Machine Learning for Landslides Prevention: A Survey

Zhengjing Ma [1] and Gang Mei [2]

[1]Affiliation not available
[2]China University of Geosciences (Beijing)

October 30, 2023

Abstract

Landslides are one of the most critical categories of natural disasters worldwide and induce severely destructive outcomes to human life and the overall economic system. To reduce its negative effects, landslides prevention has become an urgent task, which includes investigating landslide-related information and predicting potential landslides. Machine learning is a state-of-the-art analytics tool that has been widely used in landslides prevention. This paper presents a comprehensive survey of relevant research on machine learning applied in landslides prevention, mainly focusing on (1) landslides detection based on images, (2) landslides susceptibility assessment, and (3) the development of landslide warning systems. Moreover, this paper discusses the current challenges and potential opportunities in the application of machine learning algorithms for landslides prevention.

Machine Learning for Landslides Prevention: A Survey

Zhengjing Ma[a], Gang Mei[a,*]

[a]*School of Engineering and Technology, China University of Geosciences (Beijing), 100083, Beijing, China*

Abstract

Landslides are one of the most critical categories of natural disasters worldwide and induce severely destructive outcomes to human life and the overall economic system. To reduce its negative effects, landslides prevention has become an urgent task, which includes investigating landslide-related information and predicting potential landslides. Machine learning is a state-of-the-art analytics tool that has been widely used in landslides prevention. This paper presents a comprehensive survey of relevant research on machine learning applied in landslides prevention, mainly focusing on (1) landslides detection based on images, (2) landslides susceptibility assessment, and (3) the development of landslide warning systems. Moreover, this paper discusses the current challenges and potential opportunities in the application of machine learning algorithms for landslides prevention.

Keywords: Natural Disasters, Landslides Prevention, Machine Learning, Supervised Learning, Unsupervised Learning, Deep Learning

Contents

List of Abbreviations

ANN	Artificial Neural Networks
CNN	Convolutional Neural Networks
DT	Decision Tree
DBN	Deep Belief Networks
DEM	Digital Elevation Models
ELM	Extreme Learning Machine
GIS	Geographic Information Systems
GAN	Generative Adversarial Networks
GNN	Graph Neural Networks
GBDT	Gradient Boosting Decision Tree
kNN	k-Nearest Neighbors
LR	Logistic Regression
LSTM	Long Short Term Memory
NB	Naive Bayes
PCA	Principal Components Analysis
RF	Random Forest
RNN	Recurrent Neural Networks
RBN	Restricted Boltzmann Machine
SVM	Support Vector Machine

1. Introduction

Landslides are one of the most critical categories of natural disasters worldwide and induce severely destructive outcomes to human life and the overall economic system [1]. Their existence is ascribed to the geological environment and meteorological processes on earth. Some factors, including lithology, slope morphology, and unplanned urban expansions, can predispose slopes to landslides [2, 3]. Severe extreme events caused by climate change, including heavy rainfall and rapid snowmelt, could also trigger landslide occurrences [4]. With climate change has strengthened, the frequency and intensity of landslides are expected to increase rapidly as a consequence.

It is quite urgent to understand landslides to predict their occurrences and behavior, and then to adopt appropriate prevention policies and methodologies. The prevention of an incipient or potential landslide requires the recognition of the landslide and investigation of landslide-related information. Then, the region where a landslide is prone to occur in the future is predicted. Finally, the anticipation of the character and magnitude of movement may occur [5]. Therefore, common landslides prevention techniques can be divided into two categories: *detection* and *prediction*.

In the spatial extent of a landslide, an inventory collected by means of detection is often required. As a common strategy, detection can overcome the limitation of the scale and the location of landslide events to produce detailed inventories that can not only provide a better understanding of the important information about the landslide but also establish a relationship between different factors and the landslides. An accurate landslide inventory can play a part in most stages of landslides prevention, especially in landslide susceptibility assessment.

Landslide susceptibility assessment refers to the possibility of the occurrence of landslides in the spatial dimension. The principle of landslide susceptibility assessment is that a region prone to landslides can be predicted based on available data, including conditional factors and historical landslides. These data are extracted from the landslide inventory. As a static instrument, landslide susceptibility assessment has shown its value in spatial analysis [6]; however, it lacks information on the temporal landslide probability [7].

Reliable early warning systems can be used to predict the short-term behavior of landslides for preventing sudden events. Once emergency warnings are issued, people can take action before the disaster occurrences. An effective approach for achieving early warnings of landslides is to establish quantitative models of landslide evolution processes. The modeling of landslides is based on continuous monitoring of landslide-related variables [8].

Most landslides are triggered by extreme precipitation events [9]. Thus, the rainfall threshold should be regarded as a critical parameter to predict the occurrence of a landslide in the temporal domain. On the other hand, an adequate understanding of landslide deformation mechanisms is essential to develop a reliable early warning system [10]. Landslide displacement is a crucial parameter for judging the condition of the landslide, and rapid changes are generally considered a direct sign of upcoming disasters [11].

The analysis of the aforementioned stages in landslides prevention remains a challenge due to complex geodynamic and microphysical processes. Recently, as an analytics tool, machine learning methods, which can provide predictions, perform clustering, extract association features, and make decisions from given information, are coming to the fore. Various domains have successfully utilized machine learning methods to complete some demanding tasks. Likewise, the landslides prevention domain has begun to apply most major machine learning methods to accurately, efficiently, and effectively solve problems.

At least 80% of machine learning is data preprocessing, which means that the performance of machine learning methods depends on the data quality. With advances in a variety of location-aware sensors and model simulations, available data volumes in the landslides prevention domain are exponentially increasing with increasing spatial, temporal, and spectral resolutions. These continuous cumulative datasets provide applications for machine learning with more opportunities.

Overall, in the context of the spatiotemporal complexity and uncertainties, landslides prevention provides novel opportunities, challenges, and methodological demands for machine learning, which has been a hot research topic in recent years. In the following sections, this paper reviews the development of machine learning methods in landslides prevention. It mainly concentrates on various applications of machine learning in several different landslides prevention stages.

The main contributions of this survey can be summarized as follows.

(1) The paper starts with investigating the applications of machine learning in landslides prevention that comprise landslide detection and prediction. Landslide detection provides inventory data for the other prevention stages. Landslide prediction involves static and dynamic methods to predict landslide occurrences from spatial and temporal perspectives, respectively.

(2) This paper points out the potential challenges and limitations of machine learning in landslides prevention and proffers several strategies that have been utilized in other research domains to overcome or circumvent them. Moreover, this paper discusses the opportunity for the emergence of challenges, summarizes and recommends a few of the most promising deep learning methods that have been applied in other

domains, and envisages their possible applications in landslides prevention.

(3) This paper advises combining data-driven machine learning with knowledge-driven landslide mechanisms to interpret machine learning results.

The rest of this survey is organized as follows. Section 2 briefly introduces machine learning and landslides prevention. Section 3 surveys machine learning applications in landslide detection. Section 4 surveys machine learning applications in landslides susceptibility assessment. Section 5 surveys machine learning applications in landslide warning systems. Section 6 outlines major challenges and opportunities for using machine learning for landslides prevention. Finally, section 7 concludes this survey.

2. Background

For the sake of clarity, this section will briefly introduce the background concepts of machine learning and landslides prevention.

2.1. Machine Learning

2.1.1. Conventional Machine Learning Methods

As an analytics tool, one of the major objectives and tasks of machine learning is to build a model to represent complex, unknown, or incompletely understood relationships between data and target variables [12]. There are slight variations in the types of machine learning algorithms; and the machine learning algorithms can be roughly divided into two major categories according to their purpose: *supervised learning* and *unsupervised learning* algorithms.

Supervised learning refers to building a model for connecting known inputs to unknown outputs. Consequently, the output values for new data can be predicted based on those relationships learned from the previously labeled training data [13]. Supervised learning can be divided into classification and regression problems. In classification problems, the intended output is a semantic label or class. For example, to identify potential landslides, classification problems would label each pixel in an image as "landslide" or "non-landslide". Regression problems aim to predict a continuous variable.

Common supervised learning algorithms include the Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), Naive Bayes (NB), Artificial Neural Networks (ANN). Each single learning algorithm can be considered as a base learner.

Although base learners behave well, it is necessary to improve the performance further (e.g., the classification, prediction, or function approximation) of a machine learning model. Since more powerful learners

are constructed from a set of base learners, multiple learner systems (i.e., ensemble learning) have gradually gained much attention. Three representative ensemble methods are bagging, which can decrease variance, boosting, which can decrease the bias, and stacking, which improves predictions [14, 15].

Unsupervised learning methods try to find patterns in unlabeled data. One of the most common unsupervised learning algorithms is clustering, where samples are grouped based on similarity. Another typical method is dimensionality reduction that aimed at reducing the variance in a dataset and remove outliers.

Brief introductions to these methods are listed as follows.

- LR: A supervised learning algorithm that uses a logistic function to map the input variables to categorical dependent variables.

- DT: A supervised learning algorithm is commonly used in classification problems. The structure resembles a tree. The branch node represents several alternatives. Each leaf node represents a decision.

- SVM: A supervised learning algorithm is also commonly used in classification problems by constructing a separating line to distinguish between objects in a multidimensional space.

- NB: A supervised learning algorithm is based on Bayes' theorem and widely used in classification problems, which assumes that features are independent and have no correlations.

- ANN: ANN consists of a set of connected processing units that work together, can found an association of patterns among input and output.

- ELM: ELM is a feedforward neural network that can be used for classification, regression, clustering, sparse approximation, compression, and feature learning. The parameters of its hidden nodes need not be tuned.

- kNN: A supervised learning algorithm uses "feature similarity" to predict the values of new data points, in which the new data point will be assigned a value based on the distance it matches the points in the training set.

- K-Means clustering: An unsupervised learning algorithm divides all input data into k clusters, in which data in the same cluster are as similar to each other as possible.

- Boosting: An ensemble method can train weak learners sequentially, each strives to correct its predecessor.

- Bagging: An ensemble method applies the so-called Bootstrap statistical method to a high-variance machine learning algorithm. As typical bagging, RF can be structured from multiple decision trees.

- Stacking: An ensemble method can combine models of different types.

2.1.2. Deep Learning Methods

As a subdiscipline of machine learning, deep learning is an extension of ANN. Deep learning uses multilevel deep neural networks to extract features from the raw input progressively. The scale and complexity of the networks is the major difference between deep learning and traditional ANN.

A multilayer deep learning neural networks consists of some input layers, some hidden layers, and then the output layer. After loading input data into an input layer, hidden layers receive a group of weighted inputs and implement nonlinear transformations, and provide the output through activation functions.

A deep learning architecture is a multilayer stack of simple modules, all or most of which are subject to learning, and many of which compute nonlinear input-output mappings. Each module in the stack transforms its input to increase both the selectivity and the invariance of the representation. With multiple nonlinear layers, a system can implement extremely intricate functions [16].

Similar to conventional machine learning methods, common deep learning methods can be predominantly classified into two categories: supervised and unsupervised learning methods. Different categories have different architectures, which allowed them highly flexible.

CNN is a typical supervised deep learning method that achieves the best predictive performance in areas such as speech and image recognition by hierarchically composing simple local features into complex models. CNN can extract and classify features from high-dimensional data. As a variation of a multilayer perceptron, common CNN consists of one or more convolutions, max pooling, and fully connected layers [17].

The input layer is an $m \times n$ matrix in which every cell has a feature value. Each convolutional layer consists of several convolutional units, and the parameters of every unit are optimized by a back-propagation algorithm. The purpose of a convolutional manipulation is to extract different features from the input layer [18]. The first convolutional layer may only extract some low-level features such as lines, edges, and corners. Additional convolutional layers can iteratively learn more intricate representations from low-level features. Pooling is a critical manipulation technique in CNN [19]. It is a form of downsampling to reduce the dimensionality of feature maps, without altering the depth of these maps.

Since the initial development of CNN, multiple CNN architectures have been created. Some notable

examples include: VGGNet [20], ResNet [21], Inception [22], and DenseNet [23]. Each of these networks employs the same structure of convolutional layers and feature extraction but may vary in the number of layers they have, feature mapping, and efficiency [24].

RNN is mainly viewed as a supervised learning method, can be used for processing sequential data. RNN remembers the past, and their decisions are influenced by what they have learned in the past. RNN is made up of nodes, and the process of after being fed data, it outputs the result back into itself is repeated. This process allows the analysis of dynamic changes over time, where persistent information is needed [25].

LSTM is a special RNN architecture that inherits RNN's advantages of sequence learning and is able to learn time-series data with long temporal dependency [26]. With its memory block structure, LSTM models can judge whether the learned rules from the previous time step are useful or not and then determine whether the learned rules should be passed along to the next time step or abandoned. The prediction accuracy is thus not affected by the errors in some previous points.

Unsupervised learning algorithms are used to train each layer one at a time, independently, while using the previously trained layer as the input. After the pretraining step is performed on each layer, a fine-tuning step is performed on the whole network using supervised learning [24]. Common unsupervised networks include autoencoders, DBN.

Autoencoder trains a neural network so that the input and output become the same. In the same way as with general neural networks, the weights of the network are learned by stochastic gradient descent [27]. An autoencoder can extract features from unlabeled data using only a few layers. The network is symmetrical from the input to the output for dimensionality reduction and feature extraction [28, 29]. An autoencoder is capable of transforming raw data into sparse and nonlinear correlated features. Using the shallower hidden layer to obtain the optimal feature representation reduces not only the training loss but also the network complexity and network error, which reduces the amount of computation and thus speeds up the operation [30].

2.2. Landslides Prevention

2.2.1. Overview

A geohazard is a devastating phenomenon that is directly and indirectly caused by activity in the earth's interior or geological environment changes, including human activity or climate change. As one type of global geohazard, landslides are geological phenomena related to ground movements of rockfall and debris flow and can refer to the movement of a mass of rock, debris, or earth down a slope under the influences of
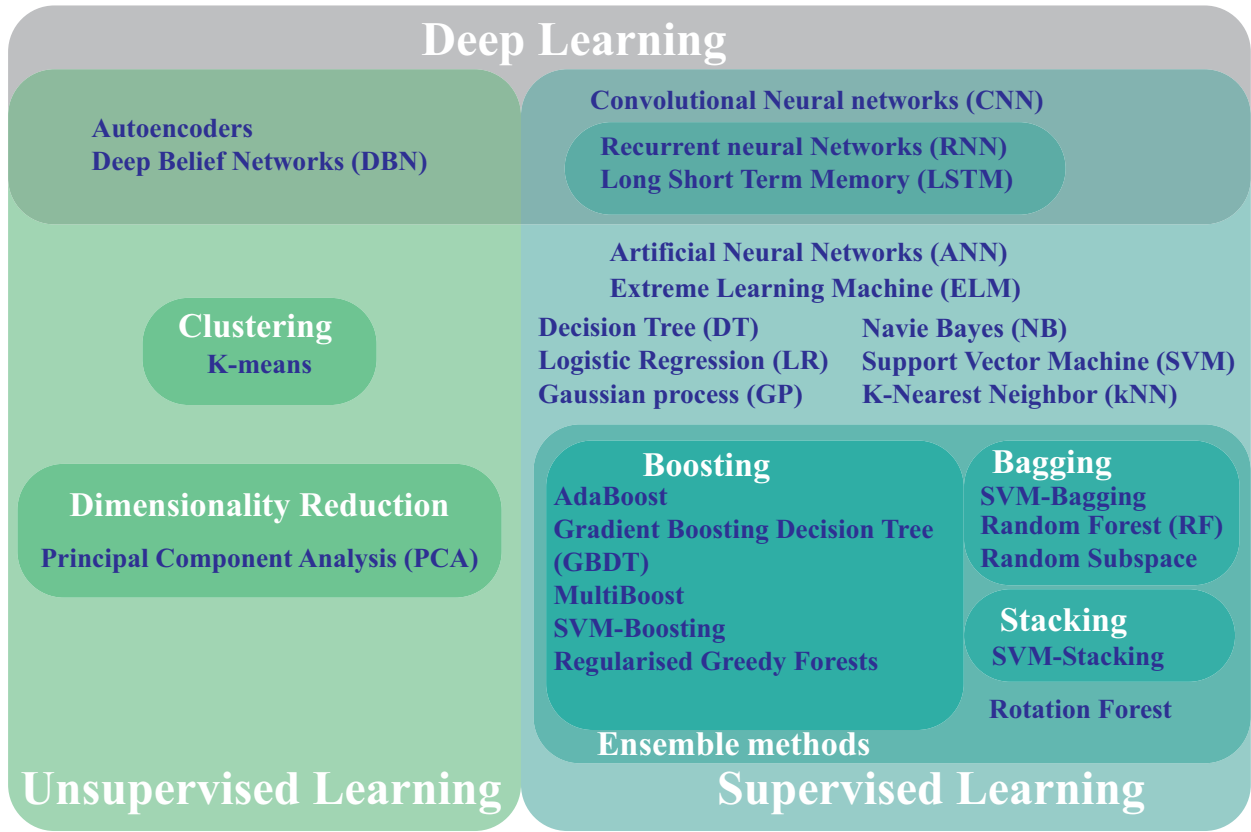
Figure 1: Main machine learning algorithms employed in landslides prevention

gravity, rainfall, and earthquake [31]. Lithology, tectonics, climate change, and anthropogenic pressure may cause slope instability that could progress to landslides [2, 32, 33]. Heavy rainfall, rapid snowmelt, or earthquakes could also trigger a landslide occurrence. Landslides are ubiquitous in any terrestrial environment with slopes.

