Suspicious Local Event Detection in Social Media and Remote Sensing: Towards a Geosocial Dataset Construction

Conference Paper · October 2020

DOI: 10.1109/ATSIP49331.2020.9231798

3 authors:

Marwen Bouabid
Ecole Nationale des Sciences de l'Informatique

Mohamed Farah
Manouba University

Imed Riadh Farah
Manouba University

Suspicious Local Event Detection in Social Media and Remote Sensing: Towards a Geosocial Dataset Construction

*

Bouabid Marwen*, Farah Mohamed† and Farah Imed Riadh‡
RIADI, LR99ES26 (SIIVT Team), National School of Computer Sciences
La Manouba, Tunisia
Email: *marwen.bouabid@ensi-uma.tn, †mohamed.farah@riadi.rnu.tn, ‡riadh.farah@ensi.rnu.tn

***Abstract*—Remote sensing is a powerful technology for earth observation. However, the spatial, spectral, and temporal resolution of the imagery are imposing various limits. Lately, with the rise of the internet and smart mobile devices, social media with location-based information has been rapidly emerging. These circumstances led to the prevailing of new scenarios where fine-grained details of social bookmarking websites are enhanced with the wide coverage of satellites. Social media and satellites are both valuable sources of data. An event-driven data, designating either normal common events or unusual suspicious ones that may threaten human lives or damage the infrastructure. In this paper, we provide an insight into the present state of knowledge to better address the task of local suspicious event detection and linking social media with satellite imagery. Also, to track suspicious local events, we treated the detection problem as a retrospective problem by training different classifiers on the crisisLexT26 dataset. Furthermore, we introduced how to use the available geo-locations in the dataset to construct a geo-social dataset by linking it with remote sensing and retrieving satellite imagery before and after the event occurrence.**

***Index Terms*—Suspicious Local event detection, Social Media, Remote Sensing, Crime, Disaster, Terror Attacks, Crisis.**

I. Introduction

Local suspicious events include out of the ordinary events associated with a specific location. Detecting those sorts of events as soon as they occur may help to save the lives that are held hostages in such situations. Responding fast and efficiently in such circumstances is decisive, as every second matter when it comes to bleeding and fatal injuries. Taking as an example, the bus accident that took place on December 1, 2019, in Tunisia that caused the injury and the death of many people. Such incidents point out the importance of efficient coordination among authorities and health organizations, which could benefit from having detailed information about the event scene.

People, in social media, tend to massively share and broadcast every event, especially the ones relevant to crises and emergencies. Such information could be diffused in online social networks ahead of any other medium [1]. Meanwhile, the extended geographic coverage and high spatial and multi-spectral resolutions provided by satellite imagery have proven their efficiency in detecting, analyzing, monitoring emergencies, and planing the recovery process [2]. One of the most popular and longest-running satellite programs is LandSat, launched to collect Earth resource data. The first Landsat mission was launched in 1972. The Landsat program continues to this date with the most recent Landsat imagery Continuity mission satellite (now Landsat 8) launched in February 2013. However, geographical coverage, resolutions, and temporal frequency are the main challenges to tackle a particular remote sensing application. Despite the importance of satellite imagery in event detection, it only gives a bird's eye view of the event.

Recently, new researches that are linking the two sources have proven to be efficient and effective in event detection. They led to the emergence of geosocial sensors. To this purpose, we present our study addressing the task of local suspicious event detection in social media and linking with remote sensing. We started by providing a little background about social media and remote sensing. Then we provided a synthesis and presented our approach. In the synthesis, we discussed the challenges, we defined the target task and explained the possible categories. In the approach section, we proposed an architecture addressing the task of local suspicious event detection as a retrospective classification problem of tweets and clarifying the enrichment of CrisisLexT26 dataset [3] with satellite imagery. This dataset was specifically chosen due to its suitability to present various types of events. We trained a set of classifiers, namely: Support vector machine SVM, Random Forrest, Logistic regression, and gradient boost. Then, we presented the results. Finally, we linked the dataset with remote sensing to get a geosocial dataset.

