ORIGINAL PAPER

Machine learning and landslide studies: recent advances and applications

Faraz S. Tehrani[1] · Michele Calvello[2] · **Zhongqiang Liu[3] · Limin Zhang[4] ·
Suzanne Lacasse[3]**

Abstract
Upon the introduction of machine learning (ML) and its variants, in the form that we know
today, to the landslide community, many studies have been carried out to explore the use-
fulness of ML in landslide research and to look at some classic landslide problems from an
ML point of view. ML techniques, including deep learning methods, are becoming popular
to model complex landslide problems and are starting to demonstrate promising predictive
performance compared to conventional methods. Almost all the studies published in the
literature in recent years belong to one of the following three broad categories: landslide
detection and mapping, landslide spatial forecasting in the form of susceptibility mapping,
and landslide temporal forecasting. In this paper, we present a brief overview of ML tech-
niques, provide a general summary of the landslide studies conducted, in recent years, in
the three above-mentioned categories, and make an attempt to critically evaluate the use of
ML methods to model landslide processes. The paper also provides suggestions for future
use of these powerful data-driven techniques in landslide studies.

Keywords AI · ML · Landslide detection · Spatial forecasting · Temporal forecasting ·
Data-driven analysis

1 Introduction

Landslides are the gravity driven motion of a mass of rock, soil and debris down a slope,
and they can cause significant fatalities and economic losses. According to the World
Bank, about 3.7 million square kilometers of inland area on earth is prone to landslide

✉ Michele Calvello
   mcalvello@unisa.it

✉ Zhongqiang Liu
   Zhongqiang.Liu@ngi.no

1   Deltares (formerly), Delft, The Netherlands

2   University of Salerno, Fisciano, SA, Italy

3   Norwegian Geotechnical Institute, Oslo, Norway

4   Hong Kong University of Science and Technology, Hong Kong, China

activities (Dilley et al. 2005), which affects the life of more than 300 million people. The fatality count of landslides between 2005 and 2016 sums up to 56,000, worldwide (Froude & Petley 2018). These facts alone show the importance of having reliable methods for landslide analysis. Particularly when it comes to the prediction of landslides, traditional physically-based methods of landslide analysis may not be sufficient for predicting the location and time of the mass movements. This is because of the high demand on real-time computations and the complexity of these phenomena, which makes physics-driven methods developed from established laws in geology, geotechnics, hydrology, and meteorology, not always able to tackle all the known and unknown factors affecting mass movements.

In recent years, with the rapid development of machine learning (ML) as a Data Science branch, and its spread over many engineering fields, many researchers have started looking into disciplinary or thematic applications of ML methods. For instance, in the geotechnical engineering community, the growing interest in ML is testified by the recent establishment of a specific technical committee (TC) of the International Society of Soil Mechanics and Geotechnical Engineering ISSMGE (TC309 "Machine Learning and Big Data") since 2018, and by many other initiatives, such as the ISSMGE TC304/309/210 Machine Learning Dialogue for Geotechnics 2019 (ISSMGE Bulletin February 2020: page 17). In this context, and also considering the increasing availability of observational data, such as remote sensing satellite data, landslide studies adopting ML algorithms and methods have been emerging, with increasing frequency, in the literature. Given the capabilities of ML and its derivatives, such as deep learning (DL), in handling large datasets and finding complex patterns hidden in the data, landslide researchers and practitioners have demonstrated that ML/DL can effectively be used in landslide-related studies. This has been the case in many research studies on landslide detection and mapping (e.g., Stumpf & Kerle 2011; Keyport et al. 2018; Prakash et al. 2020), landslide susceptibility mapping (e.g., Pourghasemi and Rahmati 2018; Chen et al. 2019; Mergadi et al. 2020) and temporal forecasting of landslides (e.g., Yoon et al. 2011; Huang and Xiang 2018; Stanley et al. 2020).

The objective of this paper is to present the most recent advances in the application of ML for landslide studies, and to discuss the challenges and opportunities provided by ML/DL for landslide researchers. To these aims, after an initial overview of ML methods, we present the ML-related researches that have been carried out in the following three areas: landslide detection and mapping, landslide spatial forecasting, and landslide temporal forecasting. We then provide a critical discussion about the use of ML methods in landslide studies, and we wrap up the paper with a conclusion section.

2 Machine learning

2.1 Background

ML algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. The term *Machine Learning* is attributed to Arthur Samuel, a pioneer in the field of computer gaming and artificial intelligence, who coined it in 1959. In his article (Samuel 1959), he said "*Two machine-learning procedures were investigated using the game of checkers. The main idea was that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program. Furthermore, it can learn to do this in a remarkably short period of time … when*

given only the rules of the game…and a redundant and incomplete list of parameters which are thought to have something to do with the game, but whose correct signs and relative weights are unknown and unspecified." Mitchell (1997) provides a formal definition of ML as follows: "*A computer program is said to learn from experience E with respect to some class of tasks T, and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.*"

One of the presumed traits of ML in comparison with humans is its comparable or superior predictive and decision-making performance. What a typical "learning machine" does is finding a rule that, when applied to a collection of inputs, produces the desired outcome. This rule also generates the correct outcome for most other inputs (distinct from the training data) on the condition that those inputs come from the same or a similar statistical distribution as the one the training data was drawn from. It can be argued that such a process is not necessarily learning (Burkov 2019) in the way humans learn, because if the inputs are slightly changed, the outcome can become completely wrong. For instance, if a machine learning algorithm is trained by "looking" at landslide images in vegetated areas, unless it is also trained to recognize landslides in bare lands, it may fail to identify such landslides. Therefore, it is reasonable to conclude that, as of today, machine learning cannot outperform humans in many fields. However, given the pace of ML/DL advances, it is probable that in the future it will revolutionize the processes of prediction and decision making, and also remarkably influence practice and research in landslide risk assessment and management.

2.2 Prediction versus explanation

There has been a long debate among statisticians about the scientific value of predictive models versus explanatory and descriptive models (e.g., Geisser 1975; Wallis 1980; Breiman 2001; Parzen 2001; Feelders 2002; Shmueli 2010). This debate has been further intensified by the emergence of ML techniques in the computer science community as powerful predictive methods compared to classical statistics-based methods. As a result, according to Breiman (2001), it can be argued that there are, at least, two cultures in data-driven analysis, namely, data modeling and algorithmic modeling, with the former aiming to gain information from data in order to predict, and the latter treating the data mechanism as unknown and only aiming at maximizing the accuracy of the predictions. As inferred from Shmueli (2010) in her thorough discussion on the difference between explanation and prediction, data modeling as explanatory methods aim to provide the truth, whereas algorithmic modeling as predictive methods aim to provide the reality based on the available data. Most ML methods fall primarily on the side of algorithmic modeling, and therefore prediction. Given these distinctions, hereinafter, we adopt the following criteria to qualify landslide studies as ML-based studies, and to distinguish them from statistics-based studies.

- Accurate prediction is the main goal of the study. Therefore, we deliberately excluded explanatory models, such as linear statistics-based methods that are used for statistical inference.
- Data are divided into training and testing datasets, and evaluating the trained model on the testing dataset is the major method for assessing the performance of the model.

2.3 Conventional machine learning versus deep learning

In general, DL can be regarded as a subset of ML. The main distinction between DL and conventional ML algorithms falls in the modality of learning from data. Besides, DL is primarily (so far) based on artificial neural networks (ANN), whereas conventional ML methods include algorithms other than, and as well as, ANN.

In conventional ML algorithms, labeled or unlabeled data come along with certain features or attributes. It might be necessary that the analyst reduces or increases these features, depending on data quantity and the utilized ML algorithm. Through the training process, a conventional ML algorithm learns to find patterns in the data based on its available features.

In DL, the input data (e.g., image, text, video or time series) is directly sent to artificial neural networks, where each network hierarchically learns specific features of the input data. The learned features are then used to find a pattern that associates the input data to a specific label, to a distinct category or to a decision. In general, DL algorithms typically require more data for training than conventional ML algorithms, given their higher number of hyperparameters.

2.4 Learning methods

Supervised, unsupervised (semi-supervised can be seen as a mix of the two), and reinforcement learning are the major ML methods.

Supervised learning algorithms are used on data that consist of a set of inputs (predictors, independent variables or features) and their corresponding outputs (target variables or labels). Training on input variables and target variables, the machine learns how to map inputs to corresponding outputs. The training process continues until the model achieves a desired level of accuracy on the training data. The validity of the model is assessed by evaluating the model on unseen data (test set). Examples of supervised learning algorithms include Decision Trees and Trees ensembles (e.g., Random Forest, Gradient Boosting algorithms such as AdaBoost, XGBoost, etc.), support vector machines and artificial neural networks including multi-layer perceptron neural nets and supervised DL algorithms.

In unsupervised learning, there is no target or label variable to predict and the goal is to "make sense" of the data. A conventional use of unsupervised learning is clustering populations of data in different groups for specific interventions. Examples of this type of unsupervised learning algorithms are hierarchical clustering, K-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN). DL algorithms can also be used for unsupervised learning. Generative DL algorithms such as autoencoders and generative adversarial networks (GANs) belong to this group.

In reinforcement learning (RL), the machine is trained to make specific decisions. An agent is exposed to an environment where it takes actions to maximize the cumulative reward (usually in episodic problems and is called return) or average reward (in continuing problems) that concerns multiple steps ahead. The agent learns from past experience and tries to capture the best possible knowledge to make accurate decisions. Markov Decision Process and Deep Reinforcement Learning are some examples of RL methods.

3 Detection and mapping

3.1 Background

Landslide detection or mapping refers to the delineation of landslide-affected areas, which include the source and the deposition zones of the moving soil or rock mass. Landslide detection is an important part of emergency response to extreme events, such as extreme rainfall and strong earthquakes, to identify hazardous areas affected by slope failure, where field surveys can be expensive, cumbersome, dangerous and involve access difficulties. Landslide detection is also useful for building landslide geomorphological inventories (historical, event-based, seasonal or multi-temporal), which help understanding the causal factors of past landslides (Guzzetti et al. 2012), and can help monitoring, predicting and mitigating future landslides. Before the widespread use of satellite remote sensing data, landslide detection was essentially done using visual inspection of aerial photographs or field surveys, a time consuming and expensive process. Mondini et al. (2011) estimated that the manual production of an event-based landslide inventory requires about 5 days per person per square kilometre, including interpretation of aerial photographs, field surveys, digitization of information and creating a geographical database. More information on landslide detection methods can be found in the review paper by Guzzetti et al. (2012).

In the past decade and by provision of high volumes of medium to high resolution satellite and airborne-based imagery, ML techniques have become attractive choices for landslide detection. The main goal in the application of ML algorithms for landslide detection is to enable the machine to detect landslide features, such as scarp and run-out track, in a similar way to humans finding these features in a set of images. This is possible primarily because a landslide makes a contrast, especially in vegetated areas, with the surrounding area by exposing fresh rock and soil, causing local change in brightness of the image. Therefore, what ML algorithms aim to achieve is a human-level capability in landslide feature detection. It is noted that landslide detection, using remote sensing data does not inevitably mean the use of ML methods, and indeed many remote sensing data are currently analysed and processed using manual and rule-based methods that require greater involvement of domain experts and setting area specific thresholds.

Identifying and reviewing 55 scholarly papers in ML-related landslide detection published over the past two decades (2007–2021) reveals that the application of ML methods in landslide detection in the literature has increased particularly in the past five years, as shown in Fig. 1.

As shown in Fig. 1a, conventional ML methods (CML), in form of pixel-based and object-based landslide detection, were more popular in the early landslide detection studies (e.g., Borghuis et al. 2007; Danneels et al. 2007; Chang et al. 2007, 2010; Gong et al. 2010; Martha et al. 2011; Stumpf & Kerle 2011; Van Den Eeckhaut et al. 2012), whereas in the past few years the interest has grown towards using DL methods (e.g., Ding et al. 2016; Chen et al. 2018a, b; Ghorbanzadeh et al. 2019a,b, c; Can et al. 2019; Bui et al. 2020a, b; Prakash et al. 2020), with some studies comparing different methods for the same test area. It should be noted that CML in landslide detection can be supervised, unsupervised or combination of the two methods. DL methods in landslide detection (up to the time of this literature review) fall primarily under the category of supervised methods.

Landslide detection is typically done either using change detection between pre- and post-landslide images or solely using feature detection in post-landslide images. In both

(a)

(b)
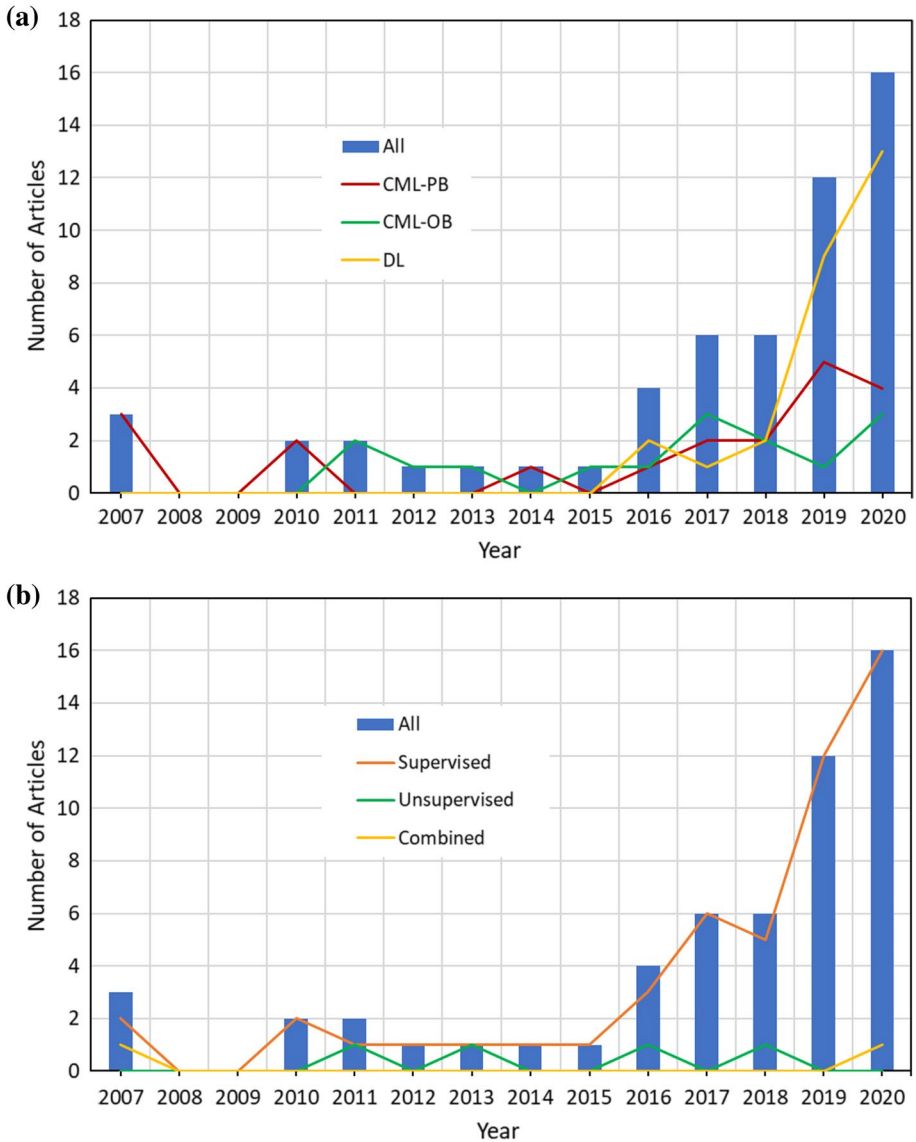
Fig. 1 Trend in application of ML algorithms in landslide detection studies (**a**) use of conventional ML in pixel-based methods (CML-PB), conventional ML in object-based methods (CML-OB) and DL methods, (**b**) use of supervised, unsupervised and combination of the two methods

cases, as shown in Fig. 1b, ML algorithms have been applied considering CML and DL as supervised learning methods (e.g., Danneels et al. 2007; Chang et al. 2010; Stumpf & Kerle, 2011; Chen et al. 2014; Pawłuszek et al. 2017; Mora et al. 2018; Chen et al. 2019; Prakash et al. 2020) and, to a lesser extent, unsupervised learning methods (e.g., Martha et al. 2011; Li et al. 2016; Keyport et al. 2018) or combined supervised and unsupervised learning (e.g., Borghuis et al. 2007; Fang et al. 2020). In Fig. 1, number of "All" articles in
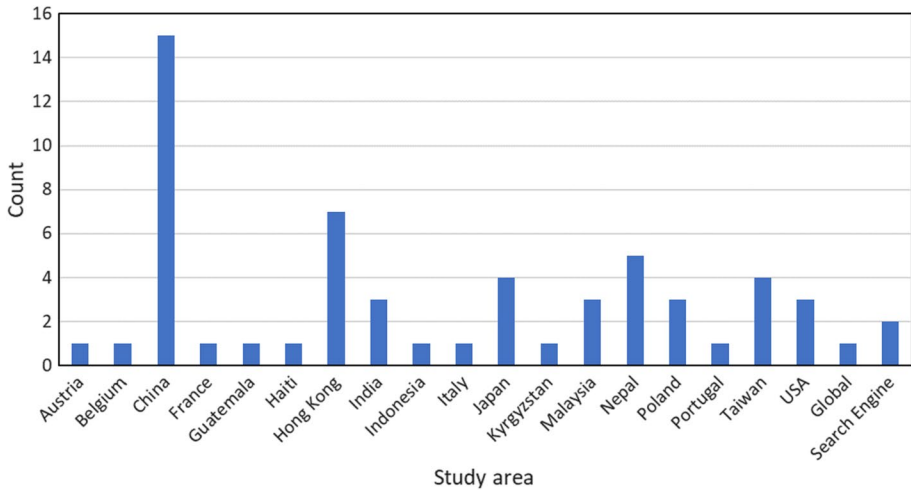
Fig. 2 Countries of the case study areas for ML-related landslide detection studies

each year is not necessarily the same as the summation of articles in each ML category as some papers consider multiple ML methods.

