UNIVERSITI PUTRA MALAYSIA

OPTIMIZED TECHNIQUES FOR LANDSLIDE DETECTION AND CHARACTERISTICS USING LiDAR DATA

MUSTAFA RIDHA MEZAAL

FK 2018 79
OPTIMIZED TECHNIQUES FOR LANDSLIDE DETECTION AND
CHARACTERISTICS USING LiDAR DATA

By

MUSTAFA RIDHA MEZAAL

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,
in Fulfilment of the Requirements for the Degree of
Doctor of Philosophy

May 2018
COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs, and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia
Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

OPTIMIZED TECHNIQUES FOR LANDSLIDE DETECTION AND CHARACTERISTICS USING LiDAR DATA

By

MUSTAFA RIDHA MEZAAL

May 2018

Chairman : Professor Biswajeet Pradhan, PhD
Faculty : Engineering

Landslides are one of the major natural disasters that cause severe damage to lives and properties worldwide. Historically, landslide occurrences are usually mapped by taking inventory of location and magnitude of the landslide in a region. This information is used to examine and manage slope failures and distributions effectively. A good landslide inventory map is a prerequisite for analyzing landslide susceptibility, hazard, and risk. Field survey, optical remote sensing, and synthetic aperture radar techniques are traditional techniques use for landslide detection in tropical regions. However, such techniques are time consuming and costly. In addition, the dense vegetation in tropical forests affects the accuracy of the landslide inventory maps. Furthermore, it is difficult to distinguish different types of landslides due to geomorphological development along hillsides using the traditional approach. so, it necessary to develop more innovative approach that can resolve the aforementioned challenges.

Therefore, in line with the objectives of this research, very- high-resolution LiDAR point cloud data and orthophotos image, have been utilized to map the landslide events in Cameron Highlands, Malaysia. The segmentation process was optimized using Fuzzy-based Segmentation Parameter. Also, six techniques: Ant Colony Optimization (ACO), Gain Ratio (GR), Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), Random forest (RF), and Correlation-based Feature Selection (CFS) were used for the feature selection. The locations of landslides were detected accurately by employing two Machine learning classifiers, namely, SVM and RF, decision rule and hierarchal rules sets were developed by applying decision tree (DT) algorithm to provide improved landslide inventory. In this task, two neural network algorithms, Recurrent Neural Networks (RNN) and Multi-Layer Perceptron Neural Networks
(MLP-NN) were used and the hyper-parameters of the network architecture was optimized based on a systematic grid search.

The performance of the outcome was validated based on the receiver operating characteristic (ROC) area under the curve (AUC) values, confusion matrix and Cross Validation method. Transferability of each of the models was verified by testing in another site for consistency. The overall accuracy of the Support Vector Machine SVM and Random Forest RF classifiers revealed that three of the six algorithms exhibited higher ranks in the landslide detection. The classification accuracy of the RF classifier is observed to be higher than that of SVM using either all features or only the optimal features. The proposed techniques performed well in detecting landslides in tropical area in Malaysia. Furthermore, the transferability indicates that the techniques can easily be extended to any region with similar characteristics.

The result show that the accuracy of shallow and deep-seated landslides were 0.80 and 0.83, respectively. The intensity derived from the LiDAR data, geometric and texture features significantly affects the accuracy of differentiating shallow from deep-seated landslides. While, the results of shallow and deep using hierarchal rules set were observed to be 87.2%, and 90% respectively, for site A (Analysis area). More so, the hierarchal rules set were evaluated using another site named site B (Test area), and the accuracies of shallow and deep seated were found to be 86.4% and 80.8% respectively. This indicates that LiDAR data are highly efficient in detecting landslide characteristics in tropical forested areas.

Furthermore, RNN and MLP-NN models in the test area showed 81.11%, and 74.56%, accuracy level, respectively. These results indicated that the proposed models with optimized hyper-parameters produced the accurate classification results. The LiDAR-derived data, orthophotos and textural features significantly affected the classification results. The results indicated that the proposed methods have the potential to produce accurate and appropriate landslide inventory in tropical regions such as Malaysia. Hopefully, this innovative method can be deployed in detecting landslide and distinguish between different types of landslides (shallow and deep-seated landslides) in the near future for landslide management due to its transferability capabilities to different environments.
Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PEMBANGUNAN TEKNIK YANG DIKENAKAN UNTUK PENGESANAN TANAH RUNTUH DAN SIFAT DARI PEMANTAUAN DATA DARI LASER UDARA DAN ORTHOFOTO

Oleh

MUSTAFA RIDHA MEZAAL

Mei 2018

Pengerusi : Profesor Biswajeet Pradhan, PhD
Fakulti : Kejuruteraan


Hasil kajian ini membentangkan beberapa kaedah untuk mengesan tanah runtuh dan keupayaan bagi membezakan antara tanah runtuh yang cetek dan dalam dengan menggunakan pengesan cahaya dan berbagai-bagai set data-LiDAR dan kepenggunaannya mengikut keadaan persekitaran yang berbeza. Antara objektif khusus kajian ini adalah: 1) mengenal pasti algoritma yang sesuai untuk pemilihan ciri bagi menambah baik proses mengesan tanah runtuh, 2) untuk mengoptimumkan ketetapan peraturan keputusan data set dan hierarki peraturan untuk membezakan ciri-

