Flood Susceptibility Prediction Using Hybrid-Based Approaches of Support Vector Regression Model and Meta-Heuristic Algorithms

Hossein Hamedi Sorajar
KN Toosi University of Technology
Ali Asghar Alesheikh
KN Toosi University of Technology
Mahdi Panahi
KIGAM: Korea Institute of Geoscience and Mineral Resources
Saro Lee (leesaro@kigam.re.kr)
Korea Institute of Geoscience and Mineral Resources https://orcid.org/0000-0003-0409-8263

Research Article

Keywords: Landslide susceptibility mapping, CNN, LSTM, Deep learning, GIS

Posted Date: September 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-876711/v1

Abstract

Landslides are one of the most destructive natural phenomena in the world, which occur mostly in mountainous areas and cause damage to the economic sectors, agricultural lands, residential areas and infrastructures of any country, and also threaten the lives and property of human beings. Therefore, landslide susceptibility mapping (LSM) can play a critical role in identifying prone areas and reducing the damage caused by landslides in each area. In the present study, deep learning algorithms including convolutional neural network (CNN) and long short-term memory (LSTM) were used to identify landslide prone areas in Ardabil province, Iran. Equql to 312 landslide locations were identified and randomly divided into train and test datasets at 70–30% ratios. Then, according to previous studies and environmental conditions in the study area, twelve factors affecting the occurrence of landslides were selected, namely altitude, slope angle, slope aspect, topographic wetness index (TWI), profile curvature, plan curvature, land-use, lithology, distance to faults, distance to rivers, distance to roads, and rainfall. The ratio of the importance of each influential factor in landslide occurrence was obtained through information gain ranking filter (IGRF) method and it was found that land-use and profile curvature had the highest and lowest impacts, respectively. Afterwards, LSMs were generated using CNN and LSTM algorithms. In the next step, the performance of the models was evaluated based on the area under curve (AUC) value of receiver operating characteristics curve and the root mean square error (RMSE) method. The AUC values for CNN and LSTM models were 0.821 and 0.832, respectively. Furthermore, the RMSE values in the CNN model for each of the training and testing dataset were 0.121 and 0.132, respectively. The RMSE values in the LSTM model for each of the training and testing dataset were 0.185 and 0.188, respectively. Therefore, it can be concluded that CNN performance is slightly better than LSTM; but in general, both models have close performance and the accuracy of both models is acceptable.

1. Introduction

Landslide is the movement of sedimentary layers on a sloping surface due to an earthquake, road construction, rainfall or gravity (Das et al., 2012). Landslides are one of the most destructive natural phenomena in the world (Yilmaz, 2010; San, 2014) which occur mostly in mountainous areas and cause damage to the economic sectors, agricultural lands, residential areas and infrastructures of any country, and threaten the lives of people living in such areas (De Blasio, 2011; Corominas et al., 2014). Governments allocate a considerable of money annually to preventing and reducing the risks of natural disasters, especially landslides (Salvati et al., 2010) Because the effects of these disasters are sometimes very irreversible; therefore, “landslide susceptibility mapping (LSM)” is one of the most important and necessary means to identifying landslide prone areas and thus reducing the damage caused by landslides in each area.

Iran is geographically located in the seismic belt of the Alps and Himalayas, which makes it more vulnerable to natural disasters, including landslides (Farrokhnia et al., 2011). On the other hand, the Iranian Plateau is located in southwest Asia which is one of the high and mountainous regions of western Asia and has many faults, folds, and valleys; so, it is an active seismic zone. Furthermore, in the

western and northern parts of the Iranian Plateau are the Zagros and Alborz Mountain Ranges, which are considered as active mountains due to their high tectonic activity and climatic diversity (Berberian, 1981, 1983a; Feizizadeh et al., 2013), which cause earthquakes and landslides in these areas. Also, natural factors such as climate, earthquakes and human activities, such as construction of roads and tunnels increase the likelihood of landslides in these areas (Wilde et al., 2018). Ardabil Province is known as one of the provinces of Iran with the highest probability of landslides due to the high-altitude difference and its slope fluctuations, which cause some of its slopes to reach a slope above 100% as well as the proximity of this province to the Caspian Sea and the presence of heavy rainfall in this region, (Iran Meteorological Organization, 2019).

According to the mentioned cases, the production of the LSM is a vital step to be. This allows us to know the location of landslides that are likely to occur in the future; and by planning and managing in these areas (e.g., construction prohibition, strengthening vegetation, or increasing soil resistance and stability) to prevent them from happening (Lacasse and Nadim, 2009; Havenith et al., 2015; Betts et al., 2017). Also, by knowing the location of landslides that have occurred before, the speed of providing assistance to these areas can increase, which will reduce human and financial losses (Haeri and Satari, 1993; Cascini et al., 2005).

There are different methods for generating LSM, which are generally divided into two categories: qualitative and quantitative approaches. Qualitative methods are mostly based on expert wisdom and knowledge, while quantitative methods are more based on numerical and statistical estimates (Aleotti and chowdhury, 1999). For example, the heuristic method and the geomorphological landslide susceptibility zoning method are considered as qualitative approaches (Aleotti and Chowdhury, 1999; Guzzetti et al., 1999). The “Analytical Hierarchy Process (AHP)” and the “Analytical Network Process (ANP)” are examples of heuristic method in which using the knowledge and experience of experts, an LSM map is developed (Feizizadeh et al., 2014; Yalcin et al., 2011; Neaupane and piantanakulchai, 2006). Quantitative methods generally consist of three groups: Statistical and mathematical methods, geotechnical techniques, and machine learning algorithms (Aleotti and Chowdhury, 1999; Thiery et al., 2007). The geotechnical method itself is divided into two parts: deterministic analysis and probabilistic approaches (van Western and Terlien, 1996; Park et al., 2013). Statistical methods strongly depend on the structure of the model and their parameters and data and are especially used for analyzing natural hazards such as landslides. In this method, classification of landslide condition factors is very important because it affects the quality of the LSM and shows the relationship between landslide occurrence and the impact of factors (Costanzo et al., 2012). Such famous statistical and mathematical approaches can be called “statistical index (SI)” (Regmi et al., 2014), “weight of evidence (WoE)” (Sifa et al., 2020), “frequency ratio (FR)” (Thanh et al., 2020), “certainty factor (CF)” (Zhao et al., 2020), and bivariate or multivariate statistics analysis (Mersha et al., 2020). Deterministic and probabilistic models usually produce more accurate results; however the process of creating models and calculating them is relatively complex and therefore may not be suitable for assessing landslide susceptibility (Akgun and Erkan, 2016). Most of these models evaluate landslide susceptibility based on slope stability theory (Xie et al., 2004). With the development of computers as well as remote sensing (RS) and geographical information

system (GIS), statistical methods have gradually given way to machine learning algorithms for evaluating and producing landslide maps (Bai et al., 2010). Machine learning algorithms solve many problems in statistical and geotechnical methods (Cracknell and Reading, 2014). In these algorithms, automatic extraction of knowledge from existing data is done with the aim of applying the learned knowledge to new data (Jordan and Mitchell, 2015). In general, machine learning algorithms have higher predictive power and have recently been used in generating a natural hazard susceptibility map. Examples of machine learning algorithms are as follows: “artificial neural networks (ANN)” (Can et al., 2019; Bragagnolo et al., 2020), “logistic regression (LR)” (Park et al., 2013), “decision trees (DT)” (Hong et al., 2020), “support vector machines (SVM)” (Hu et al., 2020), “neuro-fuzzy (NF)” (Pradhan et al., 2010; Jaafari et al., 2019), “adaptive neuro-fuzzy Inference system (ANFIS)” (Ghorbanzadeh et al., 2020), “naïve Bayes tree (NBT)” (Ali et al., 2020), “evidential belief function (EBF)” (Li et al., 2020), and “random forest (RF)” (Sun et al., 2020).

Recently, deep learning (DL) models, a subset of machine learning models, have been used by researchers to evaluate and generate the LSMs. As the name of these models suggests, the difference between these models and other models is in learning and recognizing the different patterns in the data. The DL models will be very efficient and useful when the data set size is very large (Wang et al., 2020b). Convolutional neural network (CNN) and recurrent neural network (RNN) are the most popular models of deep learning. The CNN has achieved good results in analyzing various images in computer vision and image processing (Zhu et al., 2017). The CNN has also been used for the classification and segmentation of very high resolution (VHR) remote sensing image (Maggiori et al., 2016), and semantic segmentation (Long et al., 2015), and object detection and recognition (Guirado et al., 2017; Ghorbanzadeh et al., 2019a; Ngo et al., 2020), and scene annotation using aerial images (Qayyum et al., 2017). A prerequisite for using CNN is the use of training datasets in which each set of data must be labeled in order to be used in learning model process (Ghorbanzadeh et al., 2019a; Yang et al., 2017). The CNN architecturally includes input layer, hidden layers, convolutional layers, max pooling layers, fully connected layers, and output layer (Krizhevsky et al., 2017). The larger the training dataset, the more efficient the network architecture parameters, and the faster the computer system, the better CNN performance (Ding et al., 2016). The RNN is a neural sequence model that architecturally includes input, hidden, and output layers, learnable parameters, and loss functions (Zaremba et al., 2014), that have been successful in important tasks, including: language modeling (Mikolov, 2012), and speech recognition (Graves et al., 2013), and machine translation (Kalchbrenner and Blunsom, 2013). Long short-term memory (LSTM) is a special type of RNN architecture. LSTM has a complicated dynamic that allows it to store information in memory cells for a long time in a number of time steps (Zaremba et al., 2014). LSTM has the power to decide whether to overwrite, retrieve, or hold a memory cell for the next step (Zaremba et al., 2014).

In the present study, DL algorithms, including CNN and LSTM, are utilized for evaluating and generating a landslide susceptibility map. Then the performance of these two models is compared. Based on the studies, no comparison of the performance of these two models in evaluating and producing the LSM has been done among researchers so far. The results obtained in this study can be useful for other

researchers in evaluating these two models in the production of natural hazard susceptibility maps, especially landslides.

2. Study Area

Ardabil Province is one of the provinces of Iran, which is located in the northwest of this country in the region of Azerbaijan, Iran. This province is located in the longitudes of 47° 00' to 49° 36' E and latitudes of 36° 32' to 39° 42' N, in western Caspian Sea (Fig. 1). This province is bounded on the north by the Republic of Azerbaijan, on the west by the province of East Azerbaijan and the plateau of Azerbaijan, on the east by the province of Gilan and the Talesh Mountains, and on the south by the province of Zanjan (Dodangeh et al., 2020). The area of this province is 17,953 km^2 (about 1.09 percent of the total area of the country) and its population is equal to 1,249,000 people according to the latest census. The capital of this province is Ardabil city and according to the latest divisions of the country, the province includes 10 counties, 24 cities, 27 districts, and 2,210 villages (Azizi et al., 2014).

Topographically, it is more than two third of the mountainous areasand the its height varies from 4 to 4,788 meters. Also, most of the area has an altitude between 2,000 and 3,000 meters above sea level, and the slope is between 0 ° and 63 °. In terms of climate, Ardabil Province is known as one of the coldest regions of Iran, with a variety of climates, including arid and semi-arid in the north and the Mediterranean and semi-humid in the south of this region (Dodangeh et al., 2020), which causes snowy and rainy winters along with mild and pleasant summers in this region. The average annual rainfall is about 480 mm (Iran Meteorological Organization, 2019). The average annual temperature in this region is 10.9 ° C. In terms of lithology, Ardabil Province has different types of rocks, which mainly include two groups, “Andesitic volcanic” and “low-level piedmont fan and valley terrace deposit

3. Methodology

In general, “landslide susceptibility map” production consists of five basic steps. In the first step, the landslide inventory map should be prepared using landslides that have occurred in the past, and the landslide and non-landslide points should be determined. In the second step, using landslide points identified in the previous step, our database containingtraining and testing datasets, is created. In the third step, the influence condition factors in the occurrence of landslides are opted for according to the study area and previous studies. In the fourth step, a landslide susceptibility map is generated using CNN and LSTM models. Finally, in the fifth step, the maps produced using CNN and LSTM models are validated through computing AUC and RMSE. To increase the visual perception and vision, the steps are described as a conceptual model in Figs. 2.

