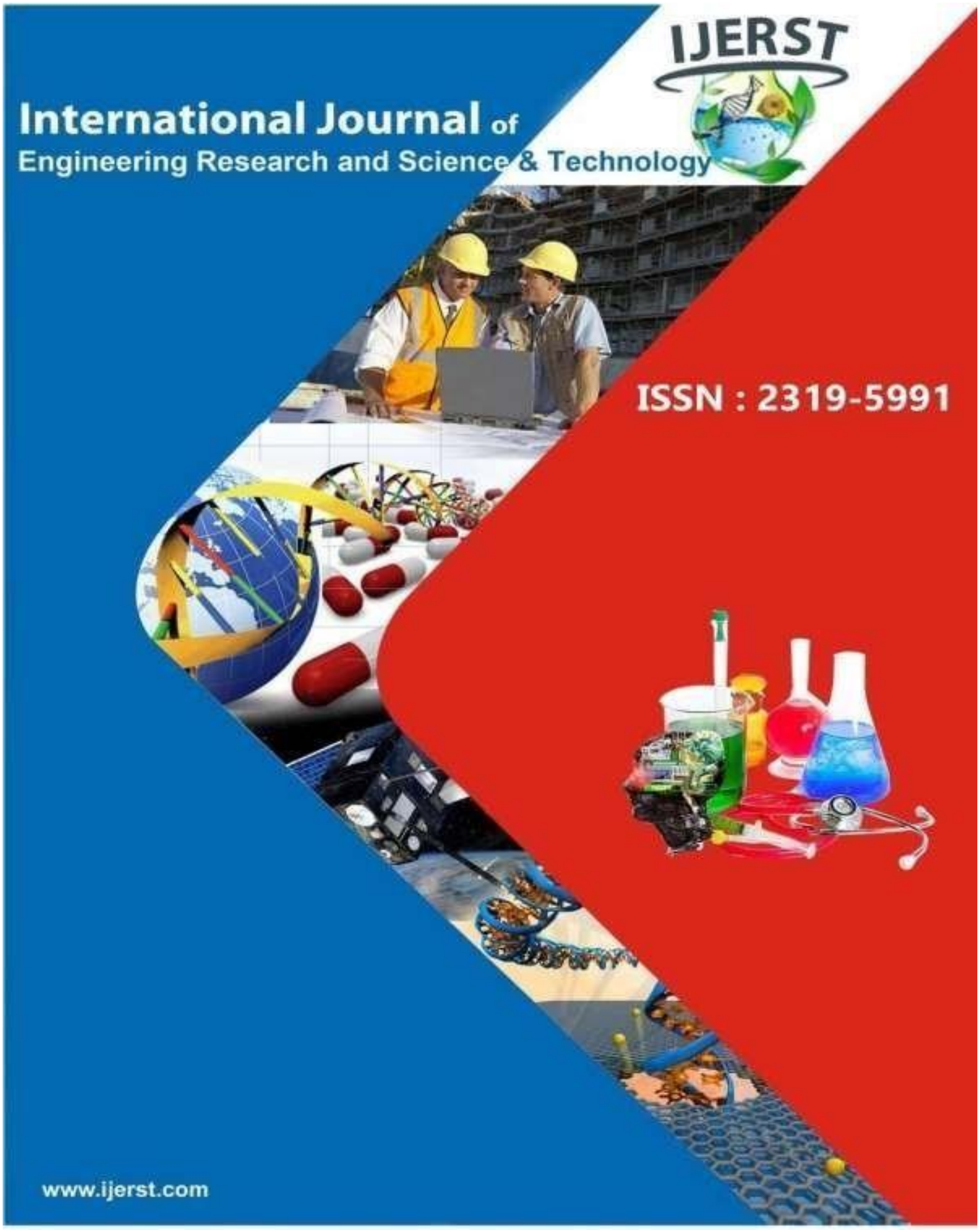

Email: editor@ijerst.com or editor.ijerst@gmail.com

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

A DEEP ATTENTIVE MULTIMODAL LEARNING APPROACH FOR DISASTER IDENTIFICATION FROM SOCIAL MEDIA POSTS

Mrs C. Chaitanya[1],Shravya[2],Maghana[3],Manasa[4],Laxmi Prasana[5],Ravali[6]