In most cases, landslide occurrence means catastrophic results.it has brought out the massive destruction of infrastructure and even thousands of fatalities every year [1]. From 2004 to 2010, 2620 fatal landslides were recorded, causing 32,322 fatalities [34]. At least 17% of all natural-hazard fatalities around the world can be attributed to landslides [35]. In the most affected areas, financial costs and countermeasures are on the order of billions of dollars [36].

Recently, as a consequence of human disturbance (e.g., deforestation, mineral mining, and intensive exploitation of land for construction) and extreme weather, the frequency and intensity of landslides have increased dramatically. With the advent of extreme natural events, the prevention of landslides has become

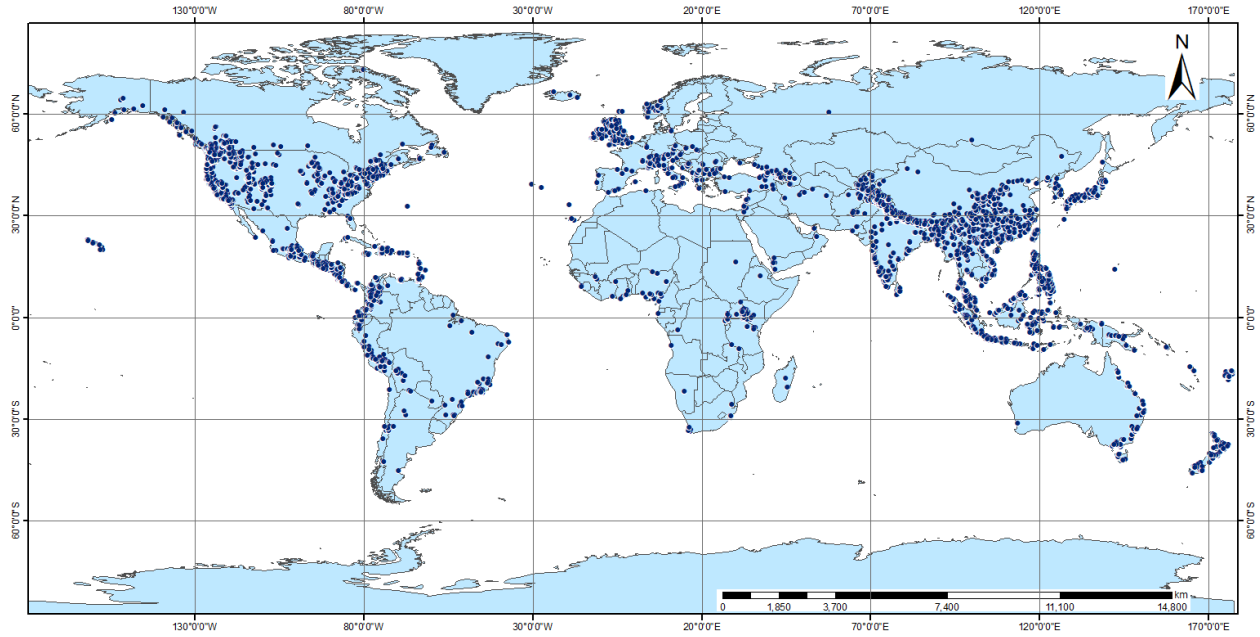Figure 2: The map below shows 2,085 reported landslides with fatalities from NASAs Global Landslide Catalog. The model showed more landslide activity in the Southern Andes, East African Rift Zone, Turkey, and Iran than was previously accounted for in the Global Landslide Catalog [37].

an urgent task. landslides prevention involves an assessment of slope instability phenomena and the change in the occurrence of slopes by means of effective geological engineering principles and other existing and emerging technologies. landslides prevention can provide valuable information for government agencies, planners, decision-makers, and local landowners to make emergency plans that reduce the negative effects on economics and human life. Typically, the study of landslides prevention is divided into two aspects: *detection* and *prediction*.

Related datasets for landslides prevention are generally obtained from three sources: (i) remotely sensed data acquired by Earth-observing satellites, (ii) data collected by in situ sensors, and (iii) data collected during fieldwork.

2.2.2. Landslide Detection

Fast and errorless detection of landslides is vital for rapid damage assessment and supporting disaster management activities and simultaneously increases the efficiency of disaster mitigation.

In general, landslide detection refers to identifying potential landslides and understanding fine-scale

landslide patterns. It is essential to quickly and accurately extract landslide information, especially in response to emergencies. Nevertheless, conducting field investigations of large landslides, especially for landslides that have just taken place, is rather dangerous and difficult [38]. Because of the risks in a field survey and the vastness of a disaster area, this type of investigation, which requires a large workforce and many material and financial resources, will be difficult to carry out. It is necessary to utilize some emergent techniques for automatically detecting landslides to avoid these disadvantages.

On the basis of detecting landslides, event-based landslide inventories should be generated as soon as possible; these inventories can provide baseline information (e.g., landslide types, location, magnitude, distribution, and boundaries) and depict the association between landslides and a single conditional factor [39]. In general, landslide inventories will be advantageous to understand the causal factors involved and to predict landslides.

2.2.3. Landslide Prediction

Significant human and economic losses push worldwide research for predicting future landslide events. Most landslide predictions follow a simple principle: the past and the present are the keys to the future. The analysis of past and current landslides will help in estimating landslide behaviors, frequencies, extents, and consequences in the future under certain conditions, which means that the spatial and temporal occurrence probability of landslides must be quantified [40]. Landslide susceptibility assessments are a static approach used to assess where landslides are most likely to occur in the future. Landslide early warning systems focus on information on the temporal landslide probability.

2.2.4. Typical Data Source for Landslides Prevention

In recent decades, the development of satellite, airborne, and ground-based remote sensing techniques has improved the ability to collect data, which mainly includes the Shuttle Radar Topography Mission (SRTM), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), or Light Detection And Ranging (LiDAR) instruments that characterize terrain morphology. Commonly collected data include visible imagery [41, 42], LiDAR data [43, 44], and Synthetic Aperture Radar (SAR) and Interferometric SAR (InSAR) data [45, 46], which can be acquired from platforms such as Google Earth Engine (GEE), the United States Geological Survey (USGS) Earth Explorer and Copernicus.

GIS technology offers a platform to integrate spatial information from these disparate sources into a single framework and to use these data for related analyses of landslides. The performance of neighborhood

operations with GIS allows the extraction of morphometrical and hydrological parameters from DEM [47]. DEM is a common type of data used in spatial analysis and are generally obtained by computing data derived from airborne or satellite sources.

In addition, mountainous regions have been monitored in various ways. A typical monitoring method is to embed different kinds of ground-based sensors related to slope movement, water table level change into the slope, and sense the dynamic change in signals. For example, local monitoring sites for landslides have been established by the USGS [48]. Rain gauges, slope movement sensors, and soil moisture probes can be used for monitoring. Wireless sensor networks are being used to achieve large-scale data collection and transmission. By employing different sensing and monitoring techniques, multidimensional and multiscale temporal and spatial data can be collected.

3. Machine Learning for Landslide Detection

3.1. Overview

Landslide inventory maps produced by image detection are usually the first step for training and testing in landslide forecasting studies [49], which means that fast and accurate landslide detection is useful for understanding landslides in a large area and predicting them in the future [50].

In detection, timely information about landslide positions, areas, and destruction levels is extremely important. With the development of satellite remote sensing technology, available datasets gradually accumulate, which can be mined for crucial information.

Previous traditional landslide image detection methods include pixel-based approaches [49] and the Object-Based Image Analysis (OBIA) [51]. To dispose of immense datasets and enhance their performance, many machine learning algorithms have been gradually introduced for these datasets. Conventional algorithms such as SVM [52, 53, 54], ANN [55, 56], and RF [57] were relatively widespread in the early years.

Recently, deep learning methods have begun to be used for solving image classification problems. For example, CNN can complete the classification of satellite imagery without requiring any preprocessing or feature extraction process [58].

3.2. Image Data Source

Remote sensing technology has proven its effectiveness in generating landslide inventories. Inventories are crucial for detecting landslides after triggering events, especially for remote or barely accessible areas

[59]. Multitemporal satellite imagery has been used to qualitatively analyze the temporal effects of phenomena and quantify the changes. The available data have become a major data source for landslide detection because of their high temporal frequency, a digital format suitable for computation, synoptic view, and a wide selection of spatial and spectral resolutions [60].

Except for satellite images, other available datasets for detecting landslides include bitemporal aerial photographs [50, 61], LiDAR [62, 63], and InSAR. Common landslide detection, according to these datasets, is as follows. (i) Optical image data (e.g., spaceborne or airborne remote sensing data) can be used for analyzing the image features of a landslide and then for recognizing the extent and location of the landslide by visual interpretation or automatic extraction methods. (ii) Radar data (e.g., SAR, InSAR, LiDAR) can be used for detecting surface deformations and deposition resulting from landslides.

3.3. Common Methods for Detection

The methods for detection based on image classification mainly include two categories: (1) pixel-based and (2) object-based methods.

3.3.1. Pixel-based Methods

An image pixel is the analytical unit of an image, which is generally assigned a value based on the detected electromagnetic energy, whose spectral characteristics are exploited to detect and measure changes without considering the spatial context [64].

By defining thresholds on pixel values, the pixels could be categorized into classes based on the spectral signature. Then, individual objects (pixels in a pixel-based approach) are allocated to the most likely class. For instance, a pixel could be classified accordingly as "landslide" or "non-landslide" by defining thresholds for pixel values in the green and red bands [65].

In the context of land cover classification and landslide detection, a variety of pixel-based studies have been conducted using different machine learning methods and earth-observation datasets. Pixel-based image classification can be treated as a typical binary classification problem, or a post-classification comparison can be performed to measure changes.

The ANN approach can provide better classification results than other methods when the classes are not normally distributed [66]. When applied to stacked multitemporal images, the SVM algorithm learns from training data and automatically finds threshold values from the spectral features [67]. For some fine-resolution images (e.g., 1 or 2 m), a Feedforward Neural Networks (FFNN) with one hidden layer and sigmoid transfer function can execute the classification [68].

15

Pixel-based methods often require extensive parametric tuning and precise geometrical correction or coregistration. Noise and outliers have significant effects on the accuracy of these methods [69].

3.3.2. Object-based Methods

OBIA is another detection method that can group neighboring pixels in regions before conducting the classification to address the limitations of pixel-based methods in high-resolution imagery. In contrast to pixel-based approaches, OBIA allows the integration of several landslide diagnostic features, such as spectral (e.g., pixel values, tones, and colors), spatial (e.g., object sizes, shapes, and patterns), and contextual features [70]. Since spectral, spatial, contextual, and morphological parameters can all be taken into account, OBIA has been employed in landslide detection from airborne and satellite data, including multitemporal data [51, 71].

Similar to other fields, machine learning methods can be used in conjunction with OBIA to select more appropriate features from VHR imagery. Some applied features include slope, local surface roughness [72], plan curvature [72, 73], slope and curvature together [74]. The most widely used algorithms include the kNN [75], RF [76] and SVM [77].

Selecting an appropriate machine learning classification method is important and requires taking into account many factors, including the spatial resolution of the data used, diverse data sources, and data types. Commonly, most machine learning steps are applied in OBIA as follows. (i) The image is segmented into homogeneous regions composed of similar pixels. (ii) Objects are classified into sets of features related to their spectral, spatial, and contextual properties [76].

For example, Stumpf et al. [71] utilized an RF-based method to evaluate the capability of a broad set of object metrics (color, texture, shape topography, and their sensitivity to changing scales of the image segmentation) and consequently removed nonrelevant features. Furthermore, they used RF-based methods to evaluate the consequences of class imbalance at each test site with different segmentation scales.

Pukar et al. [78] first used K-means clustering to obtain the Normalized Difference Vegetation Index (NDVI) threshold. Then, landslide candidate objects were separated from other vegetated surfaces. Finally, multitemporal landslide occurrence inventories were created. These inventories can be used to develop a landslide susceptibility model.

Both pixel-based and object-based methods use the spectral information of neighboring pixels or morphological properties as features, which usually require empirical parameters. Therefore, the detected spatial features strongly depend on experience and the parameter settings.

16

3.4. Deep Learning for Landslides Detection

Due to its hierarchical learning framework, deep learning methods are capable of extracting robust spatial and spectral features. It can learn the combination of satellite imagery and topographical features automatically and generalize the high-level features from low-level features layer by layer to best discriminate landslides from other objects [69, 79].

Recently, in the field of remote sensing, CNN is currently revolutionizing object detection and pattern recognition [80]. After considering multispectral bands (e.g., red, green, blue, red-edge, and near-infrared) channels, CNN enables the efficient analysis of topography, which in turn can reveal the unique spectral signatures and unique shapes required for landslides. CNN autonomously extract the contextual features of an image dataset and learn which landslide features are relevant for assigning the observations to the specified categories [81].

Using post-landslide images, Yu et al. [82] designed a CNN model that automatically learns and portrays the image data by extracting the discriminant area, boundary, and the center of the landslides as features. Thus, the landslide images can be retrieved from the test image set to achieve the objectives of feature detection and classification. The results demonstrate that the method does not need to manually design features and can achieve the intelligent detection of landslide disasters.

In other cases, both pre-landslide and post-landslide images have been utilized to detect texture changes. Ding et al. [83] removed landslide-irrelevant areas, including vegetation, water, and buildings in post-landslide images, and then built a CNN model that uses the concentration of landslides to extract with a high degree of change and consequently detected changes in image patches. The change degree is calculated by the Euclidean distance between patches in the same area in the pre- and post-landslide images. Finally, the unique texture features of the landslides were used to detect changes for accurately extracting the areas of landslides. These methods perform better at reducing the search areas and making the time period clearer than other methods.

The selection method of the sample patches plays a critical role in patch extraction by CNN. Based on the optical data from the RapidEye satellite, Ghorbanzadeh et al. [84] trained the same CNN architecture using different sample patches extracted from both random and central approaches and compared the detection results against a precise inventory dataset of landslide polygons through the mean Intersection-Over-Union (mIOU). The patches selected based on only the central areas of any landslide show an improvement in the resulting accuracy for landslide detection.