3.1. Landslide inventory map

The first step in producing an LSM is to prepare a landslide inventory map. In the landslide inventory map, the landslide locations that have occurred in the past are considered. Using this map, the location of landslides that may occur in the future can be predicted. Furthermore, the relationship between the occurrence of landslides and condition factors affecting landslides can be obtained, and the most important factors can be ranked. There are several ways to prepare a landslide inventory map. In current study, satellite images, data from the Iranian Natural Resources Management Organization and the Forest Range and Watershed Management Organization as well as information from field surveys were used to generate this map. There are two types of landslides: rotational and transitional. According to our observations, most of the landslides that have occurred in this area are of the rotational type. In total, 312 landslides were identified in the study area and randomly considered against the same number as non-landslide points (Brenning, 2005).

3.2. Training and testing dataset preparation

In in order to be able to do modeling work deep learning models, input must have more than one feature or characteristic. In the case of landslides, wheher this phenomenon occurred in the study area or not, the input of model can have two types of features: landslide points and non-landslide points. In the next step, both landslide and non-landslide points are randomly divided into two groups: The first group is the training dataset, which contains almost 70% of the points, and these points are used in building and teaching models, while the second group is the testing data set, which includes approximately 30% of the remaining points, and these points are used in validating and assessing the models. In the last step, both the training and test data sets are combined with 13 layers of landslide condition factor to obtain the value of the attribute that each of these points receives from these factors.

3.3. Landslide conditioning factors

Selecting the suitable factors is one of the basic steps for modeling landslide susceptibility. For this purpose, with respect tothe geographical, environmental and geological conditions of the region, as well as the availability of data and review of previous studies (Ayalew and Yamagishi, 2005; Yalcin, 2008), twelve factors were investigated as effective conditions in the occurrence of landslides for the study area. The factors are: altitude, slope angle, slope aspect, TWI, profile and plan curvatures, land-use, lithology, distance to faults, distance to rivers, distance to roads and rainfall. As it is known, human and natural elements play a significant role in the emergence of these factors. In this study, using Weka 3.8.4 software and utilizing the information gain ranking filter (IGRF) technique, 13 condition factors were evaluated and classified based on their degree of importance and impact on landslide occurrence. In this method, the minimum value of IGRF is zero, which indicates the least effective factor, and the highest value is 1, which represents the most effective factor. After digitizing all the layers, they were converted to a raster format using ArcGIS 10.3 software (Esri, Redlands, CA, USA). A brief description of each of these factors and the intervals associated with each is given below.

3.3.1. Altitude

Altitude is one of the most effective factors in the occurrence of landslides, because altitude changes have a direct effect on the stability or instability of the slopes due to the increase and decrease of vegetation (Feizizadeh et al., 2014). Normally in places with higher altitudes, landslides are more likely to occur. The altitude map was extracted from the digital elevation model (DEM) layer of the area in ArcGIS 10.3 software. Then the altitude map was divided into twelve classes (Fig. 3a).

3.3.2. Slope angle

One of the factors affecting slope stability is slope angle (Xie et al., 2019), because in general, the higher the slope angle, the lower the slope stability, and in addition, the lower the slope stability, the greater the probability of landslides. Of course, it should be noted that due to the different types of landslides, even on gentle slopes, the risk of landslides may be higher. In many studies (Lee et al., 2004; Sarkar and Kanungo, 2004), slope angle has been considered as one of the main factors in landslide susceptibility map assessment. The slope angle map is obtained from the DEM layer of the area in ArcGIS software and is divided into ten classes. The slope angle map is shown in Fig. 3b.

3.3.3. Slope aspect

The direction of the slope by other natural phenomena has a significant effect on the occurrence of landslides in such a waythat if for example in one direction the slope of vegetation uses more sunlight, the vegetation in that direction is strengthened and this increases the stability of the slope and reduces the occurrence of landslides and conversely (Sameen et al., 2020). Also, if rainfall in one direction is more than the other direction, it causes loosening of the soil in the area and rising humidity in that direction, resulting in increased landslides. Moreover, when there is more wind in one direction, the area becomes colder or warmer in different directions. The aspect map is extracted from the DEM layer of the area in ArcGIS software and divided into nine classes: Flat (-1°), North (337.5°-360°, 0°-22.5°), Northeast (22.5°-67.5°), East (67.5°-112.5°), Southeast (112.5°-157.5°), South (157.5°-202.5°), Southwest (202.5°-247.5°), West (247.5°-292.5°), and Northwest (292.5°-337.5°), as shown in Fig. 3c.

3.3.4. Topographic wetness index (TWI)

The TWI, also known as the “compound topographic index (CTI)”, is a steady-state wetness index. The TWI index is commonly used in hydrological processes and determines the effect of slope parameters on these processes. It also plays an essential role in controlling the topography of hydrological processes (Sörensen et al., 2006). Using this index, the percentage of water accumulation in different places (each pixel) can be measured. The TWI index is defined as follows (Moore et al., 1991):

1

where is the local upslope area draining through a certain point per unit contour length and is the slope angle at the point. The TWI map is generated from the DEM layer via SAGA–GIS software (http://saga-gis.org). The TWI map is divided into five classes as shown in Fig. 3d.

3.3.5. Profile and plan curvatures

Profile curvature is parallel to the direction of maximum slope (Smith et al., 2012). A negative value indicates that the surface is upwardly convex at that cell, in which a positive value indicates that the surface is upwardly concave at that cell, and a value of zero indicates that the surface is flat (Moore et al., 1991; Shary, 1995; Florinsky, 1998).

Plan curvature is perpendicular to the direction of maximum slope (Schmidt et al, 2003). A positive value indicates that the surface is sideward convex at that cell, a negative value indicates that the surface is sideward concave at that cell, and a value of zero indicates that the surface is flat (Zevenbergen et al., 1987; Moore et al., 1991).

Profile curvature affects the acceleration or deceleration of flow across the surface, so it affects erosion and deposition, and this effect can be observed and evaluated using a profile curvature map. Moreover, using the plan curvature map, convergence and divergence of the flow across the surface can be achieved (Zevenbergen et al., 1987). Understanding the combinations of the profile and plan curvatures is so important, because it provides a more accurate understanding of the flow across slopes. For these reasons, the results of these two factors help us to analyze the probability of the landslides occurrence. Profile and plan curvature maps are obtained from the DEM layer via SAGA–GIS software (http://saga-gis.org), and are divided into three classes: concave, convex, and flat. Profile and plan curvature maps are shown in Fig. (3e, f), respectively.

3.3.6. Land-use

Landuse means the change of natural environment or wilderness to build environments such as settlements and residential areas or semi-natural habitats such as agricultural lands, pastures, and managed woods. Land-use change causes the loss of vegetation and soil in the area, thus affecting the slope stabilization and landslide occurrence (Hadji et al., 2013), because the landslides are more likely to occur in areas with poorer vegetation. The land-use map of the study area has been obtained from the Land Management Organization of Iran and is divided into seven groups that include: Farmland, Forest, Mixed-use, Orchard, Rangeland, River, and Urban. The land-use map is shown in Fig. 3g.

3.3.7. Lithology

The lithology is a unit of rock that describes the physical properties of rocks present and visible in the area, such as color, texture, size, and composition (Allaby et al., 1991; Grana et al., 2010). As a result, according to the physical properties of rocks, different lithological units have different values of landslide susceptibility at different slopes (Lin et al., 2006). Therefore, care must be taken in grouping different lithological features, because the hardness of the rocks and the speed of their weathering affect the occurrence of landslides (Henriques et al., 2015). The lithology map was extracted from the geological map and then converted to raster format in ArcGIS software. The lithology map of the study area is divided into thirty-two groups which is presented in Fig. 3h.

3.3.8. Distance to faults

In term of geology, a fault is a planar fracture or discontinuity in a volume of rock across an area (Segall et al., 1980), in which there has been significant displacement as a result of rock-mass movement and landslide occurrence. The activity of plate tectonic forces creates the greatest boundaries between the plates. As a result, large faults occur within the Earth's crust such as subduction zones or transform faults (Lutgens et al., 2017). When the distance to faults decreases, the extent of fracturing and the degree of weathering of the rocks increases, which reduces the shear strength and increases the probability of landslide occurrence (Conforti et al., 2014). The fault map of the study area was obtained from the Geological Survey of Iran. Finally, using ArcGIS software, the distance map of the fault in eleven classes was obtained: 0-300, 300–600, 600–900, 900-1,200, 1,200-1,500, 1,500-1,800, 1,800-2,100, 2,100-2,400, 2,400-2,700, 2,700-3,000, and 3,000<, which is shown in Fig. 3i.

3.3.9. Distance to rivers

Rivers are naturally flowing watercourse, and flows to oceans, seas, lakes or other rivers. In some cases, a river flows to the surface and dries up at the end of its course without reaching any other water. In many studies, distance to rivers has been considered as a key factor in assessing landslide susceptibility (Moayedi et al., 2019; Wang et al., 2020). The shorter the distance to rivers from the slopes, the lower the stability of the slopes will be due to the erosion of the slope toes and the increase of saturation of materials on the slopes (Zhao et al., 2019). This also increases groundwater and pore water pressure and a result, landslides are more likely to occur near rivers (Dai et al., 2001). The map of rivers has been obtained from the National Cartographic Center. Then in ArcGIS software using buffer tool and Euclidean distance method was divided into eleven classes: 0-100, 100–200, 200–300, 300–400, 400–500, 500–600, 600–700, 700–800, 800–900, 900-1,000, 1,000<. The distance to rivers map is shown in Fig. 3j.

3.3.10. Distance to roads

One of the most important human activities that affect landslides is road construction, because it changes the physical condition of the slopes, hills and the topography of the area, which used to have stable conditions before construction began. On the other hand, landslides are more likely to occur along roads, highways and railways, especially in mountainous areas (Acharya and Lee, 2019). Therefore, the construction of roads in mountainous areas should be prevented as much as possible, or at least, construction should begin after studying and researching the geological conditions of the area (Nsengiyumva et al., 2019). Road map of the study area has been prepared by the Iranian Survey Organization. Then, in ArcGIS software using buffer toolbar and “Euclidean distance” method, it was divided into eleven classes. The distance to roads map is shown in Fig. 3k.

3.3.11. Rainfall

Rainfall is one of the critical factors in the occurrence and triggering of landslides, and has been used in many studies (Fang et al., 2020). Mountainous and near-sea areas with the highest rainfall are very prone to landslides. The amount, intensity, and duration of rainfall are three important factors for assessing

landslide susceptibility (Shimoda and Ochiai, 2006), and according to the type of landslide, each of these three factors could have a greater impact than the other. Rainfall has a great effect on soil moisture and weakens the stability of the slopes, as a result, the probability of landslides increases (Pham et al., 2015). Rainfall map was obtained using the average rainfall data of the Meteorological Organization. It was classified in ArcGIS software into nine classes as shown in Fig. 3l.