The identified studies have been performed in various countries across the globe as shown in Fig. 2, with China and Hong Kong being the geographical areas with most case studies. In 3 cases (i.e., Global and Search Engine), the studies used landslide data across the globe or from landslide images collected using search engines to train landslide detection algorithms.

Given the spatial extent of landslides, remote sensing technologies, including Earth Observation satellites and airborne sensors mounted on aircraft and unmanned aerial vehicles (UAV), are widely used for landslide detection. These technologies result in various data sources which typically involve medium to high and very high resolution optical, multispectral, LiDAR (light detection and ranging) and radar data. These include airborne LiDAR DEM data (e.g. Van Den Eeckhaut et al. 2012; Pawluszek-Filipiak & Borkowski 2020; Prakash et al. 2020), UAV-based optical imagery (e.g. Lei et al. 2019a, b; Catani 2021), satellite-based Synthetic Aperture Radar (SAR) data (e.g. Kamiyama et al. 2018; Mabu et al. 2020), satellite-based medium resolution multi-spectral data (e.g. Prakash et al. 2020) and satellite-based high (e.g. Bacha et al. 2020; Tavakkoli-Piralilou et al. 2019) and very high resolution multi-spectral data (e.g. Cheng et al. 2013).

3.2 Methods

Landslide detection methods are a special application of the characterization of land cover, and its change, for which there has been increasing scientific and practical interest in the remote sensing community. These methods, in general, fall within two interrelated categories: pixel-based and object-based methods.

Pixel-based landslide mapping examines each pixel in the single-band or multi-band image and determines whether it belongs to a landslide or not. This is done first by treating all input features (e.g., morphological and spectral features) as raster layers (bands), which
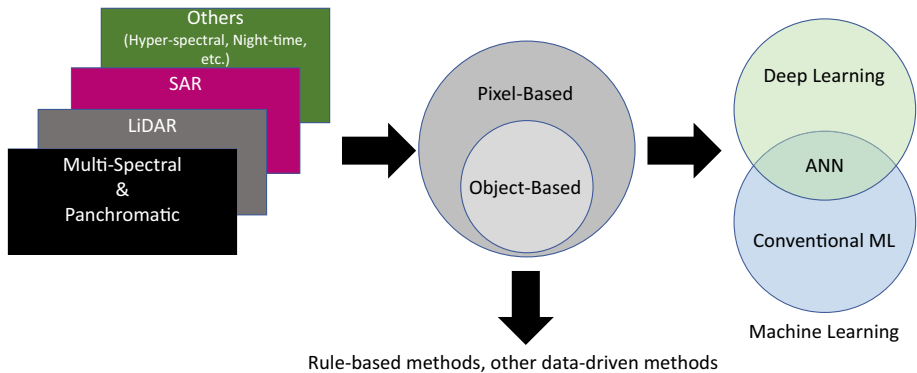
Fig. 3 Pixel-based and object-based landslide detection methods

are co-registered and re-sampled to a chosen resolution. Then feature values are extracted at given pixels and further examined to decide whether or not they represent features of a landslide.

In object-based methods, which can be seen as a subset of pixel-based methods, the spatial connection of neighbouring pixels is used to identify objects, in single-band or multi-band images, and then the objects are examined for determining if they are landslide or non-landslide segments.

Figure 3 illustrates the landslide detection and mapping methods using CML and DL as well as ruled-based and other data-driven approaches. In all these cases, change detection and feature detection methods can be used based on pre- and post-landslide images or only post-landslide images.

3.3 Pixel-based methods using CML

Pixel-based landslide detection is performed with pixels as input. In a digital image, a pixel is the basic constituent element. In general, pixel-based methods often require extensive parametric tuning and precise geometrical correction or co-registration to be applicable to large areas (Sameen & Pradhan 2019).

Based on the published works in this area, pixel-based CML covers a range of studies that use both supervised and unsupervised methods. In pixel-based supervised classification, the pixels are labelled as landslide or non-landslide by the landslide experts, and then the labelled pixels along with the corresponding signatures from bands of input images (see the left-hand side of Fig. 3) are used to train ML algorithms. Besides direct change detection methods, some authors (e.g., Si et al. 2018) used susceptibility analysis as the basis for landslide change detection by using areas with high susceptibility as candidates for applying the derived change detection thresholds for identifying new landslides.

Unsupervised classification is typically used to cluster pixels in a dataset based on their similarity with other pixels, without any user-defined label. The main limitation of unsupervised classification is that the output needs to be interpreted and manually assigned a label. K-means clustering, Gaussian Mixture Models (GMM), Markov Random Field and hierarchical clustering are the unsupervised learning algorithms in landslide detection studies considered herein (Martha et al. 2011; Cheng et al. 2013; Li et al. 2016; Keyport et al. 2018).

Table 1 Summary of pixel-based landslide detection studies using CML methods

Study	Year	Main method*	ML/DL algorithms**	Main image type	Resolution	Evaluation[+]	Case study country
Borghuis et al.	2007	CML-PB-C	MLC	Multi-spectral	2.5 m	CE, OE	Taiwan
Chang et al.	2007	CML-PB-S	GPBF, SVM, kNN	Multi-spectral	10 m	Acc	Taiwan
Danneels et al.	2007	CML-PB-S	MLC, MLP-NN	Multi-spectral	15 m	CE, OE	Kyrgyzstan
Chnag et al.	2010	CML-PB-S	MLP-NN, SVM, MLC	Multi-spectral	10 m	Vis	Taiwan
Gong et al.	2010	CML-PB-S	SVM	Multi-spectral	2.5 m	Acc, KC	China
Chen et al.	2014	CML-PB-S	RF	LiDAR	3.15 pt/sq.m	Acc	China
Li et al.	2016	CML-PB-U	GMM, HC	Red Green Blue RGB (aerial)	0.5 m	Com, Cor, Qual	Hong Kong
Pawluszek & Borkowski	2017	CML-PB-S	SVM	LiDAR	4–6 pt/sq.m	Acc, KC	Poland
Pawluszek et al.	2017	CML-PB-S	FFNN, MLC	LiDAR	4–6 pt/sq.m	Acc, KC	Poland
Keyport et al.	2018	CML-PB-U	K-means	RGB (aerial)	0.5 m	Acc, Rec, Prec, F	Guatemala
Chen et al.	2019	CML-PB-S	RF	Multi-spectral, night-time	30 m	Acc, Rec, Prec	Taiwan
Hu et al.	2019	CML-PB-S	SVM, RF, MLP-NN	Multi-spectral	30 m	Rec, Prec, F, AUC	China
Chen et al.	2020	CML-PB	MLP-NN, NBC, SVM	RGB (aerial), multi-spectral	n.a., 0.3–1.21 m	TPR, FNR, FPR	China
Pawluszek-Filipiak & Borkowski	2020	CML-PB-S, CML-OB-S	RF	Multi-spectral, LiDAR	10 m, 4–6 pt/sq.m	Acc, F, Rec, other confusion matric metrics	Poland

*CML conventional machine learning, OB object-based, PB pixel-based, S supervised, U unsupervised

**Check Appendix 1 for full names

[+]Check Appendix 2 for full names

Combining unsupervised and supervised learning methods, Borghuis et al. (2007) used a two-step unsupervised approach for landslide detection in Taiwan following typhoons Mindulle and Aere in 2004. First, they used K-means clustering for deducing spectral signatures of pixels of optical satellite images (5-m resolution SPOT-5) and then used supervised classification with Maximum Likelihood Classifier (MLC) for classification of K-means labels. Furthermore, Borghuis et al. (2007) used MLC for classification of manually labelled pixels containing spectral features from optical images and associated DEM for landslide detection.

Being a classification problem, almost all supervised CML studies in landslide detection use metrics derived from the confusion matrix for evaluating the performance of the models on test sets. Kappa coefficient and Area Under the Receiver Operating Characteristic curve (AUC) have also been used for model performance evaluation. Table 1 summarises the main features of pixel-based CML studies for landslide detection (see Appendix 1 for meaning of acronyms). It should be noted that this list only considers studies that used CML as the primary approach for pixel-based landslide detection and do not include studies where CML pixel-based methods were used for comparison with other methods.

3.4 Object-based methods using CML

Object-based methods using CML fall within the framework of Object-Based Image Analysis (OBIA) that includes two major steps: (1) image segmentation, and (2) classification of the emerged segments. ML methods can be applied in both steps. While OBIA offers extra features to distinguish landslides from other objects, it needs to optimize segmentation parameters (e.g. scale) (Myint et al. 2011) and thus the degree of automation is low compared to pixel-based methods (Sameen & Pradhan, 2019). CML methods combined with an OBIA framework have drawn the attention of the geo-informatics community for landslide detection. To this aim, both supervised and unsupervised CML have been used, and sometimes combined, under the OBIA framework.

Supervised CML methods have been frequently used in segment classification of OBIA for landslide detection. While the segmentation step in OBIA is typically performed using multi-resolution methods implemented in commercial geo-spatial analysis packages, some authors used unsupervised learning algorithms for segmentation of optical remote sensing images. For instance, in a two-step OBIA-based landslide detection in India using 5.8-m resolution multi-spectral satellite data, Martha et al. (2011) adopted K-means clustering for objective thresholding of multi-resolution image segmentation before running classification on the final segments.

The use of both supervised and unsupervised learning methods in OBIA was studied by Cheng et al. (2013), who suggested an object-based framework built on computer vision (Bag of Visual Words, BoVW) and text mining methods (probabilistic latent semantic analysis, pLSA) for detecting landslides. At the heart of these methods were K-means clustering in BoVW for clustering the pixels into visual words and kNN in pLSA. They trained and tested this approach for an area in China using 1-m resolution multi-spectral satellite data (Geoeye-1). Table 2 summarises the main features of object-based CML studies for landslide detection.

Figure 4 illustrates the difference of pixel-based and object-based methods with regard to the application of CML algorithms.

Table 2 Summary of object-based landslide detection studies using CML methods

Study	Year	Main method*	ML/DL algorithms**	Main image type	Resolution	Evaluation[+]	Case study country
Martha et al.	2011	CML-OB-U	K-means	Multi-spectral	5.8 m	Acc	India
Stumpf et al.	2011	CML-OB-S	RF	Panchromatic, multi-spectral, RGB	0.5–1 m, 2–4 m, 0.5 m	Acc	Haiti, China, France, Italy
Van Den Eeckhaut et al.	2012	CML-OB-S	SVM	LiDAR	0.05 pt/sq.m	Acc	Belgium
Cheng et al.	2013	CML-OB-S/U	kNN, K-means	Multi-spectral	1 m	Acc, TPR, FPR	China
Li et al.	2015	CML-OB-S	RF, SVM	LiDAR	3.15 pt/sq.m	Acc	China
Heleno et al.	2016	CML-OB-S	SVM	Panchromatic, multi-spectral	0.5 m, 4 m	Acc, KC, FNR, FPR	Portugal
Chen et al.	2017	CML-OB-S	RF	Panchromatic, multi-spectral	2.1 m, 5.8 m	Com, Cor, Qual	China
Mayr et al.	2017	CML-OB-S	SVM, RF	LiDAR	500 pt/sq.m	Rec, Prec, F	Austria
Rajbhandari et al.	2017	CML-OB-S	RF	Multi-spectral	5.8 m	Acc, KC	India
Chen et al.	2018	CML-OB-S	RF	Multi-spectral	30 m	CE, OE	Nepal
Mezaal et al.	2018	CML-OB-S	SVM, RF	LiDAR	8 pt/sq.m	Acc, KC, Rec, Prec, F	Malaysia
Mora et al.	2018	CML-OB-S	SVM	LiDAR	3–5 pt/sq.m	n.a	USA
Tavakkoli Piralilou et al.	2019	CML-OB-S	MLP-NN, RF, LR	Multi-spectral	3 m	Rec, Prec, F, KC	Nepal
Miandad et al.	2020	CML-OB-S	SVM	LiDAR	1–5 m [++]	Acc, KC	USA
Wang et al.	2021b	CML-OB-S	LR, SVM, RF, Discrete AdaBoost, LogitBoost, Gentle AdaBoost, CNN	LiDAR	2 m	Acc, Prec, Rec, TNR, FPR, F, MCC	Hong Kong, China

CML conventional machine learning, *OB* object-based, *S* supervised, *U* unsupervised

**Check Appendix 1 for full names

[+]Check Appendix 2 for full names

[++]Resolution of the resulting DEM from LiDAR point cloud

Pixel-based method   Object-based method
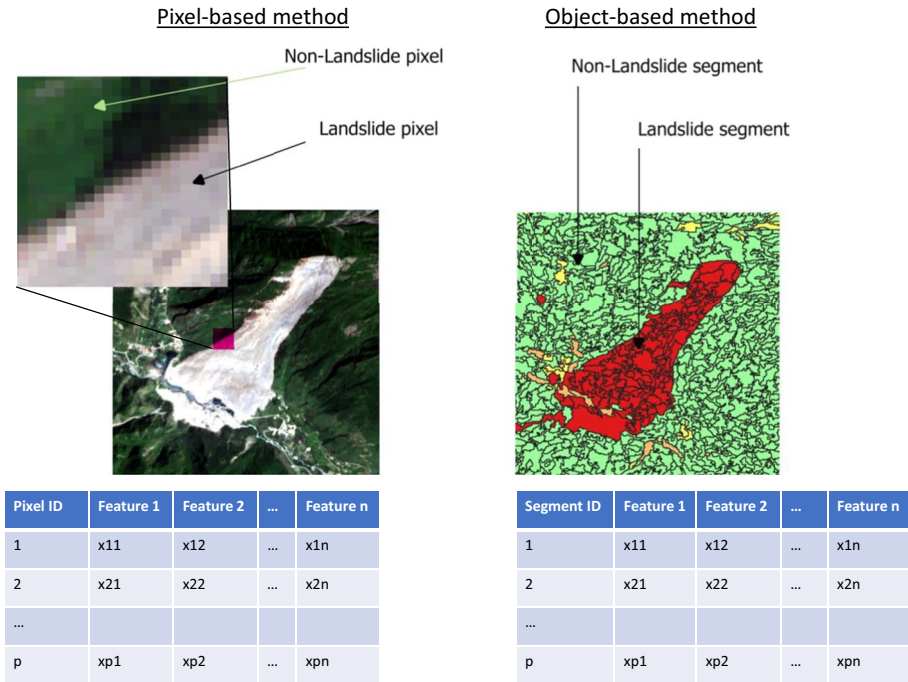
Fig. 4 Pixel-Based and Object-Based landslide detection for CML applications (modified after Tehrani et al. (2021))

3.5 DL methods

DL methods are mostly used in the context of computer vision in landslide detection studies. Unlike conventional ML methods, DL methods do not need extensive feature engineering in the preparation of the training dataset. However, in general, DL methods require more training data than conventional ML methods, given the higher number of model variables (thousands to millions) that need to be fit. This limitation is typically taken care of by data augmentation methods that involve rotation and flipping of the original images. In our literature review, we identified one work that used DL in an application outside computer vision. In this work Mezaal et al. (2017) combined fuzzy-based image segmentation (object-based) with Multi-Layer Perceptron (MLP) and Recurrent Neural Networks (RNN), a DL method, for landslide detection in Cameron Highlands in Malaysia. They used a LiDAR point cloud with a point density of 8 points/m^2 to derive 0.5 m resolution DEM for acquiring topographic features.

