalanche debris	Accuracy = 96.93%
[35]	2017	Preparedness/ disaster detection	Wildfire	Deep CNN: AlexNet, GoogLeNet and VGG-13	UAV aerial image data	Accuracy = 99%
[60]	2019	Response/ Post-disaster response	Earthquake	Unsupervised neural retrieval models that combined word-level and character-level embeddings	Twitter data about the 2015 Nepal earthquake and parts of India and the earthquake in central Italy in August 2016	Accuracy = 57%, F1 score = 19.1% (for the Nepal earth-quake data)
[59]	2018	Response/ Post-disaster response	Earthquake	LSTM	Data from Forum for Information Retrieval Evaluation 2016 (FIRE2016) and 2017 (FIRE2017) about the Nepal earthquake	Precision = 0.9234, F1 score = 91.59%
[65]	2020	Response/ Post-disaster response Case Study	Flood	Big data analytics	Satellite imagery data, census record, weather data, mobile phone calls reports, GIS data, AWS and SAR data	_____
[66]	2020	Response/ Post-disaster response Case Study	Lava Flow	CNN	Hawaii Island satellite imagery data	Precision = 81.0%, Recall = 93.0%, F1 score= 86.6%
[67]	2019	Response/ Post-disaster response Case Study	Hurricane Irma	Ridge regression, LR, Linear SVM LSTM, CNN, LDA	500 million Twitter data	Accuracy = 74.78% Precision = 0.7608, Recall = 0.7826, F1 score = 0.7514
[69]	2020	Preparedness, Response/ Disaster detection, post-disaster response Disaster Application	General	Merged Object Detection (MOD), YOLO V3, CNN	COCO and Google open image data	Accuracy = 87%
[61]	2018	Response/ Post-disaster response	General	CNN, RNN, NB	Twitter data from the CrisisLexT26 dataset	For CNN: F1 score = 89.5% (for non-natural disasters)
[62]	2020	Response/ Post-disaster response	Hurricane	ML and DL algorithms including DT, KNN, multinomial NB, LR, RF, SVM, CNN, RNN-LSTM, RNN-GRU, RNN-Attn-LSTM, BERT	Twitter data about the seven major hurricanes in the USA between 2012 and 2018	For BERT: Accuracy = 88%
[63]	2019	Response/ Post-disaster response	Hurricanes Harvey and Irma, General	Bi-LSTM and CNN	Twitter data about Hurricanes Harvey and Irma. Four different datasets from CrisisNLP and 15 different disasters data from CrisisLex	Accuracy = 93.7%, F1 score = 87.2%
[64]	2019	Response/ Post-disaster response	Flood	ResPSNet	High resolution satellite image data from the 2017 Hurricane Harvey flood in Houston, Texas	Accuracy = 94.97%, F1 score = 91.28%
[72]	2017	Response/ Post-disaster response Disaster application	Earthquake	Deep Belief Network	Big and heterogenous data including GPS data, transportation and road network data, Japan earthquake data	Accuracy = 87.8%

References

1. Centre for Research on the Epidemiology of Disasters (CRED); United Nations Office for Disaster Risk Reduction (UNDRR). Global trends and Perspectives Executive Summary. 2021. Available online: https://www.undrr.org/publication/2020-non-COVID-year-disasters (accessed on 4 October 2021).
2. Altay, N.; Green, W.G. OR/MS research in disaster operations management. *Eur. J. Oper. Res.* **2006**, *175*, 475–493. [CrossRef]

3.   Sun, W.; Bocchini, P.; Davison, B.D. Applications of Artificial Intelligence for Disaster Management. *Nat. Haz.* **2020**, *103*, 2631–2689. [CrossRef]

4.   Drakaki, M.; Tzionas, P. Investigating the impact of site management on distress in refugee sites using Fuzzy Cognitive Maps. *Int. J. Disaster Risk Reduct.* **2021**, *60*, 102282. [CrossRef]

5.   Drakaki, M.; Gören, H.G.; Tzionas, P. An intelligent multi-agent based decision support system for refugee settlement siting. *Int. J. Disaster. Risk Reduct.* **2018**, *31*, 576–588. [CrossRef]

6.   United Nations Office for Disaster Risk Reduction (UNDRR). *UNISDR Terminology on Disaster Risk Reduction*; UNISDR: Geneva, Switzerland, 2009. Available online: https://www.unisdr.org/files/7817_UNISDRTerminologyEnglish.pdf (accessed on 4 October 2021).

