


Article

Optimized Neural Architecture for Automatic Landslide Detection from High-Resolution Airborne Laser Scanning Data

Mustafa Ridha Mezaal ¹, Biswajeet Pradhan ^{1,2,*} , Maher Ibrahim Sameen ¹,
Helmi Zulhaidi Mohd Shafri ¹ and Zainuddin Md Yusoff ¹

¹ Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia; gismustafa87@gmail.com (M.R.M.); maherrsgis@gmail.com (M.I.S.); helmi@eng.upm.edu.my (H.Z.M.S.); zmy@upm.edu.my (Z.M.Y.)

² School of Systems, Management and Leadership, Faculty of Engineering and Information Technology, University of Technology Sydney, Building 11, Level 06, 81 Broadway, Ultimo NSW 2007 (P.O. Box 123), Australia

* Correspondence: Biswajeet24@gmail.com or biswajeet@lycos.com

Academic Editor: Saro Lee

Received: 29 June 2017; Accepted: 13 July 2017; Published: 16 July 2017

Abstract: An accurate inventory map is a prerequisite for the analysis of landslide susceptibility, hazard, and risk. Field survey, optical remote sensing, and synthetic aperture radar techniques are traditional techniques for landslide detection in tropical regions. However, such techniques are time consuming and costly. In addition, the dense vegetation of tropical forests complicates the generation of an accurate landslide inventory map for these regions. Given its ability to penetrate vegetation cover, high-resolution airborne light detection and ranging (LiDAR) has been used to generate accurate landslide maps. This study proposes the use of recurrent neural networks (RNN) and multi-layer perceptron neural networks (MLP-NN) in landscape detection. These efficient neural architectures require little or no prior knowledge compared with traditional classification methods. The proposed methods were tested in the Cameron Highlands, Malaysia. Segmentation parameters and feature selection were respectively optimized using a supervised approach and correlation-based feature selection. The hyper-parameters of network architecture were defined based on a systematic grid search. The accuracies of the RNN and MLP-NN models in the analysis area were 83.33% and 78.38%, respectively. The accuracies of the RNN and MLP-NN models in the test area were 81.11%, and 74.56%, respectively. These results indicated that the proposed models with optimized hyper-parameters produced the most accurate classification results. LiDAR-derived data, orthophotos, and textural features significantly affected the classification results. Therefore, the results indicated that the proposed methods have the potential to produce accurate and appropriate landslide inventory in tropical regions such as Malaysia.

Keywords: landslide detection; LiDAR; recurrent neural networks (RNN); multi-layer perceptron neural networks (MLP-NN); GIS; remote sensing

1. Introduction

Landslides are dangerous geological disasters with catastrophic effects on human lives and properties. Landslides occur with high frequency in mountainous and hilly areas, such as the Cameron Highlands in Malaysia. Landslide incidence is related to a cluster of triggering factors, such as intense rainfall, volcanic eruptions, rapid snowmelt, elevated water levels, and earthquakes. Landslide inventory maps are crucial for measuring the magnitude and analyzing the susceptibility, hazard, and

risk of earthquakes [1,2], as well as for examining distribution patterns and predicting the landscapes affected by landslide [3]. Mapping a landslide inventory in tropical areas is challenging because the dense vegetation cover in these regions obscures underlying landforms [4]. Moreover, the majority of available conventional landslide detection techniques are not rapid and accurate enough for inventory mapping given the rapid vegetation growth in tropical regions. Therefore, inventory mapping requires the use of more rapid and accurate techniques, such as light detection and ranging (LiDAR) [5], which uses active laser transmitters and receivers to acquire elevation data. In addition, LiDAR has the unique capability to penetrate densely vegetated areas [5] and provide detailed information on terrains with high point density. Moreover, it depicts ground surface features and provides useful information on topographical features in areas where landslide locations are obscured by vegetation cover [6,7].

Numerous studies have applied a multiresolution segmentation algorithm for the remote sensing of land features [8]. This algorithm requires the identification of three parameters (i.e., scale, shape, and compactness); the values of these parameters can be determined using the traditional trial-and-error method, which is very time consuming and laborious [5]. Moreover, using the algorithm to delineate the boundary of an object at different scales remains challenging [9]. Thus, optimal parameters for segmentation should be identified via semiautomatic and automatic approaches [10–12]. The automatic selection of segmentation parameters requires the use of the advanced supervised approach presented in [13].

Processing a large number of irrelevant features causes overfitting [14]. By contrast, the best classification results are obtained by selecting the most relevant feature [15]. Landslide identification in a particular area can be improved by selecting the most significant feature [15,16]. As shown in [2], selecting the most significant feature facilitates the differentiation of landslides from non-landslides. Accuracy can be improved by decreasing the number of features, as recommended in [17]. The efficiency of feature selection techniques for landslide detection has been proven in [18–20].

The neural network (NN) is effective in remote sensing applications [21], particularly in solving different image classification problems [22] specified by nonlinear mathematical fitting for function approximation. NN architectures are classified into the recurrent neural network (RNN), back-propagation neural network, probability neural network, and multilayer perceptron neural network (MLP-NN). NN-based classifiers can adapt to different types of data and inputs, and can overcome the issue of mixed pixels by providing fuzzy output and fit with multiple images [23,24]. These classifiers include parallel computation, which is superior to statistical classification approaches because it is non-parametric and does not require the prior knowledge of a distribution model for input data [25]. Moreover, NN-based classifiers can evaluate non-linear relationships between the input data and desired outputs and are distinguished by their fast generalization capability [26]. NN-based classifiers have been successfully in function approximation, prediction, pattern recognition, landslide detection, image classification, automatic control, and landslide susceptibility [27–32]. Authors of [33] found that MLP-NN can be effectively applied in landslide detection using multi-source data. The RNN model can effectively predict landslide displacement [34]. The above neural architecture techniques have not been extensively used for landslide detection using only LiDAR data. This research gap urged us to apply the RNN and MLP-NN models in landslide detection based on very high-resolution LiDAR data. To achieve this objective, we optimized multiresolution segmentation parameters via a supervised approach. Using the correlation-based feature selection (CFS) algorithm, we selected the most significant feature from high-resolution airborne laser scanning data.

2. Study Area

This study was performed in a small section of the Cameron Highlands, which is notorious for its frequent occurrence of landslides. The study area covers an area of 26.7 km². It is located on northern peninsular Malaysia within the zone comprising latitudes 4°26'3'' to 4°26'18'' and longitudes 101°23'48'' to 101°24'4'' (Figure 1). The annual average rainfall and temperature in this region are

approximately 2660 mm and 24/14 °C (daytime/nighttime temperatures), respectively. Approximately 80% of its area is forested with a flat (0°) to hilly (80°) land form.

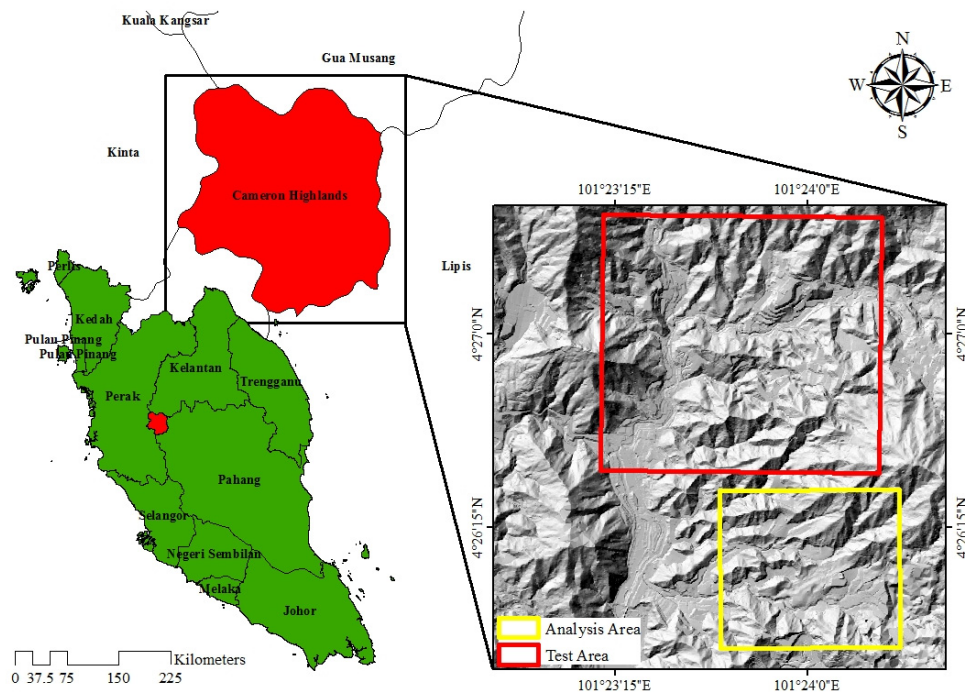


Figure 1. Location of the study area. The red boundary represents the analysis area and the yellow boundary represents the test area.

Two sites were selected to implement and test the proposed models (Figure 1). All the prerequisite considerations were taken in to account during test site selection to avoid missing any land cover classes. To obtain an accurate map of the analysis and test sites, the training sample size was measured via the stratified random sample method.

