Catena 167 (2018) 147–159

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

An improved algorithm for identifying shallow and deep-seated landslides in dense tropical forest from airborne laser scanning data

Mustafa Ridha Mezaal[a], Biswajeet Pradhan[b,c,*]

[a] Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[b] School of Systems, Management and Leadership, Faculty of Engineering and IT, University Technology Sydney, CB11.06.217, Building 11, 81 Broadway, Ultimo, NSW 2007, Australia
[c] Department of Energy and Mineral Resources Engineering, Choongmu-gwan, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea

ARTICLE INFO

Keywords:
Landslide types
Object-based approach
LiDAR
Remote sensing
GIS

ABSTRACT

Landslides are natural disasters that cause environmental and infrastructure damage worldwide. They are difficult to be recognized, particularly in densely vegetated regions of the tropical forest areas. Consequently, an accurate inventory map is required to analyze landslides susceptibility, hazard, and risk. Several studies were done to differentiate between different types of landslide (i.e. shallow and deep-seated); however, none of them utilized any feature selection techniques. Thus, in this study, three feature selection techniques were used (i.e. correlation-based feature selection (CFS), random forest (RF), and ant colony optimization (ACO)). A fuzzy-based segmentation parameter (FbSP optimizer) was used to optimize the segmentation parameters. Random forest (RF) was used to evaluate the performance of each feature selection algorithms. The overall accuracies of the RF classifier revealed that CFS algorithm exhibited higher ranks in differentiation landslide types. Moreover, the results of the transferability showed that this method is easy, accurate, and highly suitable for differentiating between types of landslides (shallow and deep-seated). In summary, the study recommends that the outlined approaches are significant to improve in distinguishing between shallow and deep-seated landslide in the tropical areas, such as; Malaysia.

1. Introduction

Cameron Highlands in Malaysia has been frequently affected due to geo-hazards such as landslides and floods. The effects include great economic damage, loss of lives and negative environmental impact (Hong et al., 2018). Landslide as one of the geo-hazards is considered as a geological phenomenon under the influence of gravity, which can occur in both onshore, offshore, and coastal environments (Pradhan and Lee, 2010). The Cameron Highlands is a steep hillside landscape with heavy vegetation cover that obscures and subdues morphologic features which are indicative of landslides (Pradhan and Mezaal, 2017). Such landscapes pose a great challenge to landslides identification using synthetic aperture radar (SAR) images, optical and aerial photographs (Bui et al., 2012; Imani et al., 2012; Pourghasemi et al., 2014), high spatial resolution multispectral images (Pradhan, 2013), very high resolution (VHR) satellite images and moderate resolution digital terrain models (DTMs) (Ardizzone et al., 2007; Chen et al., 2014; Pradhan et al., 2016; Li et al., 2015; Mezaal et al., 2017a; Bordoni et al., 2018; Sameen and Pradhan, 2017; Mezaal and Pradhan, 2018; Fanos and Pradhan, 2018).

2. Previous work

Compared with the traditional techniques, elevation data are acquired rapidly and accurately using active laser transmitters and receivers light-detection and ranging (LiDAR) data (Pradhan et al., 2016; Tarolli et al., 2009). Generally, LiDAR can penetrate dense vegetation making it a better alterative compared with other remote sensing data. In addition, other information regarding high point density terrain is provided in Mezaal et al. (2017b). Ground surface and useful information about topographic features are provided using High-resolution LiDAR-derived DEM even in landslides covered under dense vegetation (McKean and Roering, 2004). Furthermore, LiDAR imagery is capable revealing present and historic landslides and its effectiveness/vulnerability in mapping naked slopes that are formed primarily by landslides (Schulz, 2007).

Based on the depth of the surface rupture and movement features, landslides can be classified as deep-seated or shallow (Brunetti et al.,

* Corresponding author at: School of Systems, Management and Leadership, Faculty of Engineering and IT, University Technology Sydney, CB11.06.217, Building 11, 81 Broadway, Ultimo, NSW 2007, Australia.
E-mail addresses: Biswajeet.Pradhan@uts.edu.au, Biswajeet24@gmail.com (B. Pradhan).

https://doi.org/10.1016/j.catena.2018.04.038
Received 20 February 2018; Received in revised form 19 April 2018; Accepted 27 April 2018

