

**Landslide Mapping from Remote Sensing Images Using
Deep Learning: A Case Study in Lantau Island, Hong Kong**

TSANG, Ho Ming

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy

in

Earth and Atmospheric Sciences

The Chinese University of Hong Kong

August 2022

Thesis Assessment Committee

Professor CHAN, Man Nin (Chair)

Professor LIU, Lin (Thesis Supervisor)

Professor TAN, Yen Joe (Committee Member)

Abstract of thesis entitled:

Landslide Mapping from Remote Sensing Images Using Deep Learning: A Case Study in Lantau Island, Hong Kong

Submitted by TSANG, Ho Ming

for the degree of Master of Philosophy in Earth and Atmospheric Sciences

at The Chinese University of Hong Kong in August 2022

Abstract

Hong Kong, a densely populated city on hilly terrains, is prone to landslides especially after rainstorms. Landslide inventories in Hong Kong are currently created and analyzed manually. However, such manual effort is time-consuming and subject to operators' decisions. With recent advancements in deep learning networks, it is motivated to develop an efficient and objective deep learning-based framework for delineating landslide boundaries from high-resolution remote sensing imagery.

In this work, we utilize and compare four deep convolutional neural network (DCNN) architectures for semantic segmentation tasks: U-Net, DeepLabv3+, U²-Net, and our proposed network. Our proposed network has a U-structure with skipped connections at all downsampled resolutions, and an improved dilated spatial pyramid pooling (DSPP) module at each of these skipped connections. We first pre-process the digital orthophoto in 2008 over Lantau Island and the landslide inventory map in 2011 to generate a training dataset. We then train our DCNN models to obtain optimized models. Next, we perform inference on the original DOP using the optimized models. We post-process the outputs with conditional random fields to obtain a landslide probability map. Digital surface model is provided as a filter for an improved accuracy.

The best model we have obtained is our proposed model, which has achieved satisfactory accuracies by count: a precision rate 0.745, a recall rate of 0.869, and an F1 score of 0.802 (by area: precision of 0.874, recall of 0.912, F1 score of 0.893). Our proposed model can accurately delineate the boundaries of large- ($\text{area} > 346 \text{ m}^2$) and medium-scale ($142 \text{ m}^2 < \text{area} < 346 \text{ m}^2$) landslides. The digital surface model filter is effective for removing two thirds of the false predictions of non-landslide instances.

Results from this work show the potential of deep learning applications for improving the efficiency and accuracy of mapping landslides in automated manners over large areas. Future work can proceed on using similar remote sensing images acquired over Hong Kong for territory-wide landslide mapping and yearly images for more frequent updates of inventories, towards routine and effective monitoring of natural disasters.

摘要

香港是一座地處丘陵地帶的人口稠密的城市，斜坡在暴雨過後特別容易發生山泥傾瀉。香港目前的山泥傾瀉清單是由人手創建和分析。然而這種人工操作非常耗時，並且取決於操作員的決定。隨著深度學習神經網絡的進展，開發一項有效且客觀的深度學習的框架，用於從高分辨率遙感圖像中描繪山泥傾瀉的邊界，具有一定價值。

在這項研究中，我們運用並比較了四種用於語義分割任務的深度卷積神經網絡 (DCNN) 的架構，分別是：U-Net、DeepLabv3+、U²-Net 和我們提出的網絡。在我們提出的網絡中，有一個 U 型結構，在所有縮小的分辨率下都有跳過的連接，並且在每個這些跳過的連接處都有一個改良的擴張空間金字塔池 (DSPP) 模塊。我們首先對 2008 年的大嶼山數碼正射影像和 2011 年的山泥傾瀉地圖進行預處理，以生成訓練數據集。然後我們訓練 DCNN 模型以獲得優化模型。接下來，我們使用這些優化模型對原始數碼正射影像進行推理。我們使用條件隨機場對輸出進行後處理，以獲得山泥傾瀉概率圖。數碼地表模型作為過濾能提高精準度。