Kesemua prestasi tersebut telah disahkan berdasarkan nilai di bawah kurva (AUC) dalam ciri-ciri operasi penerima (ROC), matriks kekeliruan dan pengesahan silang. Setiap pemindahan model telah disahkan dengan diguna pakai di kawasan lain. Untuk objektif pertama, keseluruhan ketepatan daripada Mesin Sokongan Vektor (SVM) dan pengelas Hutan Rawak (RF) mendedahkan bahawa tiga daripada enam algoritma berjaya mempamerkan ketepatan yang lebih tinggi dalam mengesan kejadian tanah runtuh. Ketepatan pengelasan bagi pengelas (RF) adalah lebih tinggi daripada pengelas (SVM) walau sama ada dengan menggunakan kesemua ciri atau ciri optimum sahaja. Teknik yang dicadangkan ini adalah yang terbaik dalam mengesan tanah runtuh di kawasan tropika seperti Malaysia. Oleh yang demikian, teknik-teknik ini mudah untuk diguna pakai di mana-mana kawasan yang mempunyai ciri-ciri yang hampir serupa.

Objektif kedua menunjukkan bahawa ketepatan bagi mengenal pasti jenis-jenis tanah runtuh sama ada cetek dan mendalam adalah 0.80 dan 0.83. Keamatan yang diperolehi daripada tekstur dan data LiDAR dengan ketara sangat mempengaruhi ketepatan dalam membezakan tanah runtuh cetek dan mendalam. Walau bagaimanapun, didapati bahawa keputusan data bagi tanah runtuh cetek dan mendalam dengan menggunakan aturan hierarki adalah 87.2% dan 90.0%, untuk kawasan A (iaitu kawasan yang dianalisis). Sementara itu, set aturan hierarki yang sama diguna pakai di kawasan B (iaitu kawasan yang dikaji), dan ketepatan dalam mengenal pasti tanah runtuh cetek dan mendalam adalah 86.4% dan 80.8%. Oleh yang demikian, hasil kajian ini menunjukkan bahawa data LiDAR adalah sangat berkesan dalam mengesan dan mengenal pasti jenis-jenis tanah runtuh di kawasan hutan hujan tropika.

Selain itu, hasil daripada objektif ketiga pula menunjukkan bahawa ketepatan bagi model RNN dan MLP-NN yang dijalankan di kawasan uji kaji masing-masing adalah 81.11% dan 74.56%. Oleh itu, keputusan ini menunjukkan bahawa model yang dicadangkan dalam kajian ini menghasilkan hasil klasifikasi yang sangat tepat dengan
mengoptimumkan beberapa parameter-hiper. Di sini, data daripada LiDAR, orthofotos dan ciri-ciri tekstur amat mempengaruhi hasil klasifikasi kajian.

Keputusan dari hasil kajian ini menunjukkan bahawa kaedah yang dicadangkan berpotensi untuk menghasilkan inventori jenis tanah runtuh dengan tepat dan sesuai diguna pakai di kawasan hutan hujan tropika seperti Malaysia. Algoritma dan maklumat yang diperolehi daripada kajian ini dapat menyumbang kepada pengurusan tanah runtuh di negara-negara hutan hujan tropika.
ACKNOWLEDGEMENTS

I praise ALLAH for his great loving kindness, which has brought all of us to tell and encourage each other and who has pulled us from the darkness to the light. All respect for our holy prophet (Peace be upon him), who guided us to identify our creator. I also thank all my brothers and sister who answered ALLAH’s call and have made their choice to be in the straight path of ALLAH.

As always it is impossible to mention everybody who had an impact to this work however there are those whose spiritual support is even more important. I feel a deep sense of gratitude for my mother and father, who formed part of my vision and taught me good things that really matter in life. Their infallible love and support has always been my strength. Their patience and sacrifice will remain my inspiration throughout my life. I am also very much grateful to all my family members for their constant inspiration and encouragement specially my mother in law.

My heartfelt thanks to my love and my wife Doaa Talib Hashim for her guidance and moral support. She always helped me out when I got any difficulties or queries regarding all the aspect of life. Again I thank her for standing by my side and sharing a great relationship as compassionate wife. I will always remember the warmth shown by her.

I take this opportunity to express my profound gratitude and deep regards to my guide Prof. Dr. Biswajeet Pradhan for his exemplary guidance, monitoring and constant encouragement throughout the course of this thesis. The blessing, help and guidance given by him time to time shall carry me a long way in the journey of life on which I am about to embark. He created an atmosphere that encouraged innovation and shared his extraordinary experiences throughout the work. Without his unflinching encouragement, it would have been impossible for me to finish this research.