II. Related Work

Event detection history in social media could be traced back to a project called Topic Detection and Tracking (TDT) [4]. The task evolved later on into event detection. The event detection problem is either addressed as an offline

Retrospective Event Detection problem RED or as a real-time New Events Detection problem NED illustrated in fig 1. RED works on pre-collected datasets. NED, on the other hand, consists of detecting bursty keywords from online streams or via real-time crawling of social media content based on specific keywords. In both approaches, detecting the occurring of an event is insufficient. Various information about the event should be extracted. Takahashi et al. examine the use of social media, Twitter in particular, in emergencies considering several factors, such as time and location of the use [5]. After all, each event is associated with a specific location and time.

This type of event is identified as local events, and it has been gaining popularity as an active research field over the past years [6] [7] [8] [9]. Quezada et al. [10] and Watanabe et al. [11] extract geo-located events from the social bookmarking website with the main focus on geo-locating tweets and events. Sakaki et al. [9] address earthquake detection by training a classifier to estimate whether a given tweet is relevant to an earthquake or not. Li et al. [12] use a self-adaptive crawler to discover crime and disaster events. In [6] [7] [13] divers generic local event detection approaches have been proposed. Krumm et al. [7] worked on discretizing the time into equal size bins. Afterward, they extracted local events by comparing tweets number within each bin in different days. Cresci et al. [14] have proposed a crisis mapping system that relies on Support Vector Machines (SVMs), with a novel geo-parsing technique to determine the location as an attempt to address the limitations of the conventional approaches.

On the other side, satellite Imagery has been widely used in disaster monitoring and impact assessment through change detection techniques. Target detection [15], regional planning, warfare [16] are examples of applications relying on remote sensing. The latter is used to perceive observations, specifications, or for monitoring earthquakes, volcanic activity, landslides, flooding, and wildfire [2]. Several interesting approaches have been proposed to investigate the impact on the environment [17]. Different approaches are focusing on the acquisition, pre-processing, and prediction of crises [18]. Joyce et al. [2] surveyed different data types, data acquisition, and processing techniques for mapping and monitoring such events. Jaiswal et al. [19] use satellite data for the identification and mapping of fire. Eguchi et al. [20]analyze the building's destruction in urban areas due to earthquakes. Almost all the previously mentioned works are relying on satellite imagery before and after a crisis to face the problem of low temporal frequency of remote sensing data. Satellite imagery availability is increasing rapidly due to attempts made by national agencies (NASA, ESA) to publish such data openly or due to efforts of commercial startups such as PlanetLabs to provide real-time imagery.

Combining both approaches has become the trend, such a combination helps to address the limits of both approaches, Kashif et al. have introduced Jord, a system that collects social media data about natural disasters and links it to remotely sensed data. Besides, the author demonstrated the importance of using local language queries for the accuracy of the results [21]. Furthermore, Bischke et al. introduced a scalable system for satellite imagery contextual enrichment with multimedia content from social media [22]. Figure 2 illustrates the enrichment of satellite imagery with photos from social media in MediaEval 2017. Pittaras et al. introduced GeoSensor, a novel, open-source system that enriches change detection over satellite images with event content from social media [23].

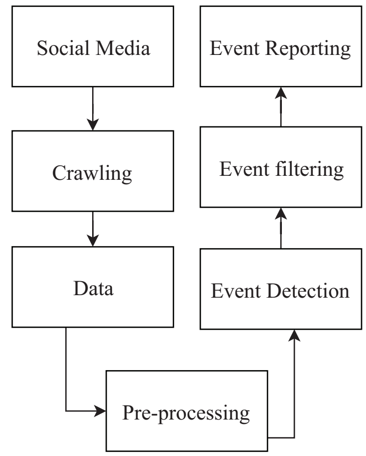

Fig. 1. Realtime event detection. The figure illustrates the steps of the online event detection approach. Starting By crawling, processing the data, detecting and filtering the events and finally reporting it.