3.4 Methods

Artificial Neural Networks (ANNs), or connecting systems, are computing systems inspired by biological neural networks (Garrett., 1994; Krenker et al., 2011). ANNs use machine computational methods to perform machine learning by examining data. In fact, the goal of ANNs is that whatever human beings can learn and do, the machine can also do and learn, making this possible by processing data and extracting their salient features. ANNs typically consist of three layers: input layer for receiving raw data, hidden layer for weighting the inputs and the connection between them, and output layer to display results (Wang, 2003). Each layer contains a group of nerve cells (neurons), similar to biological neurons in a human brain, and each neuron has a weight that is adjusted as learning progresses, and is generally associated with all neurons in other layers, unless the user restricts communication between neurons. However the neurons in each layer have no connection with other neurons in the same layer (Lee et al., 2006). In these networks, if one cell is damaged, other cells can compensate for its absence, and also participate in its reconstruction. These networks are able to learn and learning in these systems is adaptive. This network is used in various fields such as computer science (Villarrubia et al., 2018), technical and engineering sciences (Can et al., 2019), medical sciences (Hirasawa et al., 2018), experimental and biological sciences (Acheampong and Boateng, 2019), etc. Since ANNs consist of only three layers, when multidimensional data with multiple features exist, recognizing and differentiating data and the relationship between them is very complex. Deep learning neural networks are used to solving this problem, because they have more layers in the hidden layer, and this feature strengthens their ability to process multidimensional data in these networks. In this study, two specific types of deep learning neural networks called convolutional neural networks (CNN) and long short-term memory (LSTM) were used.

3.4.1. Convolutional neural network (CNN)

CNN is a class of deep neural networks commonly used for visual, speech, and text analysis in machine learning (Wang et al., 2019). CNNs are feed forward networks, meaning that the signal in these networks only travels in one direction from the input layer to the hidden layer and then to the output layer, and previous data is not stored in memory (Girshick, 2015). Architecturally, CNN has an input layer, a hidden layer, and an output layer (Shin et al., 2016). The hidden layer includes one or more convolution layers, an activation layer, max pooling, and is fully connected (Shin et al., 2016). The input layer consists of several

neurons and our data have different classes according to the number of neurons that are introduced to the network as vectors (Fang et al., 2020). For example, in this study, a landslide occurred or did not occur in one place. As a result, the number of our classes is two, and our network is two-dimensional; number 1 for landslides and zero for non-landslides. In the hidden layer, there is a convolutional layer, which can extract various features from the input layer by using kernels or filters with different dimensions. It should be noted that the kernel dimension or filters should be the same as the input dimension. The activation layer consists of a non-linear activation function after the convolutional layer, and one of the most popular activation functions is the rectified linear unit (ReLU) (Nair and Hinton, 2010). Then there is the max pooling layer, which is responsible for reducing the size and the number of CNN network parameters. The output of this layer is sent to the fully connected layer after being converted to a one-dimensional vector. As the name implies, all the neurons in this layer are connected to the neurons in the previous layer. The main task of the fully connected layer is to combine the local attribute in the lowers layer with the local attribute in the upper layers. In the classification of the last fully connected layer in the network, it combines all the features to classify the image, so the output size of this layer is equal to the number of classes recognizable by the network. For example, in this study, the number of our classes is two and it will be the size of our output. The number of convolutions, max pooling, and fully connected layers on CNN is determined by the user, and the higher the number of layers, the deeper our network becomes, resulting in more recognizable features of the input image (Sharif Razavian et al., 2014). Finally, the results of the model can be seen in the output layer. CNN's most popular network architectures include LeNet-5, AlexNet, ZFNet, VGG, GoogLeNet, and Microsof ResNet (Fang et al., 2019; Khan et al., 2020). The LeNet-5 is the first network to use convolutional filters (LeCun et al., 1998). AlexNet has five convolutional layers and three fully connected layers, which puts it in the category of shallow networks (Krizhevsky et al., 2012). The ReLU function was first introduced in this network. ZFNet does not have much of an idea, and the most important idea is visualization, a tool that uses it to determine the use of each neuron in the network to some extent and greatly helps in understanding of convolutional networks (Zeiler and Fergus, 2013). VGG is less complex than AlexNet and ZFNet because it reduces the number of hyper parameters, and because of its simplicity, it is very popular (Ren et al. 2016). GoogLeNet was introduced by Google in 2014 and designed a module called Inception module with the approach of turning network hyper parameters into a learnable parameter. In this module, three convolutional filters, which have different sizes and a pooling filter, are applied on the previous layer and put the results together as a single tensor. Microsoft ResNet was introduced by Microsoft, which stands for Residual Network. In this network, communication outside the convolutional structure between the layers is considered to transfer the inputs of the previous layer to the next layer without intermediaries, and in the back propagation stage, to transfer the error of each layer to the previous layer so that the network can be deepened and taught faster (He et al., 2016). Comparing the performance of these networks with each other has already been done by many researchers and more can be obtained by referring to Sze et al., 2017 and Khan et al., 2019.

3.4.2. Long short-term memory

Another type of deep neural network is the recurrent neural network (RNN), which is used in speech recognition, natural language processing, machine translation, sequential data processing, as well as predicting time series (such as weather forecast data, financial data, etc.) (Bandara et al., 2020). Unlike CNNs, RNNs have a feedback layer in which the network output is returned to the network with the next input (Graves et al., 2008). RNN can remember its previous input due to its internal memory and use this memory to process a sequence of inputs. In simpler terms, RNNs consist of a recurrent loop that prevents the information we have already obtained from being lost and remains in the network, which perpetuates the information. Long short-term memory (LSTM) is a type of model or structure for sequential data that emerged for the development of RNNs and is one of the most popular RNN architectures (Hochreiter et al., 1997). The term long term memory refers to the weights learned and short-term memory refers to the internal states of cells. LSTM was developed to solve the vanishing/exploding gradient problem in RNNs, the major change being the replacement of the middle layer of RNN with a block called LSTM (Cortez et al., 2018). The biggest feature of LSTM is the ability to learn long-term dependency that was not possible with RNNs. In other words, unlike RNNs, LSTMs do not have a problem with long sequences, and the mechanism designed in them allows them to work with longer sequences, which leads to better performance in comparison with other machine learning algorithms such as SVM (Hochreiter et al., 1997). The presence of LSTM networks has become very colorful in recent years. These networks are now used in advanced technologies such as Google Voice and natural hazard forecasting (Ma et al., 2019). In general, according to Christopher Olah's article published in 2015, LSTM operates as follows: the first step in LSTM is to determine on which piece of information will be deleted from the state cell. This decision is made by a sigmoid layer called the forget gate. In other words, forget gate decides what information to keep or delete from previous cells and uses the following equation (Hochreiter and Schmidhuber, 1997):

3

where , , and represent the sigmoid function, weights, and constants, respectively. In the second step, we should decide what new information to store in the state cell. This decision is made by a sigmoid layer called the gate input. Given the values of and for each number, this gate outputs a value of zero or 1 in the state cell. A value of 1 means completely transferring the current value of the state cell to , and a value of zero means completely erasing the current state cell information and not transferring any of it to , its equation is as follows:

4

The next step is a layer of hyperbolic tangent that forms a vector of values called that can be added to the state cell according to Eq. (5):

5
Then by combining the previous two steps, we update the state cell value. The new is obtained by elementwise multiplying the previous value of the state cell by (and its sum by multiplying by (in accordance withthe following equation:

6
Finally, what information isgoing as output should be determined. This output will be based on the value of the state cell, but it will pass a certain filter. First, a sigmoid layer decides what part of the state cell is to be output. Then the value of the state cell (after updating in the previous steps) is given to a layer of hyperbolic tangent (until the values are between − 1 and + 1) and its value is multiplied by the output of the previous sigmoid layer so that only the parts of information go to the output. The equations of the last step are as follows:

7

8

3.5. Generate landslide susceptibility map

In this step, after integrating all thirteen condition factor maps using the combine tool in ArcGIS software, an output from its attribute table in Excel software is prepared (Microsoft Corp., Redmond, WA, USA). Using the training dataset, the models are taught, namely CNN and LSTM. Then comparisone and evaluation of the performance of each model with a test dataset is carried out. Afterwards, with the help of the Excel file obtained from all the data of the region and each of the CNN and LSTM models, the landslide susceptibility indices (LSI) are determined for each pixel of the landslide map in the study area by Matlab programming language (MathWorks, Natick, MA, USA). Finally, the landslide susceptibility map is classified into five different classes of susceptibility using the quantile method in ArcGIS software: ‘very low, low, moderate, high, and very high”. These maps shows in which part of the area the probability of landslide occurrence is higher (Fig. 4).

3.6. Evaluating the accuracy of models and maps

After generating a landslide susceptibility map, the reliability of each of these maps should be assessed. The performance of CNN and LSTM models were evaluated using test datasets that had no role in the modeling process. Then, by obtaining the AUC value, the prediction rate and the performance of the

models are evaluated (Wu et al., 2020). Furthermore, the accuracy of the two models are determined using the RMSE.

4. Results

4.1. Evaluating the impact of condition factors

According to the results of IGRF method in Weka software, land-use has the highest importance and impact (IGRF = 0.5708), while profile curvature factor has the least importance and impact (IGRF = 0.0123) in landslide occurrence. After land-use, the order of importance of the factors is as follows: lithology (IGRF = 0.4930), slope angle (IGRF = 0.3865), altitude (IGRF = 0.2941), rainfall (IGRF = 0.1837), distance to fault (IGRF = 0.1236), TWI (IGRF = 0.1020), distance to river (IGRF = 0.0665), slope aspect (IGRF = 0.0635), distance to road (IGRF = 0.0611), and plan curvature (IGRF = 0.0176). According to the values obtained for each factor, it is clear that all factors considered for landslide susceptibility zoning affect the landslide occurrence, and this indicates the correct selection of factors in the present study.

4.2. Validation and evaluation of model results in landslide susceptibility map

One of the ways to evaluate the performance of the two applied models is to check the RMSE for each of them. The RMSE results for both CNN and LSTM models obtained using the training and testing datasets are shown in Fig. 5 and Fig. 6. Based on the results, the RMSE value in the CNN model for each of the training and testing datasets is 0.121 and 0.132, respectively. The RMSE values in the LSTM model for each of the training and testing datasets are 0.185 and 0.188, respectively. Therefore, it can be concluded that CNN performance is slightly better than LSTM, but in general, both models have close performance and the accuracy of both models is acceptable.

Additionally, each of the landslide susceptibility maps generated by CNN and LSTM models were validated using the AUC values for the test dataset (Fig. 7). The AUC values obtained from the test dataset for each of the CNN and LSTM models are 0.821 and 0.832, respectively. This number indicates that the susceptibility maps produced by these two models are reliable in terms of validation and also the accuracy of both models is close to each other. The AUC values on the CNN and LSTM models on the training dataset are 0.822 and 0.828, respectively. Comparing these two values with the AUC values obtained from the testing dataset of each model (0.821 and 0.832), it can be concluded that there is no overfitting problem in the CNN and LSTM models, and that our models perform properly.

Moreover, the CNN and LSTM models can be compared in terms of processing the time and convergence speed (Fig. 8 and Fig. 9). CNN and LSTM processing times are about 3 seconds and 4 seconds, respectively. Also, the CNN model converges in 60 iterations and the LSTM model in 500 iterations. Comparing the results, it can be seen that in terms of processing time and convergence speed, the CNN model has a better performance than the LSTM model. In general, in the study area, both models have good performance and are close to each other, but in terms of RMSE, CNN is slightly better than LSTM. In

terms of AUC, LSTM performance is slightly better than CNN, and as a result, the landslide susceptibility map produced with the LSTM model is more reliable. Finally, CNN performs better in terms of processing time and convergence speed.

According to the results obtained from the map produced by CNN model, 20.11%, 20.35%, 20.10%, 19.88%, and 19.56% of the study areas are in “very low, low, moderate, high, and very high” classes, respectively. In the map produced with the LSTM model, 19.67%, 20.14%, 20.89%, 19.74%, and 19.56% of the areas are in “very low, low, moderate, high, and very high” classes, respectively. According to the generated maps, it can be concluded that approximately 40% of the study area is in high and very high classes and therefore these areas are prone to landslide occurrence.