I am so thankful to my committee Assoc. Prof. Dr. Helmi Zulhaidi bin Mohd Shafri and Dr. Zainuddin bin Md Yusoff, for the valuable information provided by them in their respective fields. I am obliged for their cooperation during the period of my study.
I certify that a Thesis Examination Committee has met on 17 May 2018 to conduct the final examination of Mustafa Ridha Mezaal on his thesis entitled "Optimized Techniques for Landslide Detection and Characteristics Using LiDAR Data" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Farzad Hejazi, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Dato' Shattri bin Mansor, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Abdul Rashid bin Mohamed Sharif, PhD
Professor Sr.Gs.
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Hyung-Sup Jung, PhD
Professor
University of Seoul
South Korea
(External Examiner)

RUSLI HAJI ABDULLAH, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 30 July 2018
This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

**Biswaajeet Pradhan, PhD**  
Professor  
Faculty of Engineering  
Universiti Putra Malaysia  
(Chairman)

**Helmi Zulhaidi Mohd Shafri, PhD**  
Associate Professor  
Faculty of Engineering  
Universiti Putra Malaysia  
(Member)

**Zainuddin Bin Md Yusoff, PhD**  
Senior Lecturer  
Faculty of Engineering  
Universiti Putra Malaysia  
(Member)

**ROBIAH BINTI YUNUS, PhD**  
Professor and Dean  
School of Graduate Studies  
Universiti Putra Malaysia

Date
Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: _______________________ Date: __________________

Name and Matric No.: Mustafa Ridha Mezaal, GS45396
Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature: 
Name of Chairman of Supervisory Committee: Professor Dr. Biswajeet Pradhan

Signature: 
Name of Member of Supervisory Committee: Associate Professor Dr. Helmi Zulhaidi Mohd Shafri

Signature: 
Name of Member of Supervisory Committee: Dr. Zainuddin Bin Md Yusoff
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>APPROVAL</td>
<td>vii</td>
</tr>
<tr>
<td>DECLARATION</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xvii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xx</td>
</tr>
</tbody>
</table>

## CHAPTER

### 1 INTRODUCTION

1.1 Introduction                           1  
1.2 Problem Statement                      2  
1.3 Motivation Behind this Research        2  
1.4 Research Objectives                    3  
1.5 Research Questions                     3  
1.6 Scope of this Thesis                   4  
1.7 Research Hypothesis                    4  
1.8 Thesis Organization                    5  

### 2 LITERATURE REVIEW

2.1 Introduction                            6  
2.2 Landslide detection                     6  
2.2.1 Types of Landslides                  6  
2.2.1.1 Deep-seated Landslide              7  
2.2.1.2 Shallow Landslide                  9  
2.3 Traditional Based Techniques           10  
2.3.1 Geomorphological Field Mapping       10  
2.3.2 Visual Interpretation                11  
2.4 Innovative and Emerging Techniques     12  
2.4.1 Image Classification                 12  
2.4.1.1 Pixel Based Techniques             13  
2.4.1.2 Object Based Techniques            14  
2.5 Analysis of Very-High Resolution DEMs   17  
2.6 LiDAR data and Geo-Morphometric Features 18  
2.7 Machine Learning Algorithms            22  
2.7.1 Support Vector Machine (SVM)         23  
2.7.2 Random Forest (RF)                   24  
2.7.3 Decision Trees (DT)                  24  
2.7.4 Deep Neural Networks (DNN)            25  
2.7.4.1 Multi-Layer Perceptrons (MLP)      26  
2.7.4.2 Recurrent Neural Networks (RNN)     26  

xi
2.8 Hierarchal Approach
2.9 Feature Selection
  2.9.1 Ant Colony Optimization (ACO)
  2.9.2 Gain ratio (GR)
  2.9.3 Particle swarm optimization (PSO)
  2.9.4 Random Forest (RF)
  2.9.5 Genetic Algorithms (GA)
  2.9.6 Correlation Feature Selection (CFS)
2.10 Summary

3 MATERIALS AND METHODS
  3.1 Introduction
  3.2 Overall Methodology
  3.3 Study Areas
    3.3.1 Geological, and Morphological Characteristics of the Study Area
  3.4 Data Used
    3.4.1 Landslide Inventory
    3.4.2 LiDAR Data
      3.4.2.1 LiDAR Point Cloud
      3.4.2.2 Orthophoto
  3.5 Pre-Analysis
    3.5.1 Pre-processing
    3.5.2 Object Based Approach
      3.5.2.1 Image Segmentation
  3.6 Investigate the Suitable Selection Features Algorithm
    3.6.1 Overall Methodology
    3.6.2 Feature Selection Using Six Algorithms (ACO, GRA, PSO, GA, RF and CFS)
      3.6.2.1 ACO
      3.6.2.2 GR
      3.6.2.3 PSO
      3.6.2.4 RF
      3.6.2.5 GA
      3.6.2.6 CFS
  3.7 Machine Learning Algorithms
    3.7.1 SVM Classifier
    3.7.2 RF Classifier
  3.8 Optimize Rule Sets to Differentiate Various Landslide Characteristics and Compare its Performance with Hierarchical Rule-Set Classification Approach.
    3.8.1 Selecting the most Suitable Algorithm for Feature Selection
      3.8.1.1 Overall Methodology
      3.8.1.2 Feature selection using Three Algorithms (ACO, RF and CFS)
      3.8.1.3 Evaluating the Feature Selection Algorithms
3.8.2 Develop Decision Rule Sets and Hierarchical Rule Sets for Distinguishing Landslide Types (i.e. Shallow and Deep seated) 60
3.8.2.1 Overall Methodology 60
3.8.2.2 Decision Tree 61
3.8.2.3 Implementations of Hierarchal Rules Set 63