3. Methodology

3.1. Overall Methodology and Pre-Processing

LiDAR data and landslide inventories were first pre-processed to eliminate noise and outliers. A high-resolution digital elevation model (DEM) at 0.5 m was then derived from LiDAR point clouds to generate other LiDAR-derived products (i.e., slope, aspect, height or (normalized digital surface model (nDSM)), and intensity. LiDAR-derived products and orthophotos were then composited by rectifying their geometric distortions to generate one coordinate system and were finally prepared in geographic information system (GIS) for feature extraction. Suitable parameters (scale, shape, and compactness) at various levels of segmentation were obtained via a supervised approach, i.e., a fuzzy-based segmentation parameter optimizer (FbSP optimizer) [13]. The stratified random method was used to evaluate the training dataset in accordance with the procedure in [35]. The correlation-based selection algorithm (CFS) [36] was used to rank features from the most to least important. RNN and MLR-NN models were applied to detect landslide locations. The results of the models were validated using a 10-fold cross validation method. In addition, the models were evaluated in another part of the study area (i.e., the test site). Slope and aspect layers were overlaid with the results to identify other landslide characteristics (i.e., direction and run off). The study flow is illustrated in Figure 2.

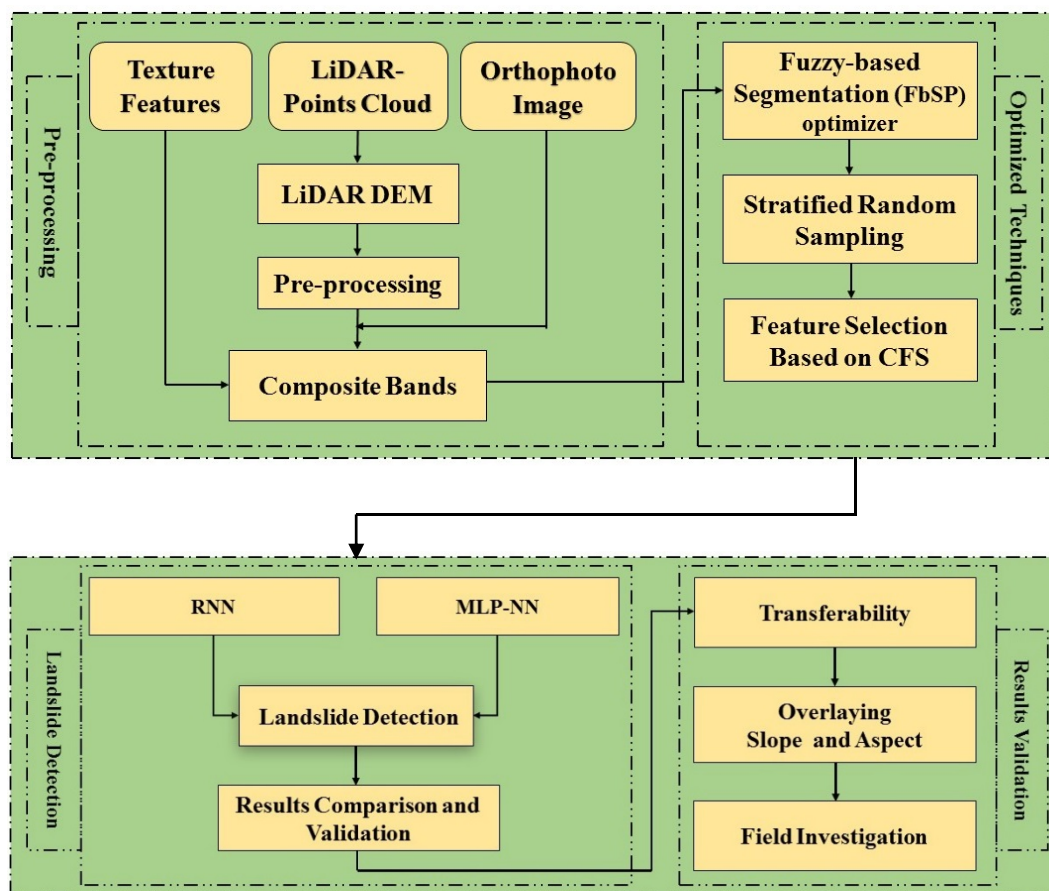


Figure 2. Overview of the proposed method. LiDAR: light detection and ranging; RNN: recurrent neural networks; MLP-NN: multi-layer perceptron neural networks; CFS: correlation-based feature selection; DEM: digital elevation model.

3.2. Landslide Inventory

The landslide inventory; produced previously by Pradhan and Lee, [39] was used to develop the proposed detection method and the total number of landslides is 21 in the study area covering 3781 m² (Figure 3).

3.3. Data

LiDAR point-cloud data were collected on 15 January 2015 at a point density of 8 points/m² and frequency pulse rate of 25,000 Hz. The absolute accuracy of the data (root-mean square errors) was restricted to 0.15 m and 0.3 m in the vertical and horizontal axes, respectively. Orthophotos were obtained using the same acquisition system that relied on the abovementioned cloud data. A DEM was derived from LiDAR point clouds with a spatial resolution of 0.5 m after non-ground points were removed using inverse distance weighting with a spatial reference of GDM2000/Peninsula RSO. Subsequently, LiDAR-based DEM was used to generate derived layers to facilitate the identification and characterization of landslide locations [37].

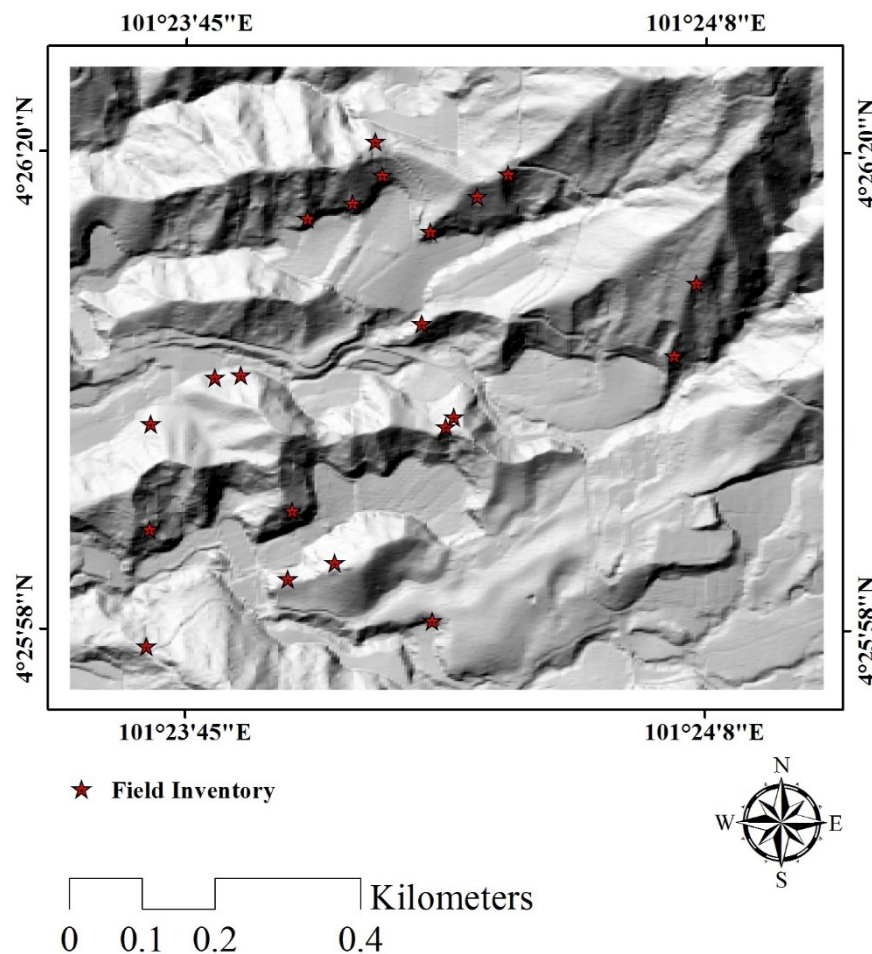


Figure 3. Shows the locations of landslide in the study area.

According to the authors of [38], slope directly and highly affects landslide phenomenology. The authors of [39] also inferred that slope is the principal factor that affects landslide occurrence. The author of [40] indicated that a hillshade map provides a good image of terrain movements, thus facilitating the development of landslide maps. Texture and geometric features are crucial for improving the classification accuracy of landslide mapping [14]. Landslide intensity and texture derived from LiDAR data are affected by the accuracy of landslide detection [9]. The accuracy and capacity of DEM to represent surface features are determined by terrain morphology, sampling density, and the interpolation algorithm [41]. In this study, hillshade, height (nDSM), slope, and aspect were generated from LiDAR-based DEM. As shown in Figure 4, landslide locations were detected using visible bands and texture features.

3.4. Image Segmentation

The sizes and shapes of image objects [42] are determined via image segmentation, the preliminary step in object-based classification. Optimal segmentation parameters depend on the environment under analysis, the selected application, and the underlying input data [8]. Previous studies have used the multiresolution segmentation algorithm with eCognition software for image segmentation [8,9]. Three parameters (scale, shape, and compactness) are defined in this algorithm. According to [5], these parameters can be obtained via the traditional trial-and-error method, which is time consuming and laborious. Therefore, the fuzzy logic supervised approach presented by [13] was adopted in this study.