[1]Assistant professor, Department of CSE, Princeton College of engineering and technology for women Narapally vijayapuri colony ghatkesar mandal, Pin code-500088

[2,3,4,5,6]UG Students,Department of CSE, Princeton College of engineering and technology for women Narapally vijayapuri colony ghatkesar mandal, Pin code-500088

Article Info

Received: 19-08-2022 *Revised: 13 -09-2022* *Accepted: 12-10-2022*

ABSTRACT

Twitter and other micro blogging platforms have become indispensable for disseminating critical information, especially in the aftermath of both natural and man-made disasters. In order to relay critical information like deaths, facility damages, and urgent needs of impacted people, people often upload multimediacomponents using photographs and/or videos. Humanitarian organisations may greatly benefit from this data in order to plan an adequate and timely response. The need for an automated method to sort through social media for actionable and non- actionable disaster-related material arises from the difficulty of extracting useful information from massive amounts of communications. Previous work mostly examined textual methods and/or used standard frequent neural networks (RNNs) or convolutional neural networks (CNNs), which might lead to efficiency degradation in the case of lengthy input sequences, although numerous studies have shown the effectiveness of integrating message and picture components for disaster recognition. Using a combination of visual and linguistic information, this article presents a multi- modal catastrophe detection system that can identify tweets by affixing salient word characteristics with aesthetic purposes. A retrained convolutional neural network (e.g., ResNet50) is used for visual attribute extraction, while a bidirectional long- lasting memory (BiLSTM) coupled with an attention device is employed for textual attribute extraction. A function combination technique and the soft max classifier are then used to accumulate visual and textual functions. The results demonstrate that the proposed multi-modal system outperforms the current baselines, which include both multi-modal and uni-modal models, by around

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

1% and 7% of performance improvement, respectively.

Keywords: CNN, LSTM, RNN, Twitter, social media.

I INTRODUCTION

systems of social media sites may play an important role in disseminating a large amount of vital information during disasters like earthquakes, floods, and storms [1]. When people utilise these social media platforms, they often establish connections across hierarchies, such as those between individuals, between businesses and the federal government, between neighbourhoods, and between the government and its citizens [2]. Tweets from catastrophe victims often detail the events of the disaster, including casualties, infrastructuredamage, and the location of affected areas. Additionally, impacted individuals are pleading for quick assistance via the publication of images, tweets, and videos. Charitable organisations may greatly benefit the harmed people if they evaluate these social media posts and derive practical conclusions in real-time [3]. However, manually analysing and extracting actionable insights from a large volume of crisis-related tweets is an incredibly challenging and time-consuming task.

The nonprofit sector of the IT industry has attempted to address theaforementioned challenge by creating automated systems that can sift through social media posts pertaining to crisesand extract relevant information. [4] For instance, researchers have developedclassifiers to categorise humanitarianfeatures (e.g., kinds of damage), article informativeness, and event categories (e.g., floods, storms) [5, 6]. Current employment opportunities are severely constrained in two ways, notwithstanding these advancements. Research on disaster response via social media has, up until recently, mostly concentrated on textual or image content assessment independently. On the other hand, new studies show that combining textual and visual information frequently yields better insights into an event and leads to more correct reasonings than just reading the text [7]. The second issue is that lengthier phrases may not be well-represented by the CNN or RNN versions used for message attribute portrayal in the few multi-modal feature- based tasks that have been published so far [7], [8].

A dependable computational approach for identifying disaster-related information via the synergistic integration of visual and textual

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

modalities is our goal in this assignment. Our primary focus is on extractingpicture functions using a pre-trained visual model, namely ResNet50. To solve the long-range dependency problem with conventional RNN and CNN architecture, we further extract textual characteristics by combining a focus mechanism with the BiLSTM network. After that, we use Deep degree fusion to combine the two sets offeatures, and then we classify the provided tweet using the soft max layer.