In computer vision applications, DL methods can be categorized into three groups, which by order of complexity are: (1) image classification, (2) object detection, (3) semantic segmentation. In image classification, the goal is to find the label of the image (e.g., landslide or non-landslide). In object detection, the aim is to identify and locate the objects that are present in an image, with the help of bounding boxes. Image semantic segmentation further moves forward, by trying to find out accurately the exact boundary of the objects in the image. In semantic segmentation, each pixel in an image is assigned to a certain class, and hence this can be thought of as a classification problem per pixel. DL-based

semantic segmentation in landslide detection is mainly about binary semantic segmentation of images at pixel level. Based on the identified landslide detection studies, it can be inferred that landslide detection using deep learning has been performed primarily either as an image classification (whole image or image patches) or as semantic segmentation.

3.5.1 DL for image classification

Image classification in landslide detection is mainly limited to classifying images as landslides or non-landslides. In DL-based image classification, it is very customary to use well-known algorithms pre-trained on massive datasets, for classification of images that are not found in those datasets. For instance, Catani (2021) used four pre-trained top performer CNN algorithms using transfer learning to train a general-purpose landslide detection from UAV and ground-based RGB (Red–Green–Blue bands) photographs found through search engines. The four pre-trained algorithms were: GoogLeNet (Szegedy et al. 2015); GoogLeNet-Places365 (Zhou et al. 2018b), a modified version of GoogLeNet specifically oriented towards the classification of the scene rather than single objects; ResNet.101, a 101-layer CNN with improved training curve based on residual learning (He et al. 2015); and Inception.V3, for classification of multi-purpose images in near real time (Szegedy et al. 2016). Table 3 summarises the main features of DL-based image classification studies for landslide detection.

3.5.2 DL for patch-wise image classification

Patch-wise image classification involves splitting the original image into multiple square patches with a width much smaller than the original image width/height and then labelling these patches as landslide or non-landslide to be used for training a CNN model or a variant of it. Once the CNN model is trained on landslide and non-landslide patches, it can be run on patches of an image and each patch can be labelled. All recognized patches put together show the extent of the landslide.

In one of the first applications of DL in landslide detection, Ding et al. (2016) performed it using patch-wise image classification (patch size = 28 pixels) for landslides occurred in 2015 in Shenzen, China. In a more recent work, Ghorbanzadeh et al. (2019b) compared machine learning methods ANN, SVM and RF (pixel-based) with different CNN-based patch-wise classification for landslide detection in Rasuwa district in Nepal. For CNNs, they used multiple square window (patch) sizes 12, 16, 22, 32, and 48 pixels in an image classification framework and found that in general smaller window size resulted in more accurate results. Their conclusion was that CNNs did not automatically outperform ANN, SVM and RF, and that the performance of CNNs strongly depended on their design, i.e., layer depth, input window sizes and training strategies. Table 4 summarises the main features of these studies.

3.5.3 DL for semantic segmentation

Semantic segmentation methods rely on pixel-wise classification. Semantic segmentation using innovative CNN architectures has gained momentum in the past few years. An example of DL methods for semantic segmentation is a fully convolutional network (FCN)

Table 3 Summary of DL image classification methods for landslide detection

Study	Year	Main method*	ML/DL algorithms**	Main image type	Resolution	Evaluation[+]	Study country
Can et al.	2019	DL-IC	CNN	RGB	n.a	Rec, Prec, F	Search engine
Bui et al.	2020	DL-IC	MobileNet.V2, AlexNet, VGGNet16, ResNet50, Inception.V3, MobileNet	Multi-spectral	10–30 m	Acc	Global
Catani	2021	DL-IC	GoogLeNet, GoogLeNet-Places365, ResNet.101, Inception.V3	RGB	n.a	Acc, Rec, Prec, F, other Confusion matrix metrics	Search engine
Ji et al.	2020	DL-IC	ResNet18,50,101,152, DenseNet121,162,201, VGG13,16,19, Inception-v3	Panchromatic, multi-spectral	0.8 m, 3.2 m	Acc, Rec, Prec, F	China

* *DL* conventional machine learning, *IC* image-classification

** Check Appendix 1 for full names

[+] Check Appendix 2 for full names

Table 4 Summary of DL patch-wise image classification methods for landslide detection

Study	Year	Main method*	ML/DL algorithms**	Main image type	Resolution	Evaluation+	Study country
Amit et al.	2016	DL–PaIC-S	CNN	RGB (aerial), multi-spectral	n.a	Rec, Prec, F	Japan
Ding et al.	2016	DL–PaIC-S	CNN	Multi-spectral	8 m	Rec, QP, CE	China
Chen et al.	2018	DL–PaIC	CNN	Multi-spectral	8–30 m	Rec, QP, CE	China
Kamiyama et al.	2018	DL–PaIC	AlexNet	SAR	10 m	Rec, Prec	Japan
Ghorbanzadeh et al.	2019	DL–PaIC	CNN	Multi-spectral	5 m	mIOU	Nepal
Ghorbanzadeh et al.	2019	DL–PaIC	CNN	RGB (aerial), multi-spectral	0.12 m	Rec, Prec, F, mIOU, OPR, UPR	India
Ghorbanzadeh et al.	2019	DL–PaIC, CML–PB–S	CNN, MLP-NN, SVM, RF	Multi-spectral	6 m	Rec, Prec, F, mIOU	Nepal
Sameen & Pradhan	2019	DL–PaIC	CNN, ResNet with DT	RGB (aerial), LiDAR	10 m, 8 pt/sq.m	Acc, F, mIOU	Malaysia
Wang et al.	2019	DL–PaIC, CML–PB–S	CNN, SVM, GBM, NBC	Panchromatic, multi-spectral	2 m, 8 m	Acc, Rec, Prec, F	China
Ye et al.	2019	DL–PaIC, CML–PB–S	DBN with LR, SVM	Hyperspectral	30 m	Acc, KC	China
Mabu et al.	2020	DL–PaIC	CNN	SAR	3 m	Acc, Rec, Prec, F	Japan

DL conventional machine learning, *PaIC* patch-wise image-classification, *CML* conventional machine learning, *PB* pixel-based, *S* supervised

**Check Appendix 1 for full names

+Check Appendix 2 for full names

(Long et al. 2015) which uses a convolutional neural network to transform image pixels to pixel categories. Unlike the convolutional neural networks previously introduced, FCN transforms the height and width of the intermediate layer feature map back to the size of input image through the transposed convolution layer, so that the predictions have a one-to-one correspondence with input image in spatial dimension (height and width). Given a position on the spatial dimension, the output of the channel dimension will be a category prediction of the pixel corresponding to the location.

U-Net is another popular CNN architecture used for semantic segmentation since it requires fewer images for training compared to conventional CNN architecture with multiple consecutive layers. Konishi & Soga (2019) used U-Net for landslide detection from pre- and post-event SAR images of the 2018 Hokkaido Eastern Iburi earthquake in Japan. They used input images of 256 pixels by 256 pixels and showed that their approach reached a better performance compared to threshold-based SAR image analysis.

To compare DL with conventional ML algorithms, Prakash et al. (2020) implemented deep learning semantic segmentation, OBIA and pixel-based algorithms for spatial mapping of hillslope landslides in the State of Oregon, USA. They used high resolution LiDAR-based DEM and Near-Infrared band of Sentinel-2 post-landslide data. The deep learning algorithm used was based on U-Net CNN with ResNet blocks, which was used for semantic segmentation and subsequent classification. Prakash et al. (2020) confirmed the observation by Ghorbanzadeh et al. (2019b) about different ML algorithms and showed that all the three methods were able to map the landslides in the testing area (with about 80% accuracy but lower recall scores), with the DL methods performing slightly better than the other two conventional methods. Other studies that used U-Net and ResNet architecture include Qi et al. (2020)*,* and Liu et al. (2020a, b, c*).*

In regions that undergo land changes other than only landslides, it is difficult to separate landslides from other land changes. To address this limitation inherent to conventional approaches for landslide detection, Fang et al. (2020) used GAN-based Siamese framework (GSF) for landslide inventory mapping. The framework comprised two cascading modules, namely, a domain adaptation module based on conditional GANs and a landslide detection module based on Siamese neural network. The domain adaptation module aims to make a cross-domain mapping between pre-landslide and post-landslide images with adversarial learning to generate a pre-landslide image as close as possible to the post-landslide image in terms of contextual image properties (lighting, atmospheric conditions, etc.) at the time of the post-landslide image. It was designed to retain only changes due to landslide activities in the generated image. The detection module aims to perform pixel-level landslide detection on the pairs of generated pre-landslide image and original post-landslide image with a Siamese neural network model. The Siamese network is used to generate an output image that reflects how similar are the pair of input images, thus identifying and detecting landslide regions. Table 5 summarises the main features of semantic segmentation studies for landslide detection.

4 Spatial forecasting

Literature studies addressing landslide spatial forecasting estimate where future landslides are likely to occur in a target region, without considering when or how frequently they will occur. In other words, data-driven methods including ML algorithms are frequently used to compute landslide susceptibility, i.e. the "likelihood of a landslide occurring in a given

Table 5 Summary of DL semantic segmentation methods for landslide detection

Study	Year	Main method*	ML/DL algorithms**	Main image type	Resolution	Evaluation+	Study country
Konishi & Suga	2019	DL-PSS	U-Net	SAR	10 m	Rec, Prec, F	Japan
Lei et al.	2019	DL-PSS	FCN-PP, CNN, FCN, U-Net	RGB (aerial)	0.5 m	Acc, Rec, Prec, F	Honk Kong
Fang et al.	2020	DL-PSS, CML-PB	GSF, GAN, FCN-PP, DSCN, CNN, SVM	Multi-spectral	0.5 m	Acc, Rec, Prec, F	Hong Kong
Liu et al.	2020	DL-PSS	U-Net, U-Net+ResNet	RGB (aerial)	0.47 m	Acc, Rec, Prec, F, mIOU	China
Lv et al.	2020	DL-PSS	DP-FCN, FCN, FCN-PP, U-Net	RGB (aerial)	0.5 m	Com, Cor, Qual, Acc	Hong Kong
Prakash et al.	2020	DL-PSS, CML-PB–S, CML-OB-S	U-Net+Res-Net, RF, ANN, LR	Multi-spectral, LiDAR	10 m, 1 m	Acc, F, Rec, other con-fusion matric metrics	USA
Qi et al.	2020	DL-PSS	U-Net, U-Net+ResNet	Panchromatic, multi-spectral	0.5 m, 1.65 m	Acc, Rec, Prec, F	China
Shi et al.	2020	DL-PSS	U-Net+Res-Net, CNN	RGB (aerial)	0.5 m	Rec, Prec, F	Hong Kong
Su et al.	2021	DL-PSS	U-Net, CNN	RGB (aerial)	0.5 m	Acc, Rec, Prec, F, mIOU	Hong Kong
Yu et al.	2020	DL-PSS	contour-based PSPNet	Multi-spectral	30 m	Rec, Prec, F	Nepal
Zhang et al.	2020	DL-PSS	U-Net, MIL	RGB	n.a	Rec, Prec, F	Hong Kong

*DL conventional machine learning, PSS pixel-wise semantic segmentation

**Check Appendix 1 for full names

+Check Appendix 2 for full names

area" (Brabb, 1984), relying on two standard key assumptions, common to statistically-based and ML approaches developed for landslide susceptibility analysis and zoning (Varnes, 1984; Reichenbach et al. 2018; Lombardo et al. 2020): (i) future landslides are more likely to occur under conditions that led to slope instabilities in the past; (ii) conditions that are directly or indirectly linked to slope failures can be collected and used to build predictive models of landslide spatial occurrence. Differently from statistical analyses, ML algorithms are able to learn the association between landslide occurrences and landslide conditioning factors without necessarily assuming a structural model in the data. The learning aspect of these methods is to develop sequences of commands or algorithms that search, in a process of iterative and gradual refinement, for associations in the data that basic descriptive statistics and the human eye may not readily detect as such (Korup and Stolle 2014).

A very recent overview of the most popular machine learning techniques available for landslide susceptibility studies is presented by Mergadi et al. (2020), who also state that "*only a handful of researchers use machine learning techniques in landslide susceptibility mapping studies*." Indeed, they identify ten authors who are responsible for approximately 47% of published landslide susceptibility studies adopting neural networks, 70% of studies adopting random forest (RF) algorithms, 83% of studies adopting decision tree (DT) algorithms, and 86% of studies adopting support vector machines (SVM) algorithms. This finding prompted us to develop the literature review for this section by mainly discussing the most recent studies from these authors, within which a comparison among different ML techniques has been performed. Nevertheless, the authors are aware that other researchers have also been dealing with such issues, both in pioneering studies (Ermini et al. 2005) and in more recent times, for instance assessing the importance of the adopted variables and the appearance of the prediction map for gaining insights into model behavior (Goetz et al. 2015), evaluating the effects of spatial autocorrelation on hyperparameter tuning and performance estimation (Schratz et al. 2019*)*, mixing training and testing set resolutions (Duric et al. 2019), exploring innovative ways of combining the results of different models (Di Napoli et al. 2020), proposing and object-based method outperforming traditional cell-based methods (Wang et al. 2021a), or combining ML algorithms with active learning strategies (Wang and Brenning, 2021). The results from these contributions will be properly considered, for their respective relevance, in the discussion session.

Table 6 shows, for each article referenced, the list of ML algorithms adopted, and the location and area of the case studies. Mergadi et al. (2020) undertook an extensive analysis and comparison among many different ML techniques using a case study from Algeria covering an area of 2760 km^2. They summarize and discuss the algorithm's accuracies, advantages and limitations using a range of evaluation criteria. As main conclusions, they highlight that tree-based ensemble algorithms achieve excellent results compared to other machine learning algorithms and that the RF algorithm offers robust performance for accurate landslide susceptibility mapping with only a small number of adjustments required before training the model. Huang et al. (2020) compared a heuristic model and two statistical models with 4 ML models (i.e. MLP-NN, BPNN, SVM, DT; see Appendix 1 for meaning of acronyms) using data from a study area of 1581 km^2 in China. They observed that ML models have higher landslide susceptibility prediction performance than general statistical and heuristic models. The main objective of the study by Bui et al. (2020a, b*)* was to introduce a deep learning neural network model (DLNN) in landslide susceptibility assessments and to compare its predictive performance with other four widely-used ML models. The efficiencies of the models were estimated for a case study in Vietnam covering an area of 6850 km^2. Results showed that the proposed DLNN model had a higher performance than the four benchmark models. Pham and Prakash (2019), Chen et al. (2019),

Table 6 Case studies, recently published by some of the most prolific authors adopting ML for landslide susceptibility analyses, comparing the performance of different ML algorithms

Study	ML algorithms*	Case study	Area (km^2)
Merghadi et al. (2020)	ANN, SVM, DT, RF, NB, KNN, GB, NF	Mila Basin, Tellian Atlas (Algeria)	2760
Huang et al. (2020)	MLP-NN, BPNN, SVM, DT	Shicheng County, Jiangxi Province (China)	1581
Bui et al. (2020a, b)	DLNN, RF, SVM, DT, MLP-NN	Kon Tum Province (Vietnam)	6850
Pham and Prakash (2019)	LBE, SVM	Tehri Garhwal District, Uttarakhand state (India)	270
Chen et al. (2019)	KLR, NB, RBFN	Long County, Baoji City (China)	2400
Pourghasemi and Rahmati (2018)	ANN, BRT, CART, GLM, GAM, MARS, NB, RF, SVM	Ghaemshahr region, Mazandaran Province (Iran)	2241
Youssef et al. (2016)	RF, BRT, CART	Wadi Tayyah Basin, Asir Region (Saudi Arabia)	630

*Check Appendix 1 for full names
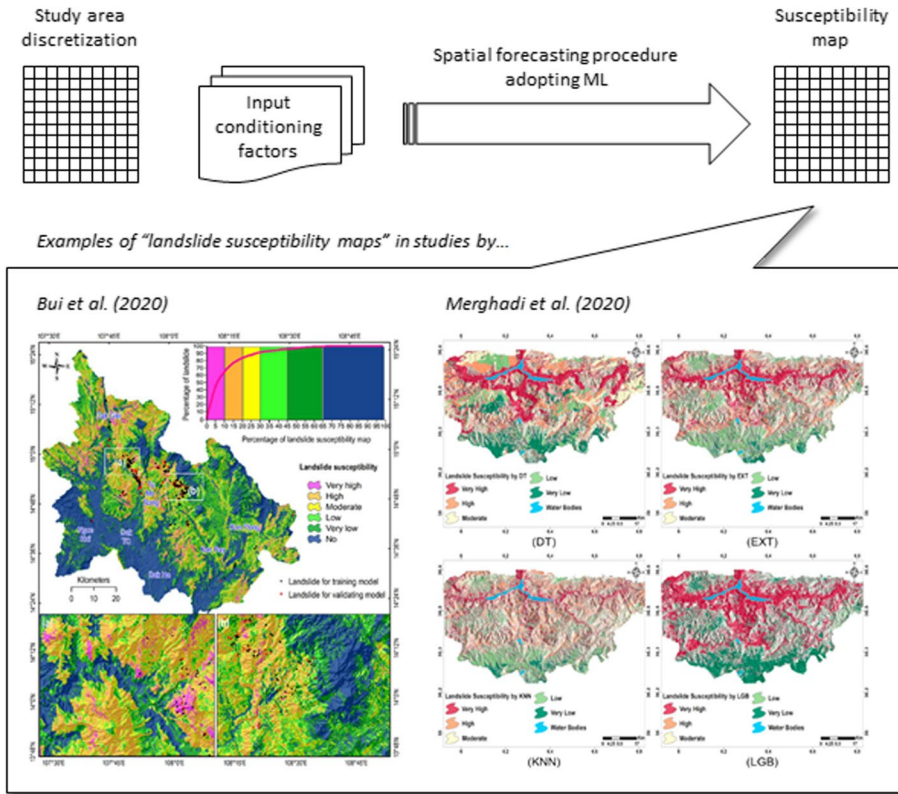
Fig. 5 Landslide susceptibility maps produced by ML spatial forecasting analyses adopting a pixel-based computational approach (modified from: Bui et al., 2020a, b; Merghadi et al. 2020)

Pourghasemi and Rahmati (2018) and Youssef et al. (2016) compared the capabilities of different ML methods for landslide prone zones in India, China, Iran and Saudi Arabia, respectively, covering areas from 270 to 2400 km^2.