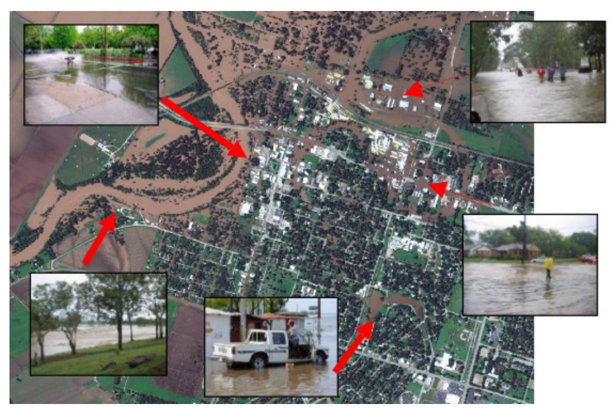

Fig. 2. Enrichment of satellite imagery with photos from social media https://tinyurl.com/utel92d

III. Synthesis and Proposed Approach

In this section, we offered a synthesis and a discussion about local suspicious event detection in social media and linking with satellite imagery. To this purpose, we provided the main challenges facing researchers in this domain. Afterward, we contributed our definition of the local suspicious events detection task, with the main focus on addressing terror attacks. Then we trained a set of classifiers for local suspicious events detection, namely Support Vector Machine(SVM), Logistic Regression, Random forest, and gradient boost. We also displayed their results in terms of precision, recall, f1 measure, and assess their performance based on accuracy.

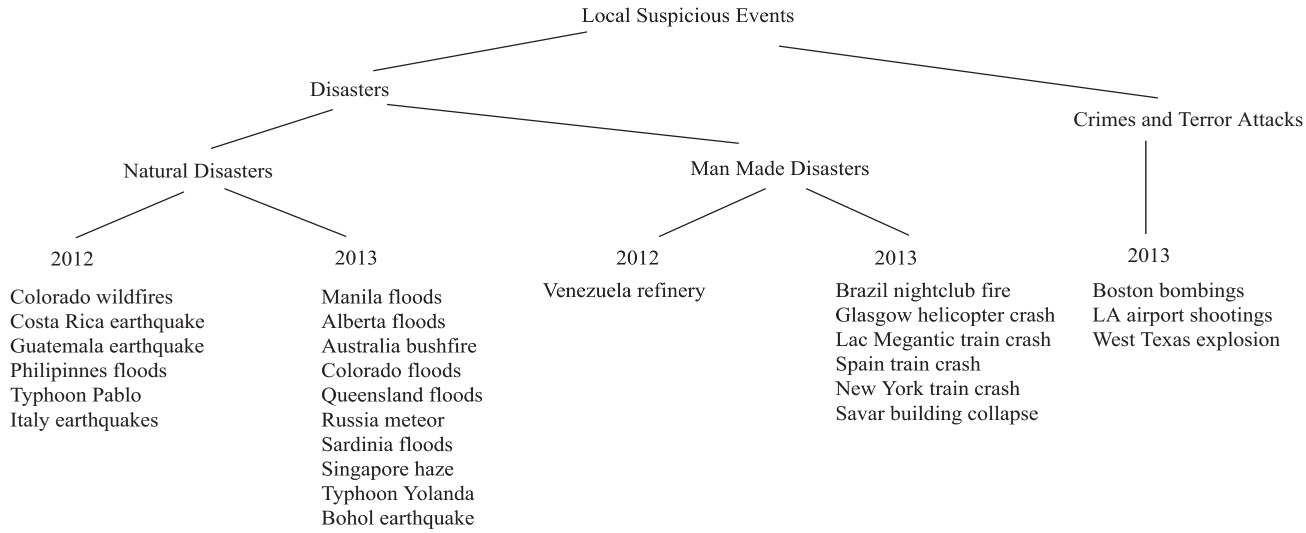

Fig. 3. Dataset structure: this graph is representing the structure of the CrisisLexT26 dataset.