7.   EM-DAT—The International Disasters Database. Guidelines. EM-DAT—Data Entry—Field Description/Definition. Available online: https://www.emdat.be/guidelines (accessed on 4 October 2021).

8.   Van Wassenhove, L.N. Blackett memorial lecture humanitarian aid logistics: Supply chain management in high gear. *J. Oper. Res. Soc.* **2006**, *57*, 475–489. [CrossRef]

9.   Arinta, R.R.; Andi, E.W.R. Natural disaster application on big data and machine learning: A review. In Proceedings of the 2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 20–21 November 2019; Volume 6, pp. 249–254. [CrossRef]

10.   Yu, M.; Yang, C.; Li, Y. Big data in natural disaster management: A review. *Geosciences* **2018**, *8*, 165. [CrossRef]

11.   Schmidhuber, J. Deep Learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]

12.   Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]

13.   Presa-Reyes, M.; Chen, S.C. Assessing Building Damage by Learning the Deep Feature Correspondence of before and after Aerial Images. In Proceedings of the 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Shenzhen, China, 6–8 August 2020. [CrossRef]

14.   Akshya, J.; Priyadarsini, P.L.K. A hybrid machine learning approach for classifying aerial images of flood-hit areas. In Proceedings of the 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 21–23 February 2019. [CrossRef]

15.   Fan, C.; Wu, F.; Mostafavi, A. A Hybrid Machine Learning Pipeline for Automated Mapping of Events and Locations from Social Media in Disasters. *IEEE Access* **2020**, *8*, 10478–10490. [CrossRef]

16.   Ben-Hur, A.; Horn, D.; Siegelmann, H.T.; Vapnik, V. A support vector clustering method. In Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain, 3–7 September 2000. [CrossRef]

17.   Hofmann, T.; Schölkopf, B.; Smola, A.J. Kernel methods in machine learning. *Ann. Statist.* **2008**, *36*, 1171–1220. [CrossRef]

18.   O'Connor, J.; Eberle, C.; Cotti, D.; Hagenlocher, M.; Hassel, J.; Janzen, S.; Narvaez, L.; Newsom, A.; Ortiz-Vargas, A.; Schuetze, S.; et al. Interconnected Disaster Risks. *UNU-EHS* **2021**, *64*. Available online: https://reliefweb.int/report/world/interconnected-disaster-risks-20202021 (accessed on 4 October 2021).

19.   Dwarakanath, L.; Kamsin, A.; Rasheed, R.A.; Anandhan, A.; Shuib, L. Automated Machine Learning Approaches for Emergency Response and Coordination via Social Media in the Aftermath of a Disaster: A Review. *IEEE Access* **2021**, *9*, 68917–68931. [CrossRef]

20.   Yuan, C.; Moayedi, H. Evaluation and comparison of the advanced metaheuristic and conventional machine learning methods for the prediction of landslide occurrence. *Eng. Comput.* **2020**, *36*, 1801–1811. [CrossRef]

21.   Sankaranarayanan, S.; Prabhakar, M.; Satish, S.; Jain, P.; Ramprasad, A.; Krishnan, A. Flood prediction based on weather parameters using deep learning. *J. Water Clim. Chang.* **2019**, *11*, 1766–1783. [CrossRef]

22.   Huang, Y.; Jin, L.; Zhao, H.S.; Huang, X.Y. Fuzzy neural network and LLE algorithm for forecasting precipitation in tropical cyclones: Comparisons with interpolation method by ECMWF and stepwise regression method. *Nat. Hazards* **2018**, *91*, 201–220. [CrossRef]

23.   Asim, K.M.; Martínez-Álvarez, F.; Basit, A.; Iqbal, T. Earthquake magnitude prediction in Hindukush region using machine learning techniques. *Nat. Hazards* **2017**, *85*, 471–486. [CrossRef]