The aforementioned CNN architectures are simple and generic. The network depth is shallow and generally has 5 ~ 8 layers. Other variants of classic networks, including the adjustments to the width or depth and the introduction of depthwise separable convolutions or group convolutions, achieve slight and trivial improvements, thus showing slight performance differences in landslide detection.
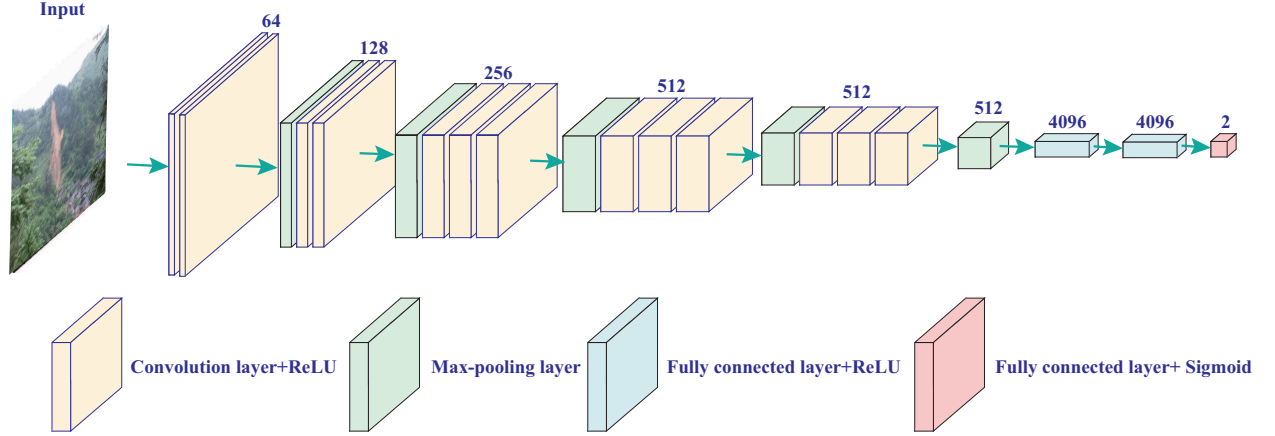

Figure 3: An example of the VGGNet model architecture in landslide detection. The VGG-16 model comprises simple convolutional blocks (i.e., convolutional layer, fully connected layer, and max-pooling layer). The ReLU and sigmoid functions are used for activation. The sigmoid function is applied to the last fully connected layer and outputs two real numbers between 0 and 1 indicating the probability belonging to landslide and non-landslide [85]

Ji et al. [85] evaluated the performance of landslide detection with different convolutional blocks, pooling strategies, reduction ratios in the multilayer perceptron, and placements of the attention module and designed a novel attention module that produced a 3D spatial and channel attention feature map. The 3D novel module emphasizes some parts of the input feature maps and suppresses the other parts in a backbone CNN; in other words, it emphasized foregrounds (i.e., landslides) and resisted noisy backgrounds (i.e., non-landslides). The experiment indicated that the attention module considerably strengthened the CNN structures, especially the ResNet-50 architecture. In contrast to other networks, such a ResNet architecture can deepen networks.

Maher et al. [69] detected landslides using ResNet architectures based on spectral bands and topographical variables. The results indicated that the architectures perform better by adding more convolutional or dense layers to the network.

In some cases, CNN employs global pooling and presumably misses some spatial information of images, thus resulting in low recognition accuracy. In general, landslide areas include serious spatial uncertainty,

resulting in difficulty extracting effective landslide features. A model can be produced that combines data from various modalities. Such models can learn abstract representations and merge them from different features.

Tao et al. [61] designed an asymmetric Fully Convolutional Network within Pyramid Pooling (FCN-PP) that is able to balance the use of context and localization accuracy, taking into account the combined features from multiple convolutional layers. After capturing five pairs of bitemporal images by using an aerial survey camera system, five areas including different types of landslides were obtained through cropping. Using these images, the proposed pyramid pooling model combined multiscale scaled features to provide a suitable feature representation of landslide areas. The final output includes changed and unchanged areas to achieve a binary classification (i.e., landslides and non-landslides).

4. Machine Learning for Landslides Susceptibility Assessment

4.1. Overview

Landslide susceptibility expresses the likelihood of a landslide event occurring in a given area based on local terrain conditions or climate conditions. It usually partitions the geographical surface into zones of varying grades of stability based on the landslide inventory [88]. The resulting output is a solely spatial distribution of the predicted categorized hazard probabilities across grid cells [89].

Machine learning methods applied for landslide susceptibility assessment represent a structured gathering of the available information extracted from landslide inventories, process/model with that information, and form a judgment about it in a transient workflow. This workflow unfolds through stages of preprocessing, implementation or modeling, and postprocessing, wherein modeling plays an essential role.

4.2. Workflow of a Machine Learning in Landslides Susceptibility Assessment

Supervised learning is by far the most widespread form of machine learning applied in landslide susceptibility assessment. The following are details about the workflow of supervised learning applied in landslide susceptibility assessment.

Initially, high-quality spatial data are collected from remotely sensed images or real-time monitoring for a landslide to produce landslide inventories [90]. A landslide inventory includes historical landslide data and other related information, such as geological data, meteorological conditions, and topographical data, which can roughly clarify the relationships between predisposing factors and landslide occurrences [91]. Based on these data, the predictive models for landslide susceptibility zonation can construct the

19
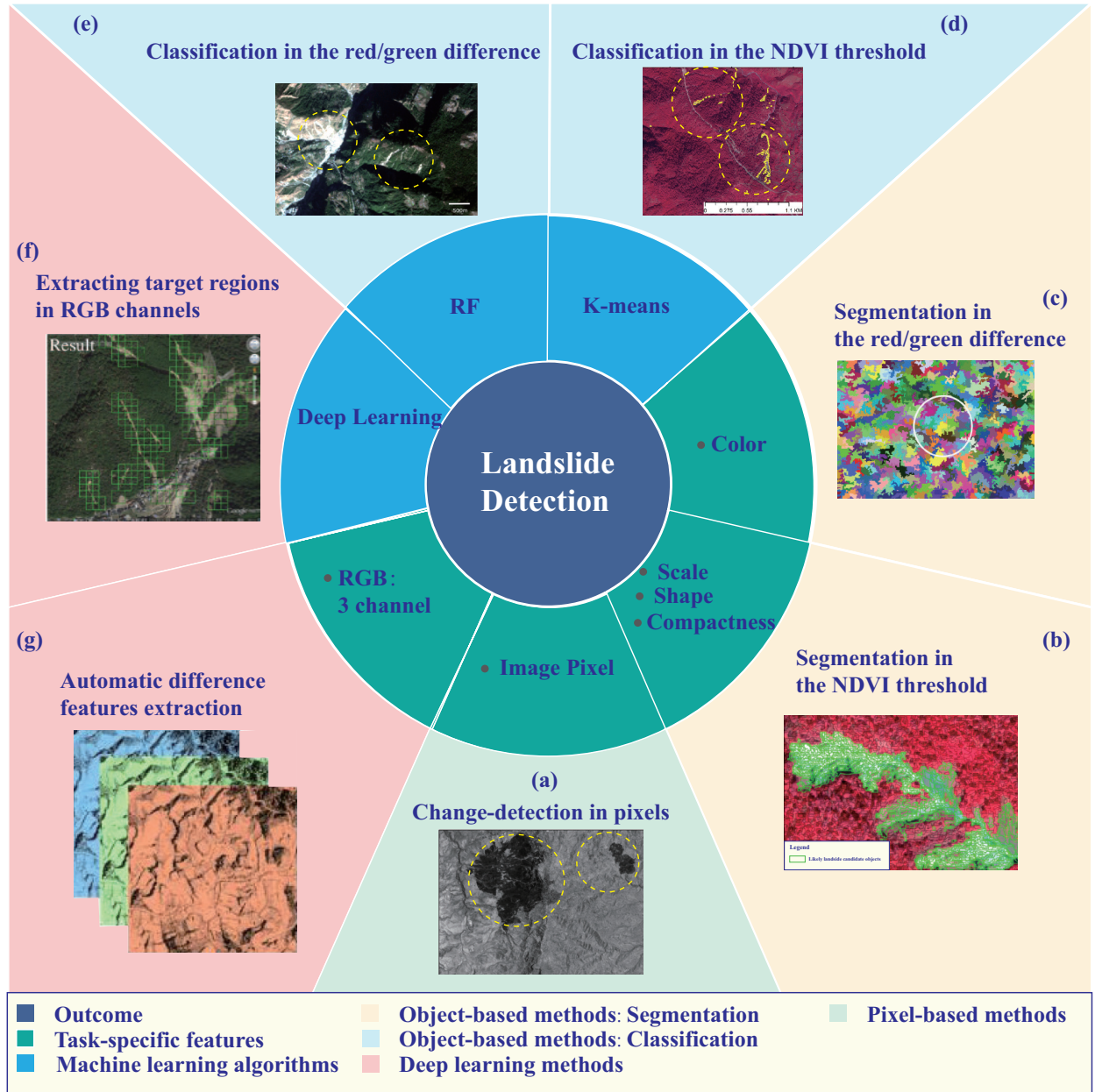
Figure 4: Examples of machine learning applications in landslide detection tasks. A pixel (a) [67] could be classified accordingly as "landslide" or "non-landslide". The first step when using object-based approaches is segmenting images basing on homogenous features, such as shape, color (b) [78], and scale (c) [86]. Then objects are classified (e) [86] (d) [78]. In deep learning, CNN autonomously extracts the contextual features of an image dataset and learn to identify landslide features (f) [87] by looking at R, G, B channels (g) [69].

relationships between the input and output variables [92]. Prior to any prediction modeling, these two types of variables should be identified. Commonly, the output consists of landslides and non-landslides. The input relates to conditioning factors of landslides.

Redundant or irrelevant factors may create noise, decreasing the overall predictive capability of the models. It is essential to choose suitable factors in landslide susceptibility assessment. Thus, the optimum conditioning factors for a study area will be subsequently determined, in which prior knowledge that refers to the analysis of the characteristics and geo-environmental conditions of the study area in relation to past landslide occurrences is needed.

To date, no universal guidelines have been agreed upon for the determination of case-specific conditioning factors [93]. Landslide conditioning factors show variation with respect to the study area and its geographical locations. Every study area has its own particular set of factors that cause landslides [94]. According to numerous studies, common landslide causal factors can be divided into two categories: (i) internal factors, which are related to geology and topography, such as the elevation, profile curvature, slope, plan curvature, distance to faults, aspect, distance from rivers, landform and lithology; and (ii) external factors, which usually cause landslides, such as rainfall, distance from roads and the seismic intensity.

To further select the appropriate input factors, one effective method involves ranking the importance of the input variables. Popular algorithms include ReliefF [91], Genetic Algorithms (GA) [95], Information Gain Ratio (IGR) [96], and symmetrical uncertainty analysis [97]. Through calculating a score for each factor, these algorithms can evaluate and rank the contributions of landslide causal factors, and the factors with lower contributions are sequentially removed. Furthermore, machine learning methods can rank these factors by their weights. The most frequently used supervised learning methods include LR [98], ANN [99], SVM [100], NB [101], and DT [102]. Unsupervised learning methods such as cluster sampling can evaluate factors by weighting the relative importance of each conditioning factor [103].

The predictive model is trained. The performance of the models is usually measured through some kind of cost function. It is also important to optimize model performance. This entails the adjustment of hyperparameters that control the training process, structure, and properties of the model [104]. For example, a validation dataset separate from the test and training sets using sampling strategies. The generic approach that was selecting the training sets is usually made by sampling 70% of all instances randomly throughout the available data. The remaining part is reserved for testing the model.

4.3. Conventional Machine Learning Methods for Landslides Susceptibility Assessment

Conventional machine learning algorithms have been applied to landslide susceptibility assessment and achieve outstanding performance and are mainly classified into single base learning algorithms and ensemble learning algorithms.

4.3.1. Single Base Learning Algorithms

The most frequently traditional single algorithms applied for landslide susceptibility assessment include (1) LR, (2) SVM, (3) DT, and (4) ANN [105, 106, 107, 108].
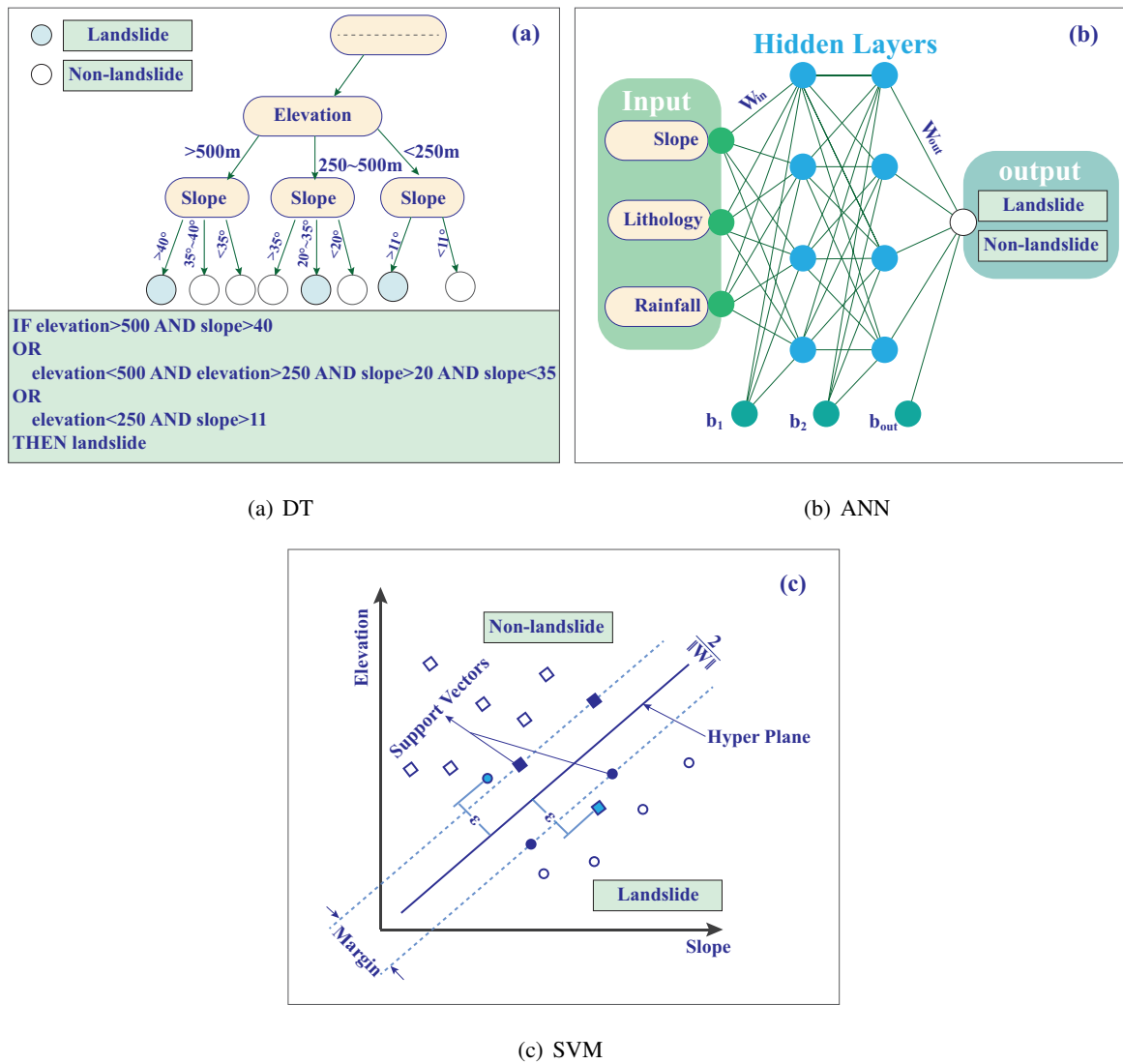

(a) DT

(b) ANN

(c) SVM

Figure 5: Three typical single algorithms that applied in landslide susceptible assessment

LR has a long tradition of application in landslide susceptibility assessment [109, 110]. A study proved that the predictive model complexity and the size of the training dataset influence the accuracy and predictive power of LR models concerning landslide susceptibility [111].

SVM can identify the optimal boundary between the training data from two classes [112]. Compared with other algorithms, the SVM algorithm achieves slightly better accuracies in shallower landslide assessment applications [113, 114, 115]. The quality of the predictive results produced by the algorithms is correlated with the input data quality.

