

AN OVERVIEW OF GEOINFORMATICS STATE-OF-THE-ART TECHNIQUES FOR LANDSLIDE MONITORING AND MAPPING

V.Yordanov^{1,2*}, L. Biagi¹, X.Q. Truong³, V.A. Tran⁴, M.A. Brovelli^{1,5}

¹ Department of Civil and Environmental Engineering (DICA)

Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan, Italy - (vasil.yordanov, maria.brovelli, ludovico.biagi)@polimi.it

² Vasil Levski National Military University, Veliko Tarnovo, Bulgaria

³ Information Technology Faculty, Hanoi University of Natural Resources and Environment, 41A Phu Dien Road, Phu Dien, North-Tu Liem district, Hanoi, Vietnam - txquang@hunre.edu.vn

⁴ Dept. of Geomatics and Land Administration, HUMG, HaNoi University of Mining and Geology, 18 Vien Street, Bac Tu Liem, Hanoi, Vietnam- tranvananh@humg.edu.vn

⁵ Istituto per il Rilevamento Elettromagnetico dell'Ambiente, CNR-IREA, via Bassini 15, 20133 Milano – maria.brovelli@polimi.it

Commission IV, WG IV/4

KEY WORDS: Landslides, Earth Observation, Artificial Intelligence, Citizen Science, Detection, Monitoring, Free and open-source approaches.

ABSTRACT:

Natural hazards such as landslides, whether they are driven by meteorologic or seismic processes, are constantly shaping Earth's surface. In large percentage of the slope failures, they are also causing huge human and economic losses. As the problem is complex in its nature, proper mitigation and prevention strategies are not straightforward to implement. One important step in the correct direction is the integration of different fields; as such, in this work, we are providing a general overview of approaches and techniques which are adopted and integrated for landslide monitoring and mapping, as both activities are important in the risk prevention strategies. Detailed landslide inventory is important for providing the correct information of the phenomena suitable for further modelling, analysing and implementing suitable mitigation measures. On the other hand, timely monitoring of active landslides could provide priceless insights which can be sufficient for reducing damages. Therefore, in this work popular methods are discussed that use remotely-sensed datasets with a particular focus on the implementation of machine learning into landslide detection, susceptibility modelling and its implementation in early-warning systems. Moreover, it is reviewed how Citizen Science is adopted by scholars for providing valuable landslide-specific information, as well as couple of well-known platforms for Volunteered Geographic Information which have the potential to contribute and be used also in the landslide studies. In addition to proving an overview of the most popular techniques, this paper aims to highlight the importance of implementing interdisciplinary approaches.

1. INTRODUCTION

Earthquake- and rainfall-triggered landslides are global natural hazards which are directly affecting lives and environment, as well as various economical aspects. The importance for mapping and monitoring, whether through ground-,air- or spaceborne techniques, landslide-prone areas and already known ones is highlighted in numerous studies, and integrated in many risk mitigation strategies. However, in the recent years several tendencies emerged from separate fields that tend to unite according to the research problem.

On one hand, Earth Observation (EO) free and open-source datasets and techniques naturally blended into Geographical Information Systems (GIS) and have found new aspects of implementation (e.g. disaster mapping, land cover changes, etc.).

On the other hand, Volunteered Geographic Information (VGI) emerged and proved as invaluable data source also in the disaster response domain. There are many practical examples of using OpenStreetMap data and volunteering collaborative mapping for risk management, relief and recovery strategies. The fusion between these different data gathering and processing methods provided information from diverse aspects and contribution even to the landslide studies.

Using EO, whether it is from optical or radar air/spaceborne sensing platforms, provided applications for detecting and mapping slope failures, even in monitoring their activity state or slow displacements, at a millimetre precision.

In addition, the use in the recent years of Artificial Intelligence (AI) and especially Machine Learning (ML) approaches, has increased in the field of Earth Observations. ML has been used for image classification applications, cloud detection and removal, enhancing the spatial resolution of satellite imagery, and many more. Naturally, scholars and decision-makers adopted machine learning techniques in their workflows and strategies for processing remotely sensed data in geohazard studies. Such implementations are already applied for landslide detection and mapping, landslide susceptibility and hazard mapping.

Even though, crowd-sourced data collection campaigns are often used in the disaster domain, both for risk mapping and disaster response, there are very few landslide-specific platforms and applications currently operational and well-known. In the paper, we will present the very few desktop and mobile-based applications and catalogues for collecting landslide-related geospatial -information. Lastly, it will be presented and discussed some current applications of AI on VGI datasets for recognition and detection of landslides. Moreover, it will be discussed how

* Corresponding author

UAV datasets could be also obtained in a citizen science manner through VGI collaborative platforms.

As all those topics can be currently considered as hot, the number of scientific publications related to landslides-focused EO, AI and citizen science applications has increased tremendously in the last few years. This paper aims to provide a general overview of the state-of-the-art open methods and techniques for monitoring and mapping landslides.

The paper will not only discuss the geoinformatics state-of-the-art landslide mapping and monitoring techniques but will also highlight the importance of combination and contribution between the different domains (Figure 1) because hardly any risk related problem is a single-aspect one. Moreover, bringing expertise from fields is a must in the hazard domain. Interdisciplinary approaches are not only bringing deeper understand of the problem but will be improving methodologies that are, and will be, yielding more accurate and time/cost efficient results for better mitigating the landslide hazard.

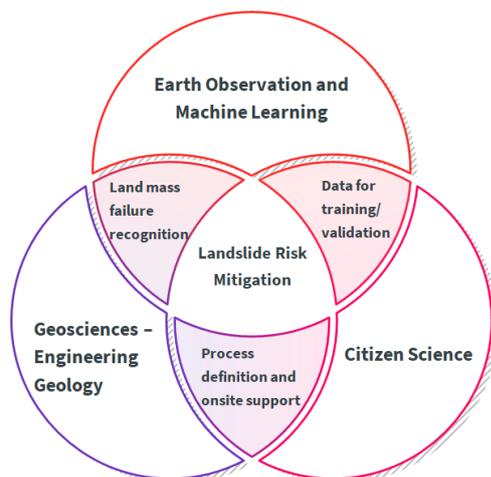


Figure 1. Integration of different disciplines for landslide risk mitigation strategies.