All the studies reported in Table 6 perform the susceptibility analyses adopting a pixel-based computational approach that can be considered "typical" of analogous studies described in the extensive literature dealing with statistically-based landslide susceptibility modeling (Reichenbach et al. 2018). They indeed discretize the study area into a regular grid whose resolution depends on the scale of the information available, i.e. a raster file in GIS environment, and they use a landslide inventory to relate a set of input conditioning factors –i.e. thematic maps– to a quantitative indicator of the model outcome – i.e. the landslide susceptibility map (Fig. 5). The jargon may be different, as input and output variables as often called features and target, respectively, in ML applications, yet the underlying principles of these data-driven landslide susceptibility analyses remain the same.

Table 7 reports the main information of the landslide susceptibility computational models adopted in the different studies, and in particular: (i) the pixel resolution of the raster maps, (ii) the number of conditioning factors used as input maps, (iii) the number and typology of landslides, (iv) the number of landslide and non-landslide cells used in the ML algorithm, (v) the percentages of training and testing data used in the ML algorithm, and

Table 7 Main characteristics of the landslide susceptibility computational models adopted in the studies reported

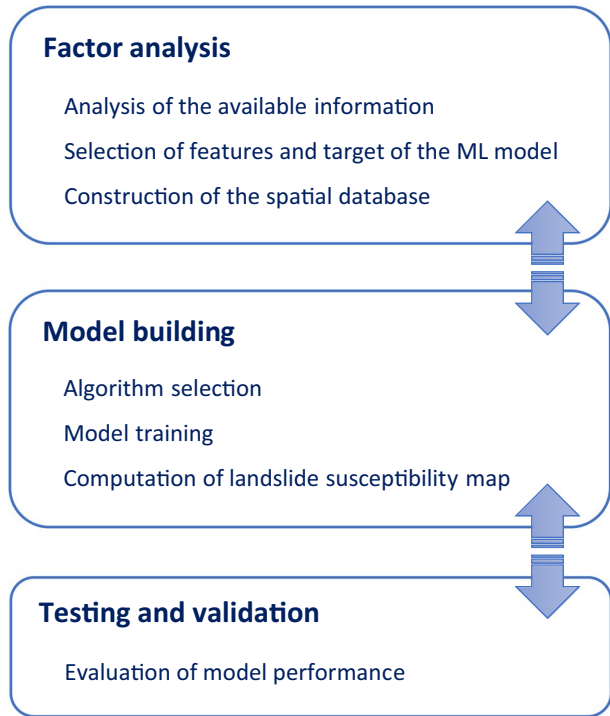| Study	Pixel resolution	No. input factors	No. landslides (and type)	No. L cells / No. NL cells	Training /Testing (%)	Susc. map classes
Merghadi et al. (2020)	30×30 m	16	578 (slide-type)	578 / 578	n/a	5
Huang et al. (2020)	30×30 m	13	369 (clay-silt slides)	2709 / 2709	70 / 30	5
Bui et al. (2020a, b)	30×30 m	9	1657 (shallow soil mixed boulder slides)	5270 / 5270	70 / 30	6
Pham and Prakash (2019)	n/a	18	391 (debris, translational, rock debris slide)	1561 / 1561	70 / 30	3
Chen et al. (2019)	30×30 m	12	171 (n/a)	171 / 171	70 / 30	5
Pourghasemi and Rahmati (2018)	20×20 m	18	276 (shallow rotational slides)	276 / 276	70 / 30	4
Youssef et al. (2016)	20×20 m	11	125 (mainly translational slides)	n/a	75 / 25	4

(vi) the number of classes adopted in the final susceptibility map. These studies adopt a medium pixel resolution ranging from 20 m×20 m to 30 m×30 m. The effect of the scale adopted to consider the landslide conditioning factors, and the DEM-derived topographic variables in particular, is obvious in terms of resolution of the information provided, yet an increase in DEM resolution does not necessarily produce a corresponding increase in the output of the landslide susceptibility analysis (Guzzetti et al. 1999). Indeed, Chang et al. (2019) stated that fine DEMs account for topography variations at the micro-scale that are not very much related to mesoscale processes like landslides, and that a 30 m resolution DEM is a good option because the minimum landslide size mapped from the satellite images is 0.1 hectare (1 hectare=100 m×100 m). At the same time, however, one must not neglect that raster files derived downscaling DEMs that are originally available as high-resolution maps derived from LiDAR or UAV surveys, can surely increase the accuracy of the susceptibility models.

The number of conditioning factors employed in the analyses is always significant, ranging from 9 to 18 in the seven considered studies. As commonly done for all the pixel-based GIS models aimed at deriving landslide susceptibility maps, they include: (i) DEM-derived topographic factors, such as elevation, slope, aspect, curvatures; (ii) geomorphological factors, such as distance to rivers, drainage density, stream power and topographic wetness indexes; (iii) geological factors, such as lithology, depth to bedrock or stratigraphy, distances to faults and to other geological boundaries; (iv) land and vegetation factors, such as land use, NDVI, solar radiation; and (v) other factors, related to natural or anthropogenic features, such as average rainfall and distance to road networks. Conditioning factors should be selected according to the considered landslide typologies. Indeed, any well-defined landslide susceptibility study should clearly focus on homogeneous landslides for which an inventory is available and for which a set of thematic information can be related to the triggering mechanisms of the considered landslides.

The main focus of the seven analyses reported are translational and rotational slide-type phenomena developing, depending on the characteristics of the study area, within different materials, ranging from clayey-silty soil to course-grained soils, like debris and boulders, to rocks. All the analyses are performed considering a random portion of the landslides reported in the inventory available, ranging from 70 to 75%, to train the ML model and the remaining landslides to test the model. Likewise, all the ML analyses consider landslide occurrences within any given cell of the study area as a binary dependent variable comprising only landslides (L cells) or non-landslides (NL cells) values, and employ an equal number of L-cells and NL-cells to run the ML model, both in the training and testing phases. To this aim, NL-cells are always selected randomly among the many cells comprising the space of the study area that is free of landslides. Some of the studies, only consider one single cell per landslide to determine the L-cells, while others consider all the cells that are included in the landslide shapes at the considered map resolution. The latter at times increases the number of L-cells used in the analyses by almost one order of magnitude compared to the number of inventoried landslides. A discussion on the influence of different sampling strategies for predicting landslide susceptibility is reported by Dou et al. (2020).

Finally, the landslide susceptibility maps are always drawn by grouping a computed landslide susceptibility index in a relatively small number of classes, ranging from a minimum of 3 to a maximum of 6 classes, and assigning to each class a susceptibility indicator such as, for instance, "very low", "low", "moderate", "high" and "very high" susceptibility when the number of classes is equal to 5.

Fig. 6 The three main phases of ML procedures for landslide susceptibility modeling

Factor analysis

Analysis of the available information

Selection of features and target of the ML model

Construction of the spatial database

Model building

Algorithm selection

Model training

Computation of landslide susceptibility map

Testing and validation

Evaluation of model performance

Different authors have been employing different operational procedures to move from the construction of the spatial database needed to feed the ML algorithm, to the generation of the landslide susceptibility map, and to the performance evaluation of the computational model. Three main common phases of analysis may be recognized in each procedure: (i) a "factor analysis" for the selection and computation of the input and output variables of the ML model; (ii) a "model building" phase that includes the ML algorithm selection, calibration and application, up to the production of the landslide susceptibility map; and (iii) a "testing and validation" analysis to evaluate the model performance. The three phases (Fig. 6) depend on each other and are done sequentially, but they often comprise sub-phases and loops, especially when the procedure proposes to compare more than one ML algorithm to define the final landslide susceptibility map for the study area.

4.1 Factor analysis

This phase is needed to analyze the thematic information available in the case study area (landslide conditioning factors and landslide inventory) and to prepare a dataset that can be used to build an ML model. Very often, the procedures adopted in this phase, to produce the optimal set of input variables (features) that can be related to the output variable (target), are based on well-known statistical methods. For instance, Mergadi et al. (2020) include two steps in their factor analysis: construction of a spatial database from the landslide inventory map and landslide conditioning factors; optimization of the landslide conditioning factors, by means of variance inflation factors and information

gain analyses. Similar procedures are proposed in other studies, for instance Huang et al. (2020) define the input–output variables adopting a frequency ratio bivariate statistical analysis from a set of conditioning factors in relation to landslide occurrences, and Chen et al. (2020) use a suite of statistical methods (i.e. normalized frequency ratio, variance inflation factors, and the chi-squared statistic) in their conditioning factor analysis.

Concerning the construction of the landslide event map, to be used as the dependent variable of the analysis, a binary raster variable is used in all the studies, comprising an equal number of L cells, defined starting from the landslide occurrences in the study area, and NL cells, identified based on a random selection of the landslide-free space.

The landslide conditioning factors to be used as input variables of the ML models may be obtained starting from the different data sources, such as available thematic maps, field investigations, reports, and remote sensing images. These factors are always processed using a GIS tool and converted to grid cell values, when they are not already provided in that format, with the desired analysis resolution. Data types may be discrete or continuous. Before using them as input variables of the ML analysis, extra data-preprocessing may be needed, such as numeric decoding of categorical variables or, most typically, grouping of the values of each continuous numerical factor in a finite number of classes. About the latter, Huang et al. (2020) state that the division of continuous conditioning factors will be rough if the attribute interval numbers are small, while the modeling processes will be complex if the attribute interval numbers are too many. There is no standard for determining the optimal number of classes to employ for computing the threshold values for the subdivision, yet analysts of ML studies typically adopt guidelines and suggestions commonly used in landslide susceptibility assessments (e.g., Guzzetti et al. 1999). In the seven studies considered herein, the number of classes adopted by the different authors for the continuous numerical variables needing reclassification varies between 3 and 9, and the adopted reclassification methods are: natural breaks, geometric intervals, frequency analyses, and heuristic assessments. The selection of these intervals, which requires significant subjective judgement, may be a key contributor to the model results.

It is worth highlighting that some procedures that adopt statistical analyses, before moving to the model building phase, aim at defining an optimal set of input variables for the training and validation datasets. Bivariate statistical methods, such as frequency or information gain ratios, are often used to evaluate the relevance of each conditioning factor on the results of the analysis, i.e. their predictive ability, and to assign weight coefficients to each class of each variable. The latter quantify, numerically, the probabilistic relation among the variable and the occurrence of landslides. The identified relevant conditioning factors are not necessarily independent from each other and, therefore, preliminary statistical analyses on these variables are also typically performed for multicollinearity (where two variables in a multiple regression model are highly linearly related) evaluation (Dormann et al. 2013), for instance by means of tolerances or variance inflation factor methods. Finally, the input variables are often scaled in the range 0–1.

4.2 Model building

As a first step of this phase, training and testing datasets must be defined, to respectively define the ML model and confirm its accuracy in the subsequent phase. To assess a model's predictive ability, after its definition and training, an independent dataset must be used for testing. Within standard multivariate statistical analyses, several procedures exist for

testing landslide prediction models (Baeza and Corominas 2001): (i) selection of a random sample to build the model and use of the remaining population to verify it; (ii) derivation of models from different random sample sizes and checking whether the function coefficients change significantly; (iii) preparation of the model from a distribution of landslides, which occurred during a specific event, and checking it with landslides triggered by a subsequent event; and (iv) development of the model in a training area, and testing it in a target area with similar characteristics. The first-mentioned procedure is widely adopted by landslide susceptibility ML studies, which use the majority of the inventoried landslides in the study area, typically more than 70%, as a training dataset, and the remaining ones as a testing dataset, to ensure that there are enough testing samples which have not been used during the training process of ML models but used to test its accuracy. Such separation ratio has been theoretically proved (Gholamy et al. 2018). However, a higher percentage for testing can also be used if the amount of raw data is large enough. As already explained, to avoid creating imbalanced datasets between L and NL grid cells, often an equal number of NL locations is randomly sampled from the landslide-free space, both during the training and testing phases. This practice, however, may create other (unwanted) biases. The second-mentioned procedure is also at times adopted. Depending on the objective of the study and availability of data, resampling strategies can indeed be nested on top of each other (Molinaro et al. 2005). To this aim, the cross-validation (CV) resampling procedure, which is based on a single parameter k that refers to the number of groups that a given data sample is to be split into, has recently emerged as a popular method in landslide susceptibility ML models. It is indeed considered a trade-off solution between speed, accuracy and computational costs (Mergadi et al. 2020).

To date, there is no consensus on a specific "optimal" ML algorithm for predicting landslide susceptibility at territorial scale, also because the performance and the predictive ability of ML models rely not only on the fundamental quality of the algorithms but also on details of their tuning, as well as on the quality of the landslide inventory and conditioning factors employed within the study area. Therefore, most of the landslide susceptibility studies published in the literature use and compare the performance of multiple ML algorithms in the same study area, thus using the same target variable derived from a given landslide
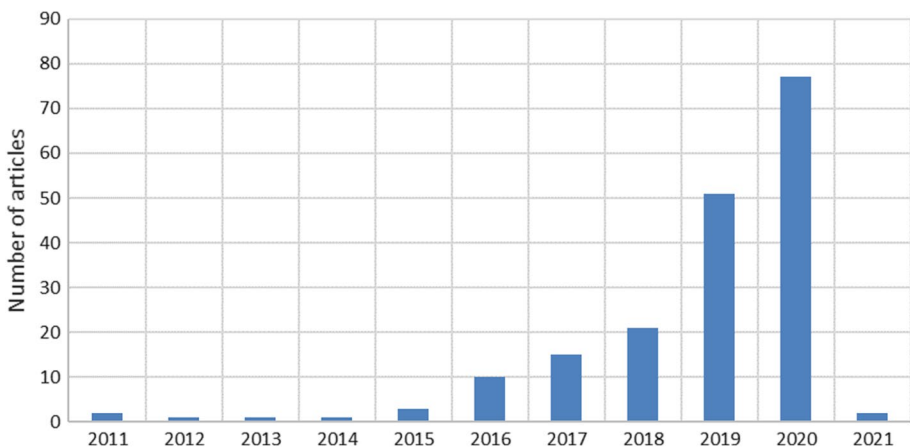
Fig. 7 Number of published journal articles dealing with ML studies for landslide susceptibility modeling (source: Scopus database, accessed 16/11/2020)

inventory, and a common set of features, derived from a suite of independent and relevant conditioning factors.