As an original tree-like structure, DT can reveal some important relations between causal factors and landslides but yield landslide susceptibility models with lower accuracy than other methods. For instance, given the slope and altitude as input variables, a DT can reveal that slope is more important than altitude [116].

The standard ANN model comprises three layers, namely, an input layer (i.e., landslide conditioning factors), hidden layers, and an output layer (i.e., landslide susceptibility) [117]. A case proved that ANN applied in landslide susceptibility assessment achieved fairly precise models [118].

In summary, several drawbacks are usually identified when utilizing the aforementioned single base learning algorithms, such as overfitting and unstable performance.

4.3.2. Ensemble Learning Algorithms

Generally, ensemble learning algorithms can enhance the performance of the single base learning algorithms and improve the robustness and generalizability.

A commonly used ensemble algorithm in landslide susceptibility assessment is RF [119, 120, 121, 122]. Usually, an RF model has a more predictive capability to identify landslide susceptibility zones than other models [123]. Since the random selection of the training dataset may affect the results of the model, a set of many trees helps to ensure the stability of the model [121].

For example, Hong et al. [124] indicated that three ensemble models (i.e., AdaBoost, bagging, and rotation forest) could significantly improve the performance of J48 DT as the base learner, and rotation forest can be considered a promising method for landslide susceptibility mapping in similar cases with better accuracy than other methods.

Other ensemble methods have been developed for landslide susceptibility assessment, including GBDT [125, 126], Random Subspace [91], Multiboost [127], and Regularized Greedy Forests [128]. These ensemble methods can reduce both the bias and variance and avoid overfitting problems compared to the base

classifiers to improve predictive capability. The main advantage of numerous ensemble models is that model construction is not complicated, and the training time was short with the data at hand.

Dou et al. [129] produced four classes of ensemble learning models (SVM-stacking, SVM, SVM-bagging, and SVM-boosting) using SVM as the base learner. The study suggests that an ensemble machine learning model does not necessarily mean excellent performance. It is better to prudently select the proper model or the base learner for the predictive models. Each leaner should be seriously considered before grouping it into the ensemble models. In addition, the interpretation of the ensemble methods is more complicated than that of the base models [124, 127, 128, 130].

In conventional machine learning applications to landslide susceptibility assessment, the correlations between factors should be first eliminated to reduce model noise [131]. Furthermore, conventional feature engineering involves a substantial amount of prior knowledge in the process of seeking the proper parameters and thresholds of each feature. These empirical methods mean that the inherent and deep features of input datasets are challenging to extract [132, 133].

4.4. Deep Learning Methods in Landslides Susceptibility Assessment

Recently, with the rapid development of deep learning, state-of-the-art learning approaches have been successfully applied in landslide susceptibility assessment in the field. Indeed, deep learning has also been commonly applied to feature extraction [24]. Deep learning can find optimal features and handle indirect relationships between features and goals and can thus simplify the feature engineering and data preprocessing steps.

4.4.1. Autoencoder

When applying an autoencoder, landslide/non-landslide labels and linear/nonlinear correlation assumptions are not needed [134]. In general, an autoencoder neural network for landslide susceptibility assessment consists of input layers for raw feature dropout, hidden layers for sparse feature encoding, output layers for sparse feature extraction, and classification for prediction. The autoencoder is trained to reconstruct the input of the landslide-influencing factors onto the output layer for feature extraction and dimensionality reduction. The methods prevent the simple copying of the data and the network [135].

Maher et al. [128] used an autoencoder as an optimized factor to learn features from a dataset in an unsupervised manner [136]. They indicated that this factor optimization strategy based on unsupervised learning improves the performance of tree-based landslide susceptibility models by reducing the dimensionality. However, the strategy requires additional experiments and statistical analysis.
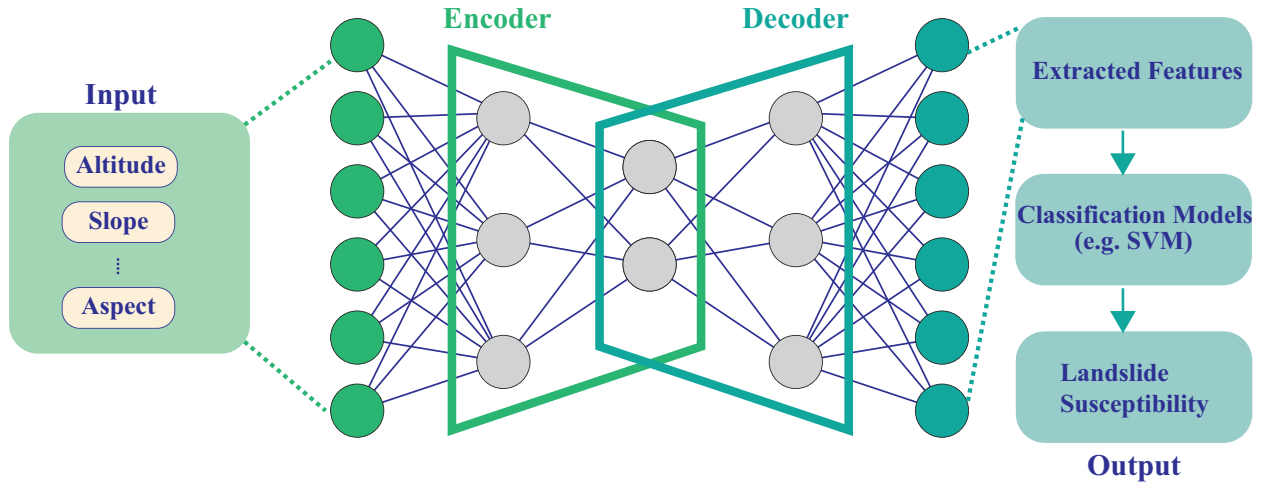
Figure 6: An autoencoder consists of two parts: an encoder and a decoder. The encoder compresses the input data into a fewer (two shown here) dimensions in the so-called bottleneck layer. The decoder tries to reconstruct the original input from the compressed data in the bottleneck layer [137]. Driven by data rather than prior knowledge, the autoencoder is capable of transforming raw data into sparse and non-linear correlated features [30].

4.4.2. CNN

As a typical deep learning method, when a CNN is applied in landslide susceptibility, whole landslide inventories can be regarded as an input matrix in which each pixel has several landslide-influencing attributes. That is, each condition factor layer can be a channel.

After selecting landslide-influencing factors, Wang et al. [138] constructed three CNN architectures (i.e., CNN-1D, CNN-2D, CNN-3D) to produce landslide susceptibility maps. More detailed results of the three architectures are as follows.

(i) CNN-1D could exploit the local correction and gradually learn more intricate representations form factor vectors to directly extract the information from landslide influencing factors and landslide susceptibility analysis.

(ii) After converting a one-dimensional input grid cell (vector) comprised of different attribute features into a two-dimensional matrix, CNN-2D can extract the valuable hidden features to reflect the probability of a landslide occurring. The output is divided into two classes: landslides and non-landslides.

(iii) CNN-3D not only learns factor representations but also extracts local spatial information. Specifically, the CNN-3D architecture extracts the influencing factor information and spatial relations to predict the probability of landslide occurrence.

25

The results indicated that CNN can effectively extract spatial information using local connections and can significantly reduce the number of network parameters by sharing weights; they can be used to produce reliable landslide susceptibility maps. Other related studies were scarce. Therefore, more research to verify different CNN architectures for landslide susceptibility assessment is needed.

Similar to other deep learning methods, CNN confronts a generalization problem: both underfitting and overfitting results in poor performance of classification models. As an optimization algorithm, gradient descent is acceptable and has been commonly used to tune CNN to minimize the cost function.

Pham et al. [139] selected a Moth-Flame Optimization (MFO) algorithm as a replacement. The regularization technique was applied by defining the search boundary (lower bound and upper bound) of the MFO algorithm to prevent the model from taking extreme values for connecting weights. In nature, moths move in groups in a spiral path towards a light source. Their positions can be modeled in dimensional space. In CNN applications, the hyperparameters of filters and weights of the fully connected layer were considered the dimensions of the moths, and the model searched for the best moth (i.e., the one with the smallest RMSE). Finally, the performance of the CNN model applied in landslide susceptibility assessment has been improved to some extent.

4.4.3. RNN

Another type of deep learning is RNN, which can use internal memory units to process arbitrary sequences of inputs. As a complicated evolution process, the formation and occurrence of landslides practically respond to the interaction of these multiple instability factors. The recurrent structure of the RNN will contribute to retaining the most critical information involved in a landslide occurrence and pass it to the next hidden state.

Wang et al. [140] sorted each landslide-influencing factor layer in descending order of importance. Landslide-influencing factors are viewed as a single-band image, and each pixel is converted into a sequential sample according to importance. Accordingly, the most important factors are sent to the RNN structure first, and the least important factors are sent last.

To express the complex relationships between landslide occurrences and continually changing factors, Xiao et al. [141] proposed a novel and dynamic model that can remember historical data using memory blocks. After collecting data and extracting features from the DEM, high-resolution remote sensing images, geologic maps, and meteorological data from January 2015 to December 2016, they built an LSTM model to solve the landslide susceptibility classification problem.
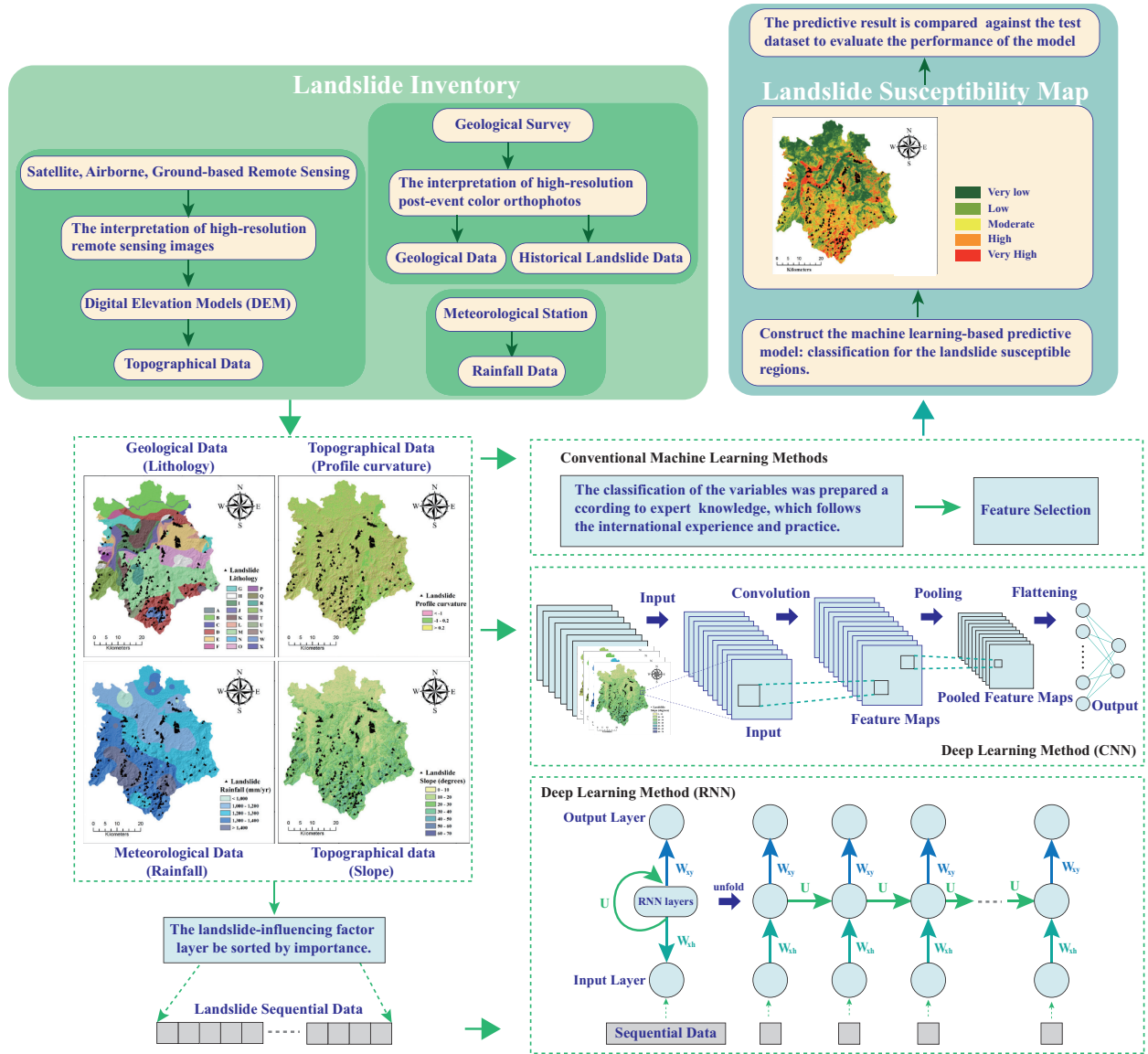
Figure 7: Workflow and general framework for machine learning methods in landslide susceptibility assessment. Conventional machine learning methods conduct hand-crafted feature selection. Deep learning methods (e.g., CNN, RNN) implement feature extraction. In CNN architectures, the landslide inventory is the input data. It can be regarded as a picture. Here, each pixel integrates information of different several influencing factor attributes. Layers of maps that converted from each factor represent a channel, respectively [138]. In RNN architectures, W_{xh} refers to weights for the connection of the input layer to the RNN layer (i.e., the hidden layer). U refers to the weights of connection in the RNN layer. W_{hy} refers to the weights for the connection of the hidden layer to the output layer [140].

27

5. Machine Learning for Landslide Warning Systems

5.1. Overview

Reliable early warning systems are a reasonable approach for risk reduction and can significantly reduce economic losses and casualties [142]. These systems are designed to predict the short-term behavior of single landslides according to information, including the potential triggers and their thresholds for issuing emergency warnings and the sliding mechanism [143, 144].

Rainfall is the most widespread and frequent trigger of landslides around the world [145]. Generally, landslide early warning systems are often based on the assumption that a critical rainfall amount exists and, when reached or exceeded, triggers shallow landslides. For estimating the minimum critical rainfall amount, a widely used method is the determination of a rainfall threshold. The rainfall threshold identifies the boundary that divides rainfall data into two categories, namely, rainfall that induced a landslide and rainfall that did not induce a landslide and thus can be mainly implemented to predict the temporal occurrence of landslides [146, 147].

In the case of a deep-seated landslide, when rainfall is intense, the infiltration rate exceeds the deep drainage rate, triggering instability conditions. An increase in the level of groundwater often induces landslides [144], which can be another important factor. It is necessary to evaluate fluctuations in the groundwater level accurately based on previous rainfall to predict the occurrence of deep-seated landslides [146]. The groundwater level is directly related to landslide displacement, which is a complex nonlinear dynamic process.

Predicting landslide displacement also plays an important role in operational early warning systems, as it will offer more detailed information for early warning systems than triggers [148, 149, 150]. Commonly, the sliding mechanism of landslides results from combining local geological and other external influencing factors. The dominant factors are precipitation and fluctuations in the groundwater level [151]. As a major result of the sliding mechanism of landslides, displacement is a key parameter to capture the interaction between landslide deformation and hydrometeorological conditions [152]. Their rapid changes are generally considered a direct sign of an upcoming landslide. Thus, to produce a new model for the prediction of landslide displacement, both the deformation mechanisms and the influencing factors should be taken into account.

Machine learning has been applied to the abovementioned various aspects of landslide warning systems, including classifiers for the analysis of rainfall thresholds themselves, cooperation between rainfall thresh-

olds and landslide susceptibility assessments, determination of rainfall thresholds in deep-seated landslides as an inverse analysis and prediction of landslide displacement.

5.2. Analysis of Rainfall Threshold

Common early warning systems need to take into account both landslides triggered by short and exceptionally intense rainstorms (e.g., shallow landslides) and landslides triggered by exceptionally prolonged rainfalls (e.g., deep-seated landslides). The decisional algorithm at the core of the warning system is based on the comparison between the thresholds and the rainfall data (recorded and forecasted) [147]. Rainfall data are collected from rain gauges. Most predictive models define rainfall thresholds based on the analysis of past rainfall events that have resulted in landslides.