In-depth reviews for landslide investigations using remote sensing have already been presented (Mondini et al., 2021; Scaioni et al., 2014), thus in the next Section 2 will briefly highlight some of more traditional remote sensing approaches for mapping and monitoring landslides. While more attention will be paid for machine learning applications (Section 3) and in Section 4 will be presented VGI approaches contributing to the field. Finally, in Section 5 we will conclude with a general discussion of the presented.

2. LANDSLIDE MAPPING AND MONITORING : LANDSLIDE PREVENTION

Landslide mapping and monitoring are parts of the general goal of landslide prevention, that involves the landslide detection (Galli et al., 2008), the modelling of landslide susceptibility (Guzzetti et al., 2006) and the final hazard assessment.

The detection of a landslide is strategic to plan immediate disaster response; moreover, historical landslide inventories are also needed to prepare landslide susceptibility and hazard maps. Susceptibility is the spatial probability of occurrence of landslides. Susceptibility maps, combined with a modelling of landslide probability in time, allow the assessment of landslide hazard, that brings to forecasting or early warning raising.

2.1 Input data in landslide prevention.

In landslide detection, multitemporal satellite imagery is the most rapid and typically the cheapest data source; moreover, aerial photographs, LiDAR and inSar can be integrated in the process. Optical or multispectral images are used to estimate the location and the size of a landslide: LiDAR and inSar can be used to estimate the deformation field. In landslide detection also Digital Terrain Models can be integrated.

Beside the above data, susceptibility modelling requires data about the local scenario. In a rather recent paper, Reichenbach et al. (2018) provide a complete discussion about the scientific history of susceptibility modelling. 565 papers of peer review international journals from 1983 to 2016 are analysed in order to classify thematic variables and statistical approaches used in modelling. For the input data, the outcome is quite dramatic: 596 different thematic variables are used in the different papers but 445 of them are used just in one paper; in each paper susceptibility is modelled by using from 2 to 22 variables. The authors classify the thematic variables in 23 classes that are then grouped into 5 thematic clusters: morphological (37.2%), land cover (18.3%), geological (16.1%), hydrological (12.8%), and 'other', like for example precipitation.

Estimates of landslide hazard and early warning systems can be implemented at several level of complexity: always, they require the evaluation and modelling of all the triggering phenomena; in the case of the monitoring of active slow landslides also the maps of deformation in time are needed (Krkac et al., 2016). Guzzetti et al. (2020) provide an extensive review of several systems that have been implemented and are presently used at the local, national and global scale.

2.2 Landslide detection and mapping

An updated and exhaustive landslide inventory is an important dataset used when modelling landslide risk and susceptibility levels (Guzzetti et al., 2012). In-situ campaigns are still in use for updating landslide databases. However, mapping landslides using satellite-based datasets is reducing the time and costs for such tasks, especially with the use of freely available images from NASA's Landsat and ESA's Sentinel missions. Due to their medium spatial resolution and relatively high revisit time, scholars are mainly implementing change detection approaches relying on presence of vegetation to map landslide extents after the failure using both multispectral and radar imagery (e.g., Plank et al., 2016; Scheip and Wegmann, 2021). However, using only optical dataset may have its limitations especially when it is present a constant cloud cover, which obstructs timely detection. To overcome it, it is more common to use SAR imagery, which is almost not affected by the weathering conditions. The use of SAR allows to be exploited its signal properties – amplitude (single and multi-polarisation) and phase, through change in the coherence or bitemporal interferometry (Mondini et al., 2021). Ground-based and aerial photogrammetry is also a widely used approach for mapping and further analysing landslides through geological means (Scaioni et al., 2018, 2014).

2.3 Landslide monitoring

Ground-based landslide monitoring is a key component for the proper implementation of early-warning systems. However, the needed sensors cannot be installed everywhere, and these represent typical cases where remote sensing landslide monitoring make an important contribution. Depending on the availability of data to be used, the most popular techniques for monitoring landslides (slow and rapid) are: Digital Image Correlation (DIC) from optical images which provides displacement into 2D space (Mazzanti et al., 2020); 3D

reconstruction and comparison from photogrammetric datasets (Scaioni et al., 2014); Differential SAR Interferometry and Persistent Scatter (PSInSAR) interferometry methods (Ferretti et al., 2001). The last two methods are yielding more than satisfactory results at sub-pixels scale for slow movements, however some limitations are found due to the presence of vegetation.

3. STATE OF THE ART OF MACHINE LEARNING FOR LANDSLIDE PREVENTION

The present section is structured in two parts: an introduction to machine learning is followed by the discussion on its application in landslide prevention and protection.

3.1 Machine learning

By Machine Learning (ML, Bishop, 2006, Hastie et al., 2011, Murphy, 2012, James et al., 2013), models are built to represent relationships between input data, or observations, and target variables, or unknowns, when these relationships cannot be described by simple parametric models or are at least partly unknown. The simplest application is a binary classification, for example for spam recognition in emails; multi class classification, ranking problems and predictions provide more complex applications: for example, a ML algorithm can be implemented to detect potential medical diseases from tomographies and other medical data of patients. From an historical point of view, the driving applications for machine learning were computer vision, speech recognition, natural language processing, medical analysis, as much as other specific scientific applications, like in Physics (Baldi et al., 2014). The application of machine learning to geophysics or geological problems started about three decades ago (Van Der Baan and Jutten, 2000, Bergen et al., 2019, Reichstein et al., 2019), for studies relevant to solid Earth as well as oceans and atmosphere. Classification problems constitute the typical applications in such disciplines: for example, Dowlal et al (1990) used artificial neural networks to discriminate signals between natural earthquakes and underground nuclear explosions. The application of machine learning to landslide prevention (Ma et al., 2020) started later but in the last two decades has experienced a disruptive growth: the analysis of landslide susceptibility provides a typical application. In this section, the main machine learning algorithms are shortly summarized, then the major applications in landslide prevention are presented.