5. Discussion

Annual natural disasters kill many people around the world, and in addition to life threats, cause great financial damage to the country's infrastructure and economy. Therefore, predicting natural disasters and reducing their damage has always been one of the most important and significant issues for different organizations in each country. Landslides are also one of the most important natural disasters; therefore obtaining an LSM and identifying landslide prone areas will reduce human and financial losses and helps relevant organizations in identifying these areas and taking the necessary measures when facing this phenomenon. Many studies have been carried out in this field and researchers have used different models for LSM in different areas. In Iran, due to the availability of data and the type of study area, different models have been used, like machine learning models in the watershed of Galikash River in Golestan Province (Arabameri et al., 2020), fuzzy methods and AHP in Guilan Province (Bahrami et al., 2020), and metaheuristic techniques & adaptive neural fuzzy inference system in Qazvin Province (Mehrabi et al., 2020). The objective of this study is using deep learning models including CNN and LSTM models to produce the LSMs in Ardabil Province. To achieve this goal, thirteen factors that are effective in the landslide occurrence were selected, and then using the IGRF method, the importance of each factor was determined.

LSMs are used for determining landslide prone areas, and using these maps accident hotspots are identified in terms of landslide occurrence and decision makers can take the necessary measures to reduce and prevent landslides. Therefore, the accuracy of the produced maps is a fundamental and an important matter, and selection of appropriate models as well as the reliability of the maps produced by these models should be considered. Due to the proper performance of deep learning models when the dataset is large (Ngo et al., 2020), the two well-known deep learning models, namely CNN and LSTM are employed to generate the LSMs.

Another thing that should be considered in preparing a landslide susceptibility map, is the selection of effective condition factors in the landslide occurrence and also evaluating the impact and importance of each of these factors in the landslide occurrence. For this purpose, in this study, according to previous studies and environmental and geographical conditions of the region, thirteen effective condition factors

were chosen. Afterward, using the IGRF method, the order of importance and impact of each of these factors on the landslide occurrence were obtained. According to IGRF results, land-use has the greatest impact on landslide occurrence. Recently, with the increase in population and development of cities, land-use change has arisen, and as mentioned before, land-use change causes vegetation loss and loosening of the soil in the area, resulting in slope instability. As a result, land-use is one of the most important factors in the landslide occurrence, and which the results obtained by the IGRF prove this in the study area. Lithology is known as the second factor influencing the landslide occurrence in the study area because according to the type of rocks in the area, their hardness and weathering speed is different, which is effective in soil erosion and creating empty spaces between the rocks and increases the landslide occurrence. Slope angle is also known as the third factor influencing the landslide occurrence, which greatly affects the stability of the slope and the lowers the stability of the slope. In other words, the greater the slope, the greater the probability of the landslide occurrence is. The results also show that the profile curvature factor has the least effect on the landslide occurrence. However, it should be noted that the order of importance of these factors in other studies may vary depending on the methods used and the study area. According to the results obtained from the IGRF, it can be concluded that all the considered factors affect the landslide occurrence with different ratios and show that the selection of condition factors for the study area has been performedcorrectly.

In the present study, CNN and LSTM deep learning algorithms have been used to generate landslide susceptibility map and model construction. The AUC values and RMSE were also utilized to evaluate the performance of the models and compare them, and the IGRF method was used for evaluating the ratio of the impact and importance of each factor on the landslide occurrence. In future studies, it is recommended to use other deep learning algorithms such as Deep Boltzmann Machine (DBM) and Deep Belief Network (DBN) in the study area, and compare these algorithms with the algorithms used in this study, namely CNN and LSTM. In addition, other available methods such as gain ratio feature evaluator (GRFE) and step-wise weight assessment ratio analysis (SWARA) can be used for investigating the impact of each factor on landslide occurrence, and each of these methods can be compared with each other and the difference in the selection of factors affecting the landslide occurrence in each of the methods could be observed, and thus selected the most effective factors are selected and eliminated the factors that have no effect on the landslide occurrence.

6. Conclusion

In this study, using CNN and LSTM algorithms, which are part of deep learning algorithms, were used to map the landslide susceptibility map in Ardabil Province, Iran. Then, using the AUC values, the performance of these models were evaluated. The AUC values for CNN and LSTM models are 0.821 and 0.832, respectively. Also, using the IGRF method, the effect and importance of each factor on the landslide occurrence was obtained. According to the results of IGRF method, land-use has the highest effect (IGRF = 0.5708) and profile curvature factor has the least effect (IGRF = 0.0123) on landslide occurrence. Finally, the study area was classified into five different classes of susceptibility and percentage of the study area in each class was determined. According to the results of CNN and LSTM

models, 39.44% and 39.30% of the study area are located in landslide prone areas, respectively. Therefore, necessary measures should be taken to control and prevent landslides in areas with high and very high classes. The results of this study can be used as a guide for organizations related to natural disasters and help them make better decisions in the event of a landslide and prevent landslides. Based on our studies, for the first time, CNN and LSTM models have been used together for landslide susceptibility mapping in the study area, and so far the performance of these two models has not been compared with each other in landslide susceptibility mapping. Therefore, the results of this study can help other researchers when using other models in the study area by evaluating and comparing the performance of different models with each other.

Declarations

Acknowledgment

This research was supported by the Basic Research Project of the Korea Institute of Geoscience and Mineral Resources (KIGAM) and Project of Environmental Business Big Data Platform and Center Construction funded by the Ministry of Science and ICT.

References

Acharya TD, Lee DH (2019) Landslide susceptibility mapping using relative frequency and predictor rate along Araniko Highway. KSCE J Civ Eng 23(2):763–776

Acheampong AO, Boateng EB (2019) Modelling carbon emission intensity: Application of artificial neural network. J Clean Prod 225:833–856

Akgun A, Erkan O (2016) Landslide susceptibility mapping by geographical information system-based multivariate statistical and deterministic models: In an artificial reservoir area at Northern Turkey. Arab J Geosci 9(2):165

Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. Bulletin of engineering geology the environment 58(1):21–44

Ali SA, Parvin F, Pham QB, Vojtek M, Vojtekova J, Costache R,.. . Ghorbani MA (2020) GIS-based comparative assessment of flood susceptibility mapping using hybrid multi-criteria decision-making approach, naïve Bayes tree, bivariate statistics and logistic regression: A case of Topľa basin, Slovakia. Ecol Ind 117:106620

Allaby A, Allaby M (1991) Concise Oxford dictionary of earth sciences. Oxford University Press

Arabameri A, Pradhan B, Rezaei K, Lee S, Sohrabi M (2020) An ensemble model for landslide susceptibility mapping in a forested area. Geocarto International 35(15):1680–1705

Arabameri A, Saha S, Roy J, Chen W, Blaschke T, Bui T, D (2020) Landslide susceptibility evaluation and management using different machine learning methods in the Gallicash River Watershed, Iran. Remote Sensing 12(3):475

Azizi A, Malekmohammadi B, Jafari HR, Nasiri H, Parsa VA (2014) Land suitability assessment for wind power plant site selection using ANP-DEMATEL in a GIS environment: case study of Ardabil province, Iran. Environ Monit Assess 186(10):6695–6709

Bahrami Y, Hassani H, Maghsoudi A (2020) Landslide susceptibility mapping using AHP and fuzzy methods in the Gilan province, Iran. *GeoJournal*, 1–20

Bai S-B, Wang J, Lü G-N, Zhou P-G, Hou S-S, Xu S-N (2010) GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. Geomorphology 115(1–2):23–31

Bandara K, Bergmeir C, Smyl S (2020) Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. Expert Syst Appl 140:112896

Berberian M (1981) Active faulting and tectonics of Iran. Zagros Hindu Kush Himalaya Geodynamic Evolution 3:33–69

Berberian M (1983) Generalized tectonic map of Iran. *Continental Deformation in the Iranian Plateau: Contribution to the Seismotectonics of Iran, part IV*, *Geol. Surv. Iran*, *52*

Betts H, Basher L, Dymond J, Herzig A, Marden M, Phillips C (2017) Development of a landslide component for a sediment budget model. Environ Model Softw 92:28–39

Bragagnolo L, da Silva R, Grzybowski J (2020) Artificial neural network ensembles applied to the mapping of landslide susceptibility. Catena 184:104240

Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat Hazards Earth Syst Sci 5(6):853–862

Bui DT, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016) Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13(2):361–378

Can A, Dagdelenler G, Ercanoglu M, Sonmez H (2019) Landslide susceptibility mapping at Ovacık-Karabük (Turkey) using different artificial neural network models: comparison of training algorithms. Bulletin of engineering geology the environment 78(1):89–102

Can A, Dagdelenler G, Ercanoglu M, Sonmez H (2019) Landslide susceptibility mapping at Ovacık-Karabük (Turkey) using different artificial neural network models: comparison of training algorithms. Bulletin of engineering geology the environment 78(1):89–102

Cascini L, Bonnard C, Corominas J, Jibson R, Montero-Olarte J (2005) Landslide hazard and risk zoning for urban planning and development. *Landslide Risk Management. Taylor and Francis, London*, 199–235

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

Corominas J, van Westen C, Frattini P, Cascini L, Malet J-P, Fotopoulou S,.. . Agliardi F (2014) Recommendations for the quantitative analysis of landslide risk. Bulletin of engineering geology the environment 73(2):209–263

Cortez B, Carrera B, Kim Y-J, Jung J-Y (2018) An architecture for emergency event prediction using LSTM recurrent neural networks. Expert Syst Appl 97:315–324

Costanzo D, Rotigliano E, Irigaray Fernández C, Jiménez-Perálvarez JD, Chacón Montero J (2012) Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: application to the river Beiro basin (Spain)

Cracknell MJ, Reading AM (2014) Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Computers geosciences 63:22–33

Dai F, Lee C, Li J, Xu Z (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. Environ Geol 40(3):381–391

Das I, Stein A, Kerle N, Dadhwal VK (2012) Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models. Geomorphology 179:116–125

De Blasio FV (2011) Introduction to the physics of landslides: lecture notes on the dynamics of mass wasting. Springer Science & Business Media

De Smith MJ, Goodchild MF, Longley P (2007) Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Troubador publishing ltd

Ding A, Zhang Q, Zhou X, Dai B (2016) *Automatic recognition of landslide based on CNN and texture change detection.* Paper presented at the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)

Dodangeh E, Choubin B, Eigdir AN, Nabipour N, Panahi M, Shamshirband S, Mosavi A (2020) Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. Sci Total Environ 705:135983

Fang W, Ding Y, Zhang F, Sheng VS (2019) DOG: A new background removal for object recognition from images. Neurocomputing 361:85–91

Fang Z, Wang Y, Peng L, Hong H (2020) Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Computers geosciences 139:104470

Farrokhnia A, Pirasteh S, Pradhan B, Pourkermani M, Arian M (2011) A recent scenario of mass wasting and its impact on the transportation in Alborz Mountains, Iran using geo-information technology. Arab J Geosci 4(7–8):1337–1349

Feizizadeh B, Blaschke T, Nazmfar H (2014) GIS-based ordered weighted averaging and Dempster–Shafer methods for landslide susceptibility mapping in the Urmia Lake Basin, Iran. Int J Digit Earth 7(8):688–708

Feizizadeh B, Blaschke T, Nazmfar H, Rezaei Moghaddam M (2013) Landslide susceptibility mapping for the Urmia Lake basin, Iran: a multi-criteria evaluation approach using GIS. International Journal of Environmental Research 7(2):319–336

Feizizadeh B, Jankowski P, Blaschke T (2014) A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. Computers geosciences 64:81–95