5.2.1. Analysis of Rainfall Threshold in Shallow Landslides

Defining rainfall thresholds requires defining a linear frontier between two categories. As a widely used two-class linear classifier, SVM has been used to determine rainfall thresholds for shallow landslides [153, 154, 155]. However, a single threshold cannot predict the exact localization of the landslides. For enhancing the spatiotemporal forecasting capability of a regional-scale landslide early warning system, a reasonable approach is to integrate rainfall thresholds into landslide susceptibility assessments.

Pradhan et al. [156] estimated the combined probability of landslides and triggering rainfall thresholds using a hazard matrix. This study is described in detail as follows. (i) Based on information on critical rainfall intensities and durations extracted from historical landslide data, rainfall threshold warning levels were prepared. (ii) A backpropagation ANN was used for landslide susceptibility assessment. (iii) Rainfall threshold warning levels and the classified shallow slide susceptibility were combined using a matrix table. After application to a practical example, the model provided reasonable results.

After complete landslide susceptibility assessment using an RF model, Segno et al. [147] integrated the results into statistical rainfall thresholds and assessed regional-scale forecasting of landslides. This study showed that the coupling of the two methodologies enhanced the forecasting effectiveness of the warning system.

Kirschbaum et al. [145] developed a near-global Landslide Hazard Assessment Model (LHASA) that combines surface susceptibility and satellite rainfall data to provide a qualitative landslide prediction. According to the DT structure, if the susceptibility values are moderate to high or very high, nowcasts are issued.

On the other hand, deep learning methods can extract features from raw rainfall data to predict a landslide. For example, Huang et al. [157] built a DBN model to train a large number of unlabeled data in an unsupervised way. The rainfall data comprised the average annual rainfall, daily rainfall, and cumulative precipitation from the previous seven days. The DBN was stacked by a series of RBM. The softmax classifier was added to the top layer of the DBN. The dropout mechanism is introduced in the RBM hidden layer structure to sample the node weights of the hidden layer with a probability of 50% to prevent overfitting. There are four labels (minor landslides, medium-sized landslides, large-scale landslides, and huge landslides) according to the disaster degrees and scales of the landslides in the data samples.

5.2.2. Analysis of Rainfall Threshold in Deep-seated Landslides

Generally, the mechanisms that induce shallow landslides and deep-seated landslides are different. Rainfall can result in rising groundwater levels. Elevated pore water pressure, induced by the change in the groundwater level, is one of the main triggers of deep-seated landslides [158]. Their interactions are complicated. For predicting deep-seated landslides, the groundwater level should be considered a relevant parameter.

To explore the relationship between rainfall and groundwater level, Qing et al. [146] developed a rainfall threshold model using a genetic algorithm SVM and a backpropagation neural network based on the antecedent rainfall data for forecasting variations in the groundwater level caused by rainfall. Then, to determine the critical threshold of the groundwater level that could reactivate the landslide, numerical computations were conducted under different values for monitoring and predicting the groundwater level. The rainfall threshold was determined in reverse. In this study, the factor of safety acquired from numerical simulations offers a threshold to divide stable/unstable landslides.

Moreover, Vallet et al. [159] used a velocity criterion method based on displacement velocity time-series data to distinguish acceleration crises (peaks) and periods of rest (troughs) to account for unstable and stable landslides, respectively. An SVM model was used to obtain the best coefficient of determination between the cumulative groundwater recharge and the landslide velocity and then to determine the rainfall threshold.

5.3. Prediction of Landslide Displacement

Forecasting landslide displacement is an important part of operational early warning systems. Landslide displacement time-series data can directly reflect landslide deformation and stability characteristics

[10]. Therefore, it has been used to develop a landslide displacement prediction model [160]. Generally, these time-series data were collected from landslide monitoring systems (e.g., GPS, InSAR, or GNSS). In addition, the triggering factor is commonly used data, which includes precipitation and the groundwater level.

Under geological conditions such as geomorphology and geological structures, a variation in the displacement over time can be viewed as an approximately monotonically increasing function on a large time scale. Under external influencing factors such as the groundwater level and rainfall intensity, the variation can be correspondingly viewed as a periodic function on a small time scale. These available sequence data provide critical input data to predict landslide displacement.

5.3.1. Conventional Machine Learning Methods for Predicting Landslide Displacement

Recently, conventional machine learning methods, including the ANN [143], SVM [161], Gaussian process [162], and ELM [163], have been applied to produce models for landslide displacement prediction. Here, the input is the landslide displacement and the triggering factor. The output is the predicted landslide displacement.

For example, Krka et al. [142] presented an RF-based methodology for the prediction of landslide movements using the landslide velocity and displacements from precipitation data. The modeling of the groundwater level change rate from the daily and historical precipitation data was followed by modeling the landslide velocity from the predicted daily groundwater level depths (calculated from the groundwater level change rates). Ultimately, the trained model was used to predict the landslide velocity for nine periods ($10 \sim 90$ days).

To optimize time-series data used as input, Li et al. [10] introduced a chaos theory-based Wavelet Analysis-Volterra filter model (chaotic WA-Volterra model) into SVM for cumulative landslide displacement prediction. The WA-Volterra model aims to decompose the cumulative displacement data into different low- and high-frequency components. Chaos theory was used to reconstruct the phase space of each frequency component. Reconstructed phase spaces were selected as the input-output data to train the SVM models. The predictive results (i.e., the predictive cumulative displacements) were obtained by summing the predictive displacements of each frequency component. This study indicates the potential for chaos characteristic identification of landslide displacements to be applied in machine learning. A certain optimization has been achieved in feature processing.

31

5.3.2. Deep Learning Methods for Predicting Landslide Displacement

The aforementioned approaches regard landslide displacement prediction as a static regression problem. On the other hand, landslides are considered a dynamic system in which the displacement continues to change. The influencing factors and displacement conditions in one moment affect the displacement and stability conditions in the next moment. To investigate the dynamic process, LSTM is an appropriate method since it is suitable for learning the temporal dynamics of sequential data.
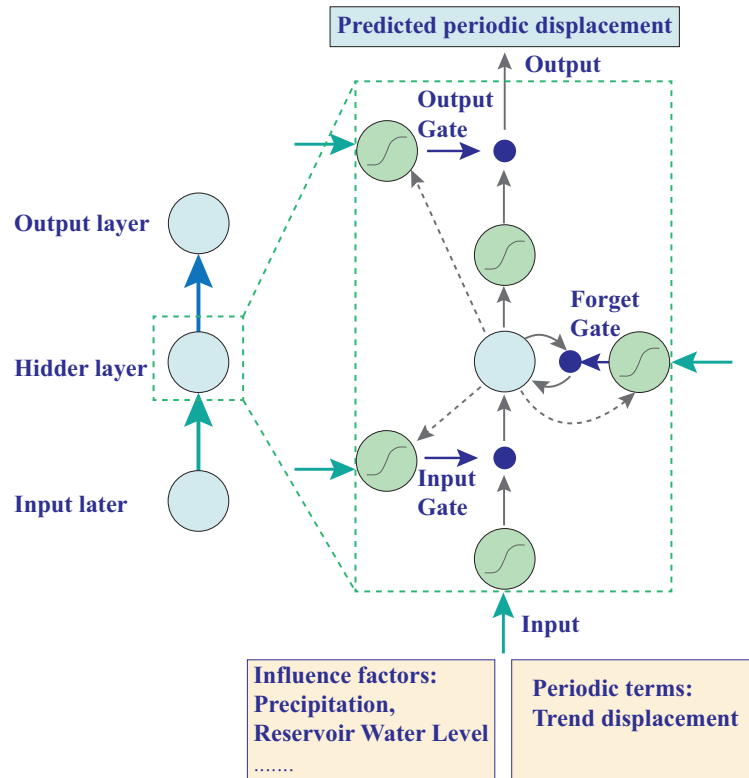

Figure 8: LTSM neural network architecture.

The general workflow for the application of LSTM in landslide displacement prediction is as follows. The measured accumulated displacement of the landslide is first divided into a trend term (i.e., a static component) and a periodic term (i.e., a dynamic component). Selected controlling factors and periodic terms will be considered input. Generally, LSTM adds loops to the architecture, receives these inputs, and outputs a predicted result. Finally, the LSTM model was validated and estimated by comparing the predicted total displacement with the monitoring results of the total displacements.

For example, Yang et al. [160] first analyzed the relationship between landslide displacement and key influencing factors (i.e., rainfall and reservoir water level). Then, an LSTM model was produced for predicting the periodic term displacement. The LSTM model can establish connections between landslide conditions at different times and learn rules from previous deformation time steps. The results indicated that the LSTM model achieved a more satisfactory performance than static SVM methods.

Xie et al. [164] adopted an LSTM model to predict dynamic landslide displacement by evaluating the dynamic characteristics with the time domain. The prediction result indicated that the rainfall intensity and the excavation-induced stress redistribution affected the periodic displacement. Moreover, the measured and predicted deformation results showed good consistency.

6. Discussion

Predictive models developed by machine learning for landslides prevention can be underconstrained. For instance, models that perform well in datasets and are consequently viewed as high quality probably deviate strongly for situations and data outside their valid local areas because of the complex physical earth system. The challenges and opportunities in the applications of machine learning for landslides prevention will be discussed below.

6.1. Challenges from Uncertain and Complex Datasets

6.1.1. Noise in Datasets

As mentioned above, selecting an appropriate machine learning model for landslides prevention depends largely on the input dataset. These datasets involve complex, nonlinear, physical earth systems that act across a vast range of spatial and temporal scales, predominately consisting of geological and meteorological environments.

As one of the critical triggering factors, rainfall is strongly nonstationary under climate change. The related data may yield a great deal of noise. Additionally, sensor interference and instrument malfunctions could also induce noise/uncertainty. Sequences of (multisensor) satellite observations also have diverse noise sources, uncertainty levels, missing data, and (often systematic) gaps (e.g., acquisition, storage, and transmission distortions) [165]. In the case of remote sensors, atmospheric (clouds and other aerosols) and surface (snow and ice) interference are constantly encountered [12].

For instance, aerial and satellite images having a varied nature of reflectance (e.g., multispectral, hyperspectral and optical images, LiDAR and drone point clouds) can generally be used as input datasets.

These datasets can contain more noise, coarser spatial resolution, more complex object geometry, and fewer samples with very large and irregular sizes.

6.1.2. Dataset Heterogeneity

Climate and ecosystem processes reveal a high level of heterogeneity due to differences in geography, topography, and climatic conditions in diverse areas of the earth. For example, some regions are mountainous, some regions are dry and experience severe, long-term droughts, and some regions are quite wet and covered with dense forests. The patterns, mechanisms, and driving forces of landslides vary among these regions.

Currently, this heterogeneity in the data emphasizes the idea that landslides prevention models primarily apply to local or regional zones. Various factors correspond to a homogeneous group of locations. Developing a universal model that can be applied to global regions remains a challenge. Moreover, another source of heterogeneity is presented in different multi-sensors, which exhibit different imaging geometries, spatial and temporal resolutions, physical meanings, contents, and statistics.

6.1.3. Deluge of Datasets

Another important issue needed to carefully consider is the data deluge. With technological development, airborne LiDAR surveys, SAR satellites, stereophotogrammetry, and mobile mapping systems are increasingly used and produce data volumes that raise computational challenges for spectral, spatial, and temporal dimensionalities.

Acquiring the effects of these multiple variables at fine spatial and temporal resolutions renders these data inherently high dimensional, where the number of dimensions can easily reach millions. For example, a dataset with a comparatively rough resolution can easily produce over 10,000 spatial grid points, where every grid point has multiple observations in time [166]. Furthermore, unlike classic computer vision applications that deal with photos with three channels (red, green, blue), hyperspectral satellite images extend to hundreds of spectral channels well beyond the visible range.

Overall, entirely data-driven approaches can present a challenge. There are several strategies for mitigating these problems.

The Persistent Scatterer InSAR (PS-InSAR) can select pixels with high coherence throughout the interferometric phase history. These pixels are less affected by spatial and temporal decorrelation noise and possess a highly stable phase history [167, 168].

To a certain extent, the utilization of deep learning techniques can mitigate data-related problems [169]. Using simulations to generate supplemental synthetic training data can mitigate reliance on large datasets. This approach uses domain adaptation to adjust for divergence in the data distribution between real and synthetic data. Domain adaptation can select some novel architectures, such as mixed-reality GAN [170].

Since complex and dynamic physical processes govern landslides, conventional machine learning methods applied to landslides prevention still need to analyze scientific datasets with some expert knowledge. Indeed, deep learning models are important for mitigating potential data problems, but the predictive result may also be physically untenable. Providing strong theoretical constraints on top of the data-driven model is needed, which can be achieved by integrating professional knowledge.

6.2. Challenges from Class Imbalance

A common hindrance in machine learning for landslides prevention is the class imbalance problem. All machine learning-based methods need to extract certain features regarding landslide and non-landslide data samples for analysis and then find a classification boundary to divide the training areas into two classes (i.e., landslides and non-landslides).

There are fewer areas in the training regions in which landslides appear than non-landslides. Such imbalances can cause a model to be biased towards classifying the susceptible areas as safe since there is a larger number of non-landslide samples. After investigating the mapping of landslides with an imbalanced training sample (i.e., the sample contained more examples of non-landslide areas than landslide areas), a study indicated that the RF method underestimated landslide occurrences [71]. For overcoming this problem, some typical solutions can be divided into three categories: data-level techniques, algorithm-level methods, and hybrid approaches [171].

6.2.1. Data-level Techniques

Data-level techniques mitigate imbalance problems through diverse sampling methods. One study selected a 1:1 sampling ratio of landslide data points to non-landslide data points by randomly selecting [172]. Beyond random oversampling, there are other sophisticated forms of resampling strategies, including the Synthetic Minority Oversampling Technique (SMOTE) [173] and Synthetic Minority Oversampling Technique-Iterative Partitioning Filter (SMOTE-IPF) [174]. Both approaches use the kNN algorithm to synthesize new instances. However, these strategies often result in drastically smaller datasets.

In addition, more strategies addressing class imbalance are needed for deep learning-based models.

Random oversampling can improve the classification of imbalanced image data [175]. Random undersampling can decrease the amount of class imbalance for pretraining a deep CNN. Dissimilar to random undersampling, which completely removes potentially useful information from the training set, the two-phase learning strategy [176] can remove some non-landslide areas as the majority group during the pretraining phase. Moreover, this removal allows landslide areas, as the minority group, to contribute more to the gradient during pretraining and still allow the model to see all of the available data during the fine-tuning phase. Arguably, the two-phase learning procedure is critical for dealing with imbalanced distributions in image data [177].

Another strategy, called the dynamic sampling technique, oversamples the low-performing classes and undersamples the high-performing classes [178]. Currently, after introducing the class imbalance problem in the landslide prediction task, a study indicated that almost all instances, both synthetic and original, had been correctly classified by the RF classifier after preprocessing data using the minority oversampling technique with iterative partitioning filter (SMOTE-IPF) [179].

6.2.2. Algorithms-level Techniques

Algorithm-level techniques to solve class imbalance problems are termed cost-sensitive learning approaches, which solve data imbalance problems by modifying the algorithms themselves by a class penalty or weight and distributes different cost values for the misclassification. For instance, the cost of misclassifying a landslide into non-landslide areas would be much higher than the opposite. The former may put human property and life at stake, while the latter just overprotects some areas identified as landslides. Commonly, the Mean Squared False Error (MSFE) can be selected to capture the errors equally.

To combat class imbalances in datasets for deep learning, Lin et al. [180] presented the Focal Loss (FL), which reshapes the Cross-Entropy (CE) loss to reduce the impact that easily classified samples have on the loss. The method can reduce the weight of non-landslide areas. Similar to the MSFE, the FL is relatively easy to integrate into existing models and hardly impacts the training time.

Gradually, some cost-sensitive deep learning methods are emerging, which learn network weight parameters and class misclassification costs during training and thus give higher importance to samples with a higher cost, such as landslide areas. Some methods include the Cost-Sensitive (CoSen) learning [181] and the Cost-Sensitive DBN with Differential Evolution (CSDBN-DE) [182].

