A Comprehensive Survey Of Methodologies In Social Media Analytics For Disaster Management

Harshit Wandhare
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai,India
harshit.wandhare@vit.edu.in

Komal Borale
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
komal.borale@vit.edu.in

Bhavesh Dhake
Department of Information Technology
Vidyalankar Institute of Technology
Mumbai, India
bhavesh.dhake1@vit.edu.in

Suraj Gadekar
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
suraj.gadekar@vit.edu.in

Amit Aylani
Department of Computer Engineering
Vidyalankar Institute of Technology
Mumbai, India
amit.aylani@vit.edu.in

Abstract- Natural disasters have increased in frequency and severity in the past few years, having a considerable negative impact on the economy, the environment, and society. The use of new advanced technologies and methodologies in disaster management has become essential to minimize the impact of such disasters. This survey paper aims to provide an overview of various methodologies used in social media for disaster management and analysis. The methodologies discussed in this paper include the CyberGIS framework, which enables integration of the social media data with other geospatial data for disaster management. Text mining in social media analytics is another methodology that allows for the analysis of large volumes of social media data to identify disaster-related information. The Quantum Geographic Information System (QGIS) tool is also an approach that allows for the analysis and visualization of geospatial data. Also included are the You Only Look Once (YOLO) and VGG-16 algorithms, which analyze photos of the disasters to determine what kind of disaster they are and what information is pertinent. The use of methods like decision tables and applied analytic hierarchy processes (AHP) in disaster management is also covered in this paper. The visual sentiment analysis of disaster images from social media using the Convolutional Neural Network (CNN) model is also explored. Finally, the use of Natural Language Processing (NLP) on Twitter data is discussed for disaster-related tweet identification. In our survey we have identified and analyzed a total of seven papers

Index Terms - ***Quantum Geographic Information System, You Only Look Once, CyberGIS, Convolutional Neural Network, Natural Language Processing.***

I. INTRODUCTION

Natural or man-made disasters can cause a great deal of damage to infrastructure, property, and human lives. The timely and effective management of disasters is crucial in reducing their impact and facilitating recovery. In the recent years, social media platforms have emerged as a valuable resource for disaster management, enabling real-time communication, information sharing, and analysis. Social media platforms like Twitter, Facebook, and Instagram have been utilized extensively during disasters to disseminate information, coordinate rescue operations, and aid in recovery efforts.

However, the huge and vast amount of data created on social media during a disaster can make it difficult to extract essential data and insights. Therefore, there is a need for advanced methodologies in social media analytics to process and analyze the huge and vast amount of data generated during disasters effectively. The use of such methodologies can help in identifying critical patterns and trends in real-time, enabling effective decision-making and response management. Our survey has identified seven research papers which are relevant to the topic of social media analytics for disaster management. We have thoroughly analyzed and discussed these papers.

In this review, we examine the various approaches used in social media analytics for disaster management, such as the CyberGIS framework, Text Mining and social media analytics, Quantum Geographic Information System (QGIS), VGG-16 and You Only Look Once (YOLO) algorithms, Visual Sentiment Analysis of disaster images in social media using CNN model, and twitter data NLP. We will talk about the approaches, advantages, and disadvantages before drawing a conclusion and making recommendations for future study in the area.

The following is how the paper is set up: A thorough literature overview of the most recent developments in this field of inquiry is provided in Section 2, Section 3 deals with the results and discussion of each methodology which is being used in different papers. In this part, approaches, advantages, and disadvantages of each methodology will be addressed. Section 4 provides a conclusion and future scope of Social Media Analytics (SMA) in disaster management. The survey's results will be outlined in this part, along with ideas for additional studies in this area. Overall, the paper aims to provide insights into the various methodologies used in social media analytics for disaster management and their potential applications in real-world scenarios.

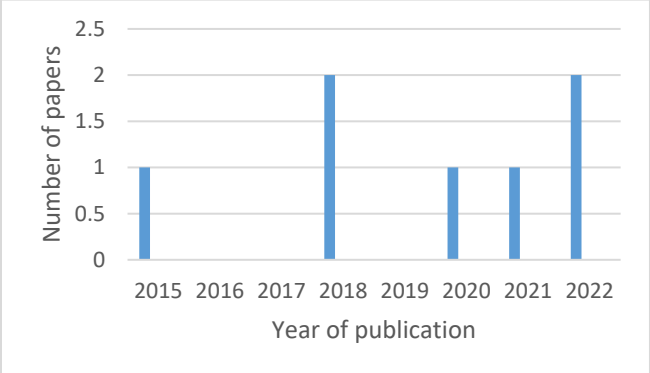

Fig. 1. Distribution of shortlisted articles by publication year.