Florinsky IV (1998) Accuracy of local topographic variables derived from digital elevation models. Int J Geogr Inf Sci 12(1):47–62

Garrett J (1994) Where and why artificial neural networks are applicable in civil engineering

Ghorbanzadeh O, Blaschke T, Aryal J, Gholaminia K (2020) A new GIS-based technique using an adaptive neuro-fuzzy inference system for land subsidence susceptibility mapping. Journal of Spatial Science 65(3):401–418

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

Girshick R (2015) *Fast r-cnn.* Paper presented at the Proceedings of the IEEE international conference on computer vision

Grana D, Della Rossa E (2010) Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. Geophysics 75(3):O21–O37

Graves A (2013) Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*

Graves A, Liwicki M, Fernández S, Bertolami R, Bunke H, Schmidhuber J (2008) A novel connectionist system for unconstrained handwriting recognition. IEEE Trans Pattern Anal Mach Intell 31(5):855–868

Guirado E, Tabik S, Alcaraz-Segura D, Cabello J, Herrera F (2017) Deep-learning convolutional neural networks for scattered shrub detection with google earth imagery. *arXiv preprint arXiv:1706.00917*

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31(1–4):181–216

Hadji R, Limani Y, Baghem M, Demdoum A (2013) Geologic, topographic and climatic controls in landslide hazard assessment using GIS modeling: a case study of Souk Ahras region, NE Algeria. Quatern Int 302:224–237

Haeri S, Satari M (1993) Great Landslides Triggered by Manjil Earthquake, 20 June 1990. *Natural Disaster Reduction Center of Iran (In Persian)*

Havenith H-B, Strom A, Torgoev I, Torgoev A, Lamair L, Ischuk A, Abdrakhmatov K (2015) Tien Shan geohazards database: Earthquakes and landslides. Geomorphology 249:16–31

He K, Zhang X, Ren S, Sun J (2016) *Deep residual learning for image recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition

Henriques C, Zêzere JL, Marques F (2015) The role of the lithological setting on the landslide pattern and distribution. Engineering geology 189:17–31

Hirasawa T, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T,.. . Fujisaki J (2018) Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer 21(4):653–660

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural computation 9(8):1735–1780

Hu Q, Zhou Y, Wang S, Wang F (2020) Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River Basin. Geomorphology 351:106975

Jaafari A, Panahi M, Pham BT, Shahabi H, Bui DT, Rezaie F, Lee S (2019) Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. Catena 175:430–445

Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. Science 349(6245):255–260

Kalchbrenner N, Blunsom P (2013) *Recurrent continuous translation models.* Paper presented at the Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing

Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53(8):5455–5516

Khan RU, Zhang X, Kumar R (2019) Analysis of ResNet and GoogleNet models for malware detection. Journal of Computer Virology Hacking Techniques 15(1):29–37

Krenker A, Bešter J, Kos A (2011) Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, 1–18

Krizhevsky A, Sutskever I, Hinton GE (2012) Advances in neural information processing systems. *Neural Information Processing Systems Foundation*, *1269*

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

Lacasse S, Nadim F (2009) Landslide risk assessment and mitigation strategy. In: Landslides–disaster risk reduction. Springer, pp 31–61

LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324

Lee S, Ryu J-H, Lee M-J, Won J-S (2006) The application of artificial neural networks to landslide susceptibility mapping at Janghung, Korea. Math Geol 38(2):199–220

Lee S, Ryu J-H, Won J-S, Park H-J (2004) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Engineering geology 71(3–4):289–302

Li Y, Chen W (2020) Landslide susceptibility evaluation using hybrid integration of evidential belief function and machine learning techniques. Water 12(1):113

Lin C-W, Liu S-H, Lee S-Y, Liu C-C (2006) Impacts of the Chi-Chi earthquake on subsequent rainfall-induced landslides in central Taiwan. Engineering geology 86(2–3):87–101

Long J, Shelhamer E, Darrell T (2015) *Fully convolutional networks for semantic segmentation.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition

Lutgens FK, Tarbuck EJ, Tasa DG (2017) Essentials of geology. Pearson

Ma J, Ding Y, Gan VJ, Lin C, Wan Z (2019) Spatiotemporal prediction of PM2. 5 concentrations at different time granularities using IDW-BLSTM. IEEE Access 7:107897–107907

Maggiori E, Tarabalka Y, Charpiat G, Alliez P (2016) Convolutional neural networks for large-scale remote-sensing image classification. IEEE Trans Geosci Remote Sens 55(2):645–657

Mehrabi M, Pradhan B, Moayedi H, Alamri A (2020) Optimizing an Adaptive Neuro-Fuzzy Inference System for Spatial Prediction of Landslide Susceptibility Using Four State-of-the-art Metaheuristic Techniques. Sensors 20(6):1723

Mersha T, Meten M (2020) GIS-based landslide susceptibility mapping and assessment using bivariate statistical methods in Simada area, northwestern Ethiopia. Geoenvironmental Disasters 7(1):1–22

Mikolov T (2012) Statistical language models based on neural networks. *Presentation at Google, Mountain View, 2nd April, 80*, 26

Moayedi H, Mehrabi M, Mosallanezhad M, Rashid ASA, Pradhan B (2019) Modification of landslide susceptibility mapping using optimized PSO-ANN technique. Engineering with Computers 35(3):967–984

Moore ID, Grayson R, Ladson A (1991) Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological processes 5(1):3–30

Nair V, Hinton GE (2010) *Rectified linear units improve restricted boltzmann machines.* Paper presented at the ICML

Neaupane KM, Piantanakulchai M (2006) Analytic network process model for landslide hazard zonation. Engineering geology 85(3–4):281–294

Ngo PTT, Panahi M, Khosravi K, Ghorbanzadeh O, Karimnejad N, Cerda A, Lee S (2020) Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geoscience Frontiers*

Nsengiyumva JB, Luo G, Amanambu AC, Mind'je R, Habiyaremye G, Karamage F,.. . Mupenzi C (2019) Comparing probabilistic and statistical methods in landslide susceptibility modeling in Rwanda/Centre-Eastern Africa. Sci Total Environ 659:1457–1472

Oh H-J, Pradhan B (2011) Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. Computers geosciences 37(9):1264–1276

Park HJ, Lee JH, Woo I (2013) Assessment of rainfall-induced shallow landslide susceptibility using a GIS-based probabilistic approach. Engineering geology 161:1–15

Park S, Choi C, Kim B, Kim J (2013) Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. Environ Earth Sci 68(5):1443–1464

Pham BT, Jaafari A, Prakash I, Bui DT (2019) A novel hybrid intelligent model of support vector machines and the MultiBoost ensemble for landslide susceptibility modeling. Bulletin of engineering geology the environment 78(4):2865–2886

Pham BT, Phong TV, Nguyen HD, Qi C, Al-Ansari N, Amini A,.. . Ly H-B (2020) A Comparative Study of Kernel Logistic Regression, Radial Basis Function Classifier, Multinomial Naïve Bayes, and Logistic Model Tree for Flash Flood Susceptibility Mapping. Water 12(1):239

Pham BT, Prakash I, Dou J, Singh SK, Trinh PT, Tran HT,.. . Shirzadi A (2020) A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. Geocarto International 35(12):1267–1292

Pham BT, Bui T, Indra D, P., & Dholakia M (2015) Landslide susceptibility assessment at a part of Uttarakhand Himalaya, India using GIS–based statistical approach of frequency ratio method. Int J Eng Res Technol 4(11):338–344

Pradhan B, Lee S (2010) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environ Earth Sci 60(5):1037–1054

Qayyum A, Malik AS, Saad NM, Iqbal M, Faris Abdullah M, Rasheed W,.. . Bin Jafaar MY (2017) Scene classification for aerial images based on CNN using sparse coding technique. Int J Remote Sens 38(8–10):2662–2685

Regmi AD, Devkota KC, Yoshida K, Pradhan B, Pourghasemi HR, Kumamoto T, Akgun A (2014) Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. Arab J Geosci 7(2):725–742

Ren S, He K, Girshick R, Sun J (2016) Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149

Salvati P, Bianchi C, Rossi M, Guzzetti F (2010) Societal landslide and flood risk in Italy. *Natural Hazards & Earth System Sciences*, *10*(3)

Sameen MI, Sarkar R, Pradhan B, Drukpa D, Alamri AM, Park H-J (2020) Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests. Computers geosciences 134:104336

San BT (2014) An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: the Candir catchment area (western Antalya, Turkey). International journal of applied earth observation geoinformation 26:399–412

Sarkar S, Kanungo D (2004) An integrated approach for landslide susceptibility mapping using remote sensing and GIS. Photogrammetric Engineering Remote Sensing 70(5):617–625

Schmidt J, Evans IS, Brinkmann J (2003) Comparison of polynomial models for land surface curvature calculation. Int J Geogr Inf Sci 17(8):797–814

Segall P, Pollard D (1980) Mechanics of discontinuous faults. Journal of Geophysical Research: Solid Earth 85(B8):4337–4350

Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S (2014) *CNN features off-the-shelf: an astounding baseline for recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition workshops

Shary PA (1995) Land surface in gravity points classification by a complete system of curvatures. Math Geol 27(3):373–390

Shimoda Y, Ochiai H (2006) Slide switch assemblies. In: Google Patents

Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I,.. . Summers RM (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35(5):1285–1298

Sifa SF, Mahmud T, Tarin MA, Haque DME (2020) Event-based landslide susceptibility mapping using weights of evidence (WoE) and modified frequency ratio (MFR) model: A case study of Rangamati district in Bangladesh. Geology Ecology Landscapes 4(3):222–235

Sörensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index:. evaluation of different methods based on field observations

Sun D, Wen H, Wang D, Xu J (2020) A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology*, 107201

Sze V, Chen Y-H, Yang T-J, Emer JS (2017) Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE, 105*(12), 2295–2329

Thanh DQ, Nguyen DH, Prakash I, Jaafari A, Nguyen V-T, Van Phong T, Pham BT (2020) GIS based frequency ratio method for landslide susceptibility mapping at Da Lat City, Lam Dong province, Vietnam. Vietnam J Earth Sci 42:55–66

Thiery Y, Malet J-P, Sterlacchini S, Puissant A, Maquaire O (2007) Landslide susceptibility assessment by bivariate methods at large scales: application to a complex mountainous environment. Geomorphology 92(1–2):38–59

Tsangaratos P, Ilia I, Hong H, Chen W, Xu C (2017) Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China. Landslides 14(3):1091–1111

Villarrubia G, De Paz JF, Chamoso P, De la Prieta F (2018) Artificial neural networks used in optimization problems. Neurocomputing 272:10–16

Wang G, Lei X, Chen W, Shahabi H, Shirzadi A (2020) Hybrid computational intelligence methods for landslide susceptibility mapping. Symmetry 12(3):325

Wang S-C (2003) Artificial neural network. In: Interdisciplinary computing in java programming. Springer, pp 81–100

Wang Y, Fang Z, Hong H (2019) Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. Sci Total Environ 666:975–993

Wang Y, Fang Z, Wang M, Peng L, Hong H (2020) Comparative study of landslide susceptibility mapping with different recurrent neural networks. Computers geosciences 138:104445

Westen Cv, Terlien M (1996) An approach towards deterministic landslide hazard analysis in GIS. A case study from Manizales (Colombia). Earth surface processes landforms 21(9):853–868

Wilde M, Günther A, Reichenbach P, Malet J-P, Hervás J (2018) Pan-European landslide susceptibility mapping: ELSUS Version 2. Journal of maps 14(2):97–104

Wu Y, Ke Y, Chen Z, Liang S, Zhao H, Hong H (2020) Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. Catena 187:104396

Xie J, Uchimura T, Chen P, Liu J, Xie C, Shen Q (2019) A relationship between displacement and tilting angle of the slope surface in shallow landslides. Landslides 16(6):1243–1251