[9] In order to determine the types of damage (such as fire, flood, and framework damage) from a group of images and tweets, we conducted extensive experiments using a multi modal damage dataset. Several baselines that do not use attention devices or multi modal functionalities are compared to our designs. The main takeaways from these trials are as follows: (i) using multi modal functions yields much better results than using uni modal features,and (ii) a focus system integrated into anRNN design may significantly outperform a design without such a device. [1]

Our main sources of income from this position are:

- We propose a multi-modal approach that classifies damage-related articles using both visual and textual information, using ResNet50 and BiLSTM recurrent neural networks with an attention mechanism.
- In this study, we evaluate the proposed model in comparison to a set of preexisting multimodal and unimodal (i.e., picture, message) categorization algorithms.
- We conducted an in-depth evaluation of the suggested model using a benchmark dataset and proved that providing emphasis improves system efficiency.
- We use quantitative and qualitative assessment to learn more about the types of mistakes, which will help us improve the model in the future.

II RELATED STUDY

Another CNN-based method for categorising tweets about disasters was proposed by Aipe et al. [22], however their focus was on multilabel classification rather than simple binary classification. Similarly, Yu et al. [3] classified the tweets related to several hurricanes into many categories using CNN, logistic regression, and support vector machines. They improved upon

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

SVM and LR with their CNN-based approach. To better capture relationships between word tokens, we examine BiLSTMs with focus systems as an alternative to CNN-based approaches.

Using the Storm Sandy and Boston Marathon fight datasets, Li et al. [4] investigated the possibility of domain name adaptation for assessing catastrophic tweets using the uninformed Bayes classifier. Graf et al. [5] aimed to make the classifier applicable to many types of disasters by focusing on cross-domain categorization. Emotional, nostalgic, and etymological functions were extracted from the damage-related tweets and used by a cross-domain classifier. Message mining and summary approaches have really been the attention of others. One example is the work of Rudra et al. [6], which summarises tweets after classifying them into several scenario courses. In their recommendation of an ESAAWTM system, Cameron et al. [7] sought to alert charitable organisations about disaster situations via the detection of beneficial damage-related Twitter posts. We just concentrate on a multi-class category problem on tweets connected to disasters, in contrast to existing systems that heavily focused on text mining and summarization.

Photos shared on social media platforms may be classified into three types of disasters: severe, medium, and no harm at all, according to a deep convolutional neural network (CNN) architecture developed by Nguyen et al. [9]. A pretrained convolutional neural network (CNN) based structure that can detect catastrophe photographs released on online platforms was also suggested by Alam et al. [10]. In order to identify the fire occurrence, Daly and Thom [31] used pretrained classifiers to filter out flicker photographs. Finally, a method to classify whether the picture shows a fire event or not was devised by Lagerstrom et al. [11]. Chen et al. [12] looked at the photos and texts and used visual attributes and socially relevant contextual attributes (e.g., time of uploading, variety of comments, retweets) to determine catastrophe details, in contrast to these works that developed binary classifiers for categorising catastrophes. Human and environmental harm were the primary foci of the damage discovery investigation by Mouzannar et al. [7]. They used a CNN style for textual characteristics and used the Inception

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

pre-trained design for visual feature extraction.

Similarly, Rizk et al. [35] proposed a multimodal approach to categorise Twitter data according to structural and natural damage types. The tweets were also categorised by Ferda et al. [8] using a multimodal method. The first category was for insightful tasks, such as useful vs. non-informative, and the second was for humanitarian jobs, such as affected persons, rescue volunteering or contribution initiatives, infrastructure and energy damage. To extract the aesthetic and linguistic functions, they used a CNN-based technique. Using the CrisisMMD [14] dataset, Gautam et al. compared unimodal and multimodal methods. Their strategy for integrating the image-tweet sets was the late combination [15] technique. When comparing the works that employ multimodal data to those that use uni-modal data, all of them found that the latter significantly improved performance.