我們得到的最好的模型是我們提出的模型，它取得了令人滿意的準確率，按數量：精確率 0.745，召回率 0.869，F1 值 0.802（按面積：精確率 0.874，召回率 0.912，F1 值 0.893）。我們提出的模型可以準確地描繪大至中等規模的山泥傾瀉的邊界。數碼地表模模型過濾可有效排除錯誤預測中，三分之二的非山泥傾瀉按例。

研究結果顯示了深度學習應用程序在提高大面積自動繪製山泥傾瀉地圖的效率方面和準確性方面的潛力。未來的工作可使用每年的香港遙感圖像進行全港性山泥傾瀉測繪，和年度更新山泥傾瀉庫存地圖，以實現對自然災害的常規和有效監測。

Acknowledgements

This thesis would not have been possible without the support of many people. Thank you to my supervisor, Professor LIU Lin, for his patience, guidance, and support. He read my numerous revisions and helped make some sense of the confusion. His encouraging words is a source of inspiration. Without his suggestion of this project, I may not find another project which best suits my engineering and mathematics background.

Thank you to, Professor CHAN Man Nin, and Professor TAN Yen Joe, for serving as my thesis assessment committee member. They generously took time out of their schedules to review my thesis and participate in my oral defense. Their thoughtful and detailed feedback have been very important to me.

Thank you to groupmates, Dr. HUANG Lingcao, for his pioneer development of codes that automate experiments, and Dr. ZHANG Enze, Dr. HU Yan, Miss. XU Xingyu (Carol), Miss. XIA Zhuoxuan (Summer), Mr. SUN Zhangyu (Joe), for their comments and suggestions.

Finally, I would like to express my deepest gratitude to my mother for her unconditional, unequivocal, and loving support.

Table of Contents

Abstract	i
摘要.....	iii
Acknowledgements	iv
Table of Contents	v
List of Tables.....	vii
List of Figures	viii
1. Introduction.....	1
1.1. Landslide Risks in Hong Kong	1
1.2. Recent Research Studies on Landslide Mapping	3
1.3. Semantic Segmentation in Remote Sensing Images	4
1.4. Motivations and Objectives of This Thesis	7
1.5. Thesis roadmap.....	7
2. Study Area and Data	8
2.1. Lantau Island	8
2.2. Digital Orthophoto in 2008	9
2.3. Territory-wide Airborne Light Detection and Ranging Survey in 2010	10
2.4. Landslide Inventory Map in 2011	11
3. Data Derivatives	13
3.1. LiDAR-derived Digital Surface Model	13
3.2. Statistics of Surface Geometry of Landslides in our Dataset	17
4. Methods	20
4.1. Overview of Automated Pipeline for Landslide Mapping	20
4.2. Pre-processing: Data Augmentation.....	22
4.3. Training and Inference: Deep Convolutional Neural Network Models	23
4.4. Post-processing: Conditional Random Fields	29
4.5. Post-processing: Digital Surface Model as Filter	30

5. Results.....	32
5.1. Overview of Inferred Landslides from DCNN Models.....	32
5.2. Accuracy Metrics of Inferred Landslides from DCNN Models	41
5.3. Comparison with Other Research Works	43
6. Discussion	44
6.1. Mapping Performance Breakdown by Size Groups	44
6.2. Success Mode: True Positive over Large Landslides and True Negative over Engineered Surfaces	45
6.3. Failure Mode: False Positive over Natural Surfaces and False Negative over Small Landslides	50
6.4. Choices of DCNN Architectures and Hyperparameters.....	54
6.5. Choices of Surface Geometry Thresholds in Post-processing.....	56
7. Conclusions and Future Work	58
Appendix	59
Bibliography.....	60

List of Tables

Table 3.1. Statistics of Surface Geometry of Landslides.....	19
Table 4.1. Empirical thresholds for filtering false landslide geometry.....	31
Table 5.1. Summary of accuracy metrics by count of four DCNN models before DSM filter post-processing.....	42
Table 5.2. Summary of accuracy metrics by count of four DCNN models after DSM filter post-processing.....	42
Table 5.3. Summary of accuracy metrics by area of four DCNN models	42
Table 5.4. Summary of results by area from other research works and this work....	43
Table 6.1. Summary of landslide instance removals by numbers by applying filters based on surface geometry thresholds.....	56
Table A.1. Full Statistics of Surface Geometry of Landslides.....	59