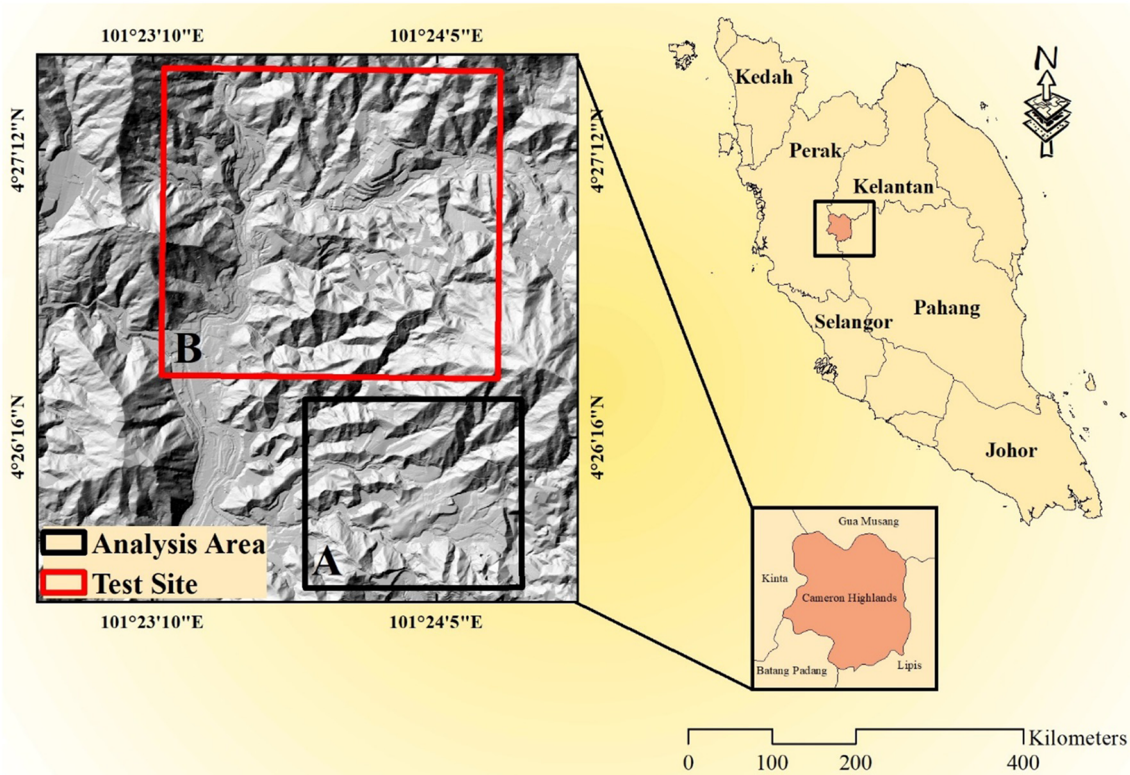

Fig. 1. Shows study are that consist of (A) Analysis area (B) Test site.

2009; Guzzetti et al., 2012). These two classifications differ in terms of damage influence, size and volume (Zêzere et al., 2005). Also, evaluation of landslide mass volume is difficult (Brunetti et al., 2009). Deep-seated landslides are usually occurred due to interaction between natural denudation process and long-term rainfall, whereas, shallow landslides are associated with short high-intensity rainfall (Zêzere et al., 2005). In literature many studies can be found which are aimed in identifying different types of landslides using LiDAR data (Chen et al., 2015; Deng and Shi, 2014; Lin et al., 2013; Rau et al., 2012; Kasai et al., 2009; Van Den Eeckhaut et al., 2005; Lashermes et al., 2007; Tarolli et al., 2009; Passalacqua et al., 2010). The different types of landslides provide significant and valuable information for the geological process. Therefore, for the purpose of investigating hillsides geomorphological development is to mitigate landslide hazards, thus, it is necessary to differentiate between the different types of landslides for better efficiency (Dou et al., 2015; Lin et al., 2013).

Object-based and pixel-based methods are the two general image analysis approaches for terrain evaluation. But object-based image analysis is becoming the most basic means of processing very high-resolution imagery. This is due to wide utilization of sub-meter imagery and availability. Furthermore, this approach is a well-known technique resulting from the recent advances in machine intelligence and computer vision, with the main purpose of automatically extracting both man-made and natural objects from remote sensing images (Akçay and Aksoy, 2008). Also, the object-based approach is a step toward replicating human interpretation process because the information content of an object is used to classify landscapes (Navulur, 2006). Finally, with the use of object-based approach, the landslides can be accurately detected by integrating contextual information to image analysis (Martha et al., 2011). This will help in reducing time and cost for developing a decent landslide inventory map especially in large areas.

Over-fitting is generally caused by processing a large number of irrelevant features (Chen et al., 2014). By contrast, in order to avoid over-fitting, the most relevant feature should be selected for best classification results (Kursa and Rudnicki, 2010). Therefore, landslide identification in any environment can be improved by selecting the most significant features (Chen et al., 2014). As shown in a study conducted by (Van Westen et al., 2008), selecting the most significant feature helps in differentiating between non-landslides and landslides. The efficiency of selecting the most significant feature for detecting landslides was proven in a study conducted by (Stumpf and Kerle, 2011). But the use LiDAR data to handle the feature selection for landslide detection is studied by few researchers (Dou et al., 2015; Li et al., 2015). Another option for feature selection is a random forest (RF) (Chen et al., 2014). More of recent, (Sameen et al., 2017) utilized the use of ant colony optimization (ACO) for feature selection. While Pradhan and Mezaal (2017) demonstrated the significance of feature selection in differentiating between the types of landslides by using correlation-based feature selection (CFS) algorithm. Although, these feature selection methods were applied in remote sensing data classification successfully. However, it was observed that there is a lack of studies on integration of correlation-based feature selection (CFS), random forest (RF), and ant colony optimization (ACO) with the object-based approach (OBA) carried out to aid in differentiating between the different types of landslides (i.e. shallow and deep-seated).

This study aims at investigating the most optimal algorithms for feature selection in order to differentiate between two types of the landslide (i.e. shallow and deep-seated) using airborne laser scanning data. To achieve this aim, it was imperative to accomplish the following objectives; 1) to optimize the multiresolution segmentation parameters, 2) to applying the three algorithms to feature selection from high-resolution airborne laser scanning data, and 3) to determine the appropriate algorithms for selecting feature by using random forest (RF) classifier. The studied algorithms have not been tested in previous studies, particularly for types of landslides detection. The advantages of novel optimization techniques may have contributed to the improvement of the differentiation between the types of landslide through a high-resolution LiDAR data and supervised random forest.