III PROPOSED SYSTEM

Our work's main contributions include the following: we provide a multimodal design that manipulates both visual and textual information to detect damage-related postings; this design makes use of ResNet50 and BiLSTM permanent semantic network with interest device. We compare the suggested model's performance to thatof many preexisting unimodal (i.e., image, text) and multimodal categorization approaches. We conducted an in-depth study to show how adding focus may improve system efficiency and then tested the proposed model on a benchmark dataset. To get a better grasp of the mistake types thatprovide guidance for future model improvements, we combine quantitative and qualitative study.

IV METHODOLOGY

Everything about the Deep Multi-level Attentive network (DMLANet) that has been suggested is laid out here. We provide a high-level description of the proposed network in this section. Our next offering is the visual attention module, which employs both spatial and channel attention to provide noteworthy bi-attentive visual characteristics. Lastly, it delves into ajoint attended multimodal learning process that leverages semantic attention to learn a combined representation for textual and visual features. This process

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

involves measuring the semantic closeness of text and visual features, and then using a self-attention mechanism to extract the crucial multimodal featuresfor sentiment classification. To create a bi-attentive visual feature map, the visual attention module uses channel- based attention to improve information- rich channels and spatial or region-based attention to hone in on emotional areas based on attended channels. Semantic attention is used in joint attended multimodal learning to quantify the emotive terms associated with the bi-attentive visual characteristics. After that, we feed the attended word features and the bi-attentive visual features into the self-attention block, and it will automatically choose the most relevant multimodal aspects to highlight.

The majority of the prior work on multimodal sentiment analysis has focused on fusion-based approaches, which integrate data from many sources and input them into a classifier [7]. To get the multimodal sentiment label, some have combined the sentiment predictions from several sources [8]. This process is called late fusion. The fact that these works don't depict the intricate relationship between the modalities is their biggest flaw. The modalities in the network's intermediate layers may be combined using intermediate fusion, which has beenused in certain experiments [9]. On the other hand, it calls for meticulous planning and could not work if part of the multimodal material is missing.

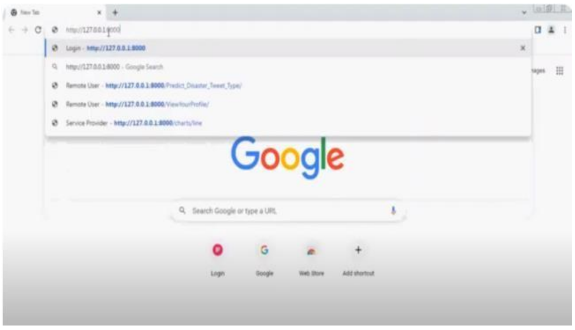

Fig.1. Web page design.

Fig.2. Admin page.

Fig.3. Login details.

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

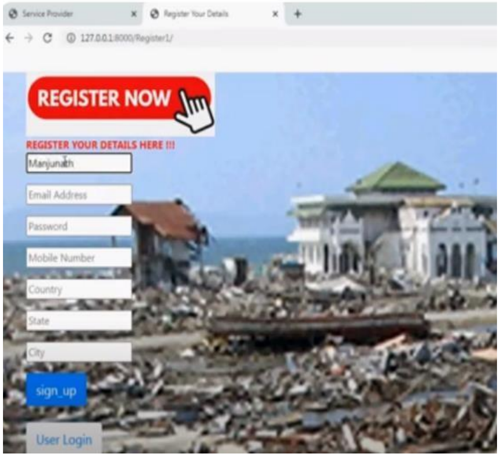

Fig.4. Registration page.

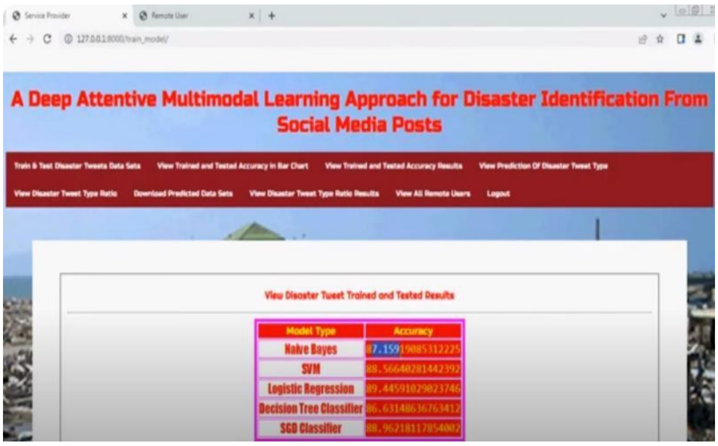

Fig.7. Accuracy of output.