Xie M, Esaki T, Zhou G (2004) GIS-based probabilistic mapping of landslide hazard using a three-dimensional deterministic model. Natural hazards 33(2):265–282

Yalcin A, Reis S, Aydinoglu A, Yomralioglu T (2011) A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. Catena 85(3):274–287

Yang HL, Lunga D, Yuan J (2017) *Toward country scale building detection with convolutional neural network using aerial images.* Paper presented at the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)

Yilmaz I (2010) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci 61(4):821–836

Zaremba W, Sutskever I, Vinyals O (2014) Recurrent neural network regularization. *arXiv preprint arXiv:1409.2329*

Zeiler MD, Fergus R (2013) Stochastic pooling for regularization of deep convolutional neural networks. *arXiv preprint arXiv:1301.3557*

Zevenbergen LW, Thorne CR (1987) Quantitative analysis of land surface topography. Earth surface processes landforms 12(1):47–56

Zhao X, Chen W (2020) GIS-based evaluation of landslide susceptibility models using certainty factors and functional trees-based ensemble techniques. Applied Sciences 10(1):16

Zhao Y, Wang R, Jiang Y, Liu H, Wei Z (2019) GIS-based logistic regression for rainfall-induced landslide susceptibility mapping under different grid sizes in Yueqing, Southeastern China. Engineering geology

259:105147

Zhu XX, Tuia D, Mou L, Xia G-S, Zhang L, Xu F, Fraundorfer F (2017) Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience Remote Sensing Magazine 5(4):8–36

Figures

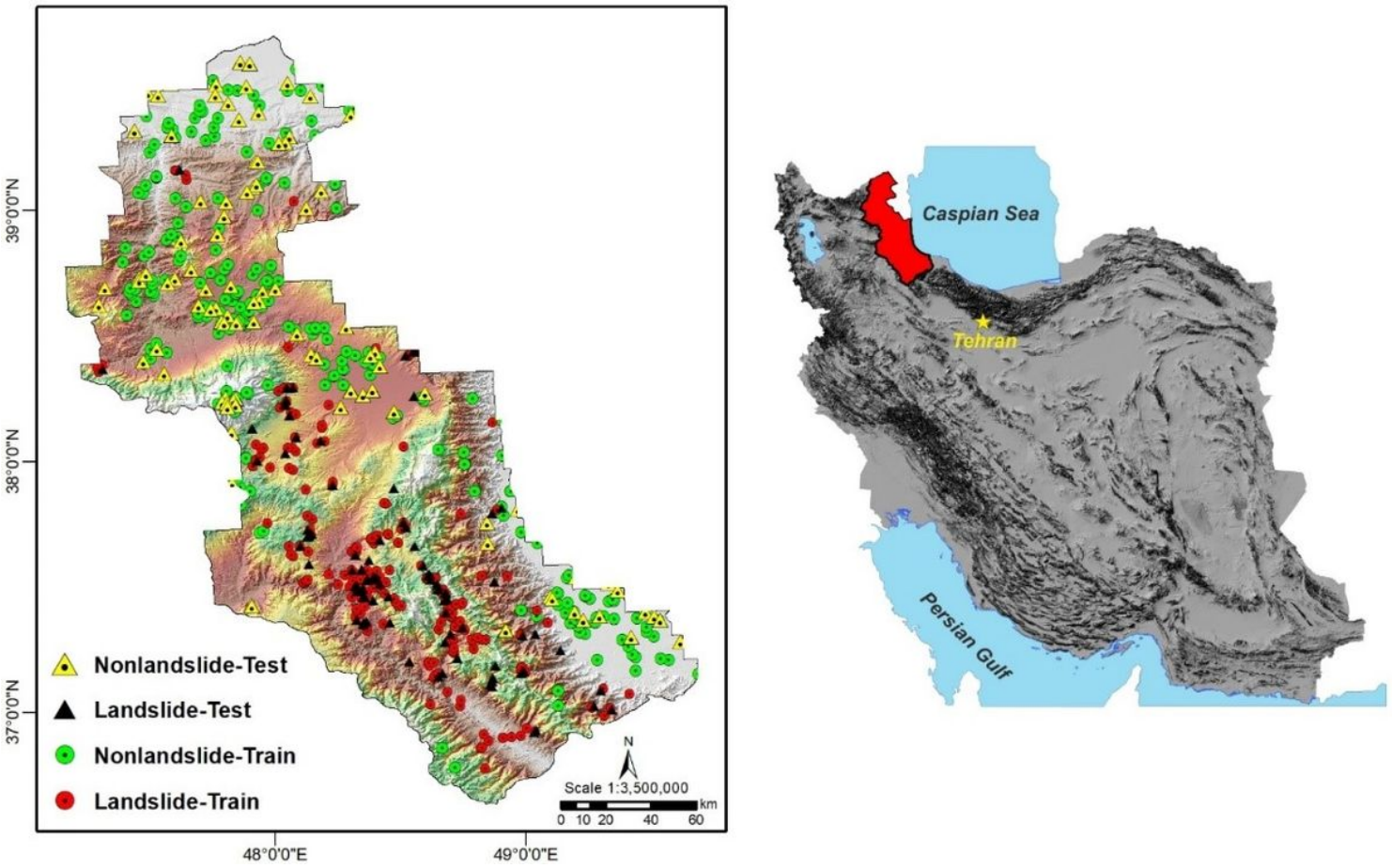

Figure 1

Location of the study area.

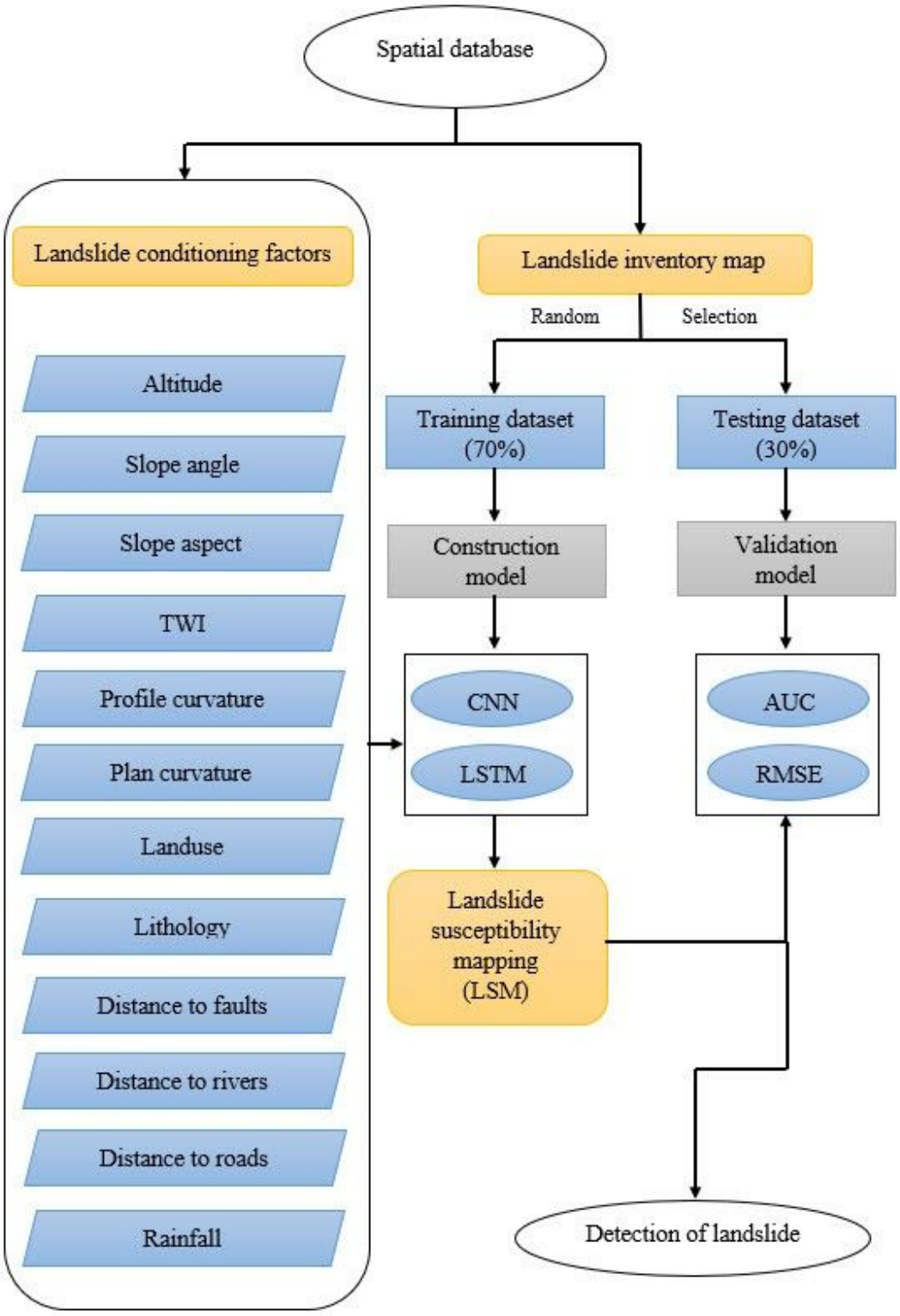

Figure 2

Flowchart of methodology.

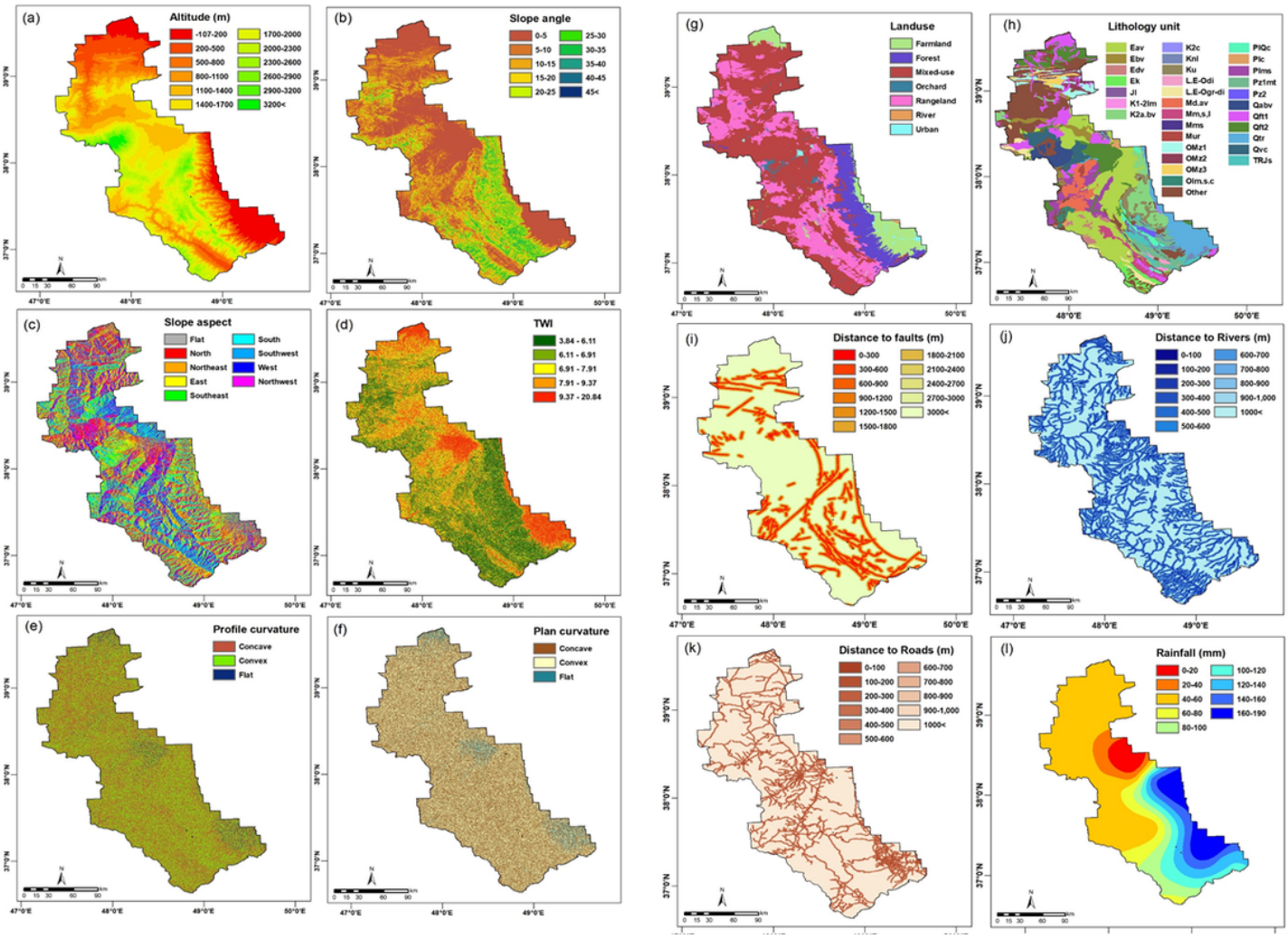

Figure 3

Thematic maps of landslide conditioning factors.