II. Literature Review

Social media has seen a rise in popularity as a tool for disaster management in recent years. During disasters, social media platforms like Twitter, Facebook, and Instagram have been heavily used to spread information, organize rescue activities, and support recovery efforts. Nevertheless, it might be difficult to gain useful, meaningful information and insights due to the enormous volume of data that is been generated on social media during disasters. To efficiently process and analyze the enormous volume of data generated during disasters, new social media analytics approaches are needed.

The objective of this literature research which is done for this survey article is to identify the various approaches utilized in social media analytics for disaster management. In conducting this survey, the following research papers were used:

Artificial Intelligence's use in social media initial findings of a systematic literature review on big data analytics for disaster management [1] The potential of artificial intelligence in big data analytics from social media for disaster management is highlighted in this research. The study examines how machine learning algorithms can be used to spot patterns and trends in social media data during emergencies.

Disaster Mapper: It is a CyberGIS platform that utilizes data from social media to manage disasters. The CyberGIS system for disaster management presented in this paper makes use of social media data. The framework enables the integration of social media data with traditional disasters [2] By creating diverse hazard maps to support various operations, activities, and decision-making processes, GIS plays a vital role in disaster management.

Earthquake impact analysis based on text mining and Social Media Analytics: The method for analyzing the effects of earthquakes proposed in this paper uses text mining and social media analytics. To extract relevant details from social media data, the suggested method makes use of techniques for natural language processing. [3] Using clawers as the data source, the earthquake-related microblogs from the Sina microblog are gathered. Subsequently, data cleaning and preprocessing are performed, including irrelevant data filtering and word segmentation. The impact of the earthquake is then assessed using a number of NLP-based analytics, including public opinion trend, sentiment, word frequency, and keyword-based text classification.

Spatial Analysis of GIS Data using Social Media Data for Disaster Management: The use of geographic information system (GIS) data for disaster management using social media data is covered in this paper. The report emphasizes how social media and GIS data may be combined to enhance disaster response coordination. [4] QGIS 2.14 Essen is open-source software of GIS, which provides various interesting functionality to make it possible to visualize and analyze the spatial data of Twitter users and get useful results to support scientific research.

Automatic analysis of social media images to identify disaster type and infer appropriate emergency response in order to determine the sort of crisis and choose the best emergency response, this research provides a technique for automatically analyzing social media photographs. [5] Deep learning techniques are used in the proposed method to analyze photographs and categorize them into different sorts of disasters.

Visual Sentiment Analysis from Disaster Images in social media: This study suggests a technique using sentiment analysis of social media disaster photographs. Convolutional neural networks (CNNs) are used in the suggested method to analyze the sentiment of disaster photographs. [6] The procedure begins with social media sites being crawled for photographs linked to disasters, and then sentiment tags and categories are chosen to be connected to the disaster-related images in the crowd-sourcing study.

Identification of Disaster-related tweets made on twitter using Natural Language Processing: This paper proposes a method for identifying disaster-related tweets using natural language processing (NLP) techniques. [7] The proposed method uses a combination of NLP techniques and machine learning algorithms to identify relevant information from disaster-related tweets.

Overall, these research papers highlight the potential of social media analytics for disaster management. The various methodologies discussed in these papers can help in identifying critical patterns and trends in real-time, enabling effective decision-making and response management. The respective strengths and limitations of each methodology is discussed in detail in Section 3 of this survey paper.

Previous Work	Objective	Technology used
[2]	To track disaster events, to produce maps, and to perform spatial and statistical analysis for disaster management	CyberGIS framework, Apache Hive, Hadoop, and Mahout
[3]	Early earthquake impact analysis using text mining	NLP, Wordcloud, pandas, matplotlib, snowNLP (a Python library for Chinese NLP tasks)
[4]	Generating heat map using twitter data, calculate the user density within a particular region to get useful information	Quantum Geographic Information System 2.14 (QGIS)
[5]	To automate the emergency response decision-making process using image classification	VGG-16 and You Only Look Once (YOLO) Analytic hierarchy processes (AHP)
[6]	Visual sentiment analysis using images and videos of disaster-prone area	Convolutional Neural Network (CNN)
[7]	Classify the tweets as "Disaster Related" or "Not Disaster Related"	Text-blob python library for text processing and NLTK for Natural Language Processing (NLP)

Fig. 2. Objectives and technology used in different papers.