A. Synthesis : Research Challenges

Although social media and satellite imagery have been proven efficiency in crisis event detection and monitoring, there are several problems associated with their use. In this section, we identified some open research challenges associated with the task of local suspicious event detection. Starting with new event detection, processing social media content to obtain relevant information involves the challenges of collecting, handling and analyzing stream content. The relevance and authenticity of content are among the main challenges associated with the task. Otherwise, if the target is offline event detection, the lack of datasets and the quality of the datasets is another challenge. The available datasets are often not large enough in terms of events numbers and types. Most of the existent works [24] [25] [26] rely on self-collected datasets. Satellite imagery usage is imposing some challenges as well. The low temporal frequency, the inability to provide a clear view of the ground due to: vegetation, lighting conditions, clouds, and even processing errors are part of those challenges. The domain lacks large-scale annotated datasets imagery for suspicious local events. Linking satellite imagery and social media has its challenges. The location estimation of the events is a serious challenge that needs addressing. Taking Twitter as an example, the location of the events could be either determined by the geolocation data available in the tweets or by analyzing the tweet text and tags searching for location hints. Also, a tweet may talk about a specific location and could be posted from another. Another challenge when detecting events from satellite imagery is associated with the detection of various types of events. Therefore to perform this task, researchers have the choice between addressing one type of event at a time or using change detection techniques and relying on social media to determine the type of change that occurred. Through our review of the literature, we have noticed that the majority of the works are only addressing specific types of crises. Therefore, this paper is trying to unfold the mystery about local suspicious event detection by focusing on the whole task instead of just dealing with particular types of events.

B. Synthesis :The Task of Local Suspicious Events Detection

The task of Local suspicious event detection could be defined as the detection of any event out of the ordinary associated with a specific location and time. An event that causes a threat to human lives or leads to infrastructure damage. Such events could have the following categories: Protest or riot, Disaster, Crimes, and terror attacks.

1) Protest or Riot : The recent events of the Arab Spring that started in Tunisia and spread across North Africa and the Middle East back in 2011 have proven the unprecedented influence of social bookmarking websites to transmit information within those countries, and from them to the outside world. Stefanidis et al. [27] discuss a framework that collects ambient geospatial data and resulting composite capabilities. These pieces of information are analyzed, to support situational awareness, as it relates to human activities. A Protest is a social event that may serve for political purposes. Therefore, it may cause damage to infrastructure, and it could labor malicious purposes. The authors in [28] present a technique portable across language boundaries and internal borders based on a dataset of 2.2 million geocoded Ukrainian tweets, used to empirically examine the theory against the observed 2014 Euromaidan protests in Ukraine. Won et al. [29] developed a novel visual model that identifies protesters, reports their activities by visual properties, and assess the level of perceived violence in an image.

2) Disasters events: Disaster event detection is another category of the task of local suspicious event detection. A disaster could be classified into two sub-categories, which are natural disasters and human-made disasters. Social media generally and twitter specifically has emerged as an influential source of information and fast communication in situations where the unavoidable can occur [30]. In such times, no matter how hard news agencies tried to get the scope of what's happening in the ground. They just could not provide information promptly or at all; Here comes the role of social media in breaking and disseminating such events [1]. For the detection of earthquakes, a method was proposed, where a graph-based clustering technique has been utilized to target geo-located communities on Twitter [31]. In many situations, twitter was used as a social sensor to capture data about a natural disaster in real-time like fire, floods, earthquakes, and typhoons [32].

3) Crimes and Terror Attacks: Crimes are another category of suspicious local events, detecting this sort of event is a challenging task, yet a few systems have been proposed to tackle this kind of events. As an example, we could mention the TEDAS system [12], which focuses on detecting crimes and disaster events. The system collects tweets related to crime events using topic-related keywords predefined by the authors. Those initial keywords are then expanded to look for semantically linked terms. Since not all the tweets containing those keywords are about crimes and disaster, a classifier is trained using twitter features, such as the use of hash-tag and some predefined pattern feature in the text content. For crime and disaster events detection in twitter text, various methods have been proposed specifically for the task [33]. Several approaches were developed to deal specifically with the task, such as the GEOBURST system, that enables effective and real-time local event detection from geo-tagged tweet streams [13].

Terror attacks are an important category of local suspicious event detection, which could be defined as the act of violence against a group of civilians, to create terror and to achieve hidden purposes(political, religious, Racism...). Some works have proven that terrorism and hate speech are both very common and closely related activities [34]. Therefore, the detection of hate speech could be beneficial for terrorism detection. For example, we cite the Christchurch mosque shootings terror attack that happened in New Zealand on March 15, 2019. The attacker live-streamed the shooting via Facebook Live service. The detection of such events is difficult due to the lack of data and the works addressing the issue. Also, the majority of malicious behaviors are already hidden within the dark web [35]. The dark web is the hidden part of the internet iceberg that isn't indexed by search engines and requires the use of a specific browser.