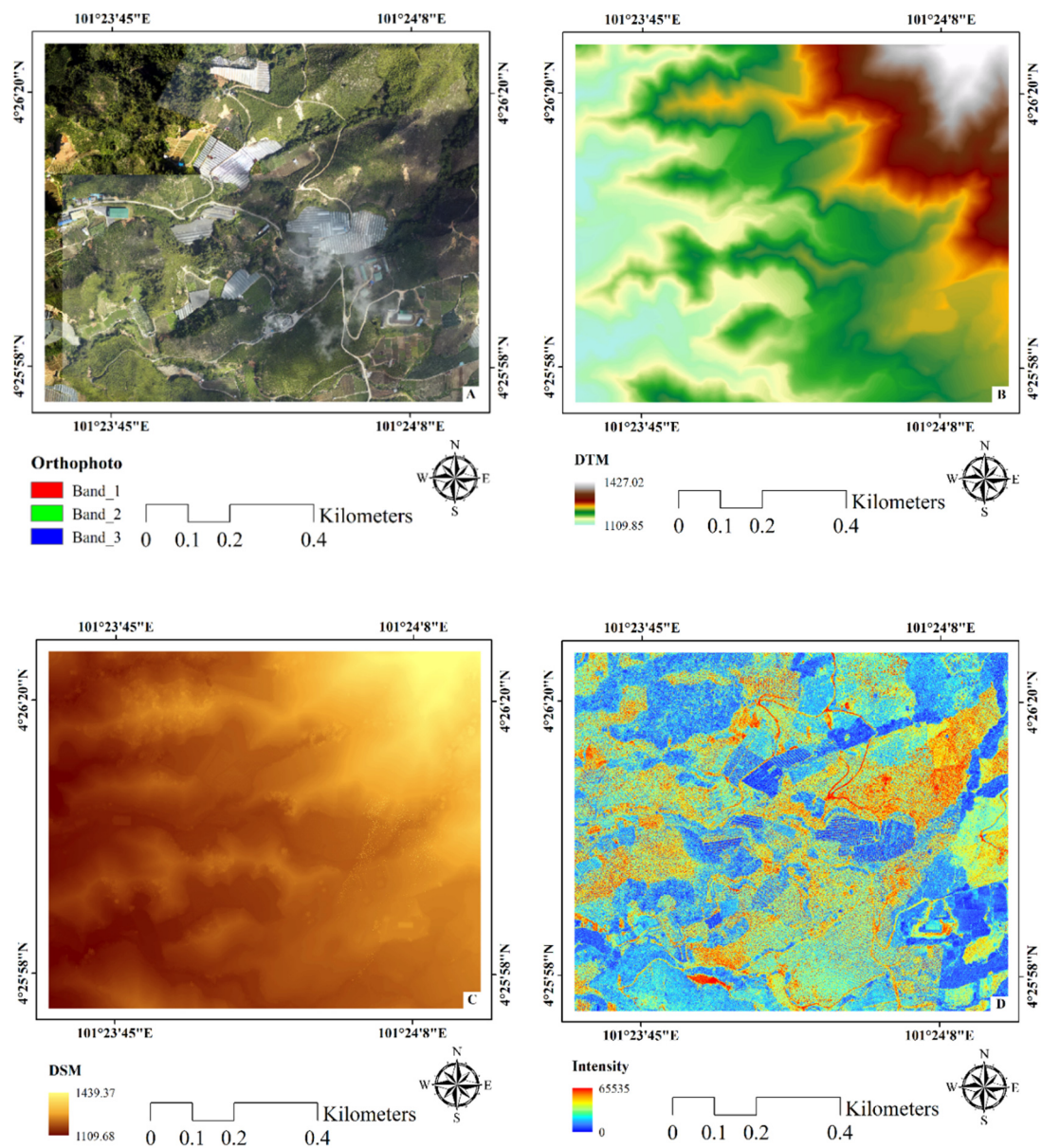


Figure 4. Cont.

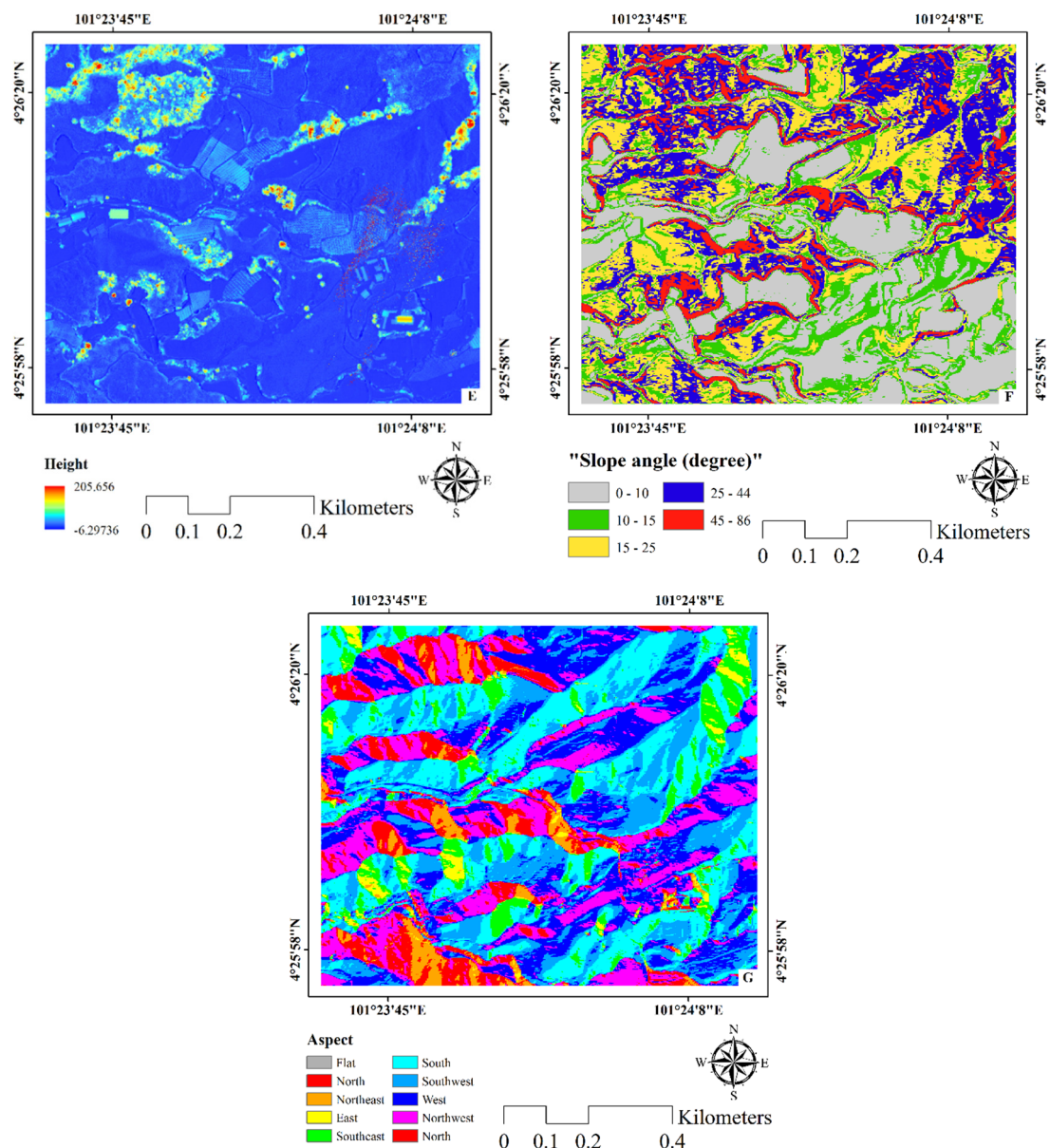


Figure 4. LiDAR-derived data; (A) Orthophotos; (B) digital terrain model (DTM); (C) digital surface model (DSM); (D) Intensity; (E) Height; (F) Slope; and (G) Aspect.

3.5. Training Sets

The authors of [35] suggested the use of stratified random sampling method to obtain an adequately sized training dataset for every class without any bias during sample selection. Accordingly, the present study adopted stratified random sampling to evaluate training samples and achieve high performance without strong bias. Four classes with different numbers of objects were set as shown in Table 1.

Stratified random sampling is a prerequisite to obtain prior knowledge of the two sites considered for landslide inventory. Hence, segmentation parameters were first optimized. Then, the landslide inventory was overlapped with the segmented layer for object labeling. ArcGIS 10.3 was used to construct sample sets automatically at each optimal scale. Subsequently, stratified random sampling was applied on the labeled objects. This process was performed 20 times at each optimal scale.

Table 1. Number of selected training objects in four classes.

Class Name	Number of the Object for Each Class
Landslide	52
Cut slope	67
Bare soil	80
Vegetation	150

3.6. Correlation-Based Feature Selection

The authors of [15] reported that the selection of only the most relevant features improves the quality of landslide identification and classification. Working with large numbers of features causes numerous problems. As reported in [43] and [14], some of these problems include the slow run time of algorithms due to the consideration of numerous resources, low accuracy when the number of features exceed the number of observation features, and overfitting when irrelevant features are used as inputs. Therefore, the most significant features should be selected to enhance the accuracy of feature extraction. In this study, relevant features were extracted using the CFS algorithm with Weka 3.7 software. Furthermore, the CFS algorithm was applied to all LiDAR-derived data, visible bands, and textural features, and was used to determine the feature subsets required to develop models for landslide identification. The CFS algorithm comprises two basic steps: the ranking of initial features and the elimination of the least important features through an iterative process.