24.   Amin, M.S.; Ahn, H. Earthquake disaster avoidance learning system using deep learning. *Cogn. Syst. Res.* **2021**, *66*, 221–235. [CrossRef]

25.   Prasad, P.; Loveson, V.J.; Das, B.; Kotha, M. Novel ensemble machine learning models in flood susceptibility mapping. *Geocarto Int.* **2021**, *26*, 1892209. [CrossRef]

26.   Nsengiyumva, J.B.; Valentino, R. Predicting landslide susceptibility and risks using GIS-based machine learning simulations, case of upper Nyabarongo catchment. *Geomat. Nat. Hazards Risk* **2020**, *11*, 1250–1277. [CrossRef]

27.   Shirzadi, A.; Bui, D.T.; Pham, B.T.; Solaimani, K.; Chapi, K.; Kavian, A.; Shahabi, H.; Revhaug, I. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* **2017**, *76*, 60. [CrossRef]

28.   Sriram, L.M.K.; Ulak, M.B.; Ozguven, E.E.; Arghandeh, R. Multi-Network Vulnerability Causal Model for Infrastructure Co-Resilience. *IEEE Access* **2019**, *7*, 35344–35358. [CrossRef]

29.   Wahab, A.M.; Ludin, A.N.M. Flood vulnerability assessment using artificial neural networks in Muar Region, Johor Malaysia. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *169*, 012056. [CrossRef]

30.   Mutlu, B.; Nefeslioglu, H.A.; Sezer, E.A.; Ali, A.M.; Gokceoglu, C. An experimental research on the use of recurrent neural networks in landslide susceptibility mapping. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 578. [CrossRef]

31. Pham, B.T.; Tien Bui, D.; Prakash, I.; Dholakia, M.B. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *Catena* **2017**, *149*, 52–63. [CrossRef]

32. Gupta, T.; Roy, S. A Hybrid Model based on Fused Features for Detection of Natural Disasters from Satellite Images. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [CrossRef]

33. Layek, A.K.; Poddar, S.; Mandal, S. Detection of flood images posted on online social media for disaster response. In Proceedings of the 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), Gangtok, India, 25–28 February 2019. [CrossRef]

34. Muhammad, K.; Ahmad, J.; Baik, S.W. Early fire detection using convolutional neural networks during surveillance for effective disaster management. *Neurocomputing* **2018**, *288*, 30–42. [CrossRef]

35. Lee, W.; Kim, S.; Lee, Y.-T.; Lee, H.-W.; Choi, M. Deep neural networks for wild fire detection with unmanned aerial vehicle. In Proceedings of the 2017 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 8–10 January 2017. [CrossRef]

36. Chin, T.-L.; Huang, C.-Y.; Shen, S.-H.; Tsai, Y.-C.; Hu, Y.H.; Wu, Y.-M. Learn to Detect: Improving the Accuracy of Earthquake Detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8867–8878. [CrossRef]

37. Li, Z.; Meier, M.A.; Hauksson, E.; Zhan, Z.; Andrews, J. Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning. *Geophys. Res. Lett.* **2018**, *45*, 4773–4779. [CrossRef]

38. Moon, S.H.; Kim, Y.H.; Lee, Y.H.; Moon, B.R. Application of machine learning to an early warning system for very short-term heavy rainfall. *J. Hydrol.* **2019**, *568*, 1042–1054. [CrossRef]

39. Gopal, L.S.; Prabha, R.; Pullarkatt, D.; Ramesh, M.V. Machine Learning based Classification of Online News Data for Disaster Management. In Proceedings of the 2020 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, USA, 29 October–1 November 2020. [CrossRef]

40. Domala, J.; Dogra, M.; Masrani, V.; Fernandes, D.; D'Souza, K.; Fernandes, D.; Carvalho, T. Automated Identification of Disaster News for Crisis Management using Machine Learning and Natural Language Processing. In Proceedings of the 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2–4 July 2020. [CrossRef]

41. Wang, T.; Tao, Y.; Chen, S.C.; Shyu, M.L. Multi-Task Multimodal Learning for Disaster Situation Assessment. In Proceedings of the 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Shenzhen, China, 6–8 August 2020. [CrossRef]