C. Proposed Approach: Event classification and linking with remote sensing

1) Proposed Architecture: We addressed the local suspicious event detection problem as a RED classification problem. We proposed an architecture for event classification and linking with remote sensing shown in Figure 4. Support vector machine SVM, Logistic regression, Random forest, Gradient boost are known as effective algorithms for text classification. Therefore, we split the tweets into unigrams and bigrams (n-

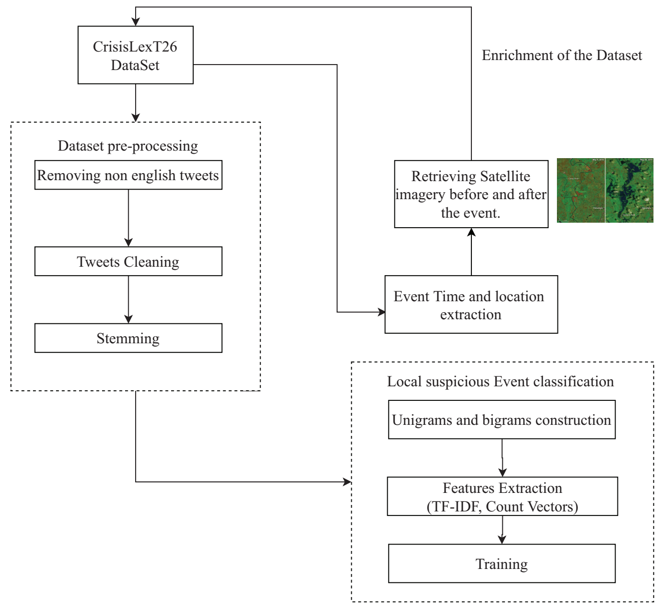

Fig. 4. The proposed architecture, showing the suspicious local event classification and linking with satellite imagery.

grams with n=1,2). Then, we constructed Count Vectors and TF-IDF Vectors. Afterward, we trained those algorithms for the task of local suspicious event classification.The training step 4 refers to the training of the previously mentioned classifiers. We fitted the training dataset on those classifiers, predicted the labels, and then we assessed each model. The target dataset was CrisisLexT26 [3].

2) Dataset: This dataset contains approximately 250k tweets posted during 26 suspicious local events in 2012 and 2013, with 2K-4K tweets in most events.28,000 tweets are labeled With the following attributes: Tweet ID, Tweet Text, Information Source, Information Type, Informativeness. We are mainly concerned with the Tweet Text and the Informativeness in the task of classification. The dataset contains different language tweets. Tweets contain retweets, URL links, usernames, and special characters(emojis). The tweets with "Not related" or "Not applicable" tags do not address suspicious local events and are hence labeled "Off-Topic". The tweets marked with "Related and Informative" and "Related but not Informative" are both labeled with the local suspicious events name. The dataset structure is shown in figure 3. It is also important to point out that the dataset contains no event related to protests or riots.

3) Dataset Pre-processing: Pre-processing the CrsisLexT26 dataset is an important step that grants the efficiency of the classifier. Therefore, we made use of the Natural Language Toolkit (NLTK). We used WordNet. We removed Non-English tweets, and we cleaned the dataset by removing stop words and punctuation, URLs, and usernames. We also performed Stemming.
Stemming is process for removing the morphological variation of words.
Stemming:
As a stemmer we used the SnowballStemmer, stemming

removes the prefix from the words such as removing the ”ing” from ”working”.

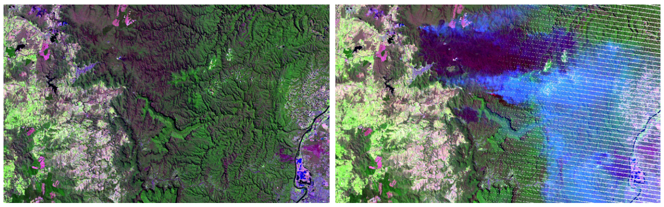

Fig. 5. A satellite image of 2013 Australia bushfire. The image on the left is the one before the event and the image on the right is the one after the event (Lansat Imagery)