3.9 Applied Deep Learning (Neural Architecture) Procedure for Landslide Mapping 64
3.9.1 MLP-NN Model 64
3.9.1.1 Architecture of the MLR-NN Model 65
3.9.2 RNN Model 66
3.9.2.1 Architecture of RNN Model 66
3.9.3 Training data for MLP-NN and RNN Models 67

3.10 Accuracy Assessment 67
3.11 Summary 69

4 RESULTS AND DISCUSSION 70
4.1 The Results of Selected Feature Techniques 70
4.1.1 Optimized Segmentation 70
4.1.2 Selection of the Relevant Features by Various Algorithms 71
4.1.3 Results of the SVM and RF Classifiers 77
4.1.4 Accuracy Assessment for Landslide Detection 78
4.1.5 Discussion 80
4.1.6 Field Investigation 81

4.2 Distinguish two Types of Landslide Namely Shallow and Deep seated 82
4.2.1 Optimizing the Boundary of the Types of Landslide 82
4.2.2 Selection Relevant Feature through three Algorithms (CFS, ACO and RF) for Different types of Landslide 83
4.2.3 RF Classifier for Differentiation between Shallow and Deep Seated Landslide in the Analysis Area 88
4.2.4 Evaluation of CFS based Feature Selection to other Site (i.e. Test site) 90
4.2.5 Accuracy Assessment 91
4.2.6 Discussion 93
4.2.7 Field Investigation 94