The number of scientific studies published in recent years on landslide susceptibility assessment adopting ML algorithms is very high, and it is growing extremely fast. A simple search performed in the Scopus database (on November 16, 2020), using the keywords "landslide" and "machine learning" and limited to journal articles, produced 286 entries, of which 186 (about 64%) are dealing with ML algorithms applied to landslide susceptibility modeling. The yearly distribution of these journal articles (Fig. 7) clearly shows that the topic has drawn growing attention in the past few years. The ML algorithms and procedures used in these studies is very heterogeneous, and tens of different algorithms are employed to the same purpose. The fact that no articles are shown before 2011 is most likely due to the fact that the expression "machine learning" started to be widely used only a few years ago to collectively identify a set of computer-based algorithms employed to find a relationship between landslide susceptibility conditioning factors, i.e. a set of features, and the presence of landslides, i.e. a single variable expressed as a dichotomous output target. Indeed, in the previous decades, starting already in the mid-1970s (Neuland 1976; Carrara 1983), the same aim has been pursued by means of heuristic or statistical analyses, among which methods like logistic regression and artificial neural networks were also included. This is confirmed by Mergadi et al. (2020), who state that LR and ANN algorithms were the earliest ML methods applied to landslide susceptibility modeling, with a total article count of 1587 and 746 since 2000, respectively. The same authors also state that the most popular methods nowadays are SVM, DT and RF algorithms, with a total article count of 342, 247 and 179 on each algorithm, respectively, since 2010.

Overall, the seven studies presented in Table 6 employ 23 different ML models to produce the landslide susceptibility maps. In the seven study areas, from a minimum of 3 (Chen et al. 2019) to a maximum of 10 (Mergadi et al. 2020) algorithms were compared, and at times also compared with other heuristic and statistical models (Huang et al. 2020). The ML algorithms adopted in more than one of these studies are: SVM (5 times); RF (4 times); DT and NB (3 times); ANN, BRT, CART, GLM, and MLP-NN (2 times). Youssef et al. (2016) and Pourghasemi and Rahmati (2018) are among the first authors in the literature to present a comprehensive comparison of the performance of many different ML techniques for landslide susceptibility modeling, respectively, 4 and 10 in the two studies. Pham and Prakash (2019) compared a hybrid ensemble approach with three single prediction models. Huang et al. (2020) chose 5 ML algorithms to compare among the ones most widely used in landslide susceptibility studies. On the other end, Chen et al. (2019) focused their comparison between NB and other two methods (KLR, RBFN) that have seldom been explored for landslide susceptibility modeling. Bui et al. (2020a, b) introduced a new deep learning neural network algorithm (DLNN) and compared its predictive performance with other four state-of-the-art ML models (RF, SVM, DT, MLP-NN). Mergadi et al. (2020) highlighted the importance of configuring and training the different ML algorithms one wants to compare for a given case study, using common hyper-parameter tuning strategies for the ML algorithms they compare.

The final step of the model building phase is the production of the landslide susceptibility maps, one for each algorithm adopted. As already mentioned, after a landslide susceptibility index is computed for each pixel of the study area, the final map is usually drawn considering a relatively small number of classes to which susceptibility indicators are attributed. The number of classes employed in the seven studies presented in Table 7 range from 3 (Pham and Prakash, 2019) to 6 (Bui et al. 2020a, b). Most commonly 4 (Youssef et al. 2016; Pourghasemi and Rahmati, 2018) or 5 (Chen et al. 2019; Huang et al.

Table 8 Validation performance metrics adopted in the studies reported in Table 6

Study	Binary classifiers							Suscep-tibility map	
	CM	AUC	OF	kappa	Other	NH	SA	SPR
Merghadi et al. (2020)	Acc	Yes	–	Yes	RD	WT	Yes	Yes
Huang et al. (2020)	–	Yes	–	–	LSI	–	–	–
Bui et al. (2020a, b)	Acc, Sp, Se, PPV, NPV	Yes	–	–	–	–	Yes	Yes
Pham and Prakash (2019)	Acc, Sp, Se	Yes	RSME	Yes	–	X2	–	–
Chen et al. (2019)	Acc, F	Yes	MAE, RSME	–	–	WT	–	Yes
Pourghasemi and Rah-mati (2018)	–	Yes	RMSE	–	–	–	–	–
Youssef et al. (2016)	–	Yes	–	–	–	–	–	Yes

*Check Appendix 2 for full names

2020; Mergadi et al. 2020) susceptibility classes are used. For instance, Bui et al. (2020a, b) acknowledged that the most common classification scheme in landslide susceptibility assessments use a five-level scale, including the "very low", "low", "moderate", "high" and "very high" susceptibility indicators (Fell et al. 2008). At the same time, however, they introduced an extra "no susceptibility" class in their study, given that a very large portion of the study area had an extremely low value of the computed landslide susceptibility index.

4.3 Testing and validation

Performance assessment for landslide susceptibility computational models can be conducted at two different levels (Table 8): (1) evaluating the quality of the classification problem with the binary model outcome of presence or absence of landslides; and (2) assessing the final landslide susceptibility map, i.e. validating the area covered by each susceptibility class against the landslide density distribution of the adopted landslide inventory map.

In relation to the first level of testing, the common performance metrics that are typically adopted in the literature include:

- Various metrics derived from a confusion matrix representation of the results (CM), including overall accuracy (Acc), specificity (Sp), sensitivity (Se), F-score (F) and others;
- The area under the ROC curve (AUC), computed as the integral over the graph that results from computing false positive rate and true positive rate for many different thresholds;
- Expressions quantifying the error of the analysis by means of an objective function (OF), like the mean absolute error (MAE) and the root mean square error (RMSE);
- The Cohen's kappa index (kappa), expressing the proportion of observed agreement beyond that expected by chance;
- Reliability diagrams (RD) and distributions of the computed landslide susceptibility indexes (LSI).

In addition to these metrics, when multiple ML algorithms are compared for a single study area, like for the case studies reported in Table 6, null-hypothesis testing (NH), such as the Wilcoxon signed-rank (WT) or the chi-square (X^2) tests, can also be conducted to assess the statistical significance of the differences between the model outcomes.

The second level of assessment for landslide susceptibility modeling is based on the assumption that a model is accurate when the landslide density ratio increases moving from low susceptibility classes to high susceptibility classes, and when the high susceptibility classes cover small extent areas (Pradhan and Lee 2010). To this aim, a necessary step is the reclassification of the landslide occurrence scores computed by the ML algorithms into a given number of classes expressing a susceptibility level, by means of an indicator, within the landslide susceptibility map. The areal extent of each susceptibility class can then be validated against the landslide density distribution from the landslide inventory map, by means of what is sometimes called a sufficiency analysis (SA). In addition to this qualitative evaluation of the output map, success and prediction rate curves (SPR) can also be drawn and the corresponding AUC computed.

5 Temporal forecasting

Temporal predictions of landslides, and more generally forecasts of the time evolution of key factors affecting the slope safety level, can be performed at global/regional scale or at slope-scale. The choice of the scale is usually linked to the choice of the monitored parameters, which is in turn related to the type of landslides. Typically, regional scale predictions are accomplished by using rainfall monitoring, geomorphological, and hydrometeorological approaches, while slope-scale predictions take advantage of a geotechnical engineering method relating displacement or other monitoring data to the time of failure (Intrieri et al. 2019). There is a relationship between monitoring parameters and types of landslide; for example, for shallow landslides that are triggered by extreme precipitation events, or by a combination of hydro-meteorological events, meteorological data dominates monitoring parameters. For slow moving deep-seated landslides, displacement monitoring can be a crucial input to assess slope behavior. New data assembling methods and Internet of Things (IoT) techniques have recently started to provide large datasets of monitoring data for landslide temporal forecasting using ML techniques. In this section, we review and discuss the main characteristics of available published studies (not very numerous up to the year 2020) that apply ML in landslide temporal forecasting.

5.1 Landslide displacement prediction at slope scale

Landslide displacement forecasting is considered an essential component for developing modern early warning systems. It can be used to set warning thresholds and to recognize when a landslide undergoes a sudden acceleration, which may lead to failure. Time series of real-time data collected from landslide monitoring systems, e.g. Geophones, Interferometric Synthetic Aperture Radar (InSAR), and Global Navigation Satellite System (GNSS), along with triggering data, e.g. water level and precipitation, provide critical inputs to ML modelling in this domain. However, the prediction of landslide displacement that changes over time is very challenging and it is inevitably linked to complex deformation mechanisms in the slope. Application of conventional ML methods, e.g. SVM, ANN, in landslide displacement forecasting, are reported in Mayoraz et al. (1996), Mayoraz & Vulliet (2002),

Ran et al. (2010), Zhu & Hu, (2012), and Du et al. (2013). By using ANN, Mayoraz et al. (1996) and Mayoraz & Vulliet (2002) predicted the velocity changes in a sliding soil mass based on meteorological and physical data and different neural network configurations. It must be noted that the future landslide velocity was the predicted parameter, instead of landslide displacement and input parameters in the multilayer perceptron neural network (MLP-NN) included daily precipitation, evaporation and pore water pressure. They showed that it is possible to obtain a reasonably good short-term (up to few days) prediction of landslide movements using a considerable number of continuous measurements. However, Mayoraz et al. (1996) concluded that MLP model yielded less precise predictions on the test set than on the training set, which is a sign of overfitting. In recent years, advances in DL algorithms and hybrid algorithms that combine different ML techniques, mainly performed on active landslides in China on the slopes of the Three Gorges Reservoir Area (TGRA, see Table 9), have shown promising results in the modelling and prediction of landslide deformations. For time series problems, advanced neural networks are generally considered as the most promising solutions since well-designed network structures could help to handle sequence dependence in the time series data (van Natijne et al. 2020).

In general, landslide displacement predictions include the following steps: (i) decomposition of the accumulated displacement, (ii) selection of conditioning factors, (iii) establishment of predicting models, and (iv) evaluation of prediction results. Wang (2003) and Du et al. (2013) proposed that the accumulated displacement (D) time series could be decomposed into three components: a trend, a periodic, and a stochastic component, i.e.

$$D = \phi + P + S \tag{1}$$

The long-term displacement, controlled by "internal" geological conditions such as lithology, geological structure and progressive weathering, is typically assumed as the driver for the trend component (ϕ). The short-term displacement, in this framework called the periodic component (P), is assumed to be influenced by "external" factors such as rainfall. The stochastic term (S) is the displacement response caused by a sudden change in the system, e.g., a raise or drop of the reservoir level (for TGRA) affecting the landslide hydraulic boundary conditions. In most of studies on landslide displacement in TGRA, the periodic and stochastic terms were not separated, or the stochastic term was completely ignored. The periodic term of displacement was believed to be caused by periodic reservoir water level fluctuations and rainfall. ML algorithms have been applied, in the literature, to predict the periodic term in the displacement time series that expresses the relationship between landslide displacement and its conditioning factors, e.g. precipitation and/or dam reservoir level.

The most recent studies on this topic are summarized in Table 9. ML algorithms have proven to be quite successful for forecasting the periodic component of landslide displacements obtained after removing the trend term from the accumulated displacement. Various ML algorithms have been tested for the prediction of periodic landslide displacement. However, in most of these studies, only one landslide case was used to verify the applicability and superiority of their proposed algorithm, which therefore may not perform well on other landslides. In some of the studies, e.g. Ma et al. (2020), Xie et al. (2019) and Krkač et al. (2017), only one ML algorithm was used for the landslide displacement prediction.

Commonly used controlling factors in the studies in TGRA include antecedent rainfall and reservoir water level over time and evolution state (e.g. Du et al. 2013; Yang et al. 2019; Zhou et al. 2018a) measured over 1 to 3 months before the event date. Not all controlling factors that may be related to landslide deformation can be used as input variables

Table 9 Recent case studies adopting ML for landslide displacement forecasting

Study	ML (or hybrid ML) algorithms*	Case study	Observed period
Li et al. (2020)	DBN, KNN, SVM, ELM, RF, BPNN	Baishuihe landslide in TGRA (China)	July 2003 and April 2015
Liu et al. (2020a, b, c)	LSTM, GRU, RF	Baishuihe, Bazimen and Baijiabao landslides in TGRA (China)	Varied periods between August 2003 and December 2013 for 3 landslides
Liu et al. (2020a, b, c)	Hybrid SVR-K-means-LSTM	Baishuihe landslide in TGRA (China)	June 2003 to June 2016
Jiang et al. (2020)	LSTM, SVM	Shengjibao landslide in TGRA (China)	January 2009 to December 2015
Ma et al. (2020)	T-GCN	Dongsheng Coal Field (China)	May 2011 to February 2012
Zhang et al. (2020a, b)	GRU, SVM	Erdaohe landslide in TGRA (China)	January 2008 to December 2013
Li et al. (2019)	ELM, LS-SVM	Baishuihe landslide in TGRA (China)	June 2005 to December 2016
Xie et al. (2019)	LSTM	Laowuji Landslide (China)	January 2016 to December 2016
Yang et al. (2019)	LSTM, SVM	Baishuihe and Bazimen landslides in TGRA (China)	Various periods between August 2003 and December 2013 for 2 landslides
Li et al. (2018)	Hybrid LASSO-ELM	Baishuihe landslide in TGRA (China)	2003 to 2015
Ma et al. (2018)	Hybrid Bootstrap-ELM-ANN	Shuping landslide in TGRA (China)	October 2003 to October 2013
Zhou et al. (2018a)	Hybrid PSO-KELM	Baishuihe landslide in TGRA (China)	January 2009 to December 2013
Krkač et al. (2017)	RF	Kostanjek landslide (Croatia)	n.a

*Check Appendix 1 for full names

Table 10 Comparison of best prediction performance of ML methods in each study reported in Table 9 for predicted periodic displacement for the Baishuihe landslide

Study	ML (or hybrid ML) algorithms*	Performance metrics**				
		MAPE/%	MAE/mm	RMSE/mm	MSE/mm	R
Li et al. (2020)	DBN	5.05	2.78	3.48	–	–
Liu et al. (2020a, b, c)	LSTM	5.0	–	7.5	–	–
Liu et al. (2020a, b, c)	Hybrid SVR-K-means-LSTM	–	3.86	6.76	–	–
Li et al. (2019)	ELM	–	–	17.41	–	0.968
Yang et al. (2019)	LSTM	10.43	–	7.11	–	–
Li et al. (2018)	Hybrid LASSO-ELM	1.30	26.44	35.63	2.28	–
Zhou et al. (2018a)	Hybrid PSO-KELM	0.083	18.104	–	–	0.983

*Check Appendix 1 for full names

**Check Appendix 2 for full names

for landslide displacement prediction in the ML models, because the ones having a low correlation with landslide deformation make the ML models complex and may reduce prediction accuracy. The controlling factors that have a strong correlation with the periodic displacement are typically selected by conducting correlation analyses, e.g. gray relational analysis (Deng 1989), and maximum information coefficient (Reshef et al. 2011).

The Baishuihe landslide, at the shores of TGRA, offers some possibilities for comparison, as multiple methods have been tested on this landslide by various authors. The Baishuihe landslide is a retrogressive landslide, where deformations occurred first at the bottom of the slope and retrogressed upwards (Du et al. 2013). The landslide reactivates frequently and have had several intense deformation periods since 2003. As indicated in Table 10, DL (e.g. DBN, LSTM) or hybrid ML methods show excellent prediction performance. However, the influence of the reservoir water level on the landslide stability, which is common to TGRA and not often present elsewhere, cannot be neglected and conclusions are therefore not easily transferable to other landslides.

5.2 Rainfall-induced landslides

For rainfall-induced landslides, a threshold defines the rainfall conditions that, when reached or exceeded, are likely to trigger a landslide. During the last decades, landslide rainfall thresholds have been mainly determined empirically or by adopting statistical methods (Segoni et al. 2018). ML methods are recently being explored to this aim. As an example, the conventional ML algorithm SVM has been used to determine rainfall thresholds by various authors (Vallet et al. 2013; Rachel and Lakshmi, 2016; Omadlao et al. 2019). At a nationwide level in Japan, Osanai et al. (2010) developed a new early-warning system for debris flow and slope-failure disasters. They used the rainfall indices of 60-min cumulative rainfall and calculated a soil–water index to set up a critical line (CL) employing a Radial Basis Function Network (RBFN). Osanai et al. (2010) state that the result of the system operation in 2009 proved its effectiveness in predicting rainfall-induced landslides. As no other references were found in the literature, we do not know if the identified thresholds have been subsequently validated.

ML methods have also been used to explore the relationship between amount of precipitation and groundwater level, a condition that is more closely linked to the pore pressure increase and shear strength reduction within the slope that leads to an instability, especially for deep-seated landslides. Yoon et al. (2011) developed two nonlinear time-series models using ANN and SVM techniques to predict groundwater level fluctuations based on data for the groundwater level, precipitation, and the tide level. Krkač et al. (2017) predicted the fluctuation of the groundwater level for the Kostanjek landslide using RF method. Huang et al. (2017) proposed a PSO-SVM model based on chaos theory to predict the daily groundwater levels of the Huayuan landslide and the weekly, monthly groundwater levels in Baijiabao in the TGRA of China. Wei et al. (2019) studied two different ML methods, i.e., the genetic algorithm back-propagation neural network (GA-BPNN) method and the genetic algorithm SVM (GA-SVM) method, for predicting the ground water level fluctuation of the Duxiantou landslide located in Zhejiang Province, China.