4) Linking with Remote sensing: The final step in our work is to use Geo-location information of the events to link the tweets with the remotely sensed images. The CrisisLexT26 dataset groups the events based on the year, location, and type of event. We made use of the available information and retrieved the exact location and time of each event. Such information is considered critical for the linking process. Therefore, we tried to retrieve the exact coordinates and time from the google engine for efficient spatial and temporal localization. We also tried to make the period separating the pre-event and the post-event as small as possible. We retrieved two satellite images before and after each event. We used those images to surround the event. Which serves to monitor and access the damage. For this purpose, we relied on Google earth’s historical imagery, Landsat imagery. Figure 5 shows an example of satellite imagery of the 2013 Australia bushfire. The image on the left is the one before the event (Lansat 8 Image), and the image on the right is the one after the event (Lansat 7 Image).

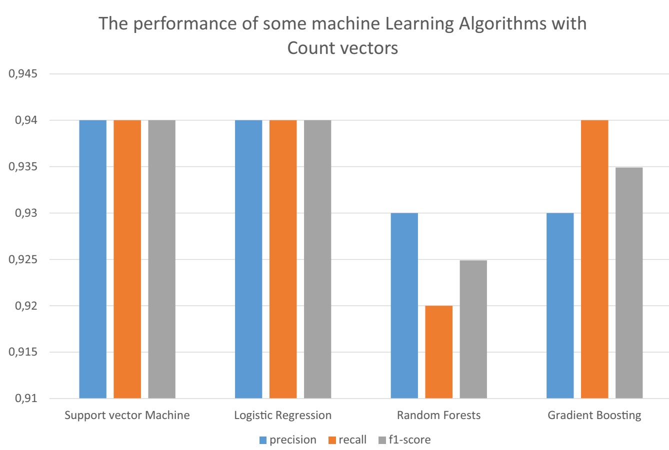

Fig. 6. This graph is representing some metrics for the different machine learning algorithms applied with count vectors.

5) Results: The results of the classification task are illustrated in terms of few metrics such as precision, recall, and f1-measure illustrated in Figure 6 for count vectors and Figure 7 for tf-idf vectors. Also, the accuracy of each algorithm is shown in table I. As it is shown the Logistic regression with

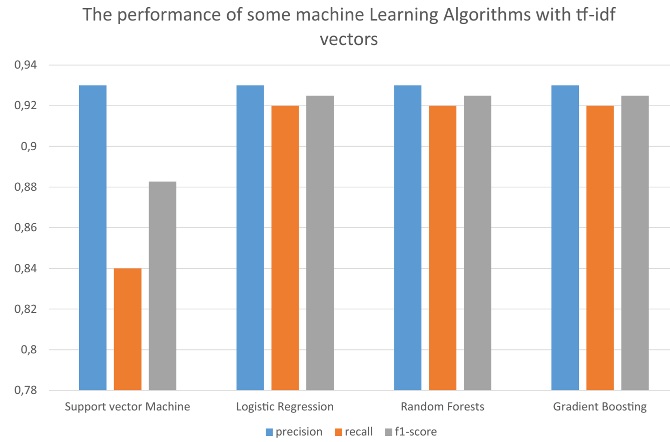

Fig. 7. This graph is representing some metrics for the different machine learning algorithms applied with tf-idf vectors.

[h]

TABLE I
PERFORMANCE COMPARISON OF CLASSIFIERS

Model	Validation Accuracy(%)
Support vector Machine SVM, Count Vectors	91.78
Support vector Machine SVM, TF-IDF Vectors	89,65
Logistic Regression, Count Vectors	92.04
Logistic Regression, TF-IDF Vectors	86.46
Random Forests, Count Vectors	90.18
Random Forests, TF-IDF Vectors	90.25
Gradient Boosting, Count Vectors	91.05
—Gradient Boosting, TF-IDF Vectors	90.32

count vectors outperforms the other models with an accuracy equal to 92.04 %, The second best result is scored by the Support vector machine SVM model with count vectors with an accuracy equal to 91.78 %. Also, we noticed that Logistic regression and Support Vector Machine SVM performs better with count vectors. In the opposite, random Forest and gradient boost perform better with TF-IDF vectors. with an accuracy respectively equal to 90.25% and 90.32 %.