3.7. MLP-NN

NNs are a family of biological learning models in machine learning. The NN model comprises interconnected neurons or nodes, which are structured into layers with random or full interconnections among successive layers [44]. The NN model comprises input, hidden, and output layers that are responsible for receiving, processing, and presenting results, respectively [44]. Each layer contains nodes connected by numeric weights and output signals. The weights are the functions of the sum of the inputs to the node modified by a simple activation function [45]. The possibility of learning is the most important feature that attracts researchers to use NNs.

Back-propagation, which was first proposed by Paul Werbos in 1974 and independently rediscovered by Rumelhart and Parker, is the most common learning algorithm used in NN. It aims to minimize the error function via the iterative approach as shown in Equation (1). NNs have been successfully used in remote sensing applications. However, this model has some limitations, specifically, high computational complexity and overlearning [46,47].

$$E = \frac{1}{2} \sum_{i=1}^L (d_i - o_i)^2 \quad (1)$$

where d_i and o_i represent the desired output and the current response of node " i " in the output layer, respectively. " L " is the number of nodes in the output layer. Corrections to weight parameters were calculated and effected with the previous values in the iterative method, as demonstrated in Equation (2):

$$\begin{cases} \Delta w_{i,j} = -\mu \frac{\partial E}{\partial w_{i,j}} \\ \Delta w_{i,j}(t+1) = \Delta w_{i,j} + \alpha \Delta w_{i,j}(t) \end{cases} \quad (2)$$

where delta rule $\Delta w_{i,j}$ is the weight parameter between nodes i and j ; μ is a positive constant that controls the amount of adjustment and is referred to as learning rate; α is the momentum factor, which takes a value between 0 and 1; and t is the iteration number. α is referred to as the stabilizing factor because it smoothens quick changes between weights [48].

3.8. RNN

RNNs are designed to model sequences in NNs with feedback connections. They are very powerful in computational analysis and are biologically more reliable than other NN techniques given their lack of internal states. The memory of past activations in RNN is very effective with feedback connections, making them suitable for learning the temporal dynamics of sequential data. RNN is very powerful when used to map input and output sequences because it uses contextual information. However, traditional RNNs face the challenge of exploding or vanishing gradients. Hochreiter and Schmidhuber [49] proposed long short-term memory (LSTM) to tackle this issue.

Hidden units in LSTM are replaced with memory blocks that contain three multiplicative units (input, output, forget gates) and self-connected memory cells to allow for reading, writing, and resetting through a memory block and behavioral control. A single LSTM unit is shown in Figure 5. c_t is the sum of inputs at time step t and its previous time step activations. LSTM updates time step i given inputs x_t , h_{t-1} , and c_{t-1} as reported in [50].

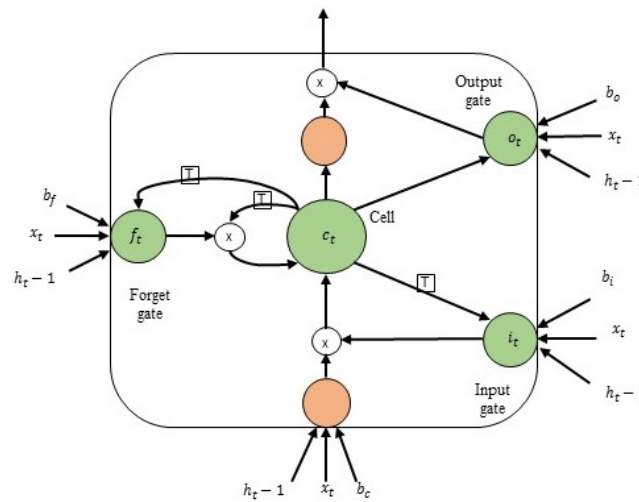


Figure 5. Structure of a memory cell in long short-term memory (LSTM)-RNN.

Input gates:

$$i_t = \sigma(W_{xi} \cdot x_t + W_{hi} \cdot h_{t-1} + W_{ci} \cdot c_{t-1} + b_i) \quad (3)$$

Forget gates:

$$f_t = \sigma(W_{xf} \cdot x_t + W_{hf} \cdot h_{t-1} + W_{cf} \cdot c_{t-1} + b_f) \quad (4)$$

Cell units:

$$c_t = i_t \cdot \tanh(W_{xc} \cdot x_t + W_{hc} \cdot h_{t-1} + b_c + b_f \cdot c_{t-1}) \quad (5)$$

Output gates:

$$o_t = \sigma(W_{xo} \cdot x_t + W_{ho} \cdot h_{t-1} + W_{co} \cdot c_{t-1} + b_o) \quad (6)$$

The hidden activation (output of the cell) is also given by a product of the two terms:

$$c_t = o_t \cdot \tanh(c_t) \quad (7)$$

where σ and \tanh are an element-wise non-linearity, such as a sigmoid function and hyperbolic tangent function, respectively; W is the weight matrix; x_t refers to input at time step t ; h_{t-1} represents the hidden state vector of the previous time step; and b_c denotes the input bias vector. The memory cell unit c_t is a sum of two terms: the previous memory cell unit c_{t-1} , which is modulated by f_t and c_t , a function of the current input, and previous hidden state, modulated by the input gate i_t due to i_t

and f_t being sigmoidal. Their values range within $[0, 1]$, and i_t and f_t can be considered as knobs that the LSTM learns to selectively forget its previous memory or consider its current input, whilst o_t is an output gate that learns how much of the memory cell to transfer to the hidden layers.

3.9. Neural Network Models

3.9.1. MLP-NN

This study proposed the network architectures RNN and MLP-NN. Figure 6 depicts the MLP-NN model architecture, which has two hidden layers of 50 hidden units. Ten features were taken as inputs in the model to detect different types of objects, such as landslide, cut slope, bare soil, and vegetation. The MLP-NN model was trained through a back-propagation technique with the Adam optimizer and a batch size of 64. The hyper-parameters used in this NN were carefully selected through grid search and a 10-fold cross validation process.

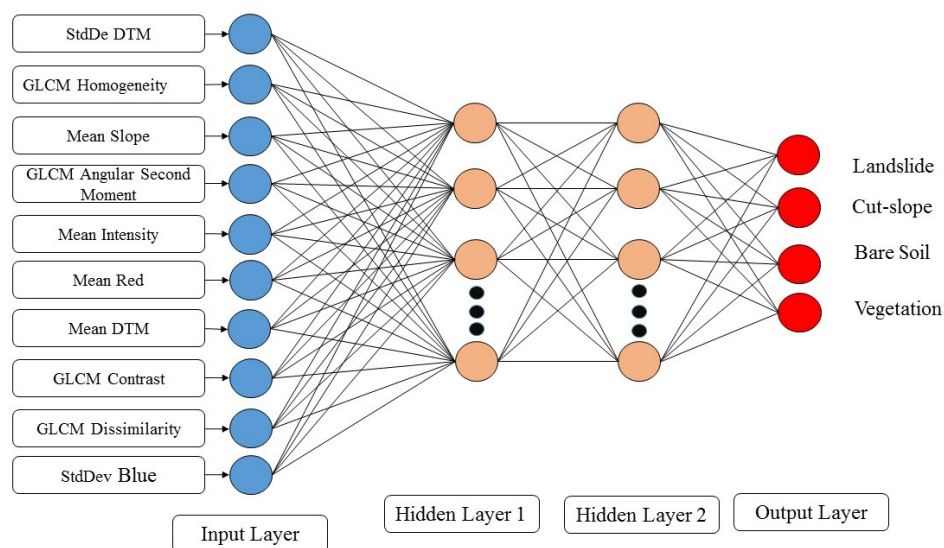


Figure 6. Architecture of the MLP-NN model; GLCM: gray level co-occurrence matrix, StdDev: standard deviation.

3.9.2. RNN

RNN is a sequence problem considered as the addition of loops to architecture. For example, in any layer under consideration, signals can be passed to each neuron and are subsequently forwarded to the next layer. The network output can be input to the network in the next input feature, and so on, as shown in Figure 7. In this study, RNN received 10 features as inputs to differentiate landslides from other objects (cut slope, bare soil, and vegetation). RNN consisted of an LSTM layer with 50 hidden units, two fully connected layers, a dropout layer, and a softmax layer. The back-propagation technique was used in trained the RNN model with Adam optimizer and a batch size of 128.

To avoid overfitting, a dropout layer was used in the RNN model and the NN learned weights from the training dataset. However, overfitting may occur when new data are inputted. The dropout layer randomly set some selected activations to zero, thus alleviating overfitting. The selected activations were used only during training and not during testing. The parameter was controlled by the number of activations that the dropout layer referred to as keep probability.

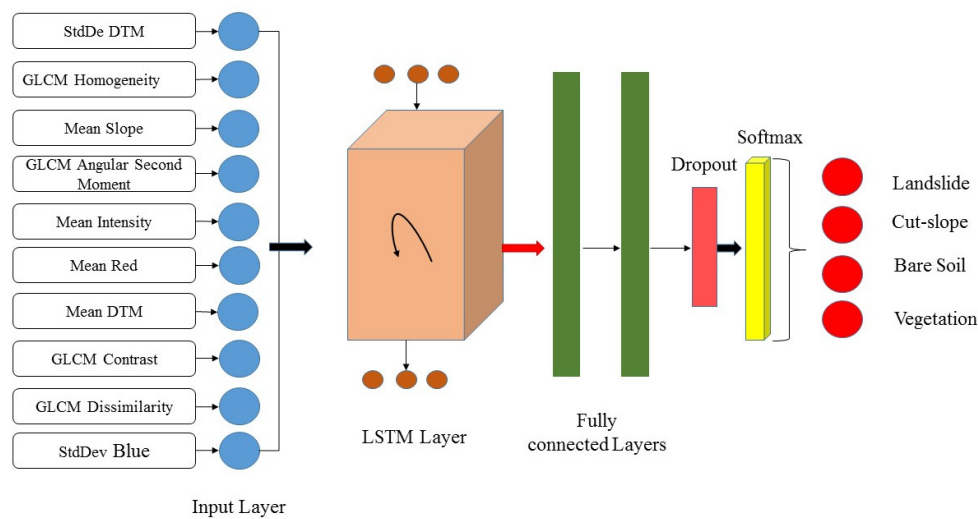


Figure 7. Architecture of the RNN model.