In general, we can write that ML algorithms map an input dataset y into a target, or label, vector x . The two main classes of ML algorithms are supervised and unsupervised (Jordan et al., 2015). Supervised algorithms require a training dataset of known couples of y and x . Training data is used to set the algorithm. If correctly trained, the algorithm can be successfully applied to new, unexplored, data. Unsupervised algorithms adopt learning approaches that are not based on the availability of training data: they are needed when no or few training data are available, for example in exploratory research. Many different ML algorithms have been proposed, both supervised and unsupervised: the choice of a proper algorithm depends on the specific application, the availability of training data, the size of the data set, the characteristics of the target vector (continuous / discrete).

Images are often processed with ML algorithms; this is particularly true in geophysical or geological applications, where the analysis always implies the use of digital maps and images. Image analysis can be either pixel oriented or object oriented. In the first approach, pixels are individually investigated on the basis of their spectral signature. Object based algorithms firstly group pixels into objects, sometimes named also superpixels: the

grouping is based on the analysis of spectral, spatial and contextual characteristics. Objects are then investigated, accordingly to the specific application.

Machine learning algorithms can be sorted by their complexity levels. Artificial neural networks (ANN) constitute the fundamental brick of most of the machine learning algorithms. Logical Regression (LR) and Bayesian are other classical and 'simple' ML algorithms that are statistically based; Support Vector Machine (SVM) and Decision Trees (DT) provide other examples.

With respect to ANN, more complicated structures of nested layers bring to convolutional and recurrent networks (CNN, RNN), that are also called deep neural networks or deep learning (Le Cun et al., 2015, Shin et al., 2016). Extreme Machine Learning (EML, Lian et al., 2014), Random Forest (RF, Breiman, 2001) represent other complex ML algorithms, that are quite popular in landslide prevention: finally, Ensemble Learning Methods (ELM) combine several individual ML algorithms. A complete discussion of each one of them is outside the scope of this review and the reader is referred to the above references.

3.2 Machine learning for landslide prevention

3.2.1 ML for landslide detection. Traditionally, classification was obtained by pixel-oriented approaches, like in Danneels et al. (2007). In time, object-oriented algorithms have proved to be more effective: this is particularly true with the new generation of high resolution images (Hussain et al., 2013, Martha et al., 2012). Stumpf and Kerle (2011) combined object oriented and supervised RF to identify landslides features after events. Four test sites were analysed in Haiti, China, French and Italy. Only post event images were used: Geoeye-1, IKONOS, Quickbird, WorldView-2 and aerial photographs, integrated with LiDAR DTM. They obtained satisfactory results but testified the need of fine tuning and iterations to face class-imbalance between landslide and non-landslide. Ji et al. (2020) apply supervised CNN to produce an historical inventory of landslides. The case study is in Bijie, northwest of Guizhou province, China, in the transitional slope zone from the Tibet Plateau to the eastern hills. Input data is an open multitemporal satellite imagery, an inventory of known landslides and the DTM of the area. Maggiori et al. (2017), Sameen et al. (2019), Ghorbanzadeh et al. (2019), Wang et al. (2021) and many other recent papers discuss other examples and comparisons on the use of object-oriented approaches combined with CNN or other kinds of deep learning in landslide detection: at the present, deep learning seems the most promising ML algorithm in landslide detection.

3.2.2 ML for landslide susceptibility. About 20 years ago, first efforts were attempted to apply simple ML algorithms to model the susceptibility: for example, Lee et al. (2003) applied ANN, Saito et al. (2009) applied DT, Heckmann et al. (2014) tested LR. Still some applications are investigated on simple ML: in 2018, Kalantar et al. compare ANN, LR and SVM on a test dataset of landslide inventories: LR provides the best results, with an average accuracy of 81.4%.

However, the research has gradually moved toward more complicated models, like RF (Krkak, 2016), deep learning and ensemble learning methods. Chen et al. (2017) compare Logical Model Tree (LMT), RF and Classification and Regression Tree (CART) on a benchmark dataset of 171 landslides in Shaanxi, China: all the models provide satisfactory results and RF outperforms the others. Pham et al. (2019) combine different EML algorithms into four new different hybrid prediction models and compare them on a test site of about 250 km² in one Himalayan region of India. All the results are satisfactory: in

particular, the algorithm named BREPT provides the most reliable ones. Wang et al. (2020) compare 4 different RNN on a test site in the western part of Jiangxi Province, China. Several accuracy and reliability indexes are investigated: all the results are satisfactory; in this case, all the developed software and training data are free and open, distributed through Github. Pham et al. (2020) investigate the use of CNN. Specifically, an algorithm, named Moth Flame Optimization is explored to train CNN. The approach is tested in a mountain area of the Lai Chau province, Vietnam. Also, in this study several statistical indexes are considered, with excellent results. Finally, Fang et al. (2021) implement four ELM, whose components are CNN and RNN, combined with SVM and LR; in the meantime, they analyse the correlation between different geomorphological parameters and landslide susceptibility. The ELM results are really encouraging with respect to the results provided by individual algorithms.

3.2.3 ML for landslide forecasting and early warning. Landslide early warning systems by ML are a relatively recent object of investigation: landslide forecasting requires the combined analysis of susceptibility and triggering effects, typically rainfalls, groundwater levels and earthquakes. Krkac et al. (2016) predict the motion of Kostanjek landslide (northern part of Zagreb city, Croatia) by applying RF to the analysis of GNSS monitoring stations, rainfalls, and groundwater levels; the prediction of the landslide velocity and its variation in time is really accurate.