IV. CONCLUSION

In this paper, we contributed a new definition of the task of local suspicious event detection. We also presented the possible categories of those events. We have chosen to approach the detection task as a retrospective classification problem. We applied machine learning on a pre-collected dataset named crisisLexT26 that contains approximately 250k tweets about various types of suspicious local events. We performed preprocessing on the dataset to ensure the efficiency of the classifier. The tweets are then split into unigrams and bigrams (n-grams with n=1, 2) and then converted into Count Vectors and TF-IDF Vectors. We evaluated the machine learning models based on precision, recall, f1-measure, and, more importantly, the accuracy. The results of the classification have shown that logistic regression with count vector outperforms the other models with an accuracy equal to 92.04 %. We also linked the social media data, available as a pre-collected twitter dataset, with remotely sensed satellite imagery based on the time and locations available in the dataset and constructed a geosocial dataset. In the future, we plan to approach the task of local suspicious event detection as an online detection problem NED. We will also try to deploy deep learning on this task.

References

[1] M. Cheong and V. Lee, "Twittering for earth: A study on the impact of microblogging activism on earth hour 2009 in australia," in *Asian Conference on Intelligent Information and Database Systems*. Springer, 2010, pp. 114–123.

[2] K. E. Joyce, S. E. Belliss, S. V. Samsonov, S. J. McNeill, and P. J. Glassey, "A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters," *Progress in physical geography*, vol. 33, no. 2, pp. 183–207, 2009.

[3] A. Olteanu, C. Castillo, F. Diaz, and S. Vieweg, "Crisislex: A lexicon for collecting and filtering microblogged communications in crises," in *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[4] G. Doddington, "The 1999 topic detection and tracking (tdt) task definition and evaluation plan," 1999.

[5] B. Takahashi, E. C. Tandoc Jr, and C. Carmichael, "Communicating on twitter during a disaster: An analysis of tweets during typhoon haiyan in the philippines," *Computers in Human Behavior*, vol. 50, pp. 392–398, 2015.

[6] H. Abdelhaq, C. Sengstock, and M. Gertz, "Eventweet: Online localized event detection from twitter," *Proceedings of the VLDB Endowment*, vol. 6, no. 12, pp. 1326–1329, 2013.

[7] J. Krumm and E. Horvitz, "Eyewitness: Identifying local events via space-time signals in twitter feeds," in *Proceedings of the 23rd sigspatial international conference on advances in geographic information systems*. ACM, 2015, p. 20.

[8] W. Feng, C. Zhang, W. Zhang, J. Han, J. Wang, C. Aggarwal, and J. Huang, "Streamcube: hierarchical spatio-temporal hashtag clustering for event exploration over the twitter stream," in *2015 IEEE 31st International Conference on Data Engineering*. IEEE, 2015, pp. 1561–1572.

[9] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 851–860.

[10] M. Quezada, V. Peña-Araya, and B. Poblete, "Location-aware model for news events in social media," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2015, pp. 935–938.

[11] K. Watanabe, M. Ochi, M. Okabe, and R. Onai, "Jasmine: a real-time local-event detection system based on geolocation information propagated to microblogs," in *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 2011, pp. 2541–2544.

[12] R. Li, K. H. Lei, R. Khadiwala, and K. C.-C. Chang, "Tedas: A twitter-based event detection and analysis system," in *2012 IEEE 28th International Conference on Data Engineering*. IEEE, 2012, pp. 1273–1276.

[13] C. Zhang, G. Zhou, Q. Yuan, H. Zhuang, Y. Zheng, L. Kaplan, S. Wang, and J. Han, "Geoburst: Real-time local event detection in geo-tagged tweet streams," in *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 2016, pp. 513–522.

[14] S. Cresci, A. Cimino, F. DellOrletta, and M. Tesconi, "Crisis mapping during natural disasters via text analysis of social media messages," in *International Conference on Web Information Systems Engineering*. Springer, 2015, pp. 250–258.

[15] Z. Li and L. Itti, "Saliency and gist features for target detection in satellite images," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2017–2029, 2010.

[16] J. B. Campbell and R. H. Wynne, *Introduction to remote sensing*. Guilford Press, 2011.