5.3 Dynamic susceptibility mapping

Landslide susceptibility mapping using ML methods has been intensively investigated by different researchers, as already mentioned. However, such studies do not intend to predict the time of occurrence of the landslides. Recently, the interest for dynamic susceptibility mapping, or spatio-temporal landslide probability assessment (*e.g.,* Lombardo et al. 2020*;* Wang et al. 2022), increased. Several works have been conducted to explore approaches for spatio-temporal landslide forecasting using conventional ML methods, e.g. SVM (Farahmand & AghaKouchak, 2013; Rachel and Lakshmi 2016; Omadlao et al. 2019), ANN (Pradhan et al. 2019), Decision Tree (Kirschbaum et al. 2015; Kirschbaum and Stanley 2018).

A few recent studies utilizing (hybrid) ML algorithms for dynamic susceptibility mapping are summarized in Table 11, showing the ML algorithms adopted, and the location and time period of the case studies. Stanley et al. (2020) identified where and when landslides were most probable, across relatively large ecoregions over the years 1976–2016, using an XGBoost model. XGBoost method was proven to be an effective method for incorporating rainfall intensity, atmospheric rivers, antecedent soil moisture, and melting snow from land data assimilation systems into a unified indicator of rainfall-triggered landslide hazard. Lee et al. (2021) proposed an MLP-NN enhanced with Gumbel distribution approach to assess the temporal probability of future landslide occurrence using the limited rainfall records and landslide inventory in a study area in Jinbu, Korea. MLP-NN was used in static landslide susceptibility analysis with the balanced pixel data. An ROC graph and the associated AUC were used to verify the accuracy of the susceptibility map by comparing actual and estimated results. Finally, the temporal probability of landslide occurrence, evaluated, using the Gumbel model, with 72-h antecedent rainfall threshold was combined with the spatial probability of landslides to determine landslide hazard. Utomo et al. (2019) proposed a hybrid model based on physically-based stability method and ADASYN (Adaptive Synthetic Sampling)—BPNN (Backpropagation Neural Network) to design an accurate early warning system. The proposed method had higher accuracy than BPNN and ADASYN-BPNN without physically-based stability analyses, but required more computational time and resources. Lombardo et al. (2020) proposed a novel Bayesian modelling framework for the spatiotemporal prediction of landslides. The spatial predictive performance of Bayesian models was quantified using a tenfold cross-validation procedure, and the temporal predictive performance using a leave-one-out cross validation procedure. Wang et al. (2022) established a space–time susceptibility model for hydromorphological (HMP) processes covering the Chinese territory from 1985 to 2015.

Table 11 Case studies, recently published adopting ML for dynamic landslide susceptibility analyses

Study	ML algorithms*	Case study	Time period
Xiao et al. (2022)	Logistic regression, CNN, Bagging, AdaBoost, RUSBoost, and Subspace method	Hong Kong	1984–2017
Wang et al. (2022)	Binomial GLM	Chinese territory	1985–2015
Lombardo et al. (2020)	Bayesian	Collazzone, Umbria, Central Italy	1941–2014
Stanley et al. (2020)	XGBoost	The Marine West Coast Forest, the Northwestern Forested Mountains, and the North American Deserts	1976–2016
Lee et al. (2021)	MLP-NN enhanced with Gumbel distribution	Jinbu, Gangwon-do (Korea)	July 14–16, 2006
Utomo et al. (2019)	ADASYN-BPNN	Shen-Mu datasets (Taiwan)	10 min
Farahmand and AghaK-ouchak (2013)	SVM	NASA global landslides	2003, 2007, 2008 and 2009

*Check Appendix 1 for full names

The space–time model was built on the basis of a binomial Generalized Linear Model (GLM), producing the mean, maximum and 95% confidence interval of the spatio-temporal susceptibility distribution per catchment, per year.

6 Discussion

6.1 Objective of landslide studies using ML

ML algorithms aim primarily at making accurate predictions, while explanation can be regarded as a secondary objective. Taking this into account, applications of ML methods in landslide studies should be mainly focused on problems where the need for predictions prevails over explanation and understanding. Such is needed when sufficient quantity of data exists and time is the key deciding factor, e.g. time to occurrence of an event or time for developing and conducting a study. Example of the former can be Landslide Early Warning Systems, where it is crucial to make a decision in a limited time based on streams of monitoring data. Example of the latter can be landslide detection in which collecting detailed field data requires many days and sufficient manpower (Mondini et al. 2011). When the objective of the landslide study is deep understanding of processes, we do not see the usefulness of a direct application of ML. However, also in these cases, features detected by ML, for instance related to the importance of conditioning factors in landslide spatial prediction studies, may help understanding landslide processes. In terms of future scenarios, we argue that ML methods are useful when interpolation is the main purpose, meaning that the machine has already learned from a broad spectrum of data and the new occasion falls within the available data space (similar statistical distribution). If the new occasion falls outside the available data space, i.e., an extrapolation problem, ML methods may not perform well.

6.2 ML and DL algorithms

There is no consensus on an "optimal" ML/DL algorithm for landslide studies, even when looking at the results of the most recent comparative studies in landslide detection or spatial and temporal forecasting. Indeed, there is a growing tendency in the literature to propose the systematic use of an ensemble of algorithms for the same study area, not only native ensemble ML algorithms such as RF but rather various different ML algorithms, and then choose the best-performing one. As indicated by Ghorbanzadeh et al. (2019b) and Prakash et al. (2020) in landslide detection studies, comparisons between conventional ML algorithms and DL methods reveal that algorithmic choice faces the so-called *No Free Lunch theorem*, which implies that there is no single "best" algorithm to look after because, on average, all algorithms will perform about the same (Wolpert 1996).

The choice between adopting conventional ML or DL algorithms primarily depends on the type and quantity of available data. In general, DL algorithms are not expected to outperform conventional ML if the size of training data is not very large. For instance, for landslide spatial prediction studies, the amount of past information on known locations of landslides is typically very low compared to the extent of the landslide susceptibility study areas. Number of features, attributes and preference over feature engineering also affect this choice. We suggest that for structured data, conventional ML algorithms are to

be preferred, whereas for unstructured data (e.g. text, video, imagery, etc.), where feature engineering can be a daunting task, DL algorithms can be more suitable.

6.3 Availability of ML/DL libraries

There is no consensus on what methods can indeed be properly called ML algorithms and some well-known inferential statistics methods, like various types of logistic regressions or discriminant analyses, are often referred to as ML algorithms. When ML/DL algorithms are used in applied science and engineering, including the landslide community, there is an overall tendency to use off-the-shelf algorithms that are already implemented in free libraries. Python libraries such as Scikit-learn for conventional ML and TensorFlow, Keras and PyTorch for DL algorithms are among them. A possible drawback is that such tendency can lead, in the long run, to ML illiteracy of the landslide community because there is no effort in implementing and deeply understanding the algorithms, which can also result in misusing them or leaning towards trial-and-error. An example that supports this claim is related to the hyperparameters of ML algorithms. In most of the studies reviewed for this paper, authors either used the default values of hyperparameters or chose them through trial and error. It can also be seen that in many DL-based landslide studies, the architecture of the DL framework is not properly explained and no efforts are spent to deeply understand why certain architectures work better than others. Another possible drawback of leaning on these implementations is that researchers will have to wait quite some time before the emergence of new promising algorithms well suited for landslide studies.

6.4 Data availability

Data-driven methods, such as ML algorithms, are not useful if the necessary data is not available. In fields such as landslide detection and landslide susceptibility mapping, where publicly available satellite images at various resolutions exist, data availability can be less problematic. However, in temporal forecasting employing monitoring data (e.g. ground-based sensors, InSAR data), good quality data do not exist freely, and this condition surely limits the application of ML algorithms. It may be expected, however, that in the near future this limitation may be overcome by the growing availability of remote sensing data and the growing competition within the remote sensing community. Datasets dedicated to ML landslide studies can thus produce a significant shift in the current way of forecasting landslide displacements. Examples of datasets already available in the ML domain can be found at: https://www.paperswithcode.com/datasets.

6.5 Code availability

The majority of the works reviewed in this study did not make the utilized computer scripts available. Within the fast-growing ML community (see https://paperswithcode.com/), availability of the script and the data used are important criteria for assessing the credibility of a study. It can be argued that such intellectual opacity in the landslide ML literature will hamper the utility of these studies because researchers, even assuming that they may have access to the original data, in a majority of cases cannot duplicate them.

6.6 Pre-trained models

In the Computer Vision community, algorithms pre-trained on large datasets exist. When it comes to applying these algorithms to a similar problem (e.g. image classification), instead of training the original algorithm from scratch, ML engineers use these pre-trained algorithms to save time and to reduce the need for more data. Such ideas can be used in landslide detection and landslide susceptibility studies, also to reach higher accuracies over time.

6.7 Physically-based methods versus ML

Compared to ML-based models, physically-based models require less data for calibration, as they are fully or partly based on well-established laws of physics. The two classes of methods are typically seen as alternatives to each other, and data-driven models, including ML algorithms, are often called upon only when the use of physically-based models is deemed unfeasible or cost prohibitive. In fact, in landslide studies we may state that ML algorithms are currently being adopted as tools for all those data-driven analyses that, in the past, would have seen researchers use, for the same purposes, statistical techniques. However, physically-based methods can help ML in various ways: (1) make ML models more explainable, (2) decrease the volume of data that is needed to train ML algorithms, (3) produce synthetic data for data-scarce problems (e.g. Jamalinia et al. 2021). The integration of ML methods in physically-based models is also a path that is currently being explored by researchers in engineering and science. Examples of this approach can be found, for instance, in the computational fluid dynamics community, where ANN has been used for solving partial differential equations used to simulate fluid dynamics problems (e.g., Kutz 2017; Schenck & Fox 2018; Clark Di Leoni et al. 2020).

6.8 Supervised, unsupervised and reinforcement learning

The majority of landslide/ML studies reviewed herein used supervised ML. Such widespread use of supervised learning is also common in other engineering and science fields. Unsupervised machine learning methods are not very popular, mainly because they do not suit labeled datasets. However, these methods can be helpful for finding anomalous data of geo-systems including natural and engineered slopes. This can be very useful, for instance, in early warning systems. Some advanced unsupervised learning methods, such as GANs and Autoencoders, have found applications in landslide detection (e.g. domain adaptation in Fang et al. 2020) and landslide susceptibility mapping. It can be foreseen that these advanced methods will receive more attention from landslide researchers in the future.

Reinforcement learning (RL) is currently mostly used in research, but the approach already shows maturity in problem-solving for game like scenarios. As suggested by Bergen et al. (2019), there have been efforts on using RL methods in earth sciences and particularly in earthquake and seismicity related studies (e.g. Delores et al. 2018). To our knowledge and up to the year 2020, however, there are no published applications of RL to landslide studies. However, it is to be expected that, due to the necessity of rapid and data-driven decision making in issues related to landslide risk assessment and management, landslide studies will adopt in the future RL techniques.

6.9 Statistics versus ML

ML and statistically-based approaches for the detection and spatial prediction of landslides over large areas share many common characteristics. Therefore, it is not strange that most of the recent spatial forecasting studies adopting ML algorithms significantly "draw" from the experience accumulated in the past decades, since the seminal publication by Varnes (1984) on bivariate and multivariate statistical techniques and procedures for landslide susceptibility assessment and zoning. The main consequence is that almost the totality of the ML literature contributions on this topic (to the Authors' knowledge) employ a "standard" pixel-based computational approach to perform the susceptibility analyses. Therefore, even if the adopted jargon may be different, the essence of the ML analyses is the same as for any other data-driven approach for deriving a landslide susceptibility map in GIS environment, starting from a set of input conditioning factors and a landslide event map. Statistical methods focus on inference, achieved through the creation and fitting of a problem specific probability model, whereas machine learning methods concentrate on prediction by using general-purpose learning algorithms to find patterns in often rich and unwieldy data (Bzdok et al. 2018). From this perspective, machine learning methods are potentially more powerful in forecasting landslide patterns. Most of the issues highlighted to explain the performance of the models are commonly treated, outside the specific ML literature, whenever geospatial data-driven analyses are performed (e.g. Goetz et al. 2015; Reichenbach et al. 2018; Lombardo et al. 2020). Examples of such specificities are: resolution of information and mapping units (*e.g.* Calvello et al. 2013), preprocessing of conditioning factors (*e.g.* Guzzetti et al. 1999), low number of landslide cells in relation to non-landslide cells (*e.g.* Tanyu et al. 2021), influence of sampling strategy (*e.g.* Wang and Brenning 2021), validation practices (*e.g.* Steger et al. 2016), number of classes of input and output variables (*e.g.* Baeza et al. 2016). A discussion of these items, which are very relevant for the implementation and the applicability of data-driven techniques for landslide spatial forecasting in operational settings, goes beyond the scope of this paper.

6.10 Generalization and evaluation of the models

Model generalization is an important aspect of ML modeling that is neglected by many researchers in ML landslide studies. For instance, in landslide susceptibility mapping, most studies verify the superiority of their proposed method(s) by comparing a small number of ML algorithms in a common area. However, the proposed models are not repeatedly tested and may not outperform other methods in areas other than training areas. In fact, there is still a lack of benchmark case studies available for testing various ML methods.

In the landslide detection and spatial forecasting studies, the evaluation of the performance of a given ML algorithm is typically done by checking the quality of the classification problem with the binary model outcome of presence or absence of landslides. To this purpose, the mostly used performance indicators are the area under the ROC curve and various metrics derived from the confusion matrix. Less common, but nevertheless used, are reliability diagrams, expressions quantifying the error of the analysis by means of an objective function, and null-hypothesis testing, which is more common in statistics-based studies. According to the recent literature on ML for landslide spatial forecasting, ML models have, typically, a higher landslide susceptibility prediction performance than statistical and heuristic models. This finding is not surprising, given that this is a pre-requisite, if not the main justification, for a scientific article to be published on this topic. However,

the case studies often report (rather suspiciously) very high values of performance indicators for many algorithms in the same area, thus the alleged "success" of an ML algorithm over the others is too often attributed because of rather small differences in the values of its performance indicators. In the landslide temporal forecasting studies, the evaluation of the performance of a given ML algorithm is typically done based on quantitative performance metrics, e.g. MAPE, MAE, RMSE, MSE, and R.

6.11 Relevance of expert opinion

ML landslide studies always need the analysist to decide much more than just the algorithm(s) to use in a given analysis. Indeed, the landslide studies discussed in all the three main sections of this paper, i.e. detection and mapping, spatial forecasting and temporal forecasting, are most often proposing procedures that fulfill the objectives by means of a combination of methods, which include ML algorithms, and a set of heuristic expert choices. This is important to recognize when evaluating, or comparing, the performances of given ML algorithms, as they cannot be easily entangled from the other elements comprised in the proposed procedures.

Expert knowledge plays a significant role in enhancing the performance of ML models. Feature selection heavily relies on expert knowledge in both spatial and temporal landslide predictions. Expert opinion is also reflected in algorithm selection and implementation.

Taking spatial prediction studies for example, the recent trend is to adopt computational procedures that combine many algorithms and methods, including standard statistical analyses, to address the different phases of the landslide susceptibility analysis. For instance, in the initial factor analysis, bivariate statistical methods are used to evaluate the relevance of each conditioning factor and to assign weight coefficients to each class of each variable; cross-validation is used for checking whether weight coefficients change significantly upon resampling; and input variables are checked for multicollinearity evaluation.

7 Conclusions and perspective

In this paper we provided a detailed overview of machine learning and ML studies pertained to landslide detection and mapping, spatial forecasting and temporal forecasting. In addition to the three sections of the paper explicitly devoted to these topics, the main general observations on different aspects of ML-based landslide studies were presented in the final Discussion section of this paper. Our review revealed that over the years the complexity of ML algorithms used in landslide studies has been matching the rapid development that is occurring in the AI/ML community. Likewise, it can be stated that ML still has a long path to follow in the landslide community.