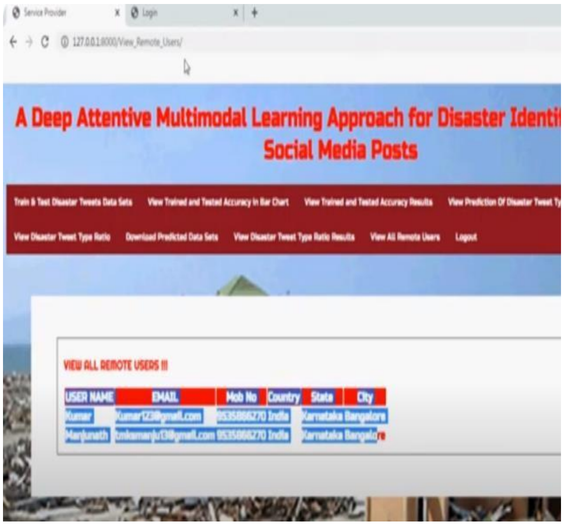

Fig.5. users' details.

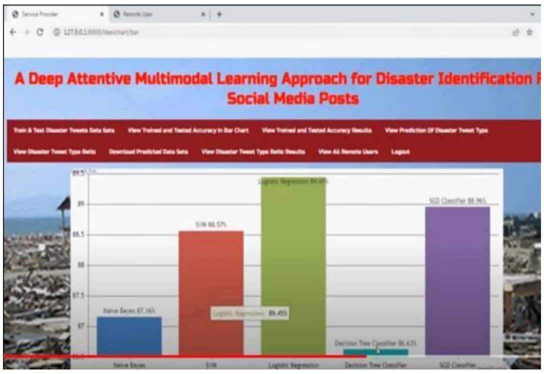

Fig.8. output results.

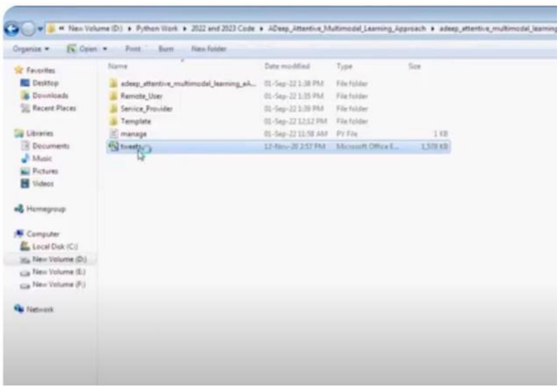

Fig.6. Upload data set.

V. CONCLUSION

In order to classify Twitter posts pertaining to damage, we have presented a multimodal approach that effectively leverages picture and message data. To extract the aesthetic features, we use the pre-trained ResNet version. To extractthe tweet functionalities, we use the attention device with a BiLSTM architecture. In order to combine the best parts of both approaches, the early fusion method is used. Also, for the baseline analysis, this study used a plethora of visual and textual

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

approaches, such as VGG19 and Inception, as well as BiLSTM, CNN, BiSTM+CNN, and BiLSTM+ Attention. With a weighted F1-score of 93:21%,the suggested model outperforms both the baseline unimodal (photo/text) and multimodal designs. In addition, the results of the comparison demonstrated that the suggested method outperforms the current state-of-the-art versions by a margin of between one percent and seven percent. It follows that the results validated the efficacy of the proposed method for catastrophe content recognition using multimodal features. Another finding from the error analysis was how difficult it is to distinguish between damaged and non-damaged materials when using a single analysis method. Simultaneously, study of inherent efficiency revealed that adding an interest system improves overall performance.