3.9.3. Optimization of Model Hyper-Parameters

The hyper-parameters of the RNN and MLP-NN models were optimized via a systematic grid search in scikit-learn [51] for 100 epochs. Despite its high computational cost, the systematic grid search provides better results because it systematically tunes the hyper-parameter values. Parameter combinations were selected for the models. The models were evaluated using a 10-fold cross-validation method. Among the evaluated parameters, the model with the highest validation accuracy was selected. Table 2 presents the most optimized parameters obtained for the models.

Table 2. Optimized model hyper-parameters; RNN: recurrent neural networks, MLP-NN: multi-layer perceptron neural networks.

Optimized Parameter	Suitable Value	Description
Minibatch size	126 (RNN) 64 (MLP-NN)	Number of training cases over which the Adam update is computed.
Loss function	categorical cross-entropy	The objective function or optimization score function is also called as multiclass legless, which is appropriate for categorical targets.
Optimizer	Adam	Adaptive moment estimation
dropout rates	0.6	Dropping out units (hidden and visible)

4. Results and Discussion

4.1. Supervised Approach for Optimizing Segmentation

The supervised approach was employed to optimize the parameters (i.e., scale, shape, and compactness) of the multiresolution segmentation algorithm for landslide identification and for differentiation from non-landslides (bare soil, cut slope, and vegetation). The optimized parameters rapidly increased the accuracy of classification to the optimum level by delineating the segmentation boundaries of the landslide. The application of optimized segmentation parameters allowed for the spatial and textural identification of features (landslide and non-slides). In our proposed method, accurate segmentation results should be first obtained prior to performing subsequent steps.

The optimal parameters of the multiresolution segmentation algorithm were obtained. The selected values for the three parameters are shown in Table 3. The initial segmentation parameters set in the supervised approach were 50, 0.1, and 0.1 for scale, shape, and compactness, respectively. After 100 iterations with these initial values, the optimal values obtained for scale, shape, and compactness

were 75.52, 0.4, and 0.5, respectively in the analysis area. Meanwhile, the test area values were 100, 0.45 and 0.74, respectively. Figure 8a,b show the initial and optimal segmentation processes. The results of optimized segmentation accurately delineated landslide objects in the analysis and test areas.

Table 3. Multi-resolution segmentation parameters.

Initial Parameters				Optimal Parameters		
Number	Scale	Shape	Compactness	Scale	Shape	Compactness
1	50	0.1	0.1	75.52	0.4	0.5
2	80	0.1	0.1	100	0.45	0.74

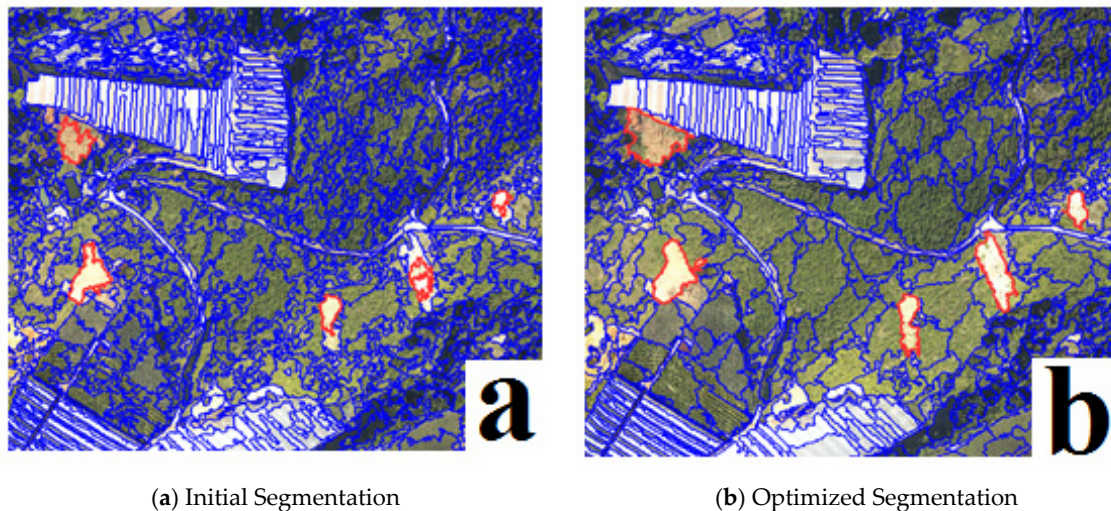


Figure 8. Parameter optimization of the multiresolution segmentation algorithm: (a) initial segmentation and (b) optimized segmentation.

4.2. Relevant Feature Subset Based on a CFS Algorithm

In this study, the feature input consisted of 39 items of LiDAR-derived data (i.e., slope, height, and intensity), texture features (i.e., GLCM StdDev and GLCM homogeneity), and visible band. The optimal combination of features was selected via ten experiments using a CFS algorithm. Selection began from (1, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 39) of the features. The most relevant feature subsets were obtained after 100 iterations in every experiment; this result is in line with the procedure proposed by Sameen et al. [52]. High classification accuracy was achieved when 10 of the features were applied, indicating that LiDAR-derived data, visible bands, and textural features were more effective in detecting the landslide location. Table 3 shows the most significant results of feature selection based on the CFS algorithm.

4.3. Results of Landslide Detection

Classification techniques affect the quality of the classification maps. Many classification algorithms have been established for each category, and each has its merits and demerits. In the present work, the RNN and MLP-NN models with optimized parameters were used for landslide detection with good accuracy. Figure 9 shows the classification results of the RNN and MLP-NN models in the analysis area. The qualitative assessment of the RNN model yielded high-quality results, as shown in Figure 9A. Well-defined landslide boundaries were detected and correctly differentiated from other objects (cut-slope, bare soil and vegetation). On the other hand, the qualitative assessment of MLP-NN produced low-quality results, as shown in Figure 9B.

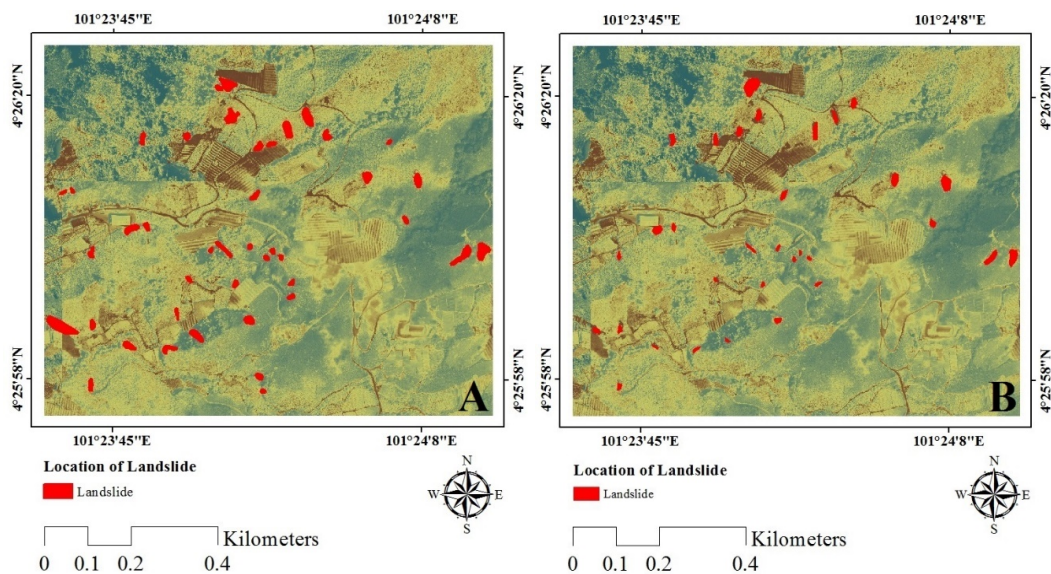


Figure 9. Results of the qualitative assessment of (A) RNN and (B) MLP-NN for the analysis area.

The proposed models were evaluated using another LiDAR dataset (test site) from the Cameron Highlands. All features (all existing objects) of the test area were carefully considered. Segmentation parameters were optimized using the FbSP optimizer. A 10-fold cross-validation approach introduced by Bartels et al. [53] was used to resolve this issue with high accuracy. Environmental conditions and differences in landslide characteristics resulted in misclassification [9]. Differences in the sensors used, illumination conditions, and the spatial resolutions of images are some of the challenges faced by the proposed NN models [54]. The results of qualitative assessment indicated that the proposed NNs with optimized techniques correctly detected landslide locations in the test site, as shown in Figure 10. The qualitative assessment of the RNN model yielded high-quality results, as shown in Figure 10A,B. On the other hand, the qualitative assessment of the MLP-NN model produced low-quality results, as shown in Figure 10.