Paradhan et al. (2019) analyse landslides triggered by rainfalls in Busan, Korea, a city of about 780 km² in a mountain area; ANN is trained and applied to estimate rainfalls thresholds for early warning. A similar research is done by Sang et al. (2019), who compare genetic algorithm back-propagation neural network and genetic algorithm support vector machine. Shiluo and Niu (2018) model the dynamic of a landslide in Zigui County, China, by considering six triggering effects, all related to hydrological and meteorological aspects. A long short-term memory neural network provides the best results; although, the accuracies are quite low and the outputs present ill conditioning. Finally, Kuradusenge et al. (2020) examine the case study of Ngororero district, Rwanda. They apply LR and RF to several thematic variables, included rainfalls, to predict landslides: LR provides the best results with an error (incorrect prediction) of less than 4%.

3.3 Final considerations on ML for landslide prevention

During the last two decades, supervised ML algorithms have been widely studied and tested in landslide detection and susceptibility modelling; more recently, research on forecasting and early warning started. In detection and susceptibility, the significant class imbalance between non landslide and landslide classes represents a problem in training and could bias the classification (Yordanov and Brovelli, 2020a,b). In susceptibility modelling, the choice of the model represents the methodological problem: indeed, the input thematic variables and their use greatly vary accordingly to the local scenario and the investigated landslide. Therefore, only local models can be tuned but no global guidelines can be given. In detection and susceptibility modelling, both the deep learning and the nested ensemble algorithms seem to be the pushing study subjects.

Landslide early warning systems are complex tasks that have been implemented and tested at various spatial scales in the world. Research on the machine learning application is now a work in progress, particularly for the analysis of triggering phenomena.

4. CITIZEN SCIENCE FOR LANDSLIDE APPLICATIONS

4.1 Citizen science

As ML became trending in the landslide domain, similar is the case with the citizen science (CitSci) contribution to the geospatial field and the attempts of scholars to “harvest” it also for the landslide-specific studies. Widely adopted term in the geospatial studies related to crowdsourced knowledge but with added geospatial reference (Schade et al., 2013) is the Volunteered Geographic Information (VGI) from Goodchild (2007), who stresses the benefits of motivating individuals to volunteer especially in emergency situations. In the current paper it would not be discussed in-detail the definitions of crowdsourcing, citizen science (Brovelli et al., 2020), nor the discussions about the data uncertainties (Fonte et al., 2017) and integrity (Juhász et al., 2020). However, it should be highlighted that usually VGI participation is more focused on some common aim and all individuals are collaborating on this specific outcome (Lee et al., 2020). An example for a notable VGI project is the OpenStreetMap which was initiated with the aim to generate and distribute free geographic data (OpenStreetMap contributors, 2017).

Scholars highlighted the disaster domain that could benefit greatly from crowdsourced contribution (Glantz and Ramírez, 2018; Goodchild, 2007) in different phases of the disaster management (Lee et al., 2020) starting from a level where the citizen are mainly contributing with data gathering to a more experienced level where they are part in the definition of the problem, information gathering and its analysis (Haklay, 2013; Kocaman et al., 2018). Interestingly, the inclusion of citizens at different levels appears to coincide with the implementation of CitSci in the time, moreover the initial usage of crowdsourced information in the disaster domain is correlated with the mass usage of mobile technologies and social networks. Where scholars were mainly exploiting the citizens as sensors, with application for earthquakes (Earle, 2010; Shan et al., 2012), floods (Kouadio and Douvina, 2015). On the other hand, the technology is used for early warning systems (Marchezini et al., 2018, 2017), crisis mapping (Norheim-Hagton and Meier, 2010), building resilience (Cieslik et al., 2019), landslide susceptibility mapping (Rohan et al., 2021).

However, in particular for the landslide domain, the incorporation of CitSci is more restricted to gathering landslide data for populating inventories. Such an example of a project on a global level is the Cooperative Open Online Landslide Repository (Juang et al., 2019) as a part of the Global Landslide Catalogue (Kirschbaum et al., 2010). On national level examples are the Great Britain National Landslide Database (Pennington et al., 2015) and Italian IdroGeo platform incorporated in the national landslide inventory (Iadanza et al., 2021). Reviewing CitSci applications for landslide data collection can be noted the two main types of crowdsourcing data collection – passive and active.

4.2 Passive crowdsourcing

Passive crowdsourcing is referring to the approach where scholars are “mining” for a specific information which has been already published/made available, or in other words an individual unknowingly contributed to a specific field. Examples of such data sources are news portals, social medias (e.g., Twitter and Flickr), search engines (some are allowing to assign alert for a specific keywords). The approach is utilized in studies at local level, but also at national and global (e.g., GBNLD and GLD). As some of the advantages could be considered that the data is

streamed almost constantly (Ghermandi and Sinclair, 2019) and it is at no cost. On the other side the data can be biased, wrong or too general which will require additional research time to be useful. For example, a news report or microblog could reveal an approximate location of a landslide, but a verification from another source or field investigation could be necessary to determine its exact location.

4.3 Active crowdsourcing

On the contrary of the passive CitSci, active crowdsourcing is when the individual is participating for a specified outcome or when it is in the spatial domain it coincides with the definition of VGI. In the active participation it is important to properly design the surveys according to proficiency level of the individuals. Possible target group is involving professionals or participants which have less to no knowledge of the topic, or both. In the former case, a VGI project would be easier for a design where less guidance and information are needed to the participants, however this may lead to a smaller group of participants, thus insufficient amount of data. In the latter case when non-professionals are involved additional efforts should be made to provide them with sufficient and straightforward informational materials prior the data collection. For example, when the aim is landslide mapping - clear and brief description of landslides, types of material and movement, parts, triggering factors, etc.

Depending on the type of data provision and tools to be used, active crowdsourcing could be further differentiated into two approaches – desktop and mobile. They differ mainly in the mean which is used for the data report. In the case of landslide data gathering, usually the fields to be filled are requiring the same type of information and are following standard geological survey questionnaires. Mainly it is requested information for the location of landslide, movement type, displaced material, date of the event, damages, mitigation measures, a photo of the landslide. The mandatory information and any other fields required are usually based on the survey designer's experience and choice.