6.2.3. Hybrid Techniques

Hybrid systems strategically combine both sampling and algorithmic methods [183]. One strategy includes performing data sampling to reduce class noise and imbalance and then apply cost-sensitive learning or thresholding further to reduce the bias towards the majority group [171].

6.3. Challenges from Lack of Interpretability

Well-trained neural networks still have the typical issue of the lack of interpretability. In other words, it is difficult to explain the results from the input data by a machine learning model. Interpretability is a useful debugging tool for detecting bias in machine learning models. Machine learning models can only be debugged and audited when they can be interpreted [184]. For example, landslide susceptibility maps produced by machine learning need to earn the trust of government agencies, planners, decision-makers, and local landowners and thus make them make emergency plans for reducing the negative effect brought by landslides.

Interpretability includes the visualization of the results for analysis by humans [165]. Obviously, relying on visualizations of neuron activations is not enough. Further research is needed to understand the patterns that models use to link to the training datasets.

Given their complexity, landslides prevention models in the earth system are often not easily traceable back to their assumptions, limiting their interpretability [165]. The analysis of inverse problems in the landslide application is an attempt to achieve interpretability [185]. Fusion approaches that integrate machine learning and landslide-related knowledge have attracted much attention in recent years. These methods might initially use deep learning methods for feature processing with subsequent reliance on handcrafted machine learning approaches for prediction, thereby leveraging domain knowledge to ensure interpretability [25, 85].

Machine learning results will predominately be viewed as hypotheses and inspiration to further studies. Complementary machine learning-based predictive results will help geohazard management agencies build trust in these approaches and outputs.

Moreover, there is another noticeable issue. It is difficult to troubleshoot developed machine learning models because of unseen inputs.

6.4. Opportunities by Using More Machine Learning Methods

With the development of machine learning methods, some state-of-the-art techniques can provide more opportunities for landslides prevention by overcoming current limitations.

6.4.1. GAN

GAN is an unsupervised learning architecture for generative modeling using deep learning methods, such as CNN, which can automatically learn the regularities or patterns based on input data.

These models can be viewed as a supervised learning problem with two submodels: the generator model and the discriminator model. The generator model aims to capture data distribution. The discriminator model aims to estimate the probability that a sample belongs to training data rather than being generated by the generative model. Iterative adversarial training is repeatedly completed between these models.

The power of abstraction in these models allows their higher levels to learn concepts and categories far more rapidly than their lower levels, owing to strong inductive biases and exposure to more data. GANs can substantially reduce the computational cost and mitigate overfitting [186].

Common analyses of landslides prevention (e.g., landslide susceptible assessment) primarily acquire data from remote sensing images. The tremendous volume of remote sensing images can make the process of labeling all the data prohibitively time consuming and expensive. GAN is an excellent opportunity since they can create training data by the generator.

Recently, some research has leveraged GAN models to learn the representation of remote sensing images using unlabeled data [187, 188, 189]. Lin et al. [188] designed a Multiple-layer Feature-Matching GAN (MARTA GAN) model. Remote sensing images with a resolution better than 256×256 were produced by adding two deconvolutional layers in the generator. The classification accuracy of remote sensing images has been improved. The model can learn interpretable representations even from challenging remote sensing datasets.

Beyond improving the imaging quality of remote images, including the temporal and spatial resolution and a high signal-to-noise ratio, GAN models may be applicable in landslide detection. Some prior studies demonstrate that GAN has the potential to solve anomaly detection problems [190].

6.4.2. GNN

While deep learning effectively captures hidden patterns in Euclidean data, there is an increasing number of applications where the data are represented in the form of graphs [191]. GNN is becoming increasingly popular, which enjoys the major advantage of incorporating a sparse and discrete dependency structure between data points. A graph allows the representation of a multitude of associations through link orientation and data points [192].

GNN is gradually being applied in some domains, including scene graph generation, point cloud clas-

sification, and segmentation. This increase means an opportunity for analyzing unstructured spatial vector data and landslide detection/prediction [193].

GNN can be used to learn discriminative features from input graph-structured data. After the acquisition by LiDAR scans, datasets such as point clouds can be classified and segmented. GNN explores the topological structure based on the data [194, 195].

As one of the predisposing factors for landslides, topographic-related data can also be considered graph- and manifold-structured data, which can be set as a graph signal for the features of each point in the point cloud. From this perspective, GNN models will offer more outlooks for landslide prediction.

Previously, landslide events were assumed to be independent. That is, interdependence has almost no effect when referring to prevention. With modern networks growing exponentially in size, variety, and complexity, various emerging types of communication networks are more useful for landslides prevention, such as the Internet of Things, wireless sensors, and cloud-based networks [196]. Complex patterns are emerging, which bring out interdependence among diverse events [197]. Incorporating the effect of interdependence into landslides prevention in some regions might present an opportunity for applying GNN. For example, within each network, a landslide location as a node is connected, via weighted links, to a number of other landslide locations.

6.4.3. Automated Approaches

Regardless of the model, it is necessary to select the proper parameters and thresholds of each feature. Automated approaches might simplify and even skip this process.

Deep learning methods applied in landslide detection can be fully automated to design stable and efficient networks that are stable and efficient using small training datasets. For example, CNN-based architectures are composed of a stack of layers or a stack of modules. Automated approaches can determine the optimal number of layers or the optimal number of modules required for a given application task, and to optimize the internal structure of a module [198].

Although the application of novel deep learning methods to landslides prevention research is still in its infancy, current applications in other domains already demonstrate that these methods have enormous potential for application in landslides prevention in the future.

7. Conclusions

A survey of machine learning algorithms applied in landslides prevention has been presented in this paper, which focuses on (1) landslides detection based on images, (2) landslides susceptibility assessment, and (3) the development of landslide warning systems. The survey shows that machine learning methods has been widely used in landslides prevention and can achieve satisfactory performance. However, there are still several challenges and limitations. First, professional knowledge is needed, which can facilitate the selection of more appropriate variables and datasets when facing increasingly complex and massive data. Second, interpretability is also a critical component in landslides prevention. The majority of established scientific theories on landslide occurrence mechanisms struggle to explain the results of machine learning models. Analyses involving machine learning results should be interpreted in combination with landslide mechanisms. Therefore, a potential research trend is to combine data-driven machine learning with expert knowledge of landslides. Gradually increasing the application of machine learning in landslides prevention will enable benefits for both the machine learning and landslide research domains.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Reference

[1] M. Froude, D. Petley, Global fatal landslide occurrence from 2004 to 2016, Natural Hazards and Earth System Sciences 18 (2018) 2161–2181. doi:10.5194/nhess-18-2161-2018.

[2] M. Jakob, S. Lambert, Climate change effects on landslides along the southwest coast of British Columbia, Geomorphology 107 (2009) 275–284. doi:10.1016/j.geomorph.2008.12.009.

[3] M. Crozier, Deciphering the effect of climate change on landslide activity: A review, Geomorphology 124 (2010) 260–267. doi:10.1016/j.geomorph.2010.04.009.

[4] R. Tichavsk, J. Ballesteros-Canovas, K. ilhn, R. Tolasz, M. Stoffel, Dry spells and extreme precipitation are the main trigger of landslides in central Europe, Scientific Reports 9 (2019) 1–10. `doi:10.1038/s41598-019-51148-2`.

[5] A. W. Root, Prevention of landslides, Landslides and engineering practice, Highway Research Board, Spec. Rep 29 (1958) 113–149.

[6] C. Westen, E. Castellanos, S. Kuriakose, Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview, Engineering Geology 102 (2008) 112–131. `doi:10.1016/j.enggeo.2008.03.010`.

[7] C. Van Westen, N. Rengers, R. Soeters, Use of geomorphological information in indirect landslide susceptibility assessment, Natural hazards 30 (3) (2003) 399–419. `doi:10.1023/B:NHAZ.0000007097.42735.9e`.

[8] W. Yao, Z. Zeng, C. Lian, H. Tang, Training enhanced reservoir computing predictor for landslide displacement, Engineering Geology 188 (2014) 101–109. `doi:10.1016/j.enggeo.2014.11.008`.

[9] S. Segoni, V. Tofani, A. Rosi, F. Catani, N. Casagli, Combination of rainfall thresholds and susceptibility maps for dynamic landslide hazard assessment at regional scale, Frontiers in Earth Science 6 (2018) 85. `doi:10.3389/feart.2018.00085`.

[10] Y. Li, R. Sun, K. Yin, Y. Xu, B. Chai, L. Xiao, Forecasting of landslide displacements using a chaos theory based wavelet analysis-volterra filter model, Scientific Reports 9 (1) (2019) 1–19. `doi:10.1038/s41598-019-56405-y`.

[11] D. Petley, M. Bulmer, W. Murphy, Patterns of movement in rotational and translational landslides, Geology 30 (2002) 719–722. `doi:10.1130/0091-7613(2002)030<0719:POMIRA>2.0.CO;2`.

[12] A. Karpatne, S. Liess, A guide to earth science data: Summary and research challenges, Computing in Science and Engineering 17 (2015) 14–18. `doi:10.1109/MCSE.2015.127`.

[13] D. Bzdok, M. Krzywinski, N. Altman, Machine learning: Supervised methods, Nature Methods 15 (2018) 5–6. `doi:10.1038/nmeth.4551`.

[14] T. Dietterich, Ensemble methods in machine learning, 2000, pp. 1–15.

[15] L. Rokach, Ensemble-based classifiers, Artif. Intell. Rev. 33 (2010) 1–39. `doi:10.1007/s10462-009-9124-7`.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444. `doi:10.1038/nature14539`.

[17] H.-c. Shin, H. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Transactions on Medical Imaging 35 (2016) 1285–1298. `doi:10.1109/TMI.2016.2528162`.

[18] A. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: An astounding baseline for recognition, Vol. 1403, 2014, pp. 806–813. `doi:10.1109/CVPRW.2014.131`.

[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, 2015, pp. 1–9. `doi:10.1109/CVPR.2015.7298594`.

[20] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv 1409.1556.

[21] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2016, pp. 770–778. `doi:10.1109/CVPR.2016.90`.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, 2016, pp. 2818–2826. `doi:10.1109/CVPR.2016.308`.

[23] G. Huang, Z. Liu, L. van der Maaten, K. Weinberger, Densely connected convolutional networks, 2017, pp. 4700–4708. `doi:10.1109/CVPR.2017.243`.

[24] M.-P. Hosseini, S. Lu, K. Kamaraj, A. Slowikowski, H. C. Venkatesh, Deep learning architectures, in: Deep Learning: Concepts and Architectures, Springer, 2020, pp. 1–24.

[25] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, S. Zhao, Applications of machine learning in drug discovery and development, Nature Reviews Drug Discovery 18 (2019) 1. `doi:10.1038/s41573-019-0024-5`.

[26] H. Sak, A. Senior, F. Beaufays, Long short-term memory recurrent neural network architectures for large scale acoustic modeling, Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH (2014) 338–342.

[27] S. Yu, J. Principe, Understanding autoencoders with information theoretic concepts, Neural Networks 117 (2019) 104–123. `doi:10.1016/j.neunet.2019.05.003`.

[28] S. Yu, J. Principe, Understanding autoencoders with information theoretic concepts, Neural Networks`doi:10.1016/j.neunet.2019.05.003`.

[29] S. Rifai, P. Vincent, X. Muller, X. Glorot, Y. Bengio, Contractive auto-encoders: Explicit invariance during feature extraction, 2011, pp. 833–840.

[30] F. Huang, J. Zhang, C. Zhou, Y. Wang, J. Huang, L. Zhu, A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction, Landslides 17 (1) (2020) 217–229. `doi:10.1007/s10346-019-01274-9`.

[31] D. Varnes, Slope Movement Types and Processes, 1978.

[32] S. L. Gariano, F. Guzzetti, Landslides in a changing climate, Earth-Science Reviews 162 (2016) 227–252. `doi:10.1016/j.earscirev.2016.08.011`.

[33] M. Stoffel, D. Tiranti, C. Huggel, Climate change impacts on mass movements - case studies from the European Alps, Science of The Total Environment 493 (2014) 12551266. `doi:10.1016/j.scitotenv.2014.02.102`.

[34] D. Petley, Global patterns of loss of life from landslides, Geology 40 (2012) 927–930. `doi:10.1130/G33217.1`.

[35] H. R. Pourghasemi, B. Pradhan, C. Gokceoglu, M. Mohammadi, H. Moradi, Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at haraz watershed, Iran, Arabian Journal of Geosciences 6 (2012) 1–15. `doi:10.1007/s12517-012-0532-7`.

[36] M. Klose, P. Maurischat, B. Damm, Landslide impacts in Germany: a historical and socioeconomic perspective, Landslides 13 (1) (2016) 183–199. `doi:10.1007/s10346-015-0643-9`.

[37] [link]. URL `https://earthobservatory.nasa.gov/images/92018/predicting-landslide-hazards-in-near-real-t`

[38] X. Yang, L.-D. Chen, Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides, International Journal of Applied Earth Observation and Geoinformation 12 (2010) 487–495. `doi:10.1016/j.jag.2010.05.006`.

[39] M. Galli, F. Ardizzone, M. Cardinali, F. Guzzetti, R. Paola, Comparing landslide inventory maps, Geomorphology 94 (2008) 268–289. `doi:10.1016/j.geomorph.2006.09.023`.

[40] J. Antonio, M. Marsella, C. Nardinocchi, J. Prez, T. Castillo, J. Chacn, C. Irigaray, Landslide detection and inventory by integrating LiDAR data in a GIS environment, Landslides 12 (2014) 1035–1050. `doi:10.1007/s10346-014-0534-5`.

[41] A. Stumpf, N. Kerle, Combining random forests and object-oriented analysis for landslide mapping from very high resolution imagery, Procedia Environmental Sciences 3 (2011) 123129. `doi:10.1016/j.proenv.2011.02.022`.

[42] J. Nichol, M. Wong, Satellite remote sensing for detailed landslide inventories using change detection and image fusion, International Journal of Remote Sensing 26 (2005) 1913–1926. `doi:10.1080/01431160512331314047`.

[43] M. Jaboyedoff, T. Oppikofer, A. Abellan, M.-H. Derron, L. Alexandre, R. Metzger, A. Pedrazzini, Use of LIDAR in landslide investigations: A review, Natural Hazards 61 (2012) 5–28. `doi:10.1007/s11069-010-9634-2`.

[44] W. Schulz, Landslide susceptibility revealed by LIDAR imagery and historical records, Seattle, Washington, Engineering Geology 89 (2007) 67–87. `doi:10.1016/j.enggeo.2006.09.019`.

[45] M. Calabro, D. Schmidt, J. Roering, An examination of seasonal deformation at the Portuguese Bend landslide, southern California, using radar interferometry, Journal of Geophysics Research 115. `doi:10.1029/2009JF001314`.

[46] A. Handwerger, J. Roering, D. Schmidt, Controls on the seasonal deformation of slow-moving landslides, Earth and Planetary Science Letters 377 (2013) 239–247. `doi:10.1016/j.epsl.2013.06.047`.

[47] M.-Y. Cheng, N.-D. Hoang, Slope collapse prediction using bayesian framework with k-nearest neighbor density estimation: Case study in taiwan, Journal of Computing in Civil Engineering 0 (2014) 0. `doi:10.1061/(ASCE)CP.1943-5487.0000456,`.

[48] [link].
URL `https://landslides.usgs.gov/monitoring/`

[49] G. Danneels, E. Pirard, H.-B. Havenith, Automatic landslide detection from remote sensing images using supervised classification methods, in: 2007 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2007, pp. 3014–3017.

[50] Z. Li, W. Shi, P. Lu, L. Yan, Q. Wang, Z. Miao, Landslide mapping from aerial photographs using change detection-based markov random field, Remote Sensing of Environment 187 (2016) 76–90.

[51] T. Martha, N. Kerle, C. Westen, V. Jetten, K. vinod Kumar, Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories, ISPRS Journal of Photogrammetry and Remote Sensing 67 (2012) 105–119. `doi:10.1016/j.isprsjprs.2011.11.004`.