42. Resch, B.; Usländer, F.; Havas, C. Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartogr. Geogr. Inf. Sci.* **2018**, *45*, 362–376. [CrossRef]

43. Scawthorn, C.; Flores, P.; Blais, N.; Seligson, H.; Tate, E.; Chang, S.; Mifflin, E.; Thomas, W.; Murphy, J.; Jones, C.; et al. HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment. *Nat. Hazards Rev.* **2006**, *7*, 72–81. [CrossRef]

44. Yang, L.; Cervone, G. Analysis of remote sensing imagery for disaster assessment using deep learning: A case study of flooding event. *Soft Comput.* **2019**, *23*, 13393–13408. [CrossRef]

45. Nguyen, D.T.; Ofli, F.; Imran, M.; Mitra, P. Damage assessment from social media imagery data during disasters. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, Sydney, Australia, 31 July–3 August 2017. [CrossRef]

46. Rizk, Y.; Jomaa, H.; Awad, M.; Castillo, C. A computationally efficient multi-modal classification approach of disaster-related Twitter images. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC '19), Limassol, Cyprus, 8–12 April 2019. [CrossRef]

47. Dotel, S.; Shrestha, A.; Bhusal, A.; Pathak, R.; Shakya, A.; Panday, S. Disaster Assessment from Satellite Imagery by Analysing Topographical Features Using Deep Learning. In Proceedings of the IVSP '20: 2020 2nd International Conference on Image, Video and Signal Processing, Singapore, 20–22 March 2020. [CrossRef]

48. Li, X.; Zhang, H.; Caragea, D.; Imran, M. Localizing and quantifying damage in social media images. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Barcelona, Spain, 28–31 August 2018.

49. Lin, A.; Wu, H.; Liang, G.; Cardenas-Tristan, A.; Wu, X.; Zhao, C.; Li, D. A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster. *Int. J. Disaster Risk Reduct.* **2020**, *49*, 101682. [CrossRef]

50. O'Neal, A.; Rodgers, B.; Segler, J.; Murthy, D.; Lakuduva, N.; Johnson, M.; Stephens, K. Training an Emergency-Response Image Classifier on Signal Data. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018. [CrossRef]

51. Li, T.; Li, Z.; Zhao, W.; Li, X.; Zhu, X.; Pan, S.; Feng, C.; Zhao, Y.; Jia, L.; Li, J. Analysis of medical rescue strategies based on a rough set and genetic algorithm: A disaster classification perspective. *Int. J. Disaster Risk Reduct.* **2019**, *42*, 101325. [CrossRef]

52. Ehara, K.; Aljehani, M.; Yokemura, T.; Inoue, M. Individual status recognition system assisted by UAV in post-disaster. In Proceedings of the 2020 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 4–6 January 2020. [CrossRef]

53. Reynard, D.; Shirgaokar, M. Harnessing the power of machine learning: Can Twitter data be useful in guiding resource allocation decisions during a natural disaster? *Transp. Res. Part D Transp. Environ.* **2019**, *77*, 449–463. [CrossRef]

54. Chaudhuri, N.; Bose, I. Exploring the role of deep neural networks for post-disaster decision support. *Decis. Support Syst.* **2020**, *130*, 113234. [CrossRef]

55. Li, H.; Caragea, D.; Caragea, C.; Herndon, N. Disaster response aided by tweet classification with a domain adaptation approach. *J. Contingencies Crisis Manag.* **2018**, *26*, 16–27. [CrossRef]

56. Bejiga, M.B.; Zeggada, A.; Nouffidj, A.; Melgani, F. A convolutional neural network approach for assisting avalanche search and rescue operations with UAV imagery. *Remote Sens.* **2017**, *9*, 100. [CrossRef]

57. Robertson, B.W.; Johnson, M.; Murthy, D.; Smith, W.R.; Stephens, K.K. Using a combination of human insights and 'deep learning' for real-time disaster communication. *Prog. Disaster Sci.* **2019**, *2*, 100030. [CrossRef]