4.3.1 Desktop-based. This approach applies when users are required to report landslides of their knowledge usually through a dedicated web-portal by compiling a form. Such a report can be done without any limitation for the current location of the user. In using a dedicated webpage for reporting there are several benefits: the user is not bounded or needed to be at the location of the landslide, as long as the provided information has a certain level of accuracy; the project designer has a wider possibility to integrate teaching materials (purely text-based or even interactive) useful for non-professionals. On the con side, not always the geographical location of the event is enough precise and further efforts are needed for determining the exact one. This approach is adopted when the aim is to populate the inventory of a large regions, as implemented for GLD, GBNLD and IdroGeo.

4.3.2 Mobile (field) based. This technique mainly relies on the relatively recent advancement of the mobile phone technologies and the integration of various sensors inside them. Scholars are adopting the use of mobile applications mainly due to the built-in GPS, compass, camera and Internet connection, which allows users to provide data with relatively high accuracy directly from the location of a landslide. Often, an option for offline data collection is present due to the nature of the hazard and its presence in location where mobile signal can be sparse. The designed survey applications usually tend to keep the user interface simple and interactive, which allows straightforward data compilation and prevents a decrease of interest of the task. It is common that the provided information is then visualized/accessible (upon validation) to a webGIS platform or online database. In the case of GBNLD, except the online report

a landslide form, is developed parallel mobile application for active mobile crowdsourcing (myHAZ - <https://vct.myhaz.app/>) – incorporating tools for data reporting, managing and exploration. Many other VGI applications for landslide reporting were developed in the recent years (Choi et al., 2018; Hennig et al., 2020; Jacobs et al., 2019; Kocaman and Gokceoglu, 2019; Žabota and Kopal, 2020). However, during the literature review phase for this manuscript, was noted a trend related to applications that were presented couple of years ago which upon a verification appear to be discontinued and not currently maintained (at the time of writing). It appears as a discouraging trend that is related to the lifetime of projects supporting such efforts. In fact, it was pointed out by Irwin (2018) as a limitation of the CitSci.

4.4 Data validation

Naturally, the provided data should be further verified before its inclusion to an inventory. Most of current available applications and forms have a validation step which aims to verify if the reporting is correct, if it is a duplicate or if it is providing additional information to an already existing one. Mainly, it is carried from experts and, when are present uncertainties, additional sources, for redundancies, are suggested. (Kirschbaum et al., 2010).

More recently, (Can et al., 2019) trained a CNN and implemented the model in a webGIS (Can et al., 2020) for landslide image classification as part of validation process from crowdsourced landslide images. They highlight that due to limitation of the training datasets the automation process needs to be further manually validated, however it shows promising interdisciplinary incorporation.

4.5 Potential platforms for landslide VGI

In the following, we would like to briefly highlight the potential of OpenStreetMap (OSM) and OpenAerialMap (OAM - a platform for contributed volunteered UAV images) for landslide mapping.

4.5.1 OpenStreetMap. The data richness of OSM relies on its contributors and assigned tags which are describing the semantic of the features of the map. According to the OSM Wiki page (<https://wiki.openstreetmap.org/wiki/Tag:natural%3Dlandslide>) there are already present tags related to landslides: *natural=landslide* – for general annotation, while the types of landslides can be assigned as *landslide=**, where the value is the specific type. However, it seems that tagging landslides is not of a particular interest since the general tag has been used only around 4,000 times. A restriction for using OSM could be that when users are contributing based on satellite imagery the datasets are not always up-to-date, therefore a presence of a landslide cannot be spotted due to the age of the imagery. However, incorporating UAV imagery may provide more recent high-resolution data.

4.5.2 OpenAerialMap. is a platform where users can explore and contribute UAV creative commons imagery (<https://openaerialmap.org/> made by Humanitarian OpenStreetMap Team). It allows citizens to share high resolution orthophotos without restriction of the topic. When it is the case of landslide imagery then can be further mapped in OSM and used from professionals. OpenDroneMap (<https://www.opendronemap.org/> - a FOSS tool for processing aerial images) has an integrated option to directly upload the output orthophoto directly into OAM.

5. CONCLUSION

In the current work we have presented a general overview of the best practices for landslide mapping and monitoring from the perspective of Earth Observation, Machine Learning and Citizen Science. Individually each of the discipline has its own advantages and disadvantages when implemented in landslide domain, but it is considered that integrating expertise from different fields is bringing firstly a better understanding of the problem, of the needed actions to be taken and of the tools for the actions to be carried out. For example, EO (whether ground-, air- or satellite-borne) can provide huge amount of data, while ML is bringing the needed means for its accurate processing, while engineering geology can provide the right problem definition and outcome interpretation. In addition, crowdsourced information could be always beneficial for the correct interpretation and bring in additional value that would polish the landslide defining algorithms.

ACKNOWLEDGEMENTS

The work is partially funded by the Italian Ministry of Foreign Affairs and International Cooperation within the project “Geoinformatics and Earth Observation for Landslide Monitoring” - CUP D19C21000480001 (Italian side) and partially funded by Ministry of Science and Technology of Vietnam (MOST) (Vietnamese side) by the Bilateral Scientific Research project between Vietnam and Italy, code: 14/2021/HĐ-NĐT.