[52] G. Camps-Valls, J. Calpe, Composite kernels for hyperspectral image classification, IEEE Geoscience and Remote Sensing Letters 3 (2006) 93–97.

[53] C. Huang, K. Song, S. Kim, J. Townshend, P. Davis, J. Masek, S. Goward, Use of a dark object concept and support vector machines to automate forest cover change analysis, Remote Sensing of Environment 112 (2008) 970–985. `doi:10.1016/j.rse.2007.07.023`.

[54] H. Nemmour, Y. Chibani, Multiple support vector machines for land cover change detection: An application for mapping urban extensions, ISPRS Journal of Photogrammetry and Remote Sensing 61 (2006) 125–133. `doi:10.1016/j.isprsjprs.2006.09.004`.

[55] X. Liu, R. Jr, Urban change detection based on an artificial neural network, International Journal of Remote Sensing 23 (2002) 2513–2518. `doi:10.1080/01431160110097240`.

[56] R. Jensen, P. Hardin, G. Yu, Artificial neural networks and remote sensing, Geography Compass 3 (2009) 630 – 646. `doi:10.1111/j.1749-8198.2008.00215.x`.

[57] J. Ham, Y. Chen, M. Crawford, J. Ghosh, Investigation of the random forest framework for classification of hyperspectral data, Geoscience and Remote Sensing, IEEE Transactions on 43 (2005) 492–501. `doi:10.1109/TGRS.2004.842481`.

[58] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, Convolutional neural networks for large-scale remote sensing image

classification, IEEE Transactions on Geoscience and Remote Sensing 55 (2016) 645–657. `doi:10.1109/TGRS.2016.2612821`.

[59] D. Hlbling, B. Friedl, C. Eisank, Correction to: An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern taiwan, Earth Science Informatics (2019) 327–335 `doi:10.1007/s12145-018-00375-8`.

[60] T. Martha, N. Kerle, C. Westen, V. Jetten, K. vinod Kumar, Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis, IEEE Transactions on Geoscience and Remote Sensing 49 (2011) 4928 – 4943. `doi:10.1109/TGRS.2011.2151866`.

[61] T. Lei, Y. Zhang, Z. Lv, S. Li, S. Liu, A. Nandi, Landslide inventory mapping from bitemporal images using deep convolutional neural networks, IEEE Geoscience and Remote Sensing Letters (2019) 1–5 `doi:10.1109/LGRS.2018.2889307`.

[62] W. Chen, X. Li, Y. Wang, G. Chen, S. Liu, Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China, Remote Sensing of Environment 152 (2014) 291301. `doi:10.1016/j.rse.2014.07.004`.

[63] P. Gorsevski, M. Brown, K. Panter, C. Onasch, S. M. Anita, J. Snyder, Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park, Ohio, Landslides (2015) 467–484 `doi:10.1007/s10346-015-0587-0`.

[64] M. Hussain, D. Chen, A. Cheng, H. Wei, D. Stanley, Change detection from remotely sensed images: From pixel-based to object-based approaches, ISPRS Journal of Photogrammetry and Remote Sensing 80 (2013) 91–106.

[65] M. Isever, C. nsalan, Two-dimensional change detection methods: Remote sensing applications.

[66] D. Lu, P. Mausel, E. Brondzio, E. Moran, Change detection techniques, International Journal of Remote Sensing 25 (2004) 2365–2401.

[67] F. Bovolo, L. Bruzzone, M. Marconcini, A novel approach to unsupervised change detection based on a semisupervised svm and a similarity measure, IEEE Transactions on Geoscience and Remote Sensing 46 (2008) 2070 – 2082. `doi:10.1109/TGRS.2008.916643`.

[68] K. Pawuszek-Filipiak, A. Borkowski, Sensitivity analysis of automatic landslide mapping: numerical experiments towards the best solution, Landslides 15 (2018) 18511865. `doi:10.1007/s10346-018-0986-0`.

[69] M. I. Sameen, B. Pradhan, Landslide detection using residual networks and the fusion of spectral and topographic information, IEEE Access 7 (2019) 114363–114373.

[70] T. R. Martha, N. Kerle, V. Jetten, C. J. van Westen, K. V. Kumar, Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods, Geomorphology 116 (1-2) (2010) 24–36.

[71] A. Stumpf, N. Kerle, Object-oriented mapping of landslides using random forests, Remote Sensing of Environment 115 (2011) 2564–2577. `doi:10.1016/j.rse.2011.05.013`.

[72] M. Eeckhaut, N. Kerle, J. Poesen, J. Hervs, Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data, Geomorphology s173-174 (2012) 30–42. `doi:10.1016/j.geomorph.2012.05.024`.

[73] G. Sofia, G. Dalla Fontana, Geomorphic features extraction from high-resolution topography: Landslide crowns and bank erosion, Natural Hazards 61 (2012) 65–83. `doi:10.1007/s11069-010-9695-2`.

[74] C. Kurtz, A. Stumpf, J.-P. Malet, P. Gancarski, A. Puissant, N. Passat, Hierarchical extraction of landslides from mul-

44

tiresolution remotely sensed optical images, ISPRS Journal of Photogrammetry and Remote Sensing 87 (2014) 122–136. `doi:10.1016/j.isprsjprs.2013.11.003`.

[75] B. Feizizadeh, T. Blaschke, D. Tiede, M. H. Rezaei Moghaddam, Evaluating fuzzy operators of an object-based image analysis for detecting landslides and their changes, Geomorphology 293 (2017) 240–254. `doi:10.1016/j.geomorph.2017.06.002`.

[76] B. Pradhan, M. Al-Zuhairi, H. Nampak, Integration of LiDAR and QuickBird Data for Automatic Landslide Detection Using Object-Based Analysis and Random Forests, 2017, pp. 69–81.

[77] D. Duro, S. Franklin, M. Dub, A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery, Remote Sensing of Environment 118 (2012) 259–272. `doi:10.1016/j.rse.2011.11.020`.

[78] P. M. Amatya, D. Kirschbaum, T. Stanley, Use of very high-resolution optical data for landslide mapping and susceptibility analysis along the karnali highway, Nepal, Remote Sensing 11 (2019) 2284. `doi:10.3390/rs11192284`.

[79] W. Zhao, S. Du, Learning multiscale and deep representations for classifying remotely sensed imagery, ISPRS Journal of Photogrammetry and Remote Sensing 113 (2016) 155–165. `doi:10.1016/j.isprsjprs.2016.01.004`.

[80] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, High-resolution aerial image labeling with convolutional neural networks, IEEE Transactions on Geoscience and Remote Sensing 55 (12) (2017) 7092–7103.

[81] T. Kattenborn, J. Eichel, F. Fassnacht, Convolutional Neural Networks enable efficient, accurate and fine- grained segmentation of plant species and communities from high-resolution UAV imagery, Scientific Reports 10 (2019) 17656. `doi:10.1038/s41598-019-53797-9`.

[82] H. Yu, Y. Ma, L. Wang, Y. Zhai, X. Wang, A landslide intelligent detection method based on CNN and RSGR (2017) 40–44.

[83] A. Ding, Q. Zhang, X. Zhou, B. Dai, Automatic recognition of landslide based on CNN and texture change detection (2016) 444–448`doi:10.1109/YAC.2016.7804935`.

[84] O. Ghorbanzadeh, T. Blaschke, Optimizing sample patches selection of CNN to improve the mIOU on landslide detection (2019) 33–40`doi:10.5220/0007675300330040`.

[85] S. Ji, Y. Dawen, C. Shen, W.-l. Li, Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks, Landslides`doi:10.1007/s10346-020-01353-2`.

[86] M. Herrera Herrera, Landslide detection using random forest classifier.

[87] S. N. K. Binti Amit, Y. Aoki, Disaster detection from aerial imagery with convolutional neural network, 2017, pp. 239–245. `doi:10.1109/KCIC.2017.8228593`.

[88] H. Hong, Y. Miao, J. Liu, A.-X. Zhu, Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping, Catena 176 (2019) 45–64. `doi:10.1016/j.catena.2018.12.035`.

[89] K. L. Na, M. Liew, A. Matori, N. A. Wan Abdullah Zawawi, Recent developments in machine learning applications in landslide susceptibility mapping, Vol. 1905, 2017. `doi:10.1063/1.5012210`.

[90] U. Djuric, M. Marjanovi, Z. Radi, B. Abolmasov, Machine learning based landslide assessment of the belgrade metropolitan area: Pixel resolution effects and a cross-scaling concept, Engineering Geology 256. `doi:10.1016/j.enggeo.2019.05.007`.

[91] X. Luo, F. Lin, Y. Chen, S. Zhu, Z. Xu, Z. Huo, M. Yu, J. Peng, Coupling logistic model tree and random subspace to

predict the landslide susceptibility areas with considering the uncertainty of environmental features, Scientific Reports 9 (2019) 15369. `doi:10.1038/s41598-019-51941-z`.

[92] F. Guzzetti, R. Paola, F. Ardizzone, M. Cardinali, M. Galli, Estimating the quality of landslide susceptibility models, Geomorphology 81 (2006) 166–184. `doi:10.1016/j.geomorph.2006.04.007`.

[93] T. Kavzoglu, E. Sahin, I. Colkesen, An assessment of multivariate and bivariate approaches in landslide susceptibility mapping: a case study of duzkoy district, Natural Hazards 76 (2014) 471–496. `doi:10.1007/s11069-014-1506-8`.

[94] C. Westen, N. Rengers, R. Soeters, Use of geomorphological information in indirect landslide susceptibility assessment, Natural Hazards 30 (2003) 399–419. `doi:10.1023/B:NHAZ.0000007097.42735.9e`.

[95] J. Dou, K.-T. Chang, S. Chen, A. P. Yunus, J.-K. Liu, X. Huan, Z. Zhu, Automatic case-based reasoning approach for landslide detection: Integration of object-oriented image analysis and a genetic algorithm, Remote Sensing 7 (2015) 4318–4342. `doi:10.3390/rs70404318`.

[96] D. Tien Bui, T. Ho, B. Pradhan, B. Pham, V.-H. Nhu, I. Revhaug, GIS-based modeling of rainfall-induced landslides using data mining based functional trees classifier with adaboost, bagging, and multiboost ensemble frameworks, Environmental Earth Sciences 75 (2016) 1101. `doi:10.1007/s12665-016-5919-4`.

[97] N. Ramaraj, A novel hybrid feature selection via symmetrical uncertainty ranking based local memetic search algorithm, Knowledge-Based Systems 23 (2010) 580–585. `doi:10.1016/j.knosys.2010.03.016`.

[98] A. Akgun, A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: A case study at zmir, Turkey, Landslides 9 (2012) 93–106. `doi:10.1007/s10346-011-0283-7`.

[99] B. Pham, D. Tien Bui, I. Prakash, M. Dholakia, Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at himalayan area (India) using GIS, Catena 149 (2016) 52–63. `doi:10.1016/j.catena.2016.09.007`.

[100] H. Hong, B. Pradhan, D. Tien Bui, C. Xu, D. A. Youssef, W. Chen, Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Suichuan area (China), Geomatics, Natural Hazards and Risk 8 (2016) 544–569. `doi:10.1080/19475705.2016.1250112`.

[101] P. Tsangaratos, I. Ilia, Comparison of a logistic regression and nave bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size, Catena 145 (2016) 164–179. `doi:10.1016/j.catena.2016.06.004`.

[102] P. Tsangaratos, I. Ilia, Landslide susceptibility mapping using a modified decision tree classifier in the xanthi perfection, Greece, Landslides (2015) 305–320`doi:10.1007/s10346-015-0565-6`.

[103] C. Melchiorre, M. Matteucci, A. Azzoni, A. Zanchi, Artificial neural networks and cluster analysis in landslide susceptibility zonation, Geomorphology 94 (2008) 379–400. `doi:10.1016/j.geomorph.2006.10.035`.

[104] J. Schmidt, M. R. Marques, S. Botti, M. A. Marques, Recent advances and applications of machine learning in solid-state materials science, npj Computational Materials 5 (1) (2019) 1–36.

[105] B. Kalantar, B. Pradhan, S. Naghibi, A. Motevalli, S. Mansor, Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN), Geomatics, Natural Hazards and Risk (2018) 49–69`doi:10.1080/19475705.2017.1407368`.

[106] T. Kavzoglu, E. Sahin, I. Colkesen, Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, sup-

port vector machines, and logistic regression, Landslides 11 (2013) 425–439. `doi:10.1007/s10346-013-0391-7`.

[107] M. Conforti, S. Pascale, G. Robustelli, F. Sdao, Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy), Catena 113 (2014) 236–250.

[108] H. Saito, D. Nakayama, H. Matsuyama, Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan, Geomorphology 109 (2009) 108–121. `doi:10.1016/j.geomorph.2009.02.026`.

[109] S. Lee, B. Pradhan, Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models, Landslides 4 (2007) 33–41. `doi:10.1007/s10346-006-0047-y`.

[110] S. Kuhn, M. Cracknell, A. Reading, Lithological mapping using Random Forests applied to geophysical and remote sensing data: a demonstration study from the Eastern Goldfields of Australia, GEOPHYSICS 83 (2018) 1–37. `doi:10.1190/geo2017-0590.1`.

[111] T. Heckmann, K. Gegg, A. Gegg, M. Becht, Sample size matters: Investigating the effect of sample size on a logistic regression susceptibility model for debris flows, Natural Hazards and Earth System Sciences 14 (2014) 259–278. `doi:10.5194/nhess-14-259-2014`.

[112] C. Cortes, V. Vapnik, Support-vector networks, Machine learning 20 (3) (1995) 273–297.

[113] I. Yilmaz, Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine, Environmental Earth Sciences 61 (2009) 821–836. `doi:10.1007/s12665-009-0394-9`.

[114] M. Marjanovi, B. Bajat, B. Abolmasov, M. Kovaevi, Machine Learning and Landslide Assessment in a GIS Environment, 2018, pp. 191–213.

[115] B. Pradhan, A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS, Computers and Geosciences 51 (2013) 350–365.

[116] S. Hwang, I. Guevarra, B. Yu, Slope failure prediction using a decision tree: A case of engineered slopes in South Korea, Engineering Geology 104 (2009) 126–134. `doi:10.1016/j.enggeo.2008.09.004`.

[117] S. Lee, J.-H. Ryu, M.-J. Lee, J.-S. Won, Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea, Environmental Geology 44 (7) (2003) 820–833.

[118] S. Lee, J.-H. Ryu, J.-S. Won, H. J. Park, Determination and application of the weights for landslide susceptibility mapping using an artificial neural network, Engineering Geology 71 (2004) 289–302. `doi:10.1016/S0013-7952(03)00142-X`.

[119] F. Catani, D. Lagomarsino, S. Segoni, V. Tofani, Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues, Natural Hazards and Earth System Sciences 13 (2013) 2815–2831. `doi:10.5194/nhess-13-2815-2013`.

[120] A. Trigila, C. Iadanza, C. Esposito, G. Scarascia-Mugnozza, Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy), Geomorphology 249 (2015) 119–136. `doi:10.1016/j.geomorph.2015.06.001`.

[121] W. Chen, X. Xie, J. Wang, B. Pradhan, H. Hong, D. Tien Bui, Z. Duan, J. Ma, A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility, Catena 151 (2017) 147–160. `doi:10.1016/j.catena.2016.11.032`.

[122] D. A. Youssef, H. R. Pourghasemi, Z. Pourtaghi, M. Al-Katheeri, 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 (2015) 839–856. `doi:10.1007/s10346-015-0614-1`.

[123] J. Goetz, A. Brenning, H. Petschko, P. Leopold, Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling, Computers and Geosciences (2015) 1–11`doi:10.1016/j.cageo.2015.04.007`.

[124] H. Hong, J. Liu, D. Tien Bui, B. Pradhan, T. Acharya, B. Pham, A.-X. Zhu, W. Chen, B. B. Ahmad, Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China), Catena 163 (2018) 399–413. `doi:10.1016/j.catena.2018.01.005`.