Despite outperforming unimodal techniques, there remains room for improvement in the suggested method. We want to investigate multitask learning strategies and other multimodal combination techniques to disaster identification in the near future. In addition, we want to use state-of-the-art aesthetic (Vision transformer), textual (BERT, XLM-R), and multimodal (VL-BERT, Visual BERT) transformer designs to better capture thecombination of visual and textualfeatures.

VI. REFERANCES

[1] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur, “Advances in social media research: Past, present and future,” Information Systems Frontiers, vol. 20, no. 3, pp. 531–558, 2018.

[2] J. Kim and M. Hastak, “Social network analysis: Characteristics of online social networks after a disaster,” International Journal of Information Management, vol. 38, no. 1, pp. 86–96, 2018.

[3] J. Son, H. K. Lee, S. Jin, and J. Lee, “Content features of tweets for effective communication during disasters: Amedia synchronicity theory perspective,” International Journal of Information Management, vol. 45, pp. 56–68, 2019.

[4] A. Elbanna, D. Bunker, L. Levine, and A. Sleigh, “Emergency management in the changing world of social media: Framing the research agenda with the stakeholders through engaged scholarship,” International Journal of

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

Information Management, vol. 47, pp. 112–120, 2019.

[5] R. Dubey, A. Gunasekaran, S. J. Childe, D. Roubaud, S. F. Wamba, M. Giannakis, and C. Foropon, "Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain," International Journal of Production Economics, vol. 210, pp. 120–136, 2019.

[6] S. Akter and S. F. Wamba, "Big data and disaster management: a systematic review and agenda for future research," Annals of Operations Research, vol. 283, no. 1, pp. 939–959, 2019.

[7] H. Mouzannar, Y. Rizk, and M. Awad, "Damage identification in social media posts using multimodal deep learning." in ISCRAM, 2018.

[8] F. Ofli, F. Alam, and M. Imran, "Analysis of social media data using multimodal deep learning for disaster response," arXiv preprint arXiv:2004.11838, 2020.

[9] F. Ofli, P. Meier, M. Imran, C. Castillo, D. Tuia, N. Rey, J. Briant, P. Millet, F. Reinhard, M. Parkan et al., "Combining human computing and machine learning to make sense of big (aerial) data for disaster response," Big data, vol. 4, no. 1, pp. 47–59, 2016.

[10] P. Jain, B. Schoen-Phelan, and R. Ross, "Automatic flood detection in sentinei-2 images using deep convolutional neural networks," in Proceedings of the 35th Annual ACM Symposium on Applied Computing, 2020, pp. 617–623.

[11] A. Kumar, J. P. Singh, Y. K. Dwivedi, and N. P. Rana, "A deep multimodal neural network for informative twitter content classification during emergencies," Annals of Operations Research, pp. 1–32, 2020.

[12] T. Ghosh Mondal, M. R. Jahanshahi, R.-T. Wu, and Z. Y. Wu, "Deep learning-based multi-class damage detection for autonomous post-disaster reconnaissance," Structural Control and Health Monitoring, vol. 27, no. 4, p. e2507, 2020.

[13] M. Imran, F. Ofli, D. Caragea, and A. Torralba, "Using ai and social media multimodal content for disaster response and management: Opportunities, challenges, and future directions," 2020.

[14] L. Dwarakanath, A. Kamsin, R. A. Rasheed, A. Anandhan, and L. Shuib, "Automated machine learning approaches for emergency response and coordination via social media in the

ISSN 2319-5991 www.ijerst.com
Vol. 18, Issue.4, Dec 2022

aftermath of a disaster: A review," IEEE Access, vol. 9, pp. 68 917–68 931, 2021.

[15] C. Castillo, M. Imran, P. Meier, J. Lucas, J. Srivastava, H. Leson, F. Ofli, and P. Mitra, "Together we stand—supporting decision in crisis response: Artificial intelligence for digital response and micromappers," by OCHA and partners. Istanbul: Tudor Rose, World Humanitarian Summit, pp. 93 95, 2016.