REFERENCES

- Baldi P., Sadowski P., Whiteson D., 2014: Searching for exotic particles in high-energy physics with deep learning. *Nature Communication* 5:4308
- Bergen K. J., Johnson P., Van de Hoop M., Beroza G.C., 2019: Machine learning for data-driven discovery in solid earth geoscience, *Science*, 363:1299
- Bishop C., 2006: *Pattern Recognition and Machine Learning*, Springer
- Breiman, L., 2001: Random Forests, *Machine Learning*, 45:5–32.
- Brovelli, M.A., Ponti, M., Schade, S., Solís, P., 2020: Citizen Science in Support of Digital Earth, in: Guo, H., Goodchild, M.F., Annoni, A. (Eds.), *Manual of Digital Earth*. Springer, Singapore, pp. 593–622. https://doi.org/10.1007/978-981-32-9915-3_18
- Can, R., Kocaman, S., Gokceoglu, C., 2019: A Convolutional Neural Network Architecture for Auto-Detection of Landslide Photographs to Assess Citizen Science and Volunteered Geographic Information Data Quality. *ISPRS International Journal of Geo-Information* 8, 300. <https://doi.org/10.3390/ijgi8070300>
- Can, R., Kocaman, S., Gokceoglu, C., 2020: DEVELOPMENT OF A CITSCI AND ARTIFICIAL INTELLIGENCE SUPPORTED GIS PLATFORM FOR LANDSLIDE DATA COLLECTION, in: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. (Volume XLIII-B5-2020) - 2020 edition, Copernicus GmbH, pp. 43–50. <https://doi.org/10.5194/isprs-archives-XLIII-B5-2020-43-2020>
- Chen W., Xie X., Wang J., Pradhan B., Hong H., Tien Bui D., Duan Z., Ma J., 2017: A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* 151:147–160.
- Choi, C.E., Cui, Y., Zhou, G.G.D., 2018: Utilizing crowdsourcing to enhance the mitigation and management of landslides. *Landslides* 15, 1889–1899. <https://doi.org/10.1007/s10346-018-1034-9>
- Cieslik, K., Shakya, P., Uprety, M., Dewulf, A., Russell, C., Clark, J., Dhital, M.R., Dhakal, A., 2019: Building Resilience to Chronic Landslide Hazard Through Citizen Science. *Front. Earth Sci.* 7. <https://doi.org/10.3389/feart.2019.00278>
- Danneels G., Pirard E., Havenith H.B., 2007: Automatic landslide detection from remote sensing images using supervised classification methods, 2007 *IEEE international geoscience and remote sensing symposium*, 3014–3017
- Dowla F.U., Taylor S.R., Anderson R.W., 1990: Seismic discrimination with artificial neural networks: Preliminary results with regional spectral data, *Bulletin of the Seismological Society of America* 80:1346–1373
- Earle, P., 2010: Earthquake Twitter. *Nature Geosci* 3, 221–222. <https://doi.org/10.1038/ngeo832>
- Fang Z., Wang Y., Peng L., Hong H., 2021: A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping, *International Journal of Geographical Information Science*, 35,2:321-347
- Ferretti, A., Prati, C., Rocca, F., 2001: Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing* 39, 8–20. <https://doi.org/10.1109/36.898661>
- Fonte, C.C., Antoniou, V., Bastin, L., Estima, J., Arsanjani, J.J., Bayas, J.-C.L., See, L., Vatsava, R., 2017: Assessing VGI Data Quality. *Ubiquity Press*. <https://doi.org/10.5334/bbf.g>
- Galli M., Ardizzone F., Cardinali M., Guzzetti F., Reichenbach P., 2008: Comparing landslide inventory maps, *Geomorphology* 94:268–289.
- Ghermandi, A., Sinclair, M., 2019: Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change* 55, 36–47. <https://doi.org/10.1016/j.gloenvcha.2019.02.003>
- Ghorbanzadeh O., Blaschke T., Gholamnia K., Meena S. R., Tiede D., Aryal J., 2019: Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection, *Remote Sensing*, 11:196
- Glantz, M.H., Ramírez, I.J., 2018: Improvisation in the Time of Disaster. *Environment: Science and Policy for Sustainable Development* 60, 4–17. <https://doi.org/10.1080/00139157.2018.1495496>
- Goodchild, M.F., 2007: Citizens as sensors: the world of volunteered geography. *GeoJournal* 69, 211–221. <https://doi.org/10.1007/s10708-007-9111-y>
- Guzzetti F., Gariano S.L., Peruccacci S., Brunetti M.T., Marchesini I., Rossi M., Melillo M., 2020: Geographical landslide early warning systems, *Earth-Science Reviews*, 200: 102973
- Guzzetti F., Paola R., Ardizzone F., Cardinali M., Galli M., 2006: Estimating the quality of landslide susceptibility models, *Geomorphology* 81:166–184.
- Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.T., 2012: Landslide inventory maps:

- New tools for an old problem, *Earth-Science Reviews*.
<https://doi.org/10.1016/j.earscirev.2012.02.001>
- Haklay, M., 2013:w Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation, in: Sui, D., Elwood, S., Goodchild, M. (Eds.), *Crowdsourcing Geographic Knowledge: VGI in Theory and Practice*. Springer Netherlands, Dordrecht, pp. 105–122.
https://doi.org/10.1007/978-94-007-4587-2_7
- Hastie T., Tibshirani R., Friedman J., 2011: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, New York
- Heckmann T., Gegg K., Gegg A., Becht M., 2014: Sample size matters: Investigating the effect of sample size on a logistic regression susceptibility model for debris flows, *Natural Hazards and Earth System Science* 14:259–278
- Hennig, S., Abad, L., Hölbling, D., 2020: Technological solutions for contributory citizen science: experiences from citizenMorph, in: *Proceedings of 5th Austrian Citizen Science Conference 2019 — PoS(ACSC2019)*, Sissa Medialab, Oberurgl, Austria, p. 005. <https://doi.org/10.22323/1.366.0005>
- Hussain M., Chen D., Cheng A., Wei H., Stanley D., 2013: Change detection from remotely sensed images: From pixel-based to object-based approaches, *ISPRS Journal of Photogrammetry and Remote Sensing* 80:91–106
- Iadanza, C., Trigila, A., Starace, P., Dragoni, A., Biondo, T., Roccisano, M., 2021: IdroGEO: A Collaborative Web Mapping Application Based on REST API Services and Open Data on Landslides and Floods in Italy. *ISPRS International Journal of Geo-Information* 10, 89. <https://doi.org/10.3390/ijgi10020089>
- Irwin, A., 2018: No PhDs needed: how citizen science is transforming research. *Nature* 562, 480–482.
<https://doi.org/10.1038/d41586-018-07106-5>
- Jacobs, L., Kabaseke, C., Bwambale, B., Katutu, R., Dewitte, O., Mertens, K., Maes, J., Kervyn, M., 2019: The geo-observer network: A proof of concept on participatory sensing of disasters in a remote setting. *Science of the Total Environment* 670, 245–261. <https://doi.org/10.1016/j.scitotenv.2019.03.177>
- James G., Witten D., Hastie T., Tibshirani R., 2013: An Introduction to Statistical Learning, Springer, New York.
- Jordan M. I., Mitchell T. M., 2015: Machine learning: Trends, perspectives, and prospects, *Science* 349:255–260
- Juang, C.S., Stanley, T.A., Kirschbaum, D.B., 2019: Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR). *PLOS ONE* 14, e0218657.
<https://doi.org/10.1371/journal.pone.0218657>
- Juhász, L., Novack, T., Hochmair, H.H., Qiao, S., 2020: Cartographic Vandalism in the Era of Location-Based Games—The Case of OpenStreetMap and Pokémon GO. *ISPRS International Journal of Geo-Information* 9, 197.
<https://doi.org/10.3390/ijgi9040197>
- Kalantar B., Pradhan B., Naghibi S., Motevalli A., Mansor S., 2018: Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN), *Geomatics, Natural Hazards and Risk* 9,1: 49-69
- Kirschbaum, D.B., Adler, R., Hong, Y., Hill, S., Lerner-Lam, A., 2010: A global landslide catalog for hazard applications: method, results, and limitations. *Nat Hazards* 52, 561–575.
<https://doi.org/10.1007/s11069-009-9401-4>
- Kocaman, S., Anbaroglu, B., Gokceoglu, C., Altan, O., 2018: A REVIEW ON CITIZEN SCIENCE (CITSCI) APPLICATIONS FOR DISASTER MANAGEMENT, in: *GeoInformation For Disaster Management*-pp.301–306.<https://doi.org/10.5194/isprs-archives-XLII-3-W4-301-2018>
- Kocaman, S., Gokceoglu, C., 2019: A CitSci app for landslide data collection. *Landslides* 16, 611–615.
<https://doi.org/10.1007/s10346-018-1101-2>
- Kouadio, S., Douvinet, J., 2015: A smartphone application to help alert in case of flash floods, in: *2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. pp. 252–257.
<https://doi.org/10.1109/ICT-DM.2015.7402045>
- Krkac M., Spoljaric´ D., Bernat Gazibara S., Mihalic´ Arbanas S., 2016: Method for prediction of landslide movements based on random forests, *Landslides*, 14:947–960
- Kuradusenge M., Kumaran S., Zennaro M., 2020: Rainfall-Induced Landslide Prediction Using Machine Learning Models: The Case of Ngororero District, Rwanda, *International Journal of Environmental Research and Public Health*, 17, 4147
- LeCun Y., Bengio Y., Hinton G., 2015: Deep Learning, *Nature*, 521:436-444
- Lee S., Ryu J.H., Lee M.J., Won J.S., 2003: Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea, *Environmental Geology* 44,7:820–833
- Lee, K.A., Lee, J.R., Bell, P., 2020: A review of Citizen Science within the Earth Sciences: potential benefits and obstacles. *Proceedings of the Geologists' Association* 131, 605–617.
<https://doi.org/10.1016/j.pgeola.2020.07.010>
- Lian C., Zeng Z., Yao W., Tang H. (2014) Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis, *Neural Computing and Applications* 24:99–107
- Ma Z., Mei G., Piccialli F., 2020: Machine learning for landslides prevention: a survey, *Neural Computing and Applications*
- Maggiori E., Tarabalka Y., Charpiat G., Alliez P., 2017: High-resolution aerial image labeling with convolutional neural networks, *IEEE Transactions on Geoscience and Remote Sensing*, 55,12:7092–7103
- Marchezini, V., Horita, F.E.A., Matsuo, P.M., Trajber, R., Trejo-Rangel, M.A., Olivato, D., 2018: A Review of Studies on Participatory Early Warning Systems (P-EWS): Pathways to Support Citizen Science Initiatives. *Front. Earth Sci.* 6. <https://doi.org/10.3389/feart.2018.00184>
- Marchezini, V., Trajber, R., Olivato, D., Muñoz, V.A., de Oliveira Pereira, F., Oliveira Luz, A.E., 2017: Participatory Early Warning Systems: Youth, Citizen Science, and Intergenerational Dialogues on Disaster Risk Reduction in Brazil. *Int J Disaster Risk Sci* 8, 390–401. <https://doi.org/10.1007/s13753-017-0150-9>
- Martha T., Kerle N., Westen C., Jetten V., Vinod Kumar K., 2012: Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories, *ISPRS Journal of Photogrammetry and Remote Sensing* 67:105–119.
- Mazzanti, P., Caporossi, P., Muzi, R., 2020: Sliding Time Master Digital Image Correlation Analyses of CubeSat Images for