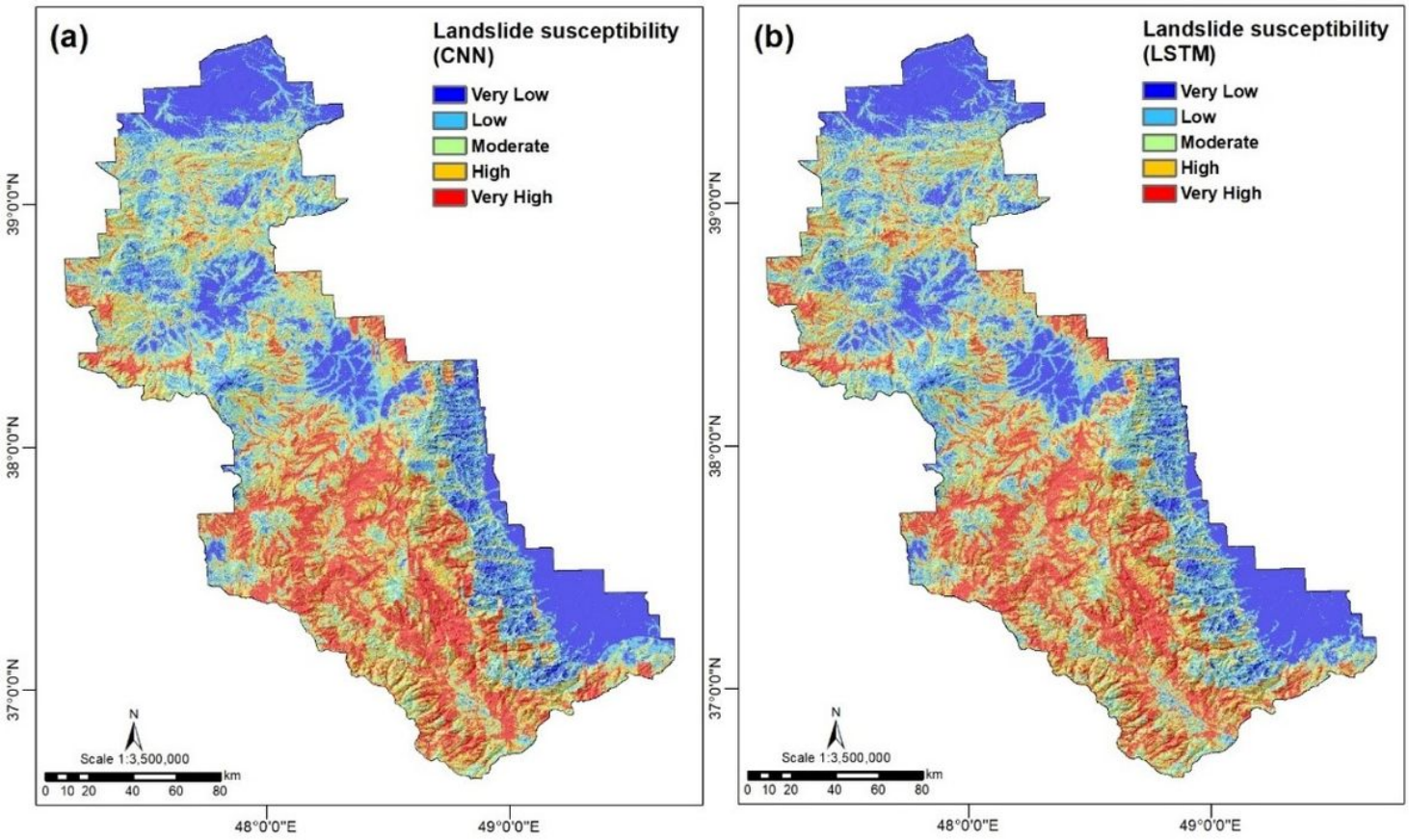

Figure 4

Landslide susceptibility maps by: (a) CNN model, (b) LSTM model.

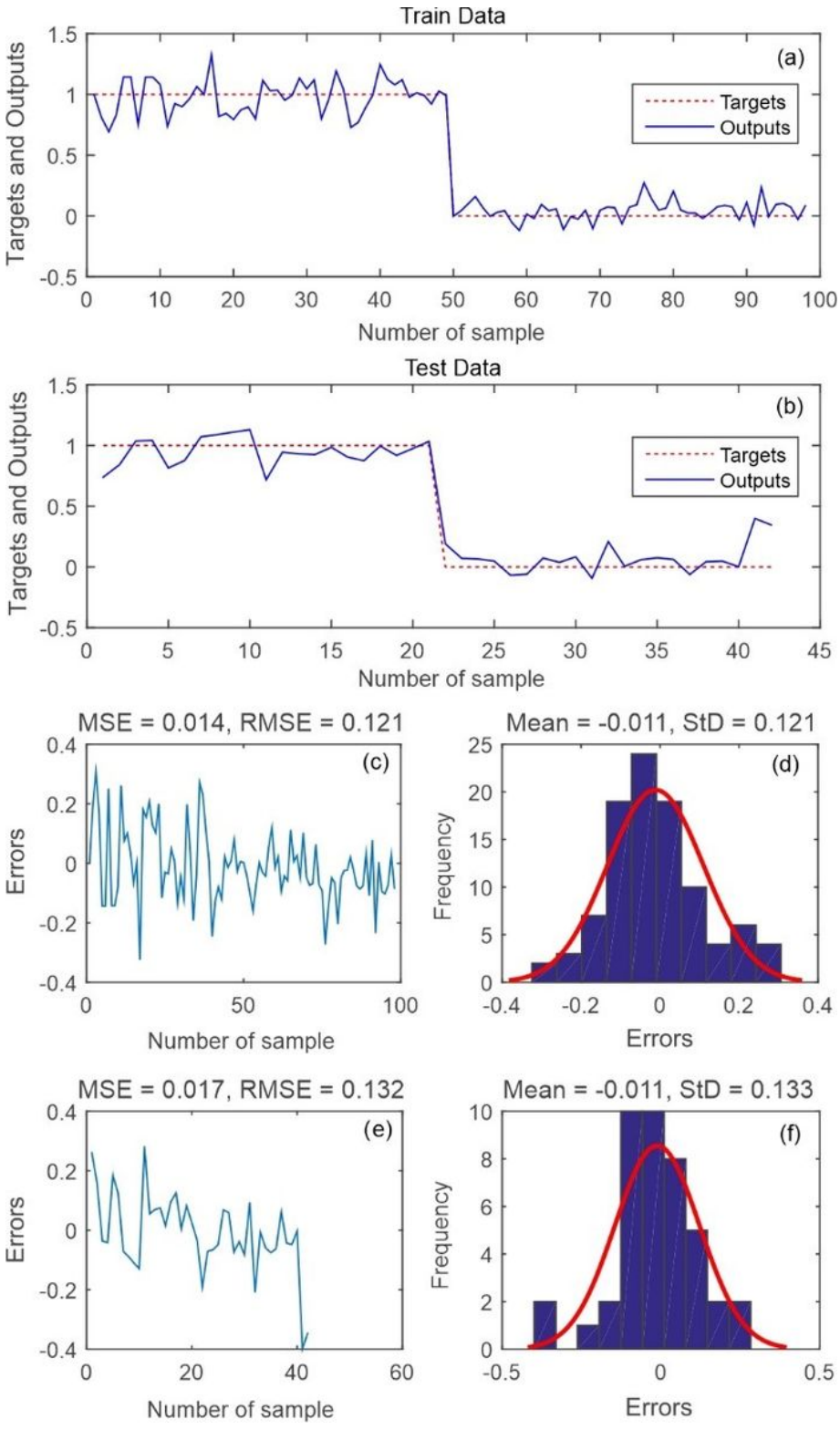

Figure 5

CNN model: (a) targets and outputs for CNN value of training data samples, (b) targets and outputs for CNN value of testing data samples, (c) MSE and RMSE value of training data samples, (d) frequency errors of training data samples, (e) MSE and RMSE value of testing data samples, (f) frequency errors of testing data samples.

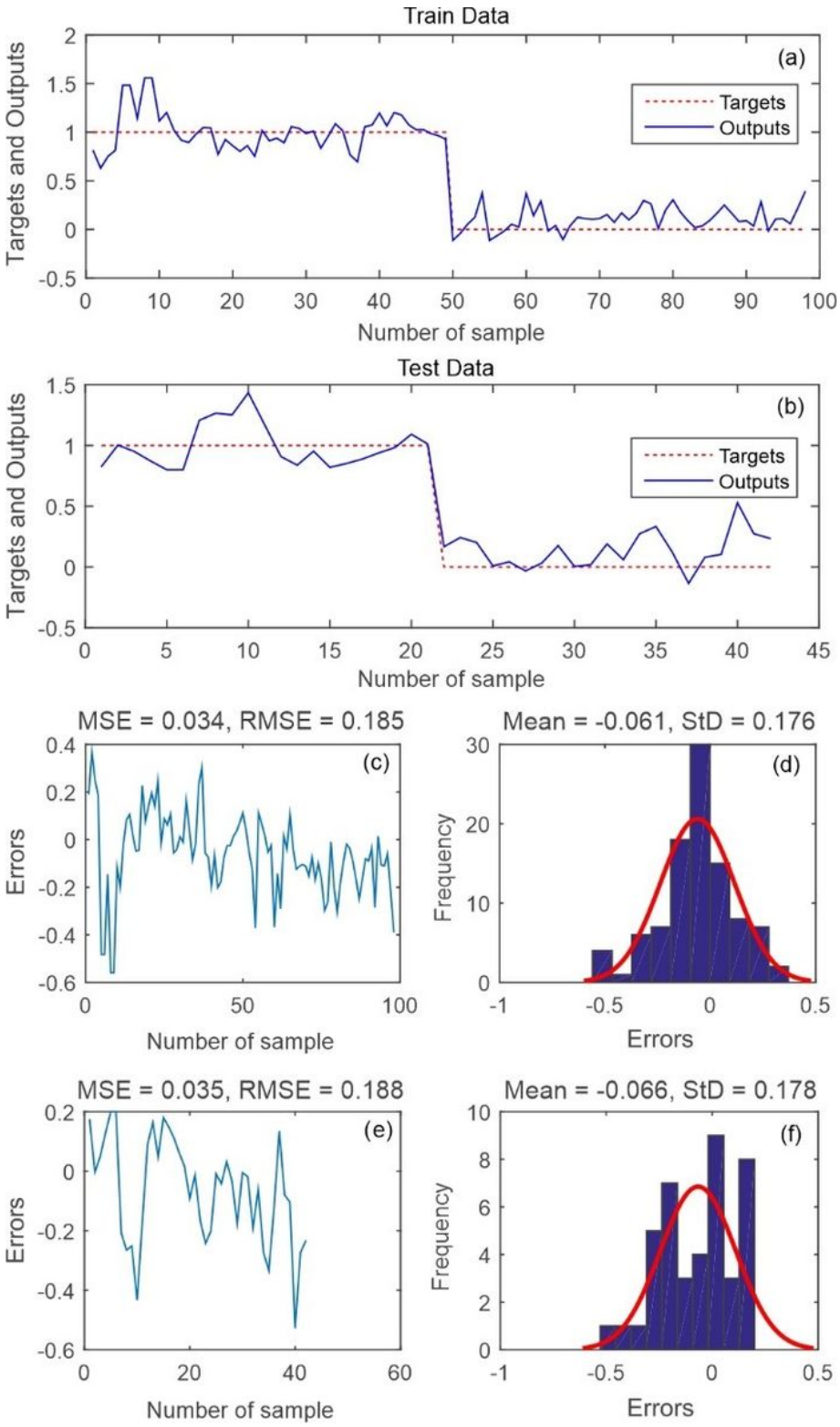

Figure 6

LSTM model: (a) targets and outputs for CNN value of training data samples, (b) targets and outputs for LSTM value of testing data samples, (c) MSE and RMSE value of training data samples, (d) frequency errors of training data samples, (e) MSE and RMSE value of testing data samples, (f) frequency errors of testing data samples.