4.3 Optimized Rule Sets for Detecting Landslide and Differentiation 95
4.3.1 Segmentation Parameters Selected Using a Fuzzy Logic Supervised Approach 95
4.3.2 Features Selected Using the CFS Method 96
4.3.3 Rule Sets Developed for Landslide Detection and Characterization 97
4.3.4 Model Transferability 98
4.3.5 Discussion 101
4.3.6 Accuracy Assessment 103
4.3.7 Field Investigation 105
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Types of Landslide with brief version</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>Overview of Techniques for the Collection of Landslide Information Obtained from (Van Westen et al., 2008)</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>The Feature Selection Using in the Current Research</td>
<td>50</td>
</tr>
<tr>
<td>3.2</td>
<td>Optimizing the Parameters of Support Vector Machine</td>
<td>56</td>
</tr>
<tr>
<td>3.3</td>
<td>Object Features used in the Current this Study</td>
<td>60</td>
</tr>
<tr>
<td>4.1</td>
<td>Multi-resolution Segmentation Parameters</td>
<td>70</td>
</tr>
<tr>
<td>4.2</td>
<td>Optimal Feature Selection for Detection Landslide using Various Algorithms</td>
<td>73</td>
</tr>
<tr>
<td>4.3</td>
<td>Correlation Coefficient for the Best Feature Selection in Landslide Detection</td>
<td>76</td>
</tr>
<tr>
<td>4.4</td>
<td>Results Comparison Based on Overall Accuracy and Kappa Coefficient for Important Features and Full Features using RF and SVM Algorithms</td>
<td>79</td>
</tr>
<tr>
<td>4.5</td>
<td>Results Comparison Based on User’s Accuracy and Producer's Accuracy for Important Features and Full Features using RF and SVM Algorithms</td>
<td>79</td>
</tr>
<tr>
<td>4.6</td>
<td>Multi-resolution Segmentation Parameters</td>
<td>83</td>
</tr>
<tr>
<td>4.7</td>
<td>The Important Features Selection through three Algorithms for Detecting Types of Landslide (i.e. Shallow and Deep Seated landslides)</td>
<td>85</td>
</tr>
<tr>
<td>4.8</td>
<td>Correlation Coefficient for the Best Feature Selection in Distinguishing Landslide Types</td>
<td>87</td>
</tr>
<tr>
<td>4.9</td>
<td>Results Comparison based on Overall Accuracy for Important and Full Features using RF Classifier</td>
<td>92</td>
</tr>
<tr>
<td>4.10</td>
<td>Results Comparison based on User’s Accuracy and Producer's Accuracy for Important and Full Features using RF Classifier</td>
<td>93</td>
</tr>
<tr>
<td>4.11</td>
<td>Optimized Segmentation Parameters</td>
<td>96</td>
</tr>
<tr>
<td>4.12</td>
<td>Results of the Features Selection using CFS Method</td>
<td>97</td>
</tr>
<tr>
<td>4.13</td>
<td>Rules Defined for Identification Landslide from Non-Landslide</td>
<td>97</td>
</tr>
<tr>
<td>4.14</td>
<td>Rules Defined for Distinguishing between Types of Landslide</td>
<td>100</td>
</tr>
<tr>
<td>4.15</td>
<td>Relevant Features Selected based on CFS Algorithm</td>
<td>107</td>
</tr>
</tbody>
</table>
4.16 Rule sets Developed by the DT Algorithm using the Important Feature Subset 109
4.17 Shows the Results of Overall Accuracy, Kappa Coefficient, user’s Accuracy and Producer’s accuracy for Analysis Area 115
4.18 Shows the Results of Overall Accuracy, Kappa Coefficient, User’s Accuracy and Producer’s Accuracy for test site “B” 115
4.19 Multi-resolution Segmentation Parameters 118
4.20 CFS Results for the most Relevant Feature Subset 119
4.21 Cross-Validation Accuracy Results of the Proposed Models 122
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Overall Methodology Flowchart</td>
<td>34</td>
</tr>
<tr>
<td>3.2</td>
<td>Shows the Location of Different Study Areas in Cameron Highland Malaysia</td>
<td>35</td>
</tr>
<tr>
<td>3.3</td>
<td>The Characteristic Maps of (A) LULC, (B) Geological, and, (C) Soil Maps of the entire Study Area</td>
<td>36</td>
</tr>
<tr>
<td>3.4</td>
<td>The Landslide Inventory Maps</td>
<td>39</td>
</tr>
<tr>
<td>3.5</td>
<td>Shows LiDAR derived data (A) Orthophotos, (B) DTM, (C) DSM, (D) Intensity, (E) Height, (F) Slope, (G) Aspect, and (H) Hillshade</td>
<td>42</td>
</tr>
<tr>
<td>3.6</td>
<td>Workflow of the Proposed FbSP Optimizer</td>
<td>46</td>
</tr>
<tr>
<td>3.7</td>
<td>Flowchart of the Proposed Methodology</td>
<td>48</td>
</tr>
<tr>
<td>3.8</td>
<td>ACO-Based Feature Selection Workflow</td>
<td>51</td>
</tr>
<tr>
<td>3.9</td>
<td>The Flowchart Designed for this research</td>
<td>59</td>
</tr>
<tr>
<td>3.10</td>
<td>An Overview of the Method Adopted in the Present Study</td>
<td>61</td>
</tr>
<tr>
<td>3.11</td>
<td>The Flowchart Illustrates the Overview of the Proposed Method</td>
<td>63</td>
</tr>
<tr>
<td>3.12</td>
<td>Architecture of the MLR-NN Model</td>
<td>65</td>
</tr>
<tr>
<td>3.13</td>
<td>Architecture of the RNN Model</td>
<td>66</td>
</tr>
<tr>
<td>4.1</td>
<td>Shows the Result of the Segmentation using Optimized Parameters (when the Scale is 75.52) for the Analysis Area. It can be Seen that the Landslide Objects were Accurately Delineated Highlighted by Red Color</td>
<td>71</td>
</tr>
<tr>
<td>4.2</td>
<td>Shows the Results of Support Vector Machine (A) Analysis area (B) Test site</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>Shows the Results of Random Forest, (A) Analysis Area, (B) Test Site</td>
<td>78</td>
</tr>
<tr>
<td>4.4</td>
<td>Landslides Locations in the Study Area</td>
<td>82</td>
</tr>
<tr>
<td>4.5</td>
<td>The Segmentation Process and Optimal Segmentation using FbSP Optimizer for the Location of Landslides. (A) And (B) Represent the Initial Segmentation, while (C) Represent the Optimized Segmentation</td>
<td>83</td>
</tr>
</tbody>
</table>
4.6 [A] Result of RF Classifier in the Analysis Area, While [B] Presented The Inventory Map And [C] Showed how the Proposed Method Differentiated between two classes (Shallow And Deep Seated) 88

4.7 Illustrates Sketch of Shallow Landslide in (Yellow Polygon) And Deep Seated Landslide in (White Polygon) in Cameron Highland 89

4.8 Result of Transferability Model Showed the Locations of Shallow and Deep Seated Landslide in the Test Site 91

4.9 Field Photos Showing the Location of Landslide During field Investigation 95

4.10 Shows the process of segmentation, (A) Initial Segmentation (left) and, (B) Optimal Segmentation (right) 96

4.11 Analysis Area for Detection the Location of the Landslide 98

4.12 Show the Results of Transferability to whole Study Area 99

4.13 Show the Location of the Shallow Landslides Represent in (Red polygon), Deep-Seated Landslides Represent in (Yellow polygon) 101

4.14 ROC Curve for, (A) Landslide Detection and their Characteristics, (B) Shallow Landslide, (C) Deep Seated Landslide 104

4.15 Landslide Photos Taken in Study Area During Field Investigation 105

4.16 Shows the Process of Optimization Segmentation, (A) Initial Segmentation and, (B) Optimized Segmentation 106

4.17 Illustrates the Structure of Different Types of Soil 108

4.18 Results of Hierarchal Rules Set Classification at Site “A” 110

4.19 Show the Digital Number (DN) Values of, (A) Texture, (B) Intensity and, (C) Average of Visible Bands, which they Contributed in Distinguish between the Classes (Shallow, Deep Seated, Cut-Slope and Bare Soil) 112