List of Figures

Figure 1.1. Post-disaster photo of Po Shan Road landslide.	2
Figure 1.2. Land cover segmentation results on ultrahigh-resolution NIR-R-G image Vaihingen dataset by Volpi & Tuia (2016).....	5
Figure 1.3. Segmentation results on Landsat images over Munich Airport during 1985 –2014 period by Lyu et al. (2018).....	6
Figure 2.1. Overview of Lantau Island on OpenStreetMap.	8
Figure 2.2. Digital Orthophoto acquired in 2008 of the study extent.	9
Figure 2.3. Overview of landslide location distribution over Lantau Island.....	11
Figure 2.4. Example locations of detailed boundaries of landslides and example locations of non-landslides.....	12
Figure 3.1. Light Detection and Ranging (LiDAR) digital surface model (DSM) in 2010 of the study extent.	13
Figure 3.2. Similar to Figure 3.1 but for the slope angles.....	15
Figure 3.3. Similar to Figure 3.1 but for the aspect angles.	15
Figure 3.4. Similar to Figure 3.1 but for the profile curvatures (Pcurv).	16
Figure 3.5. Similar to Figure 3.4 but for the tangential curvatures (Tcurv).	16
Figure 3.6. Elevation, slope, profile curvature, and tangential curvature raster maps of landslides in area A (Nga Ying Shan).	17
Figure 3.7. Distributions of Surface Geometry of Landslides.	18
Figure 4.1. Overview of Automated Landslide Mapping Pipeline.	21
Figure 4.2. Data augmentation methods.....	22
Figure 4.3. U-Net architecture (Ronneberger et al., 2015).....	24
Figure 4.4. DeepLabv3+ architecture (Chen et al., 2018).....	25
Figure 4.5. U ² -Net architecture (Qin et al., 2020).....	26
Figure 4.6. Architecture of our proposed network (This work).	27
Figure 4.7. Architecture of our proposed dilated spatial pyramid pooling (DSPP) module.....	28

Figure 5.1. Overview of landslide segmentation maps by four DCNN models.	32
Figure 5.2. Landslide segmentation maps in area A (Nga Ying Shan).....	33
Figure 5.3. Similar to Figure 5.2 but for area B (Keung Shan).	34
Figure 5.4. Similar to Figure 5.2 but for area C (Shek Pik Reservoir).	35
Figure 5.5. Similar to Figure 5.2 but for area D (Ngong Ping Market).	36
Figure 5.6. Similar to Figure 5.2 but for area E (Sze Tsz Tau Shan).	37
Figure 5.7. Similar to Figure 5.2 but for area F (Wong Lung Stream).	37
Figure 5.8. Similar to Figure 5.2 but for area G (Nei Lak Shan Angle Station).....	39
Figure 5.9. Similar to Figure 5.2 but for area H (A Po Long and Wong Kung Tin). 40	
Figure 6.1. Landslide segmentation map by This work and elevation, slope, profile curvature, tangential curvature maps in area A (Nga Ying Shan).	46
Figure 6.2. Similar to Figure 6.1 but for area B (Keung Shan).	47
Figure 6.3. Similar to Figure 6.1 but for area C (Shek Pik Reservoir)	48
Figure 6.4. Similar to Figure 6.1 but for area D (Ngong Ping Market).	49
Figure 6.5. Similar to Figure 6.1 but for area E (Sze Tsz Tau Shan).....	50
Figure 6.6. Similar to Figure 6.1 but for area F (Wong Lung Stream).	51
Figure 6.7. Similar to Figure 6.1 but for area G (Nei Lak Shan Angle Station).....	52
Figure 6.8. Similar to Figure 6.1 but for area H (A Po Long and Wong Kung Tin). 53	
Figure 6.9. Evaluation breakdown over size groups and their accuracy metrics.....	44

1. Introduction

1.1. Landslide Risks in Hong Kong

Landslides are defined as any downslope movement of soil or rock under gravity. Types of movement can be classified into falls, topples, slides, spreads, flows, and any of their combinations, and types of geo-materials can be classified into rock, debris, and earth (Cruden & Varnes, 1996). Landslides can be triggered by extreme rainfall, earthquakes, volcanoes, forest fires and other mechanisms (Chatwin, 1994). Extreme rainfall is the major cause of landslides in Hong Kong. During or after extreme rainfall, slopes are saturated with water, and soil particles lose cohesion and friction. Soil and rock masses cannot withstand the original stress state, resulting in landslides.