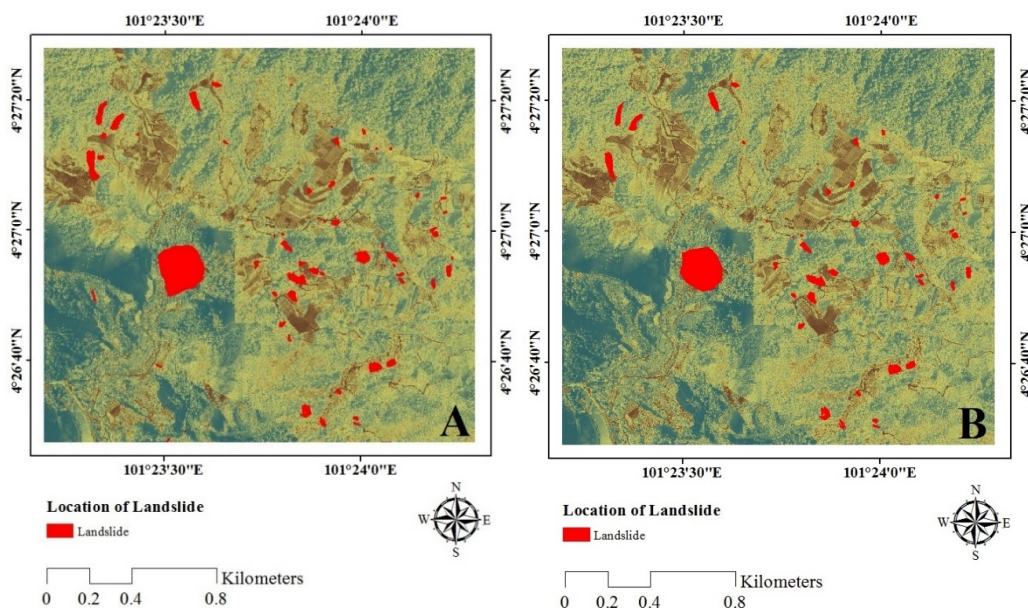


Figure 10. Results of the qualitative assessment of (A) RNN and (B) MLP-NN for the test area.

It is crucial to take the required measures to avoid the issue of the landslide separation from the bare land. The morphology characteristics of the landslide map is different from other types of land cover. For example, the shape, slope and other characteristics (i.e., dip direction, width and length) of the surface terrain may be changed after landslide occurs. Therefore, by using relevant features derived from very high resolution LiDAR, data such as texture and geometric features can be used to separate between landslides and bare land. In addition, applying different optimization techniques helped us to improve the classification accuracy in landslide detection over other landcover classes, such as bare land, man-made, etc., as described previously by Pradhan and Mezaal [9]. Their results demonstrated that using optimized techniques with very high resolution LiDAR data (0.5) enabled them to separate landslide and other types of land cover. In addition, the most relevant features in Table 4 were optimized during this study. Furthermore, authors of [16] suggested that using the object feature from LiDAR data is a suitable solution for landslide identification.

Table 4. Correlation-based feature selection (CFS) results for the most relevant feature subset at a scale of 75.52; StdDe: Standard deviation, DTM: Digital terrain model, GLCM: Gray level co-occurrence matrix.

Feature	Iteration	Rank
StdDe DTM	20	1
GLCM homogeneity	18	2
Mean slope	20	3
GLCM angular second moment	20	4
Mean intensity	17	5
Mean red	20	6
Mean DTM	20	7
GLCM contrast	18	8
GLCM dissimilarity	15	9
StdDev blue	20	10

The landslide detection results showed that the proposed model is robust. Optimizing the segmentation parameters, namely, scale, shape, and compactness, using the fuzzy logic supervised approach resulted in the effective differentiation of landslide from non-landslide (bare soil, cut slope and vegetation) objects. Creating accurate objects through the optimized segmentation process allowed the use of spatial, orthophoto, and textural features for feature detection. Landslides should be differentiated from non-landslides based on the accurate segmentation of spatial and textural features. The selection of relevant features in landslide detection relies on the experience of the analysts. Thus, a feature selection method is crucial for accurate and reliable landslide detection. The optimal features selected via the CFS method simplified landslide detection by the NN model. Computation time and reliance on the expert knowledge of the analyst were reduced. Moreover, the optimized parameters of the NN models improved the performance of the models, reduced the complexity of the models, and decreased overfitting in the training sample.

4.4. Performance of the MLP-NN and RNN

The models were implemented in Python using the open source TensorFlow deep learning framework developed by Google [26]. Meanwhile, the accuracy of the proposed NN models was tested using a 10-fold cross-validation method. The results are presented in Table 5. The best accuracy of 83.33% in the analysis area was achieved by the RNN model. The MLR-NN model achieved an accuracy of 78.38% in the same area. Furthermore, the RNN model outperformed the MLR-NN model in terms of stability of accuracy across different folds of the tested dataset. In the test area, accuracies of 81.11% and 74.56% were achieved with the RNN and MLP-NN models, respectively. These results indicated that the RNN model has better accuracy than the MLR-NN model in the analysis and test areas and indicated the high stability of the RNN model in detecting the spatial distribution of landslides.

Table 5. Cross-validation accuracy results of the proposed models.

Neural Network Model	Analysis Area	Test Area
RNN model	83.33%	81.11%
MLP-NN model	78.38%	74.56%

However, producing neural network models such as LSTM and convolution layers with fully connected networks is a crucial task. Complex networks with more hidden units and many modules often tend to have a better overfit due to the detection ability with respect to any possible interaction so the model becomes too specific to the training dataset. Thus, optimizing the network structures is very crucial for avoiding over-fitting. This study indicated that the hyperparameters in both models have a significant effect on their results. For example, the effect of learning rate varied from 0.1, 0.01, 0.05 and 0.001 in landslide detection. The highest accuracy was obtained when a learning rate reached 0.001. In contrast, increasing the learning rate to 0.1 significantly reduced the accuracy in both models. The batch size parameter in both models had significant effects on the result accuracy. The results of MLR-NN and RNN models showed high accuracies with batch sizes of 64 and 128, respectively. This indicates that RNN model achieved high accuracy with the increase of the batch size, whereas the accuracy of MLP-NN model was decreased.

Furthermore, it was revealed that the dropout rate had a substantial influence on the results of the RNN model. The RNN model showed higher accuracy when the dropout rate reached 0.6. The results of the RNN model indicated that the accuracy increased when the dropout rate parameter was increased.

The results of two models (Table 5), show that the accuracies of the RNN model outperformed the MLP-NN model in both study areas. This is due to several reasons, for example the fact that the MLP-NN model uses only local contexts and therefore it does not capture the temporal and spatial correlation in the dataset. Meanwhile, the hidden units of the RNN model contain historical information from the previous step. This indicates RNN model has more information about the data structure and accurate as compared to the MLP-NN model.

4.5. Sensitivity Analysis

The optimization of network architecture is necessary and should be considered over the use of standard parameters [28] because network architecture models are principally influenced by the analytical task and data type. Data could differ in size, relationships between independent and dependent variables, and complexity. Therefore, the neural architecture of the RNN and MLP-NN networks was enhanced using a grid search implemented in SciPy-python. The combinations of 10 parameters that can best identify landslide locations in densely vegetated areas were optimized.

The Adam optimizer is the most suitable algorithm for the optimization of the two NN models. Using the Adam optimizer with default parameters (learning rate $\mu = 0.001$, beta $\beta_1 = 0.9$, epsilon $\epsilon = 1e-08$ and weight decay = 0.0) yielded an accuracy of 0.77 and 0.825 for MLP-NN and RNN models, respectively, as shown in Figure 11. Rmsprop and Nadam optimizers also achieved excellent results for the two models. Overall, the Adam algorithm is more suitable for analyzing landslide data. However, better accuracy was obtained when Adadelata was used with the RNN model. Meanwhile, adding the weight decay in the neural network did not affect the results.

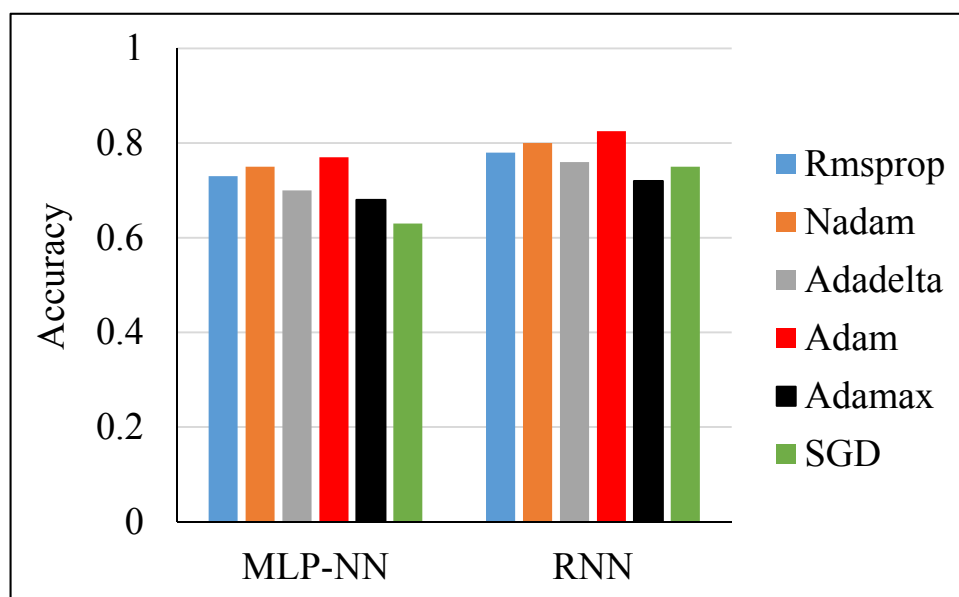


Figure 11. Impact of the optimization algorithm on the performance of MLP-NN and RNN models; SGD: Stochastic Gradient Descent.