4.20 Results of Hierarchal Rules Set Classification at Site “B” 113

4.21 Shows that the Amount of Intensity Value Involved in Landslide 114

4.22 Shows the Field Photos in Some Locations of Landslide at (A) Taman Desa Corina, Kampung Raja (B)Tanah-Rata 117
4.23 Parameter Optimization of the Multiresolution Segmentation Algorithm: (A) Initial Segmentation and (B) Optimized Segmentation

4.24 Results of the Qualitative Assessment of (A) RNN and (B) MLP-NN For the Analysis Area

4.25 Results of the Qualitative Assessment of (A) RNN And (B) MLP-NN for the Test site

4.26 Impact of the Optimization Algorithm on the Performance of MLP-NN and RNN Models

4.27 Impact of Batch Size on the Performance of The MLP-NN and RNN Models

4.28 Influence of Dropout Rate on the Performance of the RNN Model

4.29 Field Photographs Showing Landslide Locations during Field Investigation in (A) Tanah Rata (101°23'24.89"E and 4°26'24.33"N), and (B) Ringlet (101°22'55.37"E and 4°24'45.47"N)
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACO</td>
<td>Ant Colony Optimization</td>
</tr>
<tr>
<td>AUC</td>
<td>The area under curve</td>
</tr>
<tr>
<td>CFS</td>
<td>Correlation-based Feature Selection</td>
</tr>
<tr>
<td>DEM</td>
<td>Digital Elevation Model</td>
</tr>
<tr>
<td>DNN</td>
<td>Deep Neural Network</td>
</tr>
<tr>
<td>CWT</td>
<td>Continuous Wavelet Transform</td>
</tr>
<tr>
<td>DSM</td>
<td>Digital surface model</td>
</tr>
<tr>
<td>DN</td>
<td>Digital Number</td>
</tr>
<tr>
<td>EC</td>
<td>Evolutionary Computation</td>
</tr>
<tr>
<td>DT</td>
<td>Decision Tree</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>SAR</td>
<td>Synthetic Aperture Radar</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic information system</td>
</tr>
<tr>
<td>GR</td>
<td>Gain ratio</td>
</tr>
<tr>
<td>LiDAR</td>
<td>Light Detection and Ranging</td>
</tr>
<tr>
<td>MLP</td>
<td>Multi-Layer Perceptron</td>
</tr>
<tr>
<td>nDSM</td>
<td>Normalized Digital Surface Model</td>
</tr>
<tr>
<td>OBIA</td>
<td>Object-based Image Analysis</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>RF</td>
<td>Random Forest</td>
</tr>
<tr>
<td>RNN</td>
<td>Recurrent Neural Networks</td>
</tr>
<tr>
<td>ROC</td>
<td>Relative Operating Characteristic</td>
</tr>
<tr>
<td>RS</td>
<td>Remote Sensing</td>
</tr>
<tr>
<td>SPOT</td>
<td>Satellite Pour l’Observation de la Terre</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>DM</td>
<td>Data Mining</td>
</tr>
<tr>
<td>VHR</td>
<td>Very High Resolution</td>
</tr>
<tr>
<td>OOB</td>
<td>“Out-of-Bag</td>
</tr>
<tr>
<td>LULC</td>
<td>Land Use /Cover</td>
</tr>
<tr>
<td>BPNN</td>
<td>Back-Propagation neural network</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PNN</td>
<td>Probability Neural Network</td>
</tr>
<tr>
<td>UA</td>
<td>User Accuracy</td>
</tr>
<tr>
<td>PA</td>
<td>Producer Accuracy</td>
</tr>
<tr>
<td>OA</td>
<td>Overall Accuracy</td>
</tr>
<tr>
<td>Cor</td>
<td>Correlation</td>
</tr>
<tr>
<td>Stdev</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>GLCM</td>
<td>Gray-Level Co-occurrence Matrix</td>
</tr>
<tr>
<td>GLDV</td>
<td>Grey level difference vector</td>
</tr>
<tr>
<td>TP</td>
<td>True Positives,</td>
</tr>
<tr>
<td>TN</td>
<td>True Negatives</td>
</tr>
<tr>
<td>FP</td>
<td>False Positives</td>
</tr>
<tr>
<td>FN</td>
<td>False negatives</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Introduction

The rapid expansion of cities and the continuously increasing population in urban areas lead to the establishment of settlements in mountainous areas. This phenomenon has increased the impact of natural disasters, particularly landslides, in these mountainous areas. Landslides result in severe property losses, human casualties, and environmental damage. Landslides are considered one of the geological phenomena – along with earthquakes, subsidence, volcanic activity and glacial rebound – to contribute to ground deformation (Pradhan and Lee, 2010). In the near future, climate change induced by anthropogenic actions such as urban expansion and deforestation, will lead to higher landslide occurrences. Considering the wide coverage of landslide damages, planners and decision makers have to identify location of landslide to mitigate the event (Pradhan et al., 2011). Two major aspects should be considered to produce an efficient landslide inventory map. One aspect is related to the data used, whereas the other is related to the methods employed (Pradhan et al., 2016).