Hong Kong is prone to landslide hazards and there were some serious landslide incidents leading to fatalities and injuries. The Sau Mau Ping landslide on 18 June 1972 buried Sau Mau Ping squatter village, and caused 71 dead and 60 injuries. The Po Shan Road landslide on 18 June 1972, the same day, brought down the 12-storey Kotewall Court and destroyed several other buildings (Figure 1.1). The disaster caused 67 dead and 20 injuries. The Shum Wan Road landslide on 13 August 1995 damaged three shipyards and a factory. 2 people were killed and 5 were injured. All these landslides were triggered by severe rainfall.

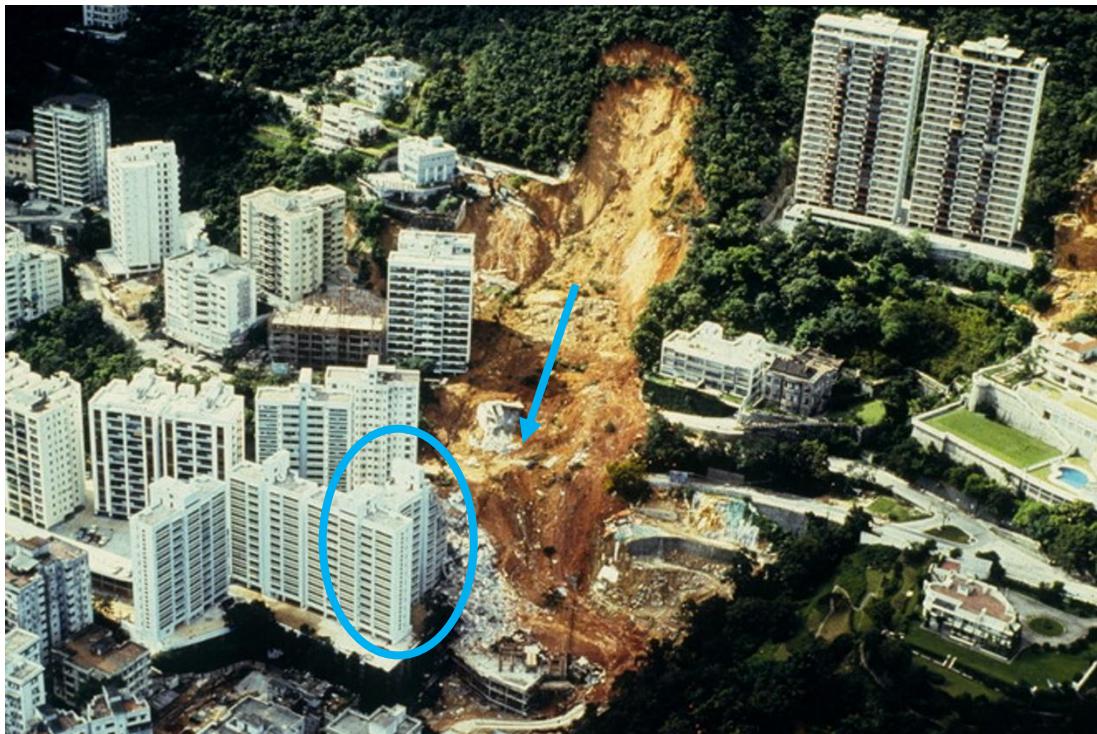


Figure 1.1. Post-disaster photo of Po Shan Road landslide. The collapse of Kotewall Court (arrow) hit and destroyed a corner of Greenview Gardens (circle). (Photo by Geotechnical Control Office (