- landslide Monitoring: The Rattlesnake Hills Landslide (USA). *Remote Sensing* 12, 592. <https://doi.org/10.3390/rs12040592>
- Mondini, A.C., Guzzetti, F., Chang, K.-T., Monserrat, O., Martha, T.R., Manconi, A., 2021: Landslide failures detection and mapping using Synthetic Aperture Radar: Past, present and future. *Earth-Science Reviews* 216, 103574. <https://doi.org/10.1016/j.earscirev.2021.103574>
- Murphy K., 2012: Machine Learning: A Probabilistic Perspective, MIT Press, Cambridge, MA
- Norheim-Hagtun, I., Meier, P., 2010: Crowdsourcing for Crisis Mapping in Haiti. *Innovations: Technology, Governance, Globalization* 5, 81–89. https://doi.org/10.1162/INOV_a_00046
- OpenStreetMap contributors, 2017: Planet dump planet.osm.org.
- Pennington, C., Freeborough, K., Dashwood, C., Dijkstra, T., Lawrie, K., 2015: The National Landslide Database of Great Britain: Acquisition, communication and the role of social media. *Geomorphology, Geohazard Databases: Concepts, Development, Applications* 249, 44–51. <https://doi.org/10.1016/j.geomorph.2015.03.013>
- Pham B., Prakash I., Singh S., Shirzadi A., Shahabi H., Tran T.T.T., Tein Bui D., 2019: Landslide susceptibility modelling using reduced error pruning trees and different ensemble techniques: hybrid machine learning approaches, *Catena* 175:203–218.
- Pham V., Nguyen Q.H., Nguyen H.D., Pham V.M., Manh V., Bui Q.T., 2020: Convolutional neural network - optimized moth flame algorithm for shallow landslide susceptible analysis, *IEEE Access*, 8:32727-32736
- Plank, S., Twele, A., Martinis, S., 2016: Landslide Mapping in Vegetated Areas Using Change Detection Based on Optical and Polarimetric SAR Data. *Remote Sensing* 8, 307. <https://doi.org/10.3390/rs8040307>
- Pradhan A., Lee S.R., Kim Y.T., 2018: A shallow slide prediction model combining rainfall threshold warnings and shallow slide susceptibility in Busan, Korea, *Landslides*, 16:647–659
- Reichenbach P., Rossi M., Malamud B.D., Mihir M., Guzzetti F., 2018: A review of statistically-based landslide susceptibility models, *Earth- science Reviews*, 180:60–91
- Reichstein M., Camps-Valls G., Stevens B., Jung M., Denzler J., Carvalhais N., Prabhat M., 2019: Deep learning and process understanding for data-driven earth system science, *Nature*, 566:195–204.
- Rohan, T.J., Wondolowski, N., Shelef, E., 2021: Landslide susceptibility analysis based on citizen reports. *Earth Surface Processes and Landforms* 46, 791–803. <https://doi.org/10.1002/esp.5064>
- Saito H., Nakayama D., Matsuyama H., 2009: Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaiishi Mountains, Japan, *Geomorphology* 109:108–121
- Sameen M.I., Pradhan B., 2019: Landslide detection using residual networks and the fusion of spectral and topographic information, *IEEE Access* 7:114363–114373
- Scaioni, M., Crippa, J., Yordanov, V., Longoni, L., Ivanov, V.I., Papini, M., 2018: Some tools to support teaching photogrammetry for slope stability assessment and monitoring. *ISPRS Archives*, pp. 453–460. <https://doi.org/10.5194/isprs-archives-XLII-3-W4-453-2018>
- Scaioni, M., Longoni, L., Melillo, V., Papini, M., 2014: Remote sensing for landslide investigations: An overview of recent achievements and perspectives, *Remote Sensing*. <https://doi.org/10.3390/rs6109600>
- Schade, S., Díaz, L., Ostermann, F., Spinsanti, L., Luraschi, G., Cox, S., Nuñez, M., De Longueville, B., 2013: Citizen-based sensing of crisis events: sensor web enablement for volunteered geographic information. *Appl Geomat* 5, 3–18. <https://doi.org/10.1007/s12518-011-0056-y>
- Scheip, C.M., Wegmann, K.W., 2021: HazMapper: a global open-source natural hazard mapping application in Google Earth Engine. *Natural Hazards and Earth System Sciences* 21, 1495–1511. <https://doi.org/10.5194/nhess-21-1495-2021>
- Shan, W., Feng, J., Chang, J., Yang, F., Li, Z., 2012: Collecting earthquake disaster area information using smart phone, in: *2012 International Conference on System Science and Engineering* pp. 310–314. <https://doi.org/10.1109/ICSSE.2012.6257197>
- Shiluo X., Niu R., 2018: Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China, *Computers and Geosciences*, 111:87–96.
- Shin H.C., Roth H., Gao M., Lu L., Xu Z., Nogues I., Yao J., Mollura D., Summers R., 2016: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, *IEEE Transactions on Medical Imaging* 35:1285–1298
- Shunping J., Dawen Y., Shen C., Li W., Qiang X., 2020: Landslide detection from an open satellite imagery and digital elevation dataset using attention boosted convolutional neural networks. *Landslides*, 17:1337–1352
- Stumpf A., Kerle N., 2011: Object-oriented mapping of landslides using random forests, *Remote Sensing of Environment*, 115:2564–2577
- Van der Baan M. and Jutten C. 2000: Neural networks in geophysical applications, *Geophysics*, 65-4: 1032-1047
- Wang H., Zhang L., Yin K., Luo H., Li J., 2021: Landslide identification using machine learning, *Geoscience Frontiers*, 12: 351–364
- Wang Y., Fang Z., Wang M., Ling P., Hong H., 2020: Comparative study of landslide susceptibility mapping with different recurrent neural networks, *Computers and Geoscience* 138:104445
- Yordanov, V., Brovelli, M.A., 2020a. Application of various strategies and methodologies for landslide susceptibility maps on a basin scale: the case study of Val Tartano, Italy. *Applied Geomatics*. <https://doi.org/10.1007/s12518-020-00344-1>
- Yordanov, V., Brovelli, M.A., 2020b. Comparing model performance metrics for landslide susceptibility mapping. *ISPRS Archives*, <https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1277-2020>
- Žabota, B., Kobal, M., 2020: A New Methodology for Mapping Past Rockfall Events: From Mobile Crowdsourcing to Rockfall Simulation Validation. *ISPRS Int. J. Geo-Inf.* 9, 514. <https://doi.org/10.3390/ijgi9090514>
- Zhenlei W., Lu Q., Sun H.Y., Sang Y., 2019: Estimating the rainfall threshold of a deep-seated landslide by integrating models for predicting the groundwater level and stability analysis of the slope, *Engineering Geology* 253:14–26.