Landslides are geological disasters with catastrophic effects on human lives and properties. Landslide incidence is related to a cluster of triggering factors, such as intense rainfall, volcanic eruptions, rapid snowmelt, elevated water levels, and earthquakes (Tehrany et al., 2014a). Landslide inventory maps are required for various purposes such as: (i) recording the magnitude of landslide in a region, (ii) performing the initial steps in analysing the susceptibility, hazard, and risk of the landslide (Van Westen et al., 2008; Guzzetti et al., 2012; Althuwaynee et al., 2014a; Althuwaynee et al., 2014b; Pradhan et al., 2016), and (iii) examining the patterns of landslide distributions and studying the evolution of landscape affected by landslides (Parker et al., 2011; Pradhan et al., 2016). Mapping of landslide susceptibility, hazards, and risks is imperative to achieve the mitigation planning (Pradhan and Buchroithner, 2010). The identification and mitigation of landslide occurrences are also crucial within the field of hazard management and mitigation research (Guzzetti et al., 2012).

According to Metternicht et al. (2005), remote sensing (RS) technology provides detail information about landslides for policy-makers and emergency managers. Selecting relevant features is important for distinguishing between landslides and non-landslides as well as classifying landslides (Van Westen et al., 2008). Mapping a landslide inventory in tropical areas is challenging because the dense vegetation cover in these regions obscures underlying landforms (Chen et al., 2015). The precision of derived maps depends not only on the methodology adopted but also on the quality of the features selection. The use of a large number of features can decrease the accuracy of landslide detection because of the presence of irrelevant features (Stumpf and Kerle, 2011). Moreover, the dense vegetation of tropical forests complicates the generation
of an accurate landslide inventory map in these regions. Therefore, the present study aims to utilize light detection and ranging (LiDAR) data and orthophoto image along with recent optimized techniques, in detecting landslides and differentiating between their characteristics.

1.2 Problem Statement

Field survey, optical remote sensing, and synthetic aperture radar techniques are traditional techniques use for landslide detection in Remote Sensing and Geographic Information System (GIS) (Pradhan et al., 2016). However, there exist some challenges that affect its accuracy and performance. In a dense vegetation cover, it is difficult to use the traditional technique due to the obstruction from mountains and shadows (Chen et al., 2014). Many researches have attempted to develop rule-set based on time trial-error method. However, the rule-set development with the trial-and-error optimization method is time consuming, and the optimum rulesets are difficult to be determined (Sameen et al., 2017). Also, many investigations have been carried out using hierarchical approach with multiresolution sets of images. However, this approach has not been adopted in differentiating between types of landslide (shallow and deep seated) using LiDAR data. Despite the viability of Neural Architecture techniques in many areas of endeavour such as GIS, the technique has not been deployed with LiDAR data in detecting landslide.

1.3 Motivation Behind this Research

Nowadays, natural hazards are common in today’s life. Increasing amounts of natural catastrophes have proved to the human the vital importance of the natural hazards issues for the safety of the environment, and the populations. Rapid urbanization and climate change are expected to raise the amount of landslide. Landslide results in sever property losses, human casualties and environmental damage. The dramatic landslides which occur in tropical countries, especially Malaysia, emphasize the extremity in climatic variations. Since large-scale landslides often cause severe damage to human lives and properties, it is desirable to identify and map them for hazard mitigation purposes (Lin et al., 2013a). This phenomenon triggers due to the unexpected variations of the state of features due to natural forces, such as rainfall, and human activities. In most cases, human beings are not capable of controlling and predicting these disasters precisely. When major natural catastrophes such as landslides, earthquakes, floods and land subsidence occur, they have a definite impact upon human lives, belongings, infrastructure, farming and environment. The influence of natural hazards varies based on the intensity and coverage of region.

Landslides are of the most commonly occurring natural catastrophes that influence human beings and their surrounding environment. Inhabitants of Asia and the Pacific regions are affected in terms of social and economic stability. According to Pradhan (2010), high percentage of the destructions related to natural catastrophes in Malaysia are produced by landslide. Furthermore, average annual landslide damage is as high
as $10 million USD. The attention towards providing proper landslide management has risen over the last few centuries. The recent reasons for recurrent landslides within some regions are mostly due to unplanned urbanization, construction and deforestation. In spite of these factors, it is possible to predict and mitigate effects of landslides with the use of various technology. The use of technology can facilitate landslide prevention actions to detect the landslide areas and to have an early warning for this catastrophe.

This thesis attempts to propose novel techniques for mapping the landslide in vulnerable locations using state-of-the-art machine learning methods. The key motivation of this research is to establish high-quality landslide inventory maps by applying semi-automatic techniques. To recognize those susceptible regions, landslide inventory maps should be generated as a basis of landslide susceptibility mapping in order to avoid more urbanization in hazardous areas. To reduce the damage and victims in case of a landslide occurrence, it is critical to locate the susceptible areas. Additionally, optimized techniques vastly improve the accuracy of classifications. Governments and planners can utilize the generated results of this study to identify safe regions for inhabitants, support the first ones to respond to emergencies, and update the urban planning strategies. Such data can decrease the requirements for performing field surveys by agencies such as departments of surveying.