III. Results and Discussion

A complete and thorough analysis of all the identified papers is given below highlighting the main purpose and objective of the paper, its advantages and limitations and the technology used along with the future scope proposed by the authors:

Nunavath and Goodwin (2018) [1] conducted a literature review to analyze the use of artificial intelligence (AI) in

social media big data analytics for disaster management. They found that AI techniques, including machine learning and natural language processing, have been used to identify patterns and trends in real-time, enabling effective decision-making and response management. The advantages of using AI include faster and more accurate analysis of data, identification of vital information, and improved accuracy over time. However, limitations include dependence on data quality, bias, and ethical concerns. The paper highlighted the use of various AI techniques and social media platforms and suggested future research focus on developing more sophisticated AI techniques to improve response efforts.

The paper by Huang et al. (2015) [2] presents DisasterMapper, a CyberGIS framework that uses social media data for disaster management. The framework integrates social media data with the geographic information system (GIS) data to enable real-time situational awareness and decision-making during disasters. DisasterMapper utilizes a variety of technologies such as web-based mapping tools, machine learning, and natural language processing to analyze and visualize social media data. The advantages of using DisasterMapper for disaster management include its ability to process and analyze huge volumes of data in real-time, enabling quick decision-making and response management. The framework also enables the identification of critical patterns and trends that may have been missed otherwise, facilitating effective disaster management. Additionally, DisasterMapper's web-based mapping tools provide an intuitive and user-friendly interface for disaster response teams. However, there are few limitations to using DisasterMapper. The accuracy of the framework depends heavily on the quality of social media data used, which may not be reliable and correct during disasters. Additionally, the framework may be susceptible to bias and may not always provide the most appropriate response in complex situations. Finally, the implementation of the framework requires significant technical expertise and resources, which may not be available in all disaster management scenarios. In terms of the technology used, DisasterMapper integrates a variety of tools such as web-based mapping, machine learning, and natural language processing. The framework utilizes the CyberGIS platform to enable real-time data integration and analysis. The authors suggest that future research should focus on developing more sophisticated machine learning algorithms and improving the accuracy of natural language processing techniques to enhance the effectiveness of DisasterMapper in disaster management.

Zheng et al. (2022) [3] published a study on text mining and social media analytics to analyze the earthquake impact. The authors sought to glean relevant data from social media platforms to assess the effects of earthquakes and the effectiveness of disaster response initiatives. The authors analyzed Twitter's data using a combination of text mining and machine learning approaches. Data was gathered during two significant earthquakes that struck China in 2017 and 2018. The findings in their investigation demonstrated that the usage of social media data can offer useful and relevant information about how earthquakes affect infrastructure, the number of fatalities they cause, and the efficiency of emergency response efforts. The advantages of using text mining and social media analytics for earthquake impact analysis include the ability to analyze large volumes of data quickly and accurately in real-time, which can inform emergency response efforts. Additionally, these techniques can identify significant patterns and trends in social media data that might offer insightful information about the tragedy which has affected nearby areas. Utilizing social media data to study earthquake impacts has several drawbacks, though. For instance, the quality of data and the prejudices of social media users can both have an impact on how accurate the study is. As not all impacted people may have access to or utilize social media, data may not give a complete picture of the disaster's effects. The technology used in this study included natural language processing and machine learning algorithms, and social media analytics tools. The authors used Python and the Natural Language Toolkit (NLTK) for text mining and employed Support Vector Machines (SVM) to classify social media messages. They also used the Google Maps API to extract location information from the data. Overall, the study by Zheng et al. (2022) highlights the potential use of text mining and social media analytics for earthquake impact analysis, but also underscores the need to address the limitations and biases inherent in data.