58. Huang, X.; Li, Z.; Wang, C.; Ning, H. Identifying disaster related social media for rapid response: A visual-textual fused CNN architecture. *Int. J. Digit. Earth* **2020**, *13*, 1017–1039. [CrossRef]

59. Kundu, S.; Srijith, P.K.; Desarkar, M.S. Classification of short-texts generated during disasters: A deep neural network based approach. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 28–31 August 2018. [CrossRef]

60. Basu, M.; Shandilya, A.; Khosla, P.; Ghosh, K.; Ghosh, S. Extracting Resource Needs and Availabilities from Microblogs for Aiding Post-Disaster Relief Operations. *IEEE Trans. Comput. Soc. Syst.* **2019**, *6*, 604–618. [CrossRef]

61. Neppalli, V.K.; Caragea, C.; Caragea, D. Deep Neural Networks versus Naïve Bayes Classifiers for Identifying Informative Tweets during Disasters. In Proceedings of the 15th ISCRAM Conference, Rochester, NY, USA, 20–23 May 2018.

62. Paul, U.; Ermakov, A.; Nekrasov, M.; Adarsh, V.; Belding, E. #Outage: Detecting Power and Communication Outages from Social Networks. In Proceedings of the Web Conference 2020, Taipei, Taiwan, 20–24 April 2020. [CrossRef]

63. Kabir, M.d.Y.; Madria, S. A Deep Learning Aproach for Tweet Classification and Rescue Scheduling for Effective Disaster Management. In Proceedings of the 27th ACM Sigspatial International Conference on Advances in Geographic Information Systems, Chicago, IL, USA, 2–5 November 2019. [CrossRef]

64. Peng, B.; Liu, X.; Meng, Z.; Huang, Q. Urban Flood Mapping with Residual Patch Similarity Learning. In Proceedings of the 3rd ACM Sigspatial International Workshop on AI for Geographic Knowledge Discovery, Chicago, IL, USA, 5 November 2019. [CrossRef]

65. Nagendra, N.P.; Narayanamurthy, G.; Moser, R. Management of humanitarian relief operations using satellite big data analytics: The case of Kerala floods. *Ann. Oper. Res.* **2020**, 1–26. [CrossRef]

66. Laverdiere, M.; Yang, L.; Tuttle, M.; Vaughan, C. Rapid Structure Detection in Support of Disaster Response: A Case Study of the 2018 Kilauea Volcano Eruption. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020. [CrossRef]

67. Sit, M.A.; Koylu, C.; Demir, I. 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. *Int. J. Digit. Earth.* **2019**, *12*, 1205–1229. [CrossRef]

68. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Li, Y.; Catani, F.; Pourghasemi, H.R. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Comput. Geosci.* **2018**, *112*, 23–37. [CrossRef]

69. Ardiansyah, M.F.; William, T.; Abdullaziz, O.I.; Wang, L.-C.; Tien, P.-L.; Yuang, M.C. EagleEYE: Aerial edge-enabled disaster relief response system. In Proceedings of the 2020 European Conference on Networks and Communications (EuCNC), Dubrovnik, Croatia, 15–18 June 2020. [CrossRef]

70. Zhang, D.Y.; Zhang, Y.; Li, Q.; Plummer, T.; Wang, D. CrowdLearn: A Crowd-AI hybrid system for deep learning-based damage assessment applications. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 7–10 July 2019. [CrossRef]

71. Alam, F.; Imran, M.; Ofli, F. 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. [CrossRef]

72. Song, X.; Shabasaki, R.; Yuan, N.J.; Xie, X.; Li, T.; Adachi, R. DeepMob: Learning Deep Knowledge of Human Emergency Behaviour and Mobility from Big and Heterogeneous Data. *ACM Transac. Inf. Sys.* **2017**, *35*, 1–19. [CrossRef]

73. Holzinger, A.; Dehmer, M.; Emmert-Streib, F.; Cucchiara, R.; Augenstein, I.; Del Ser, J.; Samek, W.; Jurisica, I.; Díaz-Rodríguez, N. Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Inf. Fusion* **2022**, *79*, 263–278. [CrossRef]