[125] Y. Song, R. Niu, X. Shiluo, R. Ye, P. Ling, T. Guo, S. Li, T. Chen, Landslide susceptibility mapping based on weighted gradient boosting decision tree in wanzhou section of the Three Gorges Reservoir Area (China), ISPRS International Journal of Geo-Information 8 (2018) 4. `doi:10.3390/ijgi8010004`.

[126] T. Chen, L. Zhu, R.-q. Niu, C. J. Trinder, L. Peng, T. Lei, Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models, Journal of Mountain Science 17 (3) (2020) 670–685.

[127] B. Pham, I. Prakash, S. Singh, A. Shirzadi, H. Shahabi, T.-T.-T. Tran, D. Tien Bui, Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: Hybrid machine learning approaches, Catena 175 (2018) 203–218. `doi:10.1016/j.catena.2018.12.018`.

[128] M. I. Sameen, R. Sarkar, B. Pradhan, D. Drukpa, A. M. Alamri, H.-J. Park, Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests, Computers and Geosciences 134 (2020) 104336.

[129] J. Dou, A. P. Yunus, D. Tien Bui, A. Merghadi, M. Sahana, Z. Zhu, C. Chen, Z. Han, B. Pham, Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan, Landslides`doi:10.1007/s10346-019-01286-5`.

[130] S. Kotsiantis, Combining bagging, boosting, rotation forest and random subspace methods, Artificial Intelligence Review 35 (2011) 223–240. `doi:10.1007/s10462-010-9192-8`.

[131] W. Chen, J. Peng, H. Hong, H. Shahabi, B. Pradhan, J. Liu, A.-X. Zhu, X. Pei, Z. Duan, Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China, Science of The Total Environment 626 (2018) 1121–1135. `doi:10.1016/j.scitotenv.2018.01.124`.

[132] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (8) (2013) 1798–1828.

[133] W. Yuankai, H. Tan, B. Ran, Z. Jiang, A hybrid deep learning based traffic flow prediction method and its understanding, Transportation Research Part C: Emerging Technologies 90 (2018) 166–180. `doi:10.1016/j.trc.2018.03.001`.

[134] J. Feng, Z.-H. Zhou, Autoencoder by forest, in: Thirty-Second AAAI Conference on Artificial Intelligence, 2018.

[135] K. Nam, F. Wang, The performance of using an autoencoder for prediction and susceptibility assessment of landslides: A case study on landslides triggered by the 2018 Hokkaido Eastern Iburi earthquake in Japan, Geoenvironmental Disasters 6 (2019) 19. `doi:10.1186/s40677-019-0137-5`.

[136] G. Hinton, R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (2006) 504–507. `doi:10.1126/science.1127647`.

[137] G. Eraslan, i. Avsec, J. Gagneur, F. Theis, Deep learning: new computational modelling techniques for genomics, Nature Reviews Genetics 20 (2019) 389–403. `doi:10.1038/s41576-019-0122-6`.

[138] Y. Wang, Z. Fang, H. Hong, Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China, Science of The Total Environment 666 (2019) 975–993. `doi:10.1016/j.scitotenv.2019.02.263`.

[139] V. Pham, Q.-H. Nguyen, H.-D. Nguyen, V.-M. Pham, V. Manh, Q.-T. Bui, Convolutional neural network optimized moth flame algorithm for shallow landslide susceptible analysis, IEEE Access 8. `doi:10.1109/ACCESS.2020.2973415`.

[140] Y. Wang, Z. Fang, M. Wang, P. Ling, H. Hong, Comparative study of landslide susceptibility mapping with different recurrent neural networks, Computers and Geosciences 138 (2020) 104445. `doi:10.1016/j.cageo.2020.104445`.

[141] L. Xiao, Y. Zhang, G. Peng, Landslide susceptibility assessment using integrated deep learning algorithm along the China-Nepal highway, Sensors 18 (2018) 4436. `doi:10.3390/s18124436`.

[142] M. Krkac, D. poljari, S. Bernat Gazibara, S. Mihali Arbanas, Method for prediction of landslide movements based on random forests, Landslides (2016) 947–960`doi:10.1007/s10346-016-0761-z`.

[143] J. Du, K. Yin, S. Lacasse, Displacement prediction in colluvial landslides, Three Gorges Reservoir, China, Landslides 10 (2013) 203–218. `doi:10.1007/s10346-012-0326-8`.

[144] F. Guzzetti, S. L. Gariano, S. Peruccacci, M. Brunetti, I. Marchesini, M. Rossi, M. Melillo, Geographical landslide early warning systems, Earth-Science Reviews (2019) 102973`doi:10.1016/j.earscirev.2019.102973`.

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

[146] Q. Lu, H.-y. Sun, W. Zhenlei, Estimating the rainfall threshold of a deep-seated landslide by integrating models for predicting the groundwater level and stability analysis of the slope, Engineering Geology 253 (2019) 14–26. `doi:10.1016/j.enggeo.2019.02.026`.

[147] S. Segoni, D. Lagomarsino, R. Fanti, S. Moretti, N. Casagli, Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system, Landslides 12 (2014) 773–785. `doi:10.1007/s10346-014-0502-0`.

[148] X. Shiluo, R. Niu, Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China, Computers and Geosciences 111 (2018) 87–96. `doi:10.1016/j.cageo.2017.10.013`.

[149] C. Lian, Z. Zeng, W. Yao, H. Tang, Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis, Neural Computing and Applications 24 (2014) 99–107. `doi:10.1007/s00521-013-1446-3`.

[150] F. Huang, J. Huang, S.-H. Jiang, C. Zhou, Landslide displacement prediction based on multivariate chaotic model and extreme learning machine, Engineering Geology 218 (2017) 173–186. `doi:10.1016/j.enggeo.2017.01.016`.

[151] Q. Siqing, J. Jiao, S. Wang, A nonlinear dynamical model of landslide evolution, Geomorphology (2002) 77–85`doi:10.1016/S0169-555X(01)00122-2`.

[152] L. Benoit, P. Briole, O. Martin, C. Thom, J.-P. Malet, P. Ulrich, Monitoring landslide displacements with the Geocube wireless network of low-cost GPS, Engineering Geology 195 (2015) 111–121. `doi:10.1016/j.enggeo.2015.05.020`.

[153] A. Vallet, D. Varron, C. Bertrand, J.-N. Mudry, Hydrogeological threshold using effective rainfall and support vector ma-

chine (SVM) applied to a deep seated unstable slope (Schilienne, French Alps), 2013.

[154] N. Rachel, M. Lakshmi, Landslide prediction with rainfall analysis using Support Vector Machine, Indian Journal of Science and Technology 9. `doi:10.17485/ijst/2016/v9i21/95275`.

[155] Z. Omadlao, N. Tuguinay, R. J. Saturay, Machine learning-based prediction system for rainfall-induced landslides in Benguet First Engineering District`doi:10.31219/osf.io/csx6r`.

[156] A. Pradhan, S.-R. Lee, Y.-T. Kim, A shallow slide prediction model combining rainfall threshold warnings and shallow slide susceptibility in Busan, Korea, Landslides (2018) 647–659`doi:10.1007/s10346-018-1112-z`.

[157] L. Huang, L.-y. Xiang, Method for meteorological early warning of precipitation-induced landslides based on deep neural network, Neural Processing Letters 48 (2018) 1243–1260. `doi:10.1007/s11063-017-9778-0`.

[158] L. Bonzanigo, E. Eberhardt, S. Loew, Long-term investigation of a deep-seated creeping landslide in crystalline rock. Part I. Geological and hydromechanical factors controlling the Campo Vallemaggia landslide, Canadian Geotechnical Journal 44 (2007) 1157–1180. `doi:10.1139/T07-043`.

[159] A. Vallet, D. Varron, C. Bertrand, O. Fabbri, J.-N. Mudry, A multi-dimensional statistical rainfall threshold for deep landslides based on groundwater recharge and support vector machines, Natural Hazards 84 (2016) 821–849. `doi:10.1007/s11069-016-2453-3`.

[160] B. Yang, K. Yin, S. Lacasse, Z. Liu, Time series analysis and long short-term memory neural network to predict landslide displacement, Landslides`doi:10.1007/s10346-018-01127-x`.

[161] C. Zhu, G. Hu, Time series prediction of landslide displacement using SVM model: Application to baishuihe landslide in Three Gorges Reservoir Area, China, Applied Mechanics and Materials 239 (2012) 1413–1420. `doi:10.4028/www.scientific.net/AMM.239-240.1413`.

[162] Z. Liu, J. Shao, W. Xu, H. Chen, C. Shi, Comparison on landslide nonlinear displacement analysis and prediction with computational intelligence approaches, Landslides 11 (2013) 889–896. `doi:10.1007/s10346-013-0443-z`.

[163] C. Lian, Z. Zeng, W. Yao, H. Tang, Multiple neural networks switched prediction for landslide displacement, Engineering Geology 186 (2015) 91–99. `doi:10.1016/j.enggeo.2014.11.014`.

[164] P. Xie, A. Zhou, B. Chai, The application of long short-term memory(LSTM) method on displacement prediction of multifactor-induced landslides, IEEE Access 7 (2019) 05–11. `doi:10.1109/ACCESS.2019.2912419`.

[165] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, M. Prabhat, Deep learning and process understanding for data-driven earth system science, Nature 566 (2019) 195–204. `doi:10.1038/s41586-019-0912-1`.

[166] A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, V. Kumar, Machine learning for the geosciences: Challenges and opportunities, IEEE Transactions on Knowledge and Data Engineering 31 (8) (2018) 1544–1554.

[167] A. Virk, A. Singh, S. Mittal, Advanced MT-InSAR landslide monitoring: Methods and trends, Journal of Remote Sensing and GIS 07. `doi:10.4172/2469-4134.1000225`.

[168] B. Kumar, G. Pandey, B. Lohani, S. C. Misra, A multi-faceted cnn architecture for automatic classification of mobile lidar data and an algorithm to reproduce point cloud samples for enhanced training, ISPRS journal of photogrammetry and remote sensing 147 (2019) 80–89.

[169] K. Bergen, P. Johnson, M. Hoop, G. Beroza, Machine learning for data-driven discovery in solid earth geoscience, Science 363 (2019) 1299. `doi:10.1126/science.aau0323`.

[170] J.-Y. Zhu, T. Park, P. Isola, A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks,

2017, pp. 2242–2251. doi:10.1109/ICCV.2017.244.

[171] B. Krawczyk, Learning from imbalanced data: open challenges and future directions, Progress in Artificial Intelligence 5 (4) (2016) 221–232.

[172] O. Althuwaynee, B. Pradhan, H. J. Park, J.-H. Lee, A novel ensemble decision tree-based chi-squared automatic interaction detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping, Landslides 11 (2014) 1063–1078. doi:10.1007/s10346-014-0466-0.

[173] N. Chawla, K. Bowyer, L. Hall, W. Kegelmeyer, SMOTE: Synthetic minority over-sampling technique, Journal of artificial intelligence research 16 (2002) 321–357. doi:10.1613/jair.953.

[174] J. A. Sez, J. Luengo, J. Stefanowski, F. Herrera, SMOTEIPF: Addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering, Information Sciences 291 (2015) 184–203. doi:10.1016/j.ins.2014.08.051.

[175] T. Shi, J. Wang, P.-f. Wang, Q.-h. Cai, Y.-c. Han, The impact of imbalanced training datasets on CNN performance in typical remote scenes classification, DEStech Transactions on Computer Science and Engineering (2018) 1–16doi:10.12783/dtcse/pcmm2018/23661.

[176] H. Lee, M. Park, J. Kim, Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning, 2016, pp. 3713–3717. doi:10.1109/ICIP.2016.7533053.

[177] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks, Medical Image Analysis 35 (2015) 18–31. doi:10.1016/j.media.2016.05.004.

[178] S. Pouyanfar, Y. Tao, A. Mohan, H. Tian, A. Kaseb, K. Gauen, R. Dailey, S. Aghajanzadeh, Y.-H. Lu, S.-C. Chen, M.-L. Shyu, Dynamic sampling in convolutional neural networks for imbalanced data classification, 2018, pp. 112–117. doi:10.1109/MIPR.2018.00027.

[179] K. Agrawal, Y. Baweja, D. Dwivedi, R. Saha, P. Prasad, S. Agrawal, S. Kapoor, P. Chaturvedi, N. Mali, V. U. Kala, et al., A comparison of class imbalance techniques for real-world landslide predictions, in: 2017 International Conference on Machine Learning and Data Science (MLDS), IEEE, 2017, pp. 1–8.

[180] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, 2017, pp. 2980–2988. doi:10.1109/ICCV.2017.324.

[181] S. Khan, M. Bennamoun, F. Sohel, R. Togneri, Cost-sensitive learning of deep feature representations from imbalanced data, IEEE Transactions on Neural Networks and Learning Systems 29 (2015) 3573–3587. doi:10.1109/TNNLS.2017.2732482.

[182] C. Zhang, K. Tan, R. Ren, Training cost-sensitive deep belief networks on imbalance data problems, 2016, pp. 4362–4367. doi:10.1109/IJCNN.2016.7727769.

[183] J. Johnson, T. Khoshgoftaar, Survey on deep learning with class imbalance, Journal of Big Data 6 (2019) 27. doi:10.1186/s40537-019-0192-5.

[184] C. Molnar, Interpretable machine learning, Lulu. com, 2019.

[185] S. Gupta, K. Kothari, M. V. de Hoop, I. Dokmani, Random mesh projectors for inverse problems, arXiv:1805.11718.

[186] J. Tenenbaum, C. Kemp, T. Griffiths, N. Goodman, How to grow a mind: Statistics, structure, and abstraction, Science 331 (2011) 1279–1285. doi:10.1126/science.1192788.

[187] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, B. Johnson, Deep learning in remote sensing applications: A meta-analysis and review, ISPRS Journal of Photogrammetry and Remote Sensing 152 (2019) 166–177. `doi:10.1016/j.isprsjprs.2019.04.015`.

[188] D. Lin, K. Fu, Y. Wang, G. Xu, X. Sun, MARTA GANs: Unsupervised representation learning for remote sensing image classification, IEEE Geoscience and Remote Sensing Letters 14 (2017) 2092–2096. `doi:10.1109/LGRS.2017.2752750`.

[189] Z. He, H. Liu, Y. Wang, J. Hu, Generative adversarial networks-based semi-supervised learning for hyperspectral image classification, Remote Sensing 9 (2017) 1042.

[190] T. Schlegl, P. Seebck, S. Waldstein, U. Schmidt-Erfurth, G. Langs, Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, 2017, pp. 146–157.

[191] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P. S. Yu, A comprehensive survey on graph neural networks, arXiv:1901.00596.

[192] D. Bacciu, F. Errica, A. Micheli, M. Podda, A gentle introduction to deep learning for graphs, arXiv:1912.12693.

[193] T. Ai, X. Yan, a graph convolution neuroal network, ISPRS Journal of Photogrammetry and Remote Sensing 150 (2019) 259–273. `doi:10.1016/j.isprsjprs.2019.02.010`.

[194] G. Te, W. Hu, A. Zheng, Z. Guo, RgCNN: Regularized graph CNN for point cloud segmentation, in: Proceedings of the 26th ACM International Conference on Multimedia, 2018, pp. 746–754.

[195] L. Landrieu, M. Simonovsky, Large-scale point cloud semantic segmentation with superpoint graphs, 2018, pp. 4558–4567. `doi:10.1109/CVPR.2018.00479`.

[196] G. Mei, N. Xu, J. Qin, B. Wang, P. Qi, A survey of Internet of Things (IoT) for geo-hazards prevention: Applications, technologies, and challenges, IEEE Internet of Things Journal 2019 (2019) 1–16. `doi:10.1109/JIOT.2019.2952593`.

[197] S. Fortunato, Community detection in graphs, Physics Reports 486 (3-5) (2010) 75–174. `doi:10.1016/j.physrep.2009.11.002`.

[198] X. Wang, Y. Zhao, F. Pourpanah, Recent advances in deep learning, International Journal of Machine Learning and Cybernetics 11 (4) (2020) 747–750. `doi:10.1007/s13042-020-01096-5`.