The goal of study by Gaur et al. (2018) [4] was to investigate how social media platforms data may be used for spatial analysis using Geographic Information System (GIS) data. The authors analysis of the social media data from Twitter during the Chennai floods in India was their main emphasis, and they employed GIS techniques to locate and display the spatial distribution of tweets about the crisis. The findings demonstrated that spatial analysis of data using GIS can offer important insights into the geographic distribution of disaster-related information. For instance, the scientists were able to pinpoint the flood-prone regions based on the volume of tweets sent from such areas. The spatial analysis also helped in identifying the most relevant hashtags used during the disaster, which can be useful in developing more effective communication strategies during such events. The advantages of using GIS-based spatial analysis in disaster management include the ability to quickly identify and prioritize response efforts in affected areas, providing real-time updates to stakeholders, and facilitate effective communication and collaboration among response teams. Decision-making during a crisis can be facilitated by GIS-based systems' ability to recognize and visualize patterns and trends in social media data. Nevertheless, there are a lot of restrictions when it comes to applying spatial analysis based on GIS for disaster management. The quality and quantity of the data available, for instance, has a significant impact on how accurate the results were. In addition, the complexity and knowledge required for GIS-based systems can be a deterrent to their wider adoption. The authors analyzed data from Twitter using GIS software. Additionally, they used spatial analysis methods including hotspot analysis and kernel density estimation to visualize the spatial distribution of tweets on the incident. The study emphasizes the potential of spatial analysis using GIS for disaster management and recommends that future research concentrate on creating more sophisticated tools and methodologies to enhance the precision and efficacy of these systems.

The paper by Asif et al. (2021) [5] aimed to develop an automatic image analysis system that could recognize the kind of crisis and determine the best emergency response based on images from social media. The scientists trained their model using machine learning techniques using an album of photos gathered from several social media networks. The study's findings demonstrated that the created system is capable of correctly identifying the type of crisis

shown in the photographs and of recommending suitable emergency measures. The system's accuracy in determining the type of calamity and in recommending the best course of action was 89.25% and 87.50%, respectively. The advantages of the developed system include faster and more accurate identification of the type of disaster and the ability to suggest appropriate emergency responses in real-time. This can help emergency responders to prioritize their efforts and provide timely assistance to those affected by the disaster. Additionally, the system can analyze a large volume of images in a short amount of time, which can be useful during disasters when time is of the essence. However, there are few limitations to the system. For instance, the accuracy of the system is heavily dependent on the quality of the images used to train it. Additionally, the system may not perform well in situations where there are multiple types of disasters occurring simultaneously or when the images are of inferior quality. In terms of technology used, the authors used machine learning techniques such as convolutional neural networks (CNNs) and transfer learning to develop the system. They also used various software tools such as Python, Keras, and TensorFlow for data preprocessing, model development, and evaluation. The study highlights the potential of using AI-based image analysis systems for disaster management. However, further research is needed to improve the accuracy and robustness of these systems and to address the ethical and privacy concerns associated with their use.

The objective of Hassan et al.'s paper from 2022 [6] was to analyze the visual sentiment of disaster photographs found on social media. The authors proposed an innovative method for categorizing photographs according to their tone and assessing the magnitude of destruction caused by the catastrophe. Using crowdsourcing approaches, they gathered an album of pictures from various social media sites and annotated them according to how they felt. In order to categorize the photographs into different sentiment categories, such as positive, negative, and neutral, they next utilized a convolutional neural network (CNN). The advantages of using visual sentiment analysis in disaster management include the ability to quickly analyze large volumes of visual data and identify critical information that can aid in decision-making. This technique can also provide insights into public perception and sentiment regarding a disaster, which can help improve the effectiveness of response efforts. However, there are also few limitations to using visual sentiment analysis in disaster management. For example, the accuracy of the analysis depends on the quality of the images used, and the sentiment analysis may not always reflect the true emotions of the people affected by the disaster. Additionally, the annotation of images can be subjective and prone to bias. In terms of technology used, the authors employed CNN for visual sentiment analysis, which is a popular technique for image classification. They also utilized crowdsourcing techniques for image annotation, which is an effective way to collect substantial amounts of labeled data. The authors suggested that to improve the analysis of data during disasters, future research should mainly focus on increasing the precision of sentiment analysis and including more data sources, such as text data.