1.4 Research Objectives

The present thesis proposes and applies various new methods that clearly contribute to the gap in the literature. These methods are simple, repeatable, and comprehensive. The following are the main objectives of the thesis:

1. To investigate the suitable selection features algorithms in landslide detection.
2. To optimize rule sets to distinguish various landslide characteristics and compare its performance with hierarchical rule-set classification approach.
3. To apply deep learning (Neural Architecture) classifier for landslide mapping.

1.5 Research Questions

This thesis comprehensively addresses the following research questions:

1. What are the results of the selected feature techniques using high resolution LiDAR data? How do the selected relevant features contribute to detection of landslides and differentiation of two landslide types in a highly dense, vegetated area?
2. What is the result of the optimize rule data with high density LiDAR data? How can this be helpful in detecting landslides and differentiating between two types of landslide?
3. How effective is the supervised approach classification method in defining the landslide prone areas?
4. How do the LiDAR derived data aid in differentiation of deep seated and shallow landslides?
5. How does the optimized segmentation aid in delineating the boundary of landslides and improve in its exploitation?
6. How do the decision rule sets aid to distinguish between two types of landslides?
7. How do the hierarchal rules set contribute to differentiation of two types of landslides?
8. How do deep neural networks help in detecting landslides within tropical areas?

1.6 Scope of this Thesis

This study is carried out to develop a method for detecting and distinguish landslide areas using remote sensing and GIS techniques in a tropical environment like Cameron Highland, Malaysia. Also, transferability capability to similar geological and geomorphological areas are investigated and tested in other study areas. In this research, airborne laser scanning techniques is used to detect landslides and to distinguish between shallow and deep-seated landslides. Six (6) algorisms; Ant Colony Optimization (ACO), Gain Ratio (GR), Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), Random forest (RF), and Correlation-based Feature Selection (CFS) are used for the feature selection. The performance of the feature selection algorism is evaluated using two Machine learning classifiers, namely support vector machine (SVM) and random forest (RF). Decision rule and hierarchal rules set are developed by applying decision tree (DT) to provide improved landslide inventory map. In this task, two neural network algorithms, Recurrent Neural Networks (RNN) and Multi-Layer Perceptron Neural Networks (MLP-NN) are used and the hyper-parameters of the network architecture was optimized based on a systematic grid search.

1.7 Research Hypothesis

1. The optimized techniques for both segmentation parameters and feature selection may improve and facilitate the generation of landslide inventory map with high density LiDAR data.
2. Decision rule and hierarchal rule sets along with relevant features may improve significantly in differentiating between the landslide types.
3. Deep neural network may contribute in improving accuracy of a landslide inventory map.
1.8 Thesis Organization

This thesis is organized into five chapters. The summary of each chapter is as follows:

(i) CHAPTER 1: INTRODUCTION
This chapter introduces the problem statement of the study, aim, objectives and scope of the study. Furthermore, this chapter highlights the research questions proposed for the thesis as well as the significant contribution of new knowledge and the overall structure of the thesis.

(ii) CHAPTER 2: LITERATURE REVIEW
This chapter provides an overview of landslide status in various regions and previous work of using remote sensing and GIS to detect landslides. The two types of landslide detection techniques were outlined followed by the traditional and innovative and emerging techniques for detecting landslide prone areas. Next, the methodology used to detect and distinguish landslides using different techniques were discussed. Finally, validation methods used to assess the accuracy of maps produced are summarized.

(iii) CHAPTER 3: MATERIALS AND METHODOLOGY
This chapter details the characteristics of the study area followed by the following details in order: materials, data, methodology, optimized segmentation, evaluation of the training sample, features selection algorithms, detection, differentiation of the types of landslide (i.e. deep seated and shallow), and various remote sensing techniques.

(iv) CHAPTER 4: RESULTS AND DISCUSSION
This chapter focuses on the results of the study, which include the analysis results, optimized segmentation, and selected feature algorithms using LiDAR data – supported by photos, tables, diagram and charts. This chapter also discusses the results of optimized segmentation. Afterwards, decision rule and hierarchal rules were used in differentiating between two types landslides with regards to their characteristics. Lastly, the results of the deep neural network – using only the LiDAR derived data – were employed for detecting landslides and differentiating between the two types of landslides are discussed.

(v) CHAPTER 5: CONCLUSION AND FUTURE WORK RECOMMENDATIONS
This chapter provides the overall conclusion from this study, recommendation and further research for the study area.
REFERENCES


134


Chen, Y., Su, W., Li, J., Sun, Z. (2009). Hierarchical object based classification using very high resolution imagery and LiDAR data over urban areas. Advances in Space Research, 43(7), 1101-1110.


Donati, L., Turrini, M.C. (2002). An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy). *Engineering Geology*, 63(3), 277-289.


Han, J., Pei, J., Kamber, M. Data Mining: Concepts and Techniques; Elsevier: Amsterdam, The Netherlands, 2011.