Batch size, which refers to the number of training examples computed during optimization, has substantial effects on model accuracy. The results of the RNN and MLP-NN models are shown in Figure 12 and depict how increasing batch size from 2 to 128 (by $\times 2$) affected model accuracy. The MLP-NN and RNN models exhibited the best accuracies with batch sizes of 64 and 128, respectively.

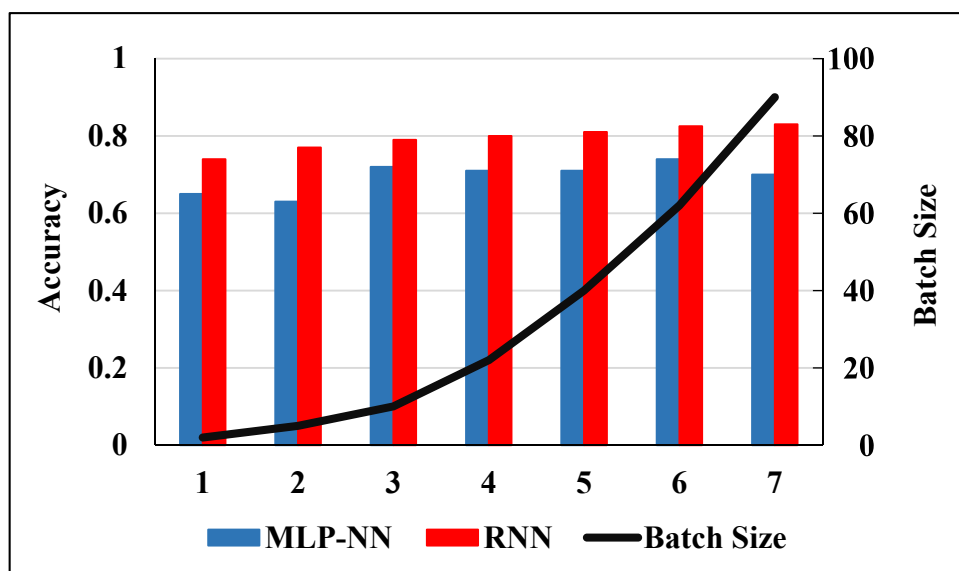


Figure 12. Impact of batch size on the performance of the MLP-NN and RNN models.

Overfitting can be avoided when dropouts are controlled through the number of parameters in the RNN model. Figure 13 illustrates the sensitivity analysis of the effects of dropout rate with various keep probability parameters on the RNN model. The results showed that the appropriate dropout rate is 0.6 for the RNN model. The selected dropout rate considerably affects the performance of NN models. The keep probability was selected in each dataset and analysis was conducted via a systematic grid search.

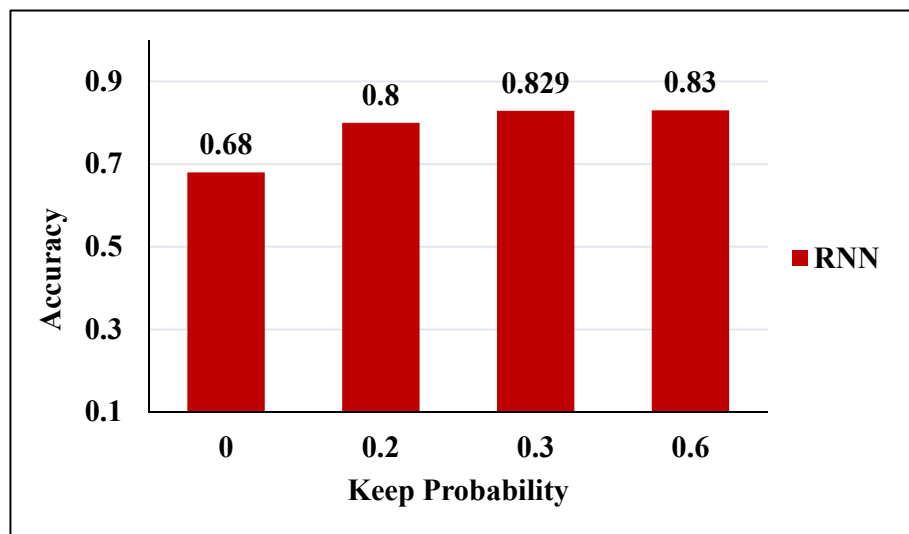


Figure 13. Influence of dropout rate on the performance of the RNN model.

4.6. Field Investigation

The reliability of the proposed methods was validated via field investigation using a handheld Global Position System (GPS) device (GeoExplorer 6000) to locate landslides (Figure 14) and to produce a precise and reliable inventory map of the Cameron Highlands. The more detailed information (landslide extent, source area, deposition, and volume) was obtained from in situ measurements which ultimately demonstrate the reliability of the produced inventory map in the field through use of a GeoExplorer 6000 handheld GPS. The results illustrated that the neural network techniques were able to detect true landslide locations which occurred in past years. Therefore, the results of this study verified that the proposed models can detect landslide locations and generate a reliable landslide inventory map.



Figure 14. Field photographs showing landslide locations during field investigation in (A) Tanah Rata and (B) Tanah Runtuh.

5. Conclusions

The Cameron Highlands, Malaysia form an ideal site for testing the feasibility of RNN and MLP-NN models for landslide detection based on high-resolution LiDAR data. The optimization of segmentation parameters is crucial for improving model performance and computational efficiency with different spatial subsets in the Cameron Highlands. Furthermore, optimization is essential for feature selection to improve the classification accuracy and the computational efficiency of the proposed

methodology. The optimization of NN model parameters helped improve the performance of the model by reducing model complexity and preventing overfitting in the training sample. The RNN model exhibited better accuracy in the analysis and test areas than the MLR-NN model. This investigation showed that network architectures based on optimized techniques, very high resolution (VHR) airborne LiDAR-derived data, and spatial features could be used to effectively identify landslide locations in tropical regions. Therefore, this proposed automatic landslide detection method is a potential geospatial solution for managing landslide hazards and conducting landslide risk assessments.

Given that the proposed RNN model is more efficient than the MLP-NN model and has the potential to process the most relevant features, further studies should be conducted to fully optimize network structures for higher flexibility and eligibility for landslide detection. More theoretical tasks are recommended to enhance the representation of variables and data structure by the RNN model and the storage capacity of the data. Faster and more accurate NN techniques for landslide detection should be developed to overcome all the limitations related to accuracy and time. In addition, the RNN model can be integrated with other NN techniques to help improve other landslide applications.

Author Contributions: Mustafa Ridha Mezaal and Biswajeet Pradhan conceived and designed the experiments; Mustafa Ridha Mezaal, Biswajeet Pradhan and Maher Ibrahim Sameen performed the experiment and analyzed the data; Mustafa Ridha Mezaal and Biswajeet Pradhan contributed reagents/materials/analysis tools; Mustafa Ridha Mezaal and Biswajeet Pradhan wrote the paper. Biswajeet Pradhan is the main supervisor of PhD candidate Mustafa Ridha Mezaal; Helmi Zulhaidi Mohd Shafri and Zainuddin Md Yusoff supported as part of the co-supervisory committee of the first author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guzzetti, F.; Mondini, A.C.; Cardinali, M.; Fiorucci, F.; Santangelo, M.; Chang, K.-T. Landslide inventory maps: New tools for an old problem. *Earth-Sci. Rev.* **2012**, *112*, 42–66. [[CrossRef](#)]
2. Van Westen, C.J.; Castellanos, E.; Kuriakose, S.L. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Eng. Geol.* **2008**, *102*, 112–131. [[CrossRef](#)]
3. Parker, R.N.; Densmore, A.L.; Rosser, N.J.; De Michele, M.; Li, Y.; Huang, R.; Petley, D.N. Mass wasting triggered by the 2008 Wenchuan earthquake is greater than orogenic growth. *Nat. Geosci.* **2011**, *4*, 449–452. [[CrossRef](#)]
4. Chen, R.-F.; Lin, C.-W.; Chen, Y.-H.; He, T.-C.; Fei, L.-Y. Detecting and Characterizing Active Thrust Fault and Deep-Seated Landslides in Dense Forest Areas of Southern Taiwan Using Airborne LiDAR DEM. *Remote Sens.* **2015**, *7*, 15443–15466. [[CrossRef](#)]
5. Pradhan, B.; Jebur, M.N.; Shafri, H.Z.M.; Tehrany, M.S. Data Fusion Technique Using Wavelet Transform and Taguchi Methods for Automatic Landslide Detection From Airborne Laser Scanning Data and QuickBird Satellite Imagery. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1610–1622. [[CrossRef](#)]
6. McKean, J.; Roering, J. Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry. *Geomorphology.* **2004**, *57*, 331–351. [[CrossRef](#)]
7. Whitworth, M.; Giles, D.; Murphy, W. Airborne remote sensing for landslide hazard assessment: A case study on the Jurassic escarpment slopes of Worcestershire, UK. *Q. J. Eng. Geol. Hydrogeol.* **2005**, *38*, 285–300. [[CrossRef](#)]
8. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [[CrossRef](#)]
9. Pradhan, B.; Mezaal, M.R. Optimized Rule Sets for Automatic Landslide Characteristic Detection in a Highly Vegetated Forests. In *Laser Scanning Applications in Landslide Assessment*; Springer: New York, NY, USA, 2017; pp. 51–68.
10. Anders, N.S.; Seijmonsbergen, A.C.; Bouten, W. Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. *Remote Sens. Environ.* **2011**, *115*, 2976–2985. [[CrossRef](#)]
11. Belgiu, M.; Drăguț, L. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 67–75. [[CrossRef](#)] [[PubMed](#)]