[17] D. Arvor, L. Durieux, S. Andrés, and M.-A. Laporte, "Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 82, pp. 125–137, 2013.

[18] T. W. Gillespie, J. Chu, E. Frankenberg, and D. Thomas, "Assessment and prediction of natural hazards from satellite imagery," *Progress in Physical Geography*, vol. 31, no. 5, pp. 459–470, 2007.

[19] R. K. Jaiswal, S. Mukherjee, K. D. Raju, and R. Saxena, "Forest fire risk zone mapping from satellite imagery and gis," *International Journal of Applied Earth Observation and Geoinformation*, vol. 4, no. 1, pp. 1–10, 2002.

[20] R. T. Eguchi, C. K. Huyck, B. J. Adams, B. Mansouri, B. Houshmand, and M. Shinozuka, "Resilient disaster response: using remote sensing technologies for post-earthquake damage detection," *Earthquake Engineering to Extreme Events (MCEER), Research Progress and Accomplishments*, vol. 2003, p. 2003, 2001.

[21] K. Ahmad, K. Pogorelov, M. Riegler, N. Conci, and P. Halvorsen, "Social media and satellites," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 2837–2875, 2019.

[22] B. Bischke, P. Bhardwaj, A. Gautam, P. Helber, D. Borth, and A. Dengel, "Detection of flooding events in social multimedia and satellite imagery using deep neural networks." in *MediaEval*, 2017.

[23] N. Pittaras, G. Papadakis, G. Stamoulis, G. Argyriou, E. K. Taniskidou, E. Thanos, G. Giannakopoulos, L. Tsekouras, and M. Koubarakis, "Geosensor: semantifying change and event detection over big data," in *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. ACM, 2019, pp. 2259–2266.

[24] H. Shekhar and S. Setty, "Disaster analysis through tweets," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2015, pp. 1719–1723.

[25] H. To, S. Agrawal, S. H. Kim, and C. Shahabi, "On identifying disaster-related tweets: Matching-based or learning-based?" in *2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*. IEEE, 2017, pp. 330–337.

[26] B. Truong, C. Caragea, A. Squicciarini, and A. H. Tapia, "Identifying valuable information from twitter during natural disasters," *Proceedings of the American Society for Information Science and Technology*, vol. 51, no. 1, pp. 1–4, 2014.

[27] A. Stefanidis, A. Crooks, and J. Radzikowski, "Harvesting ambient geospatial information from social media feeds," *GeoJournal*, vol. 78, no. 2, pp. 319–338, 2013.

[28] S. L. Wilson, "Detecting mass protest through social media," *The Journal of Social Media in Society*, vol. 6, no. 2, pp. 5–25, 2017.

[29] D. Won, Z. C. Steinert-Threlkeld, and J. Joo, "Protest activity detection and perceived violence estimation from social media images," in *Proceedings of the 25th ACM international conference on Multimedia*. ACM, 2017, pp. 786–794.

[30] J. Yin, S. Karimi, A. Lampert, M. Cameron, B. Robinson, and R. Power, "Using social media to enhance emergency situation awareness," in *Twenty-fourth international joint conference on artificial intelligence*, 2015.

[31] M. Bakillah, R.-Y. Li, and S. H. Liang, "Geo-located community detection in twitter with enhanced fast-greedy optimization of modularity: the case study of typhoon haiyan," *International Journal of Geographical Information Science*, vol. 29, no. 2, pp. 258–279, 2015.

[32] A. Crooks, A. Croitoru, A. Stefanidis, and J. Radzikowski, "# earthquake: Twitter as a distributed sensor system," *Transactions in GIS*, vol. 17, no. 1, pp. 124–147, 2013.

[33] C. Li, A. Sun, and A. Datta, "Twevent: segment-based event detection from tweets," in *Proceedings of the 21st ACM international conference on Information and knowledge management*. ACM, 2012, pp. 155–164.

[34] N. Chetty and S. Alathur, "Hate speech review in the context of online social networks," *Aggression and violent behavior*, vol. 40, pp. 108–118, 2018.

[35] S. A. Ríos and R. Muñoz, "Dark web portal overlapping community detection based on topic models," in *Proceedings of the ACM SIGKDD workshop on intelligence and security informatics*. ACM, 2012, p. 2.