Out of the three landslide subfields investigated herein, it seems that landslide detection studies are the ones which benefited the most from ML progresses, whereas it appears that spatial and temporal forecasting still did not get a clear and distinct advantage from incorporating ML algorithms in their studies. This is mainly because landslide detection is essentially a Computer Vision (CV) problem, for which there is an active community within the AI community, where many developments are carried out. Those developments, as well as the fact that landslide detection does not require much physical understanding compared to other landslide research areas, encourage implementation of robust ML

algorithms for more accurate landslide detection. It also can be expected that the application of DL to this aim will further increase in the coming years, thus replacing more traditional methods such as OBIA. Within DL methods, it is expected that more advanced CV algorithms will replace conventional ones. Methods such as Graph Neural Networks and various generative modeling methods such as GANs are foreseen to find more applications in landslide studies. In landslide spatial forecasting, the number of publications adopting ML algorithms in landslide susceptibility studies has been growing at a very fast pace in recent years, with a trend that is resembling the growth shown, in the previous two decades, by multivariate statistical studies conducted with the same purposes. In this area, we expect that the current trend that focuses on the use of different ML algorithms and compares their performance within a common area, will remain the main not-too-innovative procedural strategy explored by researchers, at least in the near future. Nevertheless, given the redundancy of these studies, we can surely hope that a new trend will emerge, possibly combining ML and process-based methods for a more robust and generalized assessment and understanding of landslide susceptibility at regional scale. Also in landslide temporal forecasting, it may be expected that procedures will be developed that combine ML algorithms with physically-based methods, such as computational geomechanics models. For the temporal prediction of slope failure processes, probabilistic ML/DL, such as Bayesian DL, may also possibly be an emerging trend.

In conclusion, we can confidently state that ML is a vibrant field with expanding interest and rapid advancement. We do not encourage landslide researchers to follow the same pace of ML progress in implementing ML algorithms for landslide studies, as it can jeopardize the deep understanding of both processes and ML methods. However, we do encourage the landslide community to closely observe ML upgrades and get inspiration for implementing innovative data-driven methods in landslide studies. It is surely a challenge to use ML algorithms appropriately to advance the field of landslide studies, yet the growing interest shown in recent years for such endeavors is promising. There is potential for a wider use in practice and consultancy in the future, but further research is surely needed to this aim.

Appendix 1: Machine learning acronyms

ADASYN	Adaptive synthetic sampling
ANN	Artificial neural network
BPNN	Back-propagation neural network
BRT	Boosted regression tree
CART	Classification and regression trees
CNN	Convolutional neural network
DBN	Deep belief network
DLNN	Deep learning neural network
DP-FCN	Dual path FCN
DSCN	Deep Siamese neural network
DT	Decision tree
ELM	Extreme learning machine
FCN-PP	Fully connected network with pyramid pooling
FFNN	Feed-forward neural network
GA	Genetic algorithm

GAM	Generalized additive model
GAN	Generative adversarial network
GBM	Gradient boosting machines
GBM	Gradient boosting
GLM	Generalized linear model
GMM	Gaussian mixture model
GPBF	Generalized positive Boolean function
GRU	Gated recurrent unit
GSF	GAN-based Siamese framework
HC	Hierarchical clustering
KELM	Kernel extreme learning machine
KLR	Kernel logistic regression
K-Means	k-means clustering
kNN	k nearest neighbour
LASSO	Least absolute shrinkage and selection operator
LBE	LogitBoost ensemble
LDA	linear discriminant analysis
LR	logistic regression
LS-SVM	Least squares support vector machine
LSTM	Long short term memory
MARS	Multivariate adaptive regression splines
MIL	Multi-instance learning
MLC	Maximum likelihood classifier
MLP-NN	Multi-layer perceptron neural network
MPNN	Multi-perceptron neural networks
NB	Naïve bayes
NBC	Naïve bayes classifier
NF	Neuro-fuzzy
PSO	Particle swarm optimization
PSPNet	Pyramid scene parsing network
QDA	Quadratic discriminant analysis
RBFN	Radial basis function network
ResNet	Residual neural network
RF	Random forest
RNN	Recurrent neural network
SVM	Support vector machine
SVR	Support vector regression
T-GCN	Temporal graph convolutional networks
XGBoost	eXtreme gradient boosting

Appendix 2: ML metric acronyms

Acc	Accuracy
AUC	Area under the receiver operating characteristic curve
CE	Commission error
CM	Confusion matrix
Com	Completeness
Cor	Correctness
CT	Contingency table
F	F score
FNR	False negative rate
FPR	False positive rate
FPR	False positive rate
kappa	Cohen's kappa index
KC	Kappa coefficient
MAE	Mean absolute error
MAPE	Mean absolute percentage error
mIOU	Mean intersection over union
MSE	Mean square error
NH	Null-hypothesis
NPV	Negative prediction value
OA	Overall accuracy
OE	Omission error
OF	Objective function
OPR	Over prediction rate
PPV	Positive prediction value
Prec	Precision
QP	Quality percentage
Qual	Quality
R	Coefficient of determination
RD	Reliability diagrams
Rec	Recall
RMSE	Root mean square error
SA	Sufficiency analysis
Se	Sensitivity
Sp	Specifity
SPR	Success and prediction rate curves
TPR	True positive rate
UPR	Unpredicted presence rate
Vis	Visual
WT	Wilcoxon signed-ran

Appendix 3: Other acronyms

C	combined
CL	Critical line
CML	Conventional machine learning
DL	Deep learning
IC	Image classification
IoT	Internet of things
LSI	Landslide susceptibility indexes
nCV	Non computer vision
OB	Object-based
PaIC	Patch-wise image classification
PB	Pixel-based
PSS	Pixel-wise semantic segmentation
S	Supervised
TGRA	Three gorges reservoir area
U	Unsupervised

Declarations

Conflict of interest Not applicable.

References

Alvioli M, Mondini AC, Fiorucci F, Cardinali M, Marchesini I (2018) Topography-driven satellite imagery analysis for landslide mapping. Geomat Nat Haz Risk 9(1):544–567

Amit SNKB, Shiraishi S, Inoshita T, Aoki Y (2016) Analysis of satellite images for disaster detection. In: 2016 IEEE international geoscience and remote sensing symposium (IGARSS), IEEE, pp 5189–5192

Bacha AS, Van Der Werff H, Shafique M, Khan H (2020) Transferability of object-based image analysis approaches for landslide detection in the Himalaya Mountains of northern Pakistan. Int J Remote Sens 41(9):3390–3410

Baeza C, Corominas J (2001) Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. Earth Surf Proc Landf 26:1251–1263

Baeza C, Lantada M, Amorim S (2016) Statistical and spatial analysis of landslide susceptibility maps with different classification systems. Environ Earth Sci 75:1318

Bergen KJ, Johnson PA, Maarten V, Beroza GC (2019) Machine learning for data-driven discovery in solid Earth geoscience. Science 363(6433):eaau0323

Borghuis AM, Chang K, Lee HY (2007) Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. Int J Remote Sens 28(8):1843–1856

Brabb EE (1984) Innovative approaches to landslide hazard mapping. In: Proceedings of the 4th international symposium landslides, vol 1, pp 307-323

Breiman L (2001) Statistical modeling: The two cultures (with comments and a rejoinder by the author). Stat Sci 16(3):199–231

Bui DT, Tsangaratos P, Nguyen VT, Liem NV, Trinh PT (2020a) Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment. CATENA 188:104426

Bui TA, Lee PJ, Lum KY, Loh C, Tan K (2020b) Deep learning for landslide recognition in satellite architecture. IEEE Access 8:143665–143678

Burkov A (2019) The hundred-page machine learning book. Andriy Burkov, Quebec City, pp 70–82

Bzdok D, Altman N, Krzywinski M (2018) Statistics versus machine learning. Nat Methods 15(4):233–234

Calvello M, Cascini L, Mastroianni S (2013) Landslide zoning over large areas from a sample inventory by means of scale-dependent terrain units. Geomorphology 182:33–48

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 Int J Geo Inf 8(7):300

Carrara A (1983) Multivariare models for landslide hazard evaluation. Math Geol 15:403–426

Casagli N, Cigna F, Bianchini S, Hölbling D, Füreder P, Righini G, Vlcko J (2016) Landslide mapping and monitoring by using radar and optical remote sensing: examples from the EC-FP7 project SAFER. Remote Sens Appl Soc Environ 4:92–108

Catani F (2021) Landslide detection by deep learning of non-nadiral and crowdsourced optical images. Landslides 18:1025–1044

Chang YL, Liang LS, Han CC, Fang JP, Liang WY, Chen KS (2007) Multisource data fusion for landslide classification using generalized positive Boolean functions. IEEE Trans Geosci Remote Sens 45(6):1697–1708

Chang KT, Liu JK, Chang YM, Kao CS (2010) An Accuracy Comparison for the Landslide Inventory with the BPNN and SVM Methods. Proceeding of Gi4DM 2010

Chang et al (2019) Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. Sci Rep 9:12296

Chen W, Li X, Wang Y, Chen G, Liu S (2014) Forested landslide detection using LiDAR data and the random forest algorithm: a case study of the three gorges, China. Remote Sens Environ 152:291–301

Chen F, Yu B, Li B (2018a) A practical trial of landslide detection from single-temporal Landsat8 images using contour-based proposals and random forest: a case study of national Nepal. Landslides 15(3):453–464

Chen Z, Zhang Y, Ouyang C, Zhang F, Ma J (2018b) Automated landslides detection for mountain cities using multi-temporal remote sensing imagery. Sensors 18(3):821

Chen et al (2019) Spatial prediction of landslide susceptibility using data mining-based kernel logistic regression, naive Bayes and RBFNetwork models for the Long County area (China). Bull Eng Geol Environ 78:247–266

Chen S, Xiang C, Kang Q, Zhong W, Zhou Y, Liu K (2020) Accurate landslide detection leveraging UAV-based aerial remote sensing. IET Commun 14(15):2434–2441

Cheng K, Wei C, Chang S (2004) Locating landslides using multi-temporal satellite images. Adv Space Res 33:296–301

Cheng G, Guo L, Zhao T, Han J, Li H, Fang J (2013) Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA. Int J Remote Sens 34(1):45–59

Clark Di Leoni P, Meneveau C, Karniadakis G, Zaki T (2020) Deep operator neural networks (DeepONets) for prediction of instability waves in high-speed boundary layers. Bulletin of the American Physical Society

Danneels, G., Pirard, E., & Havenith, H. B. (2007, July). Automatic landslide detection from remote sensing images using supervised classification methods. In: 2007 IEEE international geoscience and remote sensing symposium. IEEE, pp 3014–3017

Deng JL (1989) Introduction to grey system theory. J Grey Syst 1(1):1–24

Di Napoli M, Carotenuto F, Cevasco A, Confuorto P, Di Martire D, Firpo M, Pepe G, Raso E, Calcaterra D (2020) Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. Landslides 17:1897–1914

Dilley M, Chen RS, Deichmann U, Lerner-Lam A, Arnold M, Agwe J, Buys P, Kjekstad O, Lyon B, Yetman G (2005) Natural disaster hotspots: a global risk analysis (English). World Bank, Washington, DC, pp 1–132

Ding A, Zhang Q, Zhou X, Dai B (2016) Automatic recognition of landslide based on CNN and texture change detection. In: 2016 31st youth academic annual conference of Chinese association of automation (YAC), IEEE, pp 444–448

Dormann et al (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46

Dou J, Yunus AP, Merghadi A, Shirzadi A, Nguyen H, Hussain Y, Avtar R, Chen Y, Pham BT, Yamagishi H (2020) Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. Sci Total Environ 720:137320

Dou, J., Paudel, U., Oguchi, T., Uchiyama, S., & Hayakavva, Y. S. (2015). Shallow and deep-seated landslide differentiation using support vector machines: a case study of the Chuetsu Area, Japan. Terr Atmos Ocean Sci, 26(2).

Draelos TJ, Peterson MG, Knox HA, Lawry BJ, Phillips-Alonge KE, Ziegler AE, Faust A (2018) Dynamic tuning of seismic signal detector trigger levels for local networksdynamic tuning of seismic signal detector trigger levels for local networks. Bull Seismol Soc Am 108(3A):1346–1354

Du J, Yin K, Lacasse S (2013) Displacement prediction in colluvial landslides, three gorges reservoir, China. Landslides 10:203–218

Duric U, Marjanović M, Radića Z, Abolmasov B (2019) Machine learning based landslide assessment of the Belgrade metropolitan area: Pixel resolution effects and a cross-scaling concept. Eng Geol 256:23–38

Ermini L, Catani F, Casagli N (2005) Artificial neural networks applied to landslide susceptibility assessment. Geomorphology 66:327–343

Fang B, Chen G, Pan L, Kou R, Wang L (2020) GAN-based siamese framework for landslide inventory mapping using bi-temporal optical remote sensing images. IEEE Geosci Remote Sens Lett 18(3):391–395

Farahmand A, AghaKouchak A (2013) A satellite-based global landslide model. Nat Hazard 13(5):1259–1267

Feelders A (2002) Data mining in economic science. Dealing with the data flood, 166–175

Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savagef WZ, JTC-1 Joint Technical Committee on Landslides and Engineered Slopes (2008) Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Eng Geol 102(3–4):85–98

Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. Nat Hazard 18(8):2161–2181

Geisser S (1975) The predictive sample reuse method with applications. J Am Stat as 70(350):320–328

Gholamy A, Kreinovich V, Kosheleva O (2018) Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation. Departmental Technical Reports (CS). 1209

Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena SR, Tiede D, Aryal J (2019a) Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sens 11(2):196

Ghorbanzadeh O, Meena SR, Blaschke T, Aryal J (2019b) UAV-based slope failure detection using deep-learning convolutional neural networks. Remote Sens 11(17):2046

Ghorbanzadeh O, Blaschke T (2019) Optimizing sample patches selection of CNN to improve the mIOU on landslide detection. In GISTAM, pp 33–40

Goetz JN, Brenning A, Petschko H, Leopold P (2015) Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput Geosci 81:1–11

Gong J, Wang D, Li Y et al (2010) Earthquake-induced geological hazards detection under hierarchical stripping classification framework in the Beichuan area. Landslides 7:181–189

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31:181–216

Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang K-T (2012) Landslide inventory maps: New tools for an old problem. Earth Sci Rev 112(1–2):42–66

He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision, pp 1026–1034

Heleno S, Matias M, Pina P, Sousa AJ (2016) Semiautomated object-based classification of rain-induced landslides with VHR multispectral images on Madeira Island. Nat Hazards Earth Syst Sci 16:1035–1048

Hölbling D, Füreder P, Antolini F, Cigna F, Casagli N, Lang S (2012) A semi-automated object-based approach for landslide detection validated by persistent scatterer interferometry measures and landslide inventories. Remote Sensing 4(5):1310–1336

Hu Q, Zhou Y, Wang S, Wang F, Wang H (2019) Improving the accuracy of landslide detection in "off-site" area by machine learning model portability comparison: a case study of Jiuzhaigou Earthquake. China Remote Sens 11(21):2530

Huang L, Xiang LY (2018) Method for meteorological early warning of precipitation-induced landslides based on deep neural network. Neural Process Lett 48:1243–1260

Huang et al (2020) Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. CATENA 191:104580

Huang F, Huang J, Jiang S, Zhou C (2017) Groundwater levels prediction using evidence of chaos and support vector machine. J Hydroinf 19(4):jh2017102

Intrieri E, Carlà T, Gigli G (2019) Forecasting the time of failure of landslides at slope-scale: a literature review. Earth Sci Rev 193:333–349

Jamalinia E, Tehrani FS, Steele-Dunne SC, Vardon PJ (2021) A data-driven surrogate approach for the temporal stability forecasting of vegetation covered dikes. Water 13(1):107

Ji S, Yu D, Shen C, Li W, Xu Q (2020) Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks. Landslides 17:1337–1352

Jiang HM, Li YY, Zhou C, Hong HY, Glade T, Yin KL (2020) landslide displacement prediction combining LSTM and SVR algorithms: a case study of Shengjibao Landslide from the three gorges reservoir Area. Appl Sci 10:7830

Kamiyama J, Noro T, Sakagami M, Suzuki Y, Yoshikawa K, Hikosaka S, Hirata I (2008) Detection of Landslide Candidate Interference Fringes in DInSAR Imagery Using Deep Learning

Karantanellis E, Marinos V, Vassilakis E, Christaras B (2020) Object-based analysis using unmanned aerial vehicles (UAVs) for site-specific landslide assessment. Remote Sens 12(11):1711

Keyport RN, Oommen T, Martha TR, Sajinkumar KS, Gierke JS (2018) A comparative analysis of pixel- and object-based detection of landslides from very high-resolution images. Int J Appl Earth Obs Geoinf 64:1–11

Kirschbaum D, Stanley T (2018) Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness. Earth's Future 6:505–523. https://doi.org/10.1002/2017EF000715

Kirschbaum D, Stanley T, Simmons J (2015) A dynamic landslide hazard assessment system for Central America and Hispaniola. Nat Hazards Earth Syst Sci 15:2257–2272

Konishi T, Suga Y (2019). Landslide detection with ALOS-2/PALSAR-2 data using convolutional neural networks: a case study of 2018 Hokkaido Eastern Iburi earthquake. In Active and passive microwave remote sensing for environmental monitoring III (vol 11154, p 111540H). International society for optics and photonics