12. Drăguț, L.; Tiede, D.; Levick, S.R. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [CrossRef]
13. Zhang, Y.; Maxwell, T.; Tong, H.; Dey, V. Development of a supervised software tool for automated determination of optimal segmentation parameters for ecognition. In Proceedings of the ISPRS TC VII Symposium–100 Years, ISPRS, Vienna, Austria, 5–7 July 2010.
14. Chen, W.; Li, X.; Wang, Y.; Chen, G.; Liu, S. Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China. *Remote Sens. Environ.* **2014**, *152*, 291–301. [CrossRef]
15. Kursa, M.B.; Rudnicki, W.R. Feature selection with the Boruta package. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
16. Li, X.; Cheng, X.; Chen, W.; Chen, G.; Liu, S. Identification of forested landslides using LiDAR data, object-based image analysis, and machine learning algorithms. *Remote Sens.* **2015**, *7*, 9705–9726. [CrossRef]
17. Stumpf, A.; Kerle, N. Object-oriented mapping of landslides using Random Forests. *Remote Sens. Environ.* **2011**, *115*, 2564–2577. [CrossRef]
18. Borghuis, A.; Chang, K.; Lee, H. Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. *Int. J. Remote Sens.* **2007**, *28*, 1843–1856. [CrossRef]
19. Danneels, G.; Pirard, E.; Havenith, H.B. Automatic landslide detection from remote sensing images using supervised classification methods. In *Geoscience and Remote Sensing Symposium*; IEEE: Hoboken, NJ, USA, July 2007; pp. 3014–3017.
20. Moine, M.; Puissant, A.; Malet, J.-P. Detection of Landslides From Aerial and Satellite Images With a Semi-Automatic Method. Application to the Barcelonnette Basin (Alpes-de-Hautes-Provence, France). February 2009, pp. 63–68. Available online: <https://halshs.archives-ouvertes.fr/halshs-00467545/document> (accessed on 16 July 2017).
21. Pratola, C.; Del Frate, F.; Schiavon, G.; Solimini, D.; Licciardi, G. Characterizing land cover from X-band COSMO-SkyMed images by neural networks. In Proceedings of the Urban Remote Sensing Event (JURSE), Munich, Germany, 11–13 April 2011; pp. 49–52.
22. Singh, A.; Singh, K.K. Satellite image classification using Genetic Algorithm trained radial basis function neural network, application to the detection of flooded areas. *J. Vis. Commun. Image Represent.* **2017**, *42*, 173–182. [CrossRef]
23. Singh, K.K.; Singh, A. Detection of 2011 Sikkim earthquake-induced landslides using neuro-fuzzy classifier and digital elevation model. *Nat. Hazards* **2016**, *83*, 1027–1044. [CrossRef]
24. Mehrotra, A.; Singh, K.K.; Nigam, M.J.; Pal, K. Detection of tsunami-induced changes using generalized improved fuzzy radial basis function neural network. *Nat. Hazards* **2015**, *77*, 367–381. [CrossRef]
25. Benediktsson, J.A.; Sveinsson, J.R. Feature extraction for multisource data classification with artificial neural networks. *Int. J. Remote Sens.* **1997**, *18*, 727–740. [CrossRef]
26. Yuan, H.; Van Der Wiele, C.F.; Khorram, S. An automated artificial neural network system for land use/land cover classification from Landsat TM imagery. *Remote Sens.* **2009**, *1*, 243–265. [CrossRef]
27. Fu, G.; Liu, C.; Zhou, R.; Sun, T.; Zhang, Q. Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network. *Remote Sens.* **2017**, *9*, 498. [CrossRef]
28. Sameen, M.I.; Pradhan, B. Severity Prediction of Traffic Accidents with Recurrent Neural Networks. *Appl. Sci.* **2017**, *7*, 476. [CrossRef]
29. Gorsevski, P.V.; Brown, M.K.; Panter, K.; Onasch, C.M.; Simic, A.; Snyder, J. Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: A case study in the Cuyahoga Valley National Park, Ohio. *Landslides* **2016**, *13*, 467–484. [CrossRef]
30. Chang, K.T.; Liu, J.K.; Chang, Y.M.; Kao, C.S. An Accuracy Comparison for the Landslide Inventory with the BPNN and SVM Methods. *Gi4DM 2010*; Turino, Italy, 2010. Available online: https://www.researchgate.net/profile/Jin_King_Liu/publication/267709454_An_Accuracy_Comparison_for_the_Landslide_Inventory_with_the_BPNN_and_SVM_Methods/links/5511773f0cf29a3bb71de12c.pdf (accessed on 20 March 2017).
31. Robert, H.N. *Neurocomputing*; Addison-Wesley Pub. Co.: Boston, CA, USA, 1990; pp. 21–42.
32. Zurada, J.M. *Introduction to Artificial Neural Systems*; West Pub. Co.: Eagan, MN, USA, 1992; pp. 163–248.

33. Chang, K.T.; Liu, J.K. Landslide features interpreted by neural network method using a high-resolution satellite image and digital topographic data. In Proceedings of the XXth ISPRS Congress, Istanbul, Turkey, 12–23 July 2004.
34. Chen, H.; Zeng, Z.; Tang, H. Landslide deformation prediction based on recurrent neural network. *Neural Process. Lett.* **2015**, *41*, 169–178. [CrossRef]
35. Ma, H.-R.; Cheng, X.; Chen, L.; Zhang, H.; Xiong, H. Automatic identification of shallow landslides based on Worldview2 remote sensing images. *J. Appl. Remote Sens.* **2016**, *10*, 016008. [CrossRef]
36. Hall, M.A. Correlation-based Feature Selection for Machine Learning. Doctoral dissertation, The University of Waikato, Hamilton, New Zealand, April 1999.
37. Miner, A.; Flentje, P.; Mazengarb, C.; Windle, D. Landslide Recognition Using LiDAR Derived Digital Elevation Classifiers-Lessons Learnt from Selected Australian Examples. Available online: <http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1590&context=engpapers> (accessed on 10 June 2017).
38. Martha, T.R.; Kerle, N.; van Westen, C.J.; Jetten, V.; Kumar, K.V. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4928–4943. [CrossRef]
39. Pradhan, B.; Lee, S. Regional landslide susceptibility analysis using back-propagation neural network classifier at Cameron Highland, Malaysia. *Landslides* **2010**, *7*, 13–30. [CrossRef]
40. Olaya, V. Basic land-surface parameters. *Dev. Soil Sci.* **2009**, *33*, 141–169.
41. Barbarella, M.; Fiani, M.; Lugli, A. Application of LiDAR-derived DEM for detection of mass movements on a landslide. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *1*, 89–98. [CrossRef]
42. Duro, D.C.; Franklin, S.E.; Dubé, M.G. Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. *Int. J. Remote Sens.* **2012**, *33*, 4502–4526. [CrossRef]
43. Kohavi, R.; John, G.H. Wrappers for feature subset selection. *Artif. Intell.* **1997**, *97*, 273–324. [CrossRef]
44. Mokhtarzade, M.; Zoj, M.V. Road detection from high-resolution satellite images using artificial neural networks. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 32–40. [CrossRef]
45. Gardner, M.W.; Dorling, S.R. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [CrossRef]
46. Baczyński, D.; Parol, M. Influence of artificial neural network structure on quality of short-term electric energy consumption forecast. *IEE Proc.-Gener., Transm. Distrib.* **2004**, *151*, 241–245. [CrossRef]
47. Mia, M.M.A.; Biswas, S.K.; Urmi, M.C.; Siddique, A. An algorithm for training multilayer perceptron (MLP) for Image reconstruction using neural network without overfitting. *Int. J. Sci. Techno. Res.* **2015**, *4*, 271–275.
48. Yang, G.Y.C. *Geological Mapping from Multi-Source Data Using Neural Networks*; Geomatics Engineering, University of Calgary: Calgary, AB, Canada, 1995.
49. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
50. Donahue, J.; Anne Hendricks, L.; Guadarrama, S.; Rohrbach, M.; Venugopalan, S.; Saenko, K.; Darrell, T. Long-term recurrent convolutional networks for visual recognition and description. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 2625–2634.
51. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Vanderplas, J. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
52. Sameen, M.I.; Pradhan, B.; Shafri, H.Z.; Mezaal, M.R.; bin Hamid, H. Integration of Ant Colony Optimization and Object-Based Analysis for LiDAR Data Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2055–2066. [CrossRef]
53. Bartels, M.; Wei, H. Threshold-free object and ground point separation in LIDAR data. *Pattern Recognit. Lett.* **2010**, *31*, 1089–1099. [CrossRef]
54. Rau, J.Y.; Jhan, J.P.; Rau, R.J. Semiautomatic object-oriented landslide recognition scheme from multisensor optical imagery and DEM. *IEEE Trans. Geosci. Remote Sens.* **2011**, *52*, 1336–1349. [CrossRef]