The aim of Goswami and Raychaudhuri's research from 2020 [7] was to use natural language processing (NLP) methods to find tweets about disasters. A dataset of tweets about three disasters was compiled by the authors - Hurricane Harvey, Hurricane Irma, and Hurricane Maria - and applied NLP techniques to classify them as disaster-related or not. The authors used various NLP techniques such as stemming, stop word removal, and part-of-speech tagging to preprocess the tweets before classifying them. The results showed that the proposed approach achieved high accuracy in identifying disaster-related tweets, they achieved an F1 score of 0.91 for Hurricane Harvey, 0.86 for Hurricane Irma, and 0.83 for Hurricane Maria. The authors also compared their approach with existing approaches and found that their approach outperformed them in terms of accuracy. The advantages of using NLP techniques for disaster-related tweet identification include faster and more accurate identification of disaster-related information, which can help in timely response and disaster management. NLP techniques can also help in identifying specific types of disasters and their severity, which can further aid in effective disaster management. However, there are also some limitations to using NLP techniques in disaster-related tweet identification. For instance, the accuracy of NLP-based systems depends heavily on the quality of data used to train them. Additionally, these systems can be vulnerable to bias and may not always provide the most appropriate response in complex situations. In terms of technology used, the authors used various NLP techniques such as stemming, stop word removal, and part-of-speech tagging. They also used machine learning algorithms such as Naive Bayes, Support Vector Machines, and Random Forests for classification. The authors suggested that further research should focus on developing more sophisticated NLP techniques that can manage complex data and improve the accuracy of disaster-related tweet identification.

IV. Conclusion and Future Work

In conclusion, this survey paper has provided a comprehensive overview of the role of Social Media Analytics (SMA) in disaster management. We can conclude by saying all these different methodologies should be implemented together to create a good disaster management system. The analysis of the selected papers shows that AI techniques such as machine learning, deep learning, natural language processing, and sentiment analysis are widely used to analyze social media data during disasters. The advantages of using AI techniques in social media big data analytics which is used disaster management include faster and more accurate analysis of large volumes of data, identification of essential information that may have been missed otherwise, and the ability to prioritize response efforts based on the severity of the disaster. However, there are also some limitations such as data quality issues, bias, and ethical concerns that need to be taken care of.

In the areas of crisis management and social media analytics, there is still considerable work to be done. Future study could focus on a number of areas, including:

- Developing more sophisticated AI techniques that can manage complex data and improve the accuracy of responders in disaster management.
- Exploring the use of emerging technologies such as blockchain and edge computing to improve data security and enhance the speed of disaster response.

- Analyzing how AI-based disaster management systems perform in relation to various sorts of disasters, including floods, earthquakes, and wildfires.
- Examining the potential for using AI techniques to predict and prevent disasters by analyzing historical data and identifying patterns that may indicate an impending disaster.

Overall, the use of social media analytics in disaster management is a rapidly evolving field, and there is much potential for future innovation and improvement.

REFERENCES

[1] V. Nunavath and M. Goodwin, "The Role of Artificial Intelligence in Social Media Big data Analytics for Disaster Management -Initial Results of a Systematic Literature Review," 2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), Sendai, Japan, 2018, pp. 1-4, doi: 10.1109/ICT-DM.2018.8636388.

[2] Huang, Qunying & Cervone, Guido & Jing, Duangyang & Chang, Chaoyi. (2015). DisasterMapper: A CyberGIS framework for disaster management using social media data. 1-6. 10.1145/2835185.2835189.

[3] Zheng, Zhe & Shi, Hong-Zheng & Zhou, Yu-Cheng & Lu, Xinzheng & Lin, Jia-Rui. (2022). Earthquake Impact Analysis Based on Text Mining and Social Media Analytics.

[4] Neelima Gaur, Farheen Fauziya , Ajit Kumar Jain , " Spatial Analysis of GIS Data for Social Media Data Disaster Management" , International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) , Volume 7, Issue 2, March - April 2018 , pp. 095-099 , ISSN 2278-6856.

[5] Asif, A., Khatoon, S., Hasan, M.M. et al. Automatic analysis of social media images to identify disaster type and infer appropriate emergency response. J Big Data 8, 83 (2021).

[6] Hassan, S.Z.; Ahmad, K.; Hicks, S.; Halvorsen, P.; Al-Fuqaha, A.; Conci, N.; Riegler, M. Visual Sentiment Analysis from Disaster Images in Social Media. Sensors 2022, 22, 3628.

[7] Goswami, Shriya and Raychaudhuri, Debaditya, Identification of Disaster-Related Tweets Using Natural Language Processing: International Conference on Recent Trends in Artificial Intelligence, IOT, Smart Cities & Applications (ICAISC-2020) (May 26, 2020).