Korup O, Stolle A (2014) Landslide prediction from machine learning. GeologyToday 30(1):26–33

Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90

Krkač M, Špoljarić D, Bernat S (2017) Method for prediction of landslide movements based on random forests. Landslides 14:947–960

Kutz JN (2017) Deep learning in fluid dynamics. J Fluid Mech 814:1–4

Lee JH, Kim H, Park HJ, Heo JH (2021) Temporal prediction modeling for rainfall-induced shallow landslide hazards using extreme value distribution. Landslides 18:321–338

Lei T, Zhang Y, Lv Z, Li S, Liu S, Nandi AK (2019b) Landslide inventory mapping from bitemporal images using deep convolutional neural networks. IEEE Geosci Remote Sens Lett 16(6):982–986

Lei T, Zhang Q, Xue D, Chen T, Meng H, Nandi AK (2019a) End-to-end change detection using a symmetric fully convolutional network for landslide mapping. In: ICASSP 2019a–2019a IEEE international conference on acoustics, speech and signal processing (ICASSP) IEEE, pp 3027–3031

Li X, Cheng X, Chen W, Chen G, Liu S (2015) Identification of forested landslides using LiDar data, object-based image analysis, and machine learning algorithms. Remote Sens 7(8):9705–9726

Li Z, Shi W, Lu P, Yan L, Wang Q, Miao Z (2016) Landslide mapping from aerial photographs using change detection-based Markov random field. Remote Sens Environ 187:76–90

Li H, Xu Q, He Y, Deng J (2018) Prediction of landslide displacement with an ensemble-based extreme learning machine and copula models. Landslides 15(10):2047–2059

Li D, Sun YQ, Yin KL, Miao FS, Glade T, Leo C (2019) Displacement characteristics and prediction of Baishuihe landslide in the three Gorges reservoir. J Mt Sci 16(9):2203–2214

Li H, Xu Q, He Y, Fan X, Li S (2020) Modeling and predicting reservoir landslide displacement with deep belief network and EWMA control charts: a case study in three Gorges reservoir. Landslides 17:693–707

Liu Y, Xu C, Huang B, Ren X, Liu C, Hu B, Chen Z (2020a) Landslide displacement prediction based on multi-source data fusion and sensitivity states. Eng Geol 271(2020):105608

Liu ZQ, Guo D, Lacasse S, Li JH, Yang BB, Choi JC (2020b) Algorithms for intelligent prediction of landslide displacements. J Zhejiang Univ Sci A 21(6):412–429

Liu P, Wei Y, Wang Q, Chen Y, Xie J (2020c) Research on Post-earthquake landslide extraction algorithm based on improved U-Net model. Remote Sens 12(5):894

Lombardo L, Opitz T, Ardizzone F, Guzzetti F, Huser R (2020) Space-time landslide predictive modelling. Earth Sci Rev 209:103318

Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 2015, pp 3431-3440

Lu P, Stumpf A, Kerle N, Casagli N (2011) Object-oriented change detection for landslide rapid mapping. IEEE Geosci Remote Sens Lett 8(4):701–705

Lu P, Qin Y, Li Z, Mondini AC, Casagli N (2019) Landslide mapping from multi-sensor data through improved change detection-based Markov random field. Remote Sens Environ 231:111235

Lv ZY, Shi W, Zhang X, Benediktsson JA (2018) Landslide inventory mapping from bitemporal high-resolution remote sensing images using change detection and multiscale segmentation. IEEE J Sel Top Appl Earth Obs Remote Sens 11(5):1520–1532

Lv Z, Liu T, Kong X, Shi C, Benediktsson JA (2020) Landslide inventory mapping with bitemporal aerial remote sensing images based on the dual-path full convolutional network. IEEE J Sel Topics Appl Earth Obs Remote Sens 13:4575–4584

Ma J, Tang H, Liu X, Wen T, Zhang J, Tan Q, Fan Z (2018) Probabilistic forecasting of landslide displacement accounting for epistemic uncertainty: a case study in the Three Gorges Reservoir area. China Landslides 15(6):1145–1153

Ma Z, Mei G, Zhang Z, Xu N (2020) A deep learning approach using graph convolutional networks for slope deformation prediction based on time-series displacement data. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.12987995.v1

Mabu S, Nakayama Y, Kuremoto T (2020) Landslide classification from synthetic aperture radar images using convolutional neural network with multichannel information. J Signal Process 24(2):61–73

Martha TR, Kerle N, van Westen CJ, Jetten V, Kumar KV (2011) Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. IEEE Trans Geosci Remote Sens 49(12):4928–4943

Mayoraz F, Cornu T, Vulliet L (1996) Using neural networks to predict slope movements. In: Proceedings of VII international symposium on landslides, Trondheim, Balkema, pp 295–300

Mayoraz F, Vulliet L (2002) Neural networks for slope movement prediction. Int J Geomech 2:153–173

Merghadi et al (2020) Machine learning methods for landslide susceptibility studies: a comparative overview of algorithm performance. Earth Sci Rev 207:103225

Mezaal MR, Pradhan B, Sameen MI, Mohd Shafri HZ, Yusoff ZM (2017) Optimized neural architecture for automatic landslide detection from high-resolution airborne laser scanning data. Appl Sci 7(7):730

Miandad J, Darrow MM, Hendricks MD, Daanen RP (2020) Landslide mapping using multiscale LiDAR digital elevation models. Environ Eng Geosci 26(4):405–425

Mitchell T (1997) Machine Learning. McGraw Hill, New York

Molinaro AM, Simon R, Pfeiffer RM (2005) Prediction error estimation: a comparison of resampling methods. Bioinformatics 21:3301–3307

Mondini AC, Guzzetti F, Reichenbach P, Rossi M, Cardinali M, Ardizzone F (2011) Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. Remote Sens Environ 115(7):1743–1757

Mora OE, Lenzano MG, Toth CK, Grejner-Brzezinska DA, Fayne JV (2018) Landslide change detection based on multi-temporal Airborne LiDAR-derived DEMs. Geosciences 8(1):23

Myint SW, Gober P, Brazel A, Grossman-Clarke S, Weng Q (2011) Per-pixel vs object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens Environ 115(5):1145–1161

Neuland H (1976) A prediction model of landslips. CATENA 3(2):215–230

Omadlao Z, Tuguinay N, Saturay RJ (2019) Machine learning based prediction system for rainfall-induced landslides in Benguet First Engineering District. https://doi.org/10.31219/osf.io/csx6r

Osanai N, Shimizu T, Kuramoto K, Kojima S, Noro T (2010) Japanese early-warning for debris flows and slope failures using rainfall indices with radial basis function network. Landslides 7(3):325–338

Parzen E (2001) [Statistical modeling: the two cultures]: comment. Stat Sci 16(3):224–226

Pawłuszek K, Borkowski A (2017) automatic landslides mapping in the principal component domain. In: Workshop on world landslide forum Springer, Cham, pp 421–428

Pawłuszek K, Borkowski A, Tarolli P (2017) Towards the optimal pixel size of dem for automatic mapping of landslide areas. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42

Pawluszek-Filipiak K, Borkowski A (2020) On the importance of train-test split ratio of datasets in automatic landslide detection by supervised classification. Remote Sens 12(18):3054

Pham BT, Prakash I (2019) Evaluation and comparison of LogitBoost ensemble, fisher's linear discriminant analysis, logistic regression and support vector machines methods for landslide susceptibility mapping. Geocarto Int 34:316–333

Pourghasemi HR, Rahmati O (2018) Prediction of the landslide susceptibility: Which algorithm, which precision? CATENA 162:177–192

Pradhan B, Lee S (2010) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environ Earth Sci 60:1037–1054

Pradhan AM, Lee SR, Kim YT (2019) A shallow slide prediction model combining rainfall threshold warnings and shallow slide susceptibility in Busan, Korea. Landslides 16:647–659

Prakash N, Manconi A, Loew S (2020) Mapping landslides on EO data: performance of deep learning models vs. traditional machine learning models. Remote Sens 12(3):346

Qi W, Wei M, Yang W, Xu C, Ma C (2020) Automatic mapping of landslides by the ResU-Net. Remote Sens 12(15):2487

Rachel N, Lakshmi M (2016) Landslide prediction with rainfall analysis using support vector machine. Indian J Sci Technol. https://doi.org/10.17485/ijst/2016/v9i21/95275

Ran YF, Xiong GC, Li SS, Ye LY (2010) Study on deformation prediction of landslide based on genetic algorithm and improved BP neural network. Kybernetes 39:1245–1254

Rau JY, Jhan JP, Rau RJ (2013) Semiautomatic object-oriented landslide recognition scheme from multisensor optical imagery and DEM. IEEE Trans Geosci Remote Sens 52(2):1336–1349

Reichenbach et al (2018) A review of statistically-based landslide susceptibility models. Earth Sci Rev 180:60–91

Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, Lander ES, Mitzenmacher M, Sabeti PC (2011) Detecting novel associations in large data sets. 511 Science 334:1518–1524

Sameen MI, Pradhan B (2019) Landslide detection using residual networks and the fusion of spectral and topographic information. IEEE Access 7:114363–114373

Samuel AL (1959) Some studies in machine learning using the game of checkers. IBM J Res Dev 3(3):210–229

Schenck C, Fox D (2018) Spnets: differentiable fluid dynamics for deep neural networks. In: Conference on robot learning, PMLR, pp 317–335

Schratz P, Muenchow J, Iturritxa E, Richter J, Brenning A (2019) Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. Ecol Model 406:109–120

Segoni S, Piciullo L, Gariano SL (2018) A review of the recent literature on rainfall thresholds for landslide occurrence. Landslides 15:1483–1501

Shi W, Zhang M, Ke H, Fang X, Zhan Z, Chen S (2020) Landslide recognition by deep convolutional neural network and change detection. IEEE Trans Geosci Remote Sens 59(6):4654–4672

Shmueli G (2010) To explain or to predict? Stat Sci 25(3):289–310

Si A, Zhang J, Tong S, Lai Q, Wang R, Li N, Bao Y (2018) Regional landslide identification based on susceptibility analysis and change detection. ISPRS Int J Geo Inf 7(10):394

Soares LP, Dias HC, Grohmann CH (2020) Landslide segmentation with U-Net: evaluating different sampling methods and patch sizes. Preprint arXiv:2007.06672

Stanley TA, Kirschbaum DB, Sobieszczyk S, Jasinski MF, Borak JS, Slaughter SL (2020) Building a landslide hazard indicator with machine learning and land surface models. Environ Model Softw 129:104692

Steger S, Brenning A, Bell R, Petschko H, Glade T (2016) Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology 262:8–23

Stumpf A, Kerle N (2011) Object-oriented mapping of landslides using Random Forests. Remote Sens Environ 115(10):2564–2577

Su Z, Chow JK, Tan PS, Wu J, Ho YK, Wang YH (2021) Deep convolutional neural network–based pixel-wise landslide inventory mapping. Landslides 18:1421–1443

Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9.

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826

Tanyu BF, Abbaspour A, Alimohammadlou Y, Tecuci G (2021) Landslide susceptibility analyses using Random Forest, C4.5, and C5.0 with balanced and unbalanced datasets. CATENA 203:1053555

Tavakkoli Piralilou S, Shahabi H, Jarihani B, Ghorbanzadeh O, Blaschke T, Gholamnia K, Aryal J (2019) Landslide detection using multi-scale image segmentation and different machine learning models in the higher himalayas. Remote Sens 11(21):2575

Tehrani FS, Santinelli G, Herrera MH (2021) Multi-regional landslide detection using combined unsupervised and supervised machine learning. Geomatics, Natural Hazards and Risk (accepted)

Ullo SL, Mohan A, Sebastianelli A, Ahamed SE, Kumar B, Dwivedi R, Sinha GR (2020) A new mask R-CNN based method for improved landslide detection. Preprint arXiv:2010.01499

Utomo D, Chen SF, Hsiung PA (2019) Landslide prediction with model switching. Appl Sci-Basel 9(9):1839

Vallet A, Varron D, Bertrand C, Mudry JN (2013) Hydrogeological threshold using effective rainfall and support vector machine (SVM) applied to a deep seated unstable slope. Se´chilienne, French Alps

Van Den Eeckhaut M, Kerle N, Poesen J, Hervás J (2012) Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data. Geomorphology 173:30–42

van Natijne AL, Lindenbergh RC, Bogaard T (2020) Machine learning: new potential for local and regional deep-seated landslide nowcasting. Sensor 20(5):1425

Varnes DJ, IAEG Commission on Landslides and other Mass-Movements (1984) Landslide Hazard zonation: a review of principles and practice. The UNESCO Press, Paris, p 63

Wallis WA (1980) The statistical research group, 1942–1945. J Am Stat Assoc 75(370):320–330

Wang JF (2003) Quantitative prediction of landslide using S-curve. Chin J Rock Mech Eng 14:1–8

Wang Z, Brenning A (2021) Active-learning approaches for landslide mapping using support vector machines. Remote Sens 13:2588

Wang Y, Wang X, Jian J (2019) Remote sensing landslide recognition based on convolutional neural network. Math Probl Eng 2019:8389368

Wang HJ, Zhang LM, Luo HY, He J, Cheung RWM (2021a) AI-powered landslide susceptibility assessment in Hong Kong. Eng Geol 288(3):106103

Wang HJ, Zhang LM, Yin KS, Luo HY, Li JH (2021b) Landslide identification using machine learning. Geosci Front 12:351–364

Wang N, Cheng W, Marconcini M, Bachofer F, Liu C, Xiong J, Lombardo L (2022) Space-time susceptibility modeling of hydro-morphological processes at the Chinese national scale. Eng Geol 301:106586

Wei ZL, Qing L, Sun HY, Shang YQ (2019) Estimating the rainfall threshold of a deep-seated landslide by integrating models for predicting the groundwater level and stability analysis of the slope. Eng Geol 253:14–26

Wolpert DH (1996) The lack of a priori distinctions between learning algorithms. Neural Comput 8(7):1341–1390

Xiao T, Zhang LM, Cheung RWM, Lacasse S (2022) Predicting spatio-temporal man-made slope failures induced by rainfall in Hong Kong using machine learning techniques. Geotechnique. https://doi.org/10.1680/jgeot.21.00160

Xie PH, Zhou AG, Chai B (2019) The application of long short-term memory (LSTM) method on displacement prediction of multifactor-induced landslides. IEEE Access 7:54305–54311

Xu Q, Ouyang C, Jiang T, Fan X, Cheng D (2019) DFPENet-geology: a deep learning framework for high precision recognition and segmentation of co-seismic landslides. Preprint arXiv:1908.10907

Yang BB, Yin KL, Lacasse S, Liu ZQ (2019) Time series analysis and long short-term memory neural network to predict landslide displacement. Landslides 16(4):677–694

Ye C, Li Y, Cui P, Liang L, Pirasteh S, Marcato J, Li J (2019) Landslide detection of hyperspectral remote sensing data based on deep learning with constrains. IEEE J Sel Topics Appl Earth Obs Remote Sens 12(12):5047–5060

Yoon H, Jun S, Hyun Y, Bae G, Lee K (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J Hydrol 396(1):128–138

Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM (2016) Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear

models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides 13:839–856

Yu B, Chen F, Xu C (2020) Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015. Comput Geosci 135:104388

Yu H, Ma Y, Wang L, Zhai Y, Wang X (2017) A landslide intelligent detection method based on CNN and RSG_R. In: 2017 IEEE international conference on mechatronics and automation (ICMA), IEEE, pp 40–44

Zhang L, Zhang L, Du B (2016) Deep learning for remote sensing data: a technical tutorial on the state of the art. IEEE Geosci Remote Sens Mag 4(2):22–40

Zhang M, Shi W, Chen S, Zhan Z, Shi Z (2020a) Deep multiple instance learning for landslide mapping. IEEE Geosci Remote Sens Lett 99:1–5

Zhang X, Han L, Han L, Zhu L (2020b) How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery? Remote Sens 12(3):417

Zhou C, Yin K, Cao Y, Intrieri E, Ahmed B, Catani F (2018a) A novel displacement prediction method using gated recurrent unit model with time series analysis in the Erdaohe landslide. Nat Hazards 15(11):2211–2225

Zhou B, Lapedriza A, Khosla A, Oliva A, Torralba A (2018b) Places: A 10 million image database for scene recognition. IEEE Trans Pattern Anal Mach Intell 40(6):1452–1464

Zhu C, Hu G (2012) Time series prediction of landslide displacement using SVM model: Application to baishuihe landslide in Three Gorges Reservoir Area, China. Appl Mech Mater 239:1413–1420

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.