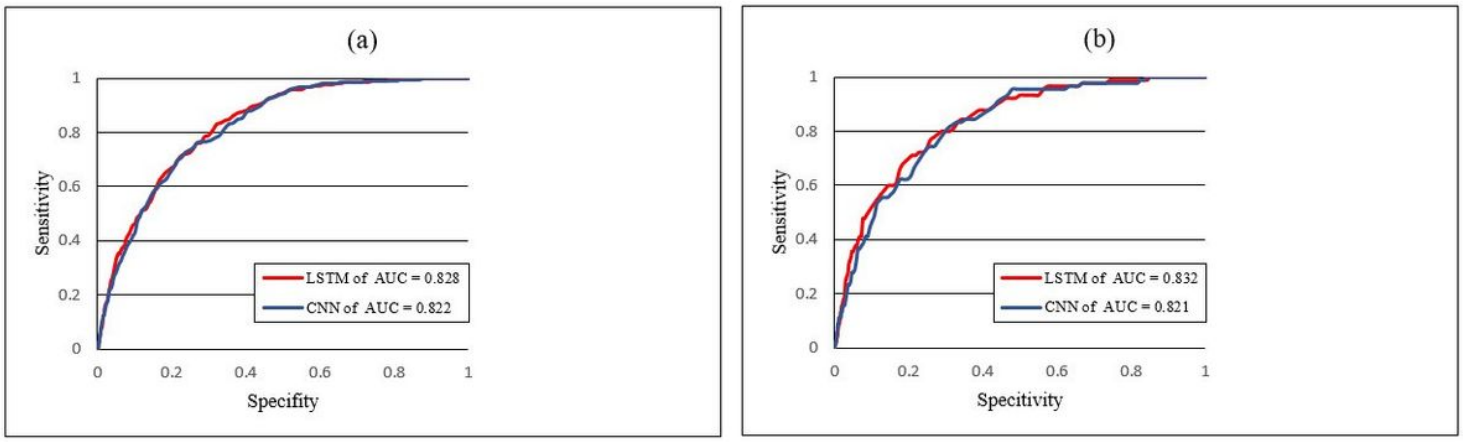

Figure 7

Validation of CNN and LSTM models using AUC values: (a) training datasets, (b) testing datasets.

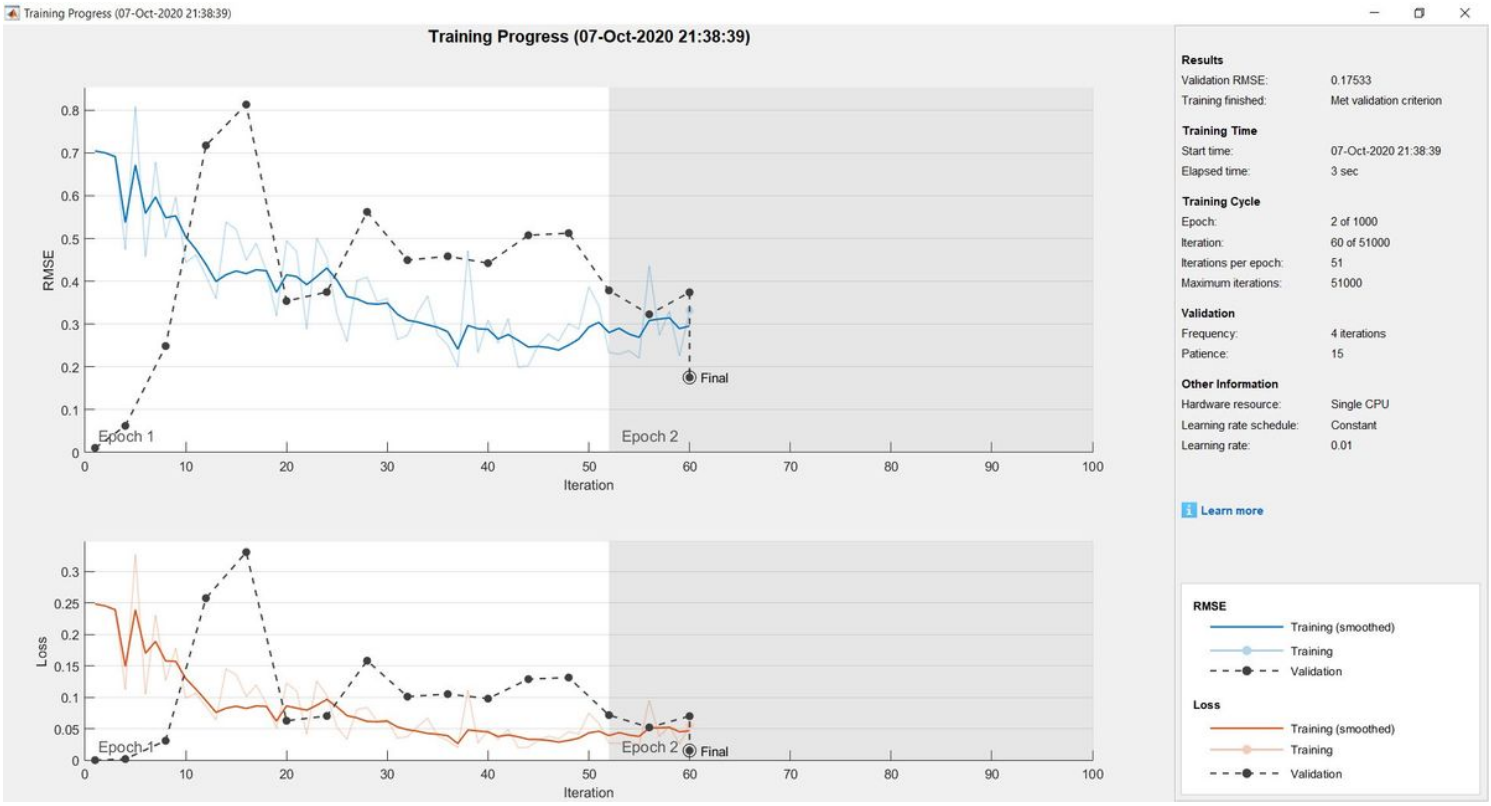

Figure 8

Processing time and convergence speed in CNN model.

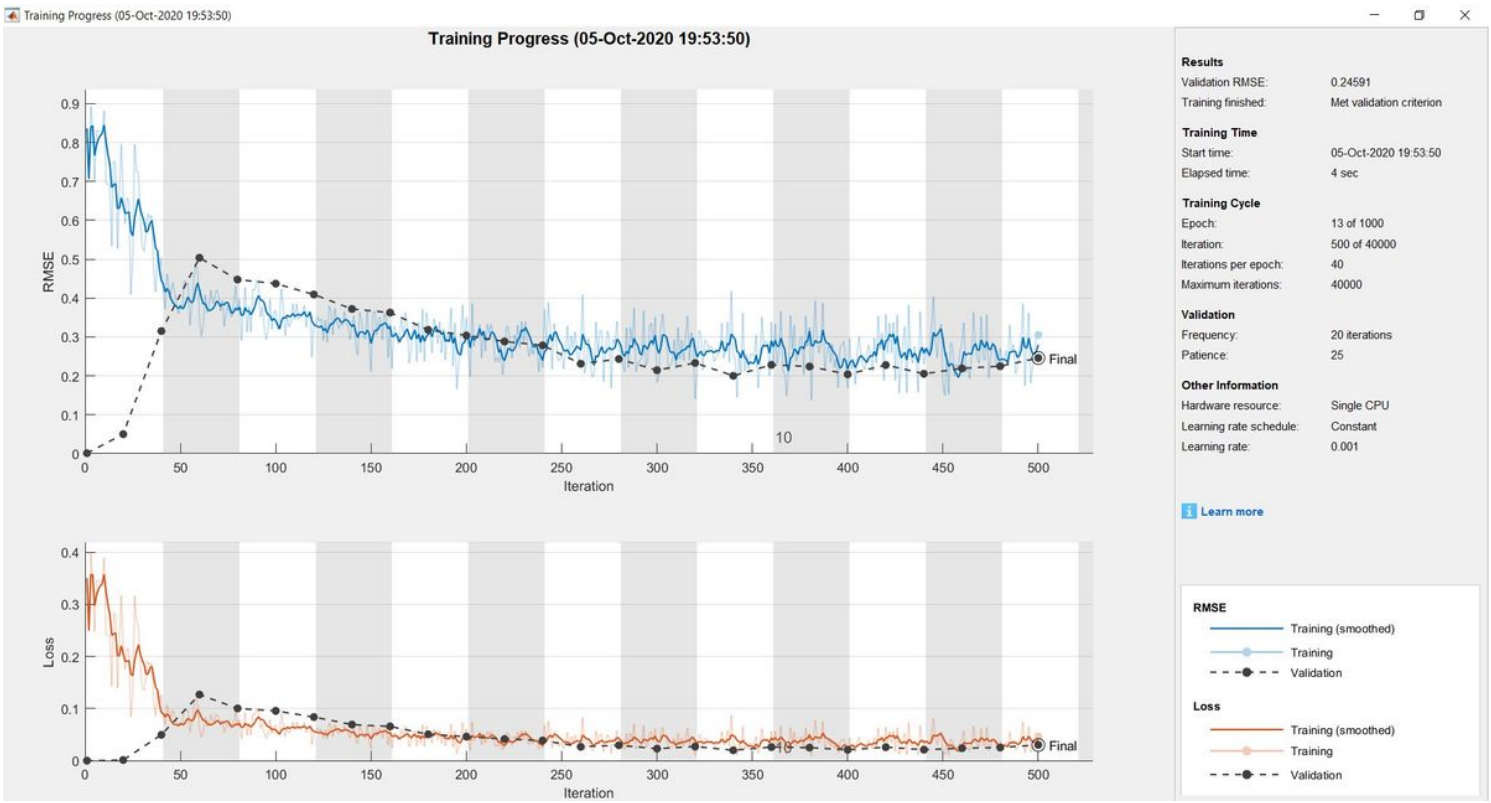

Figure 9

Processing time and convergence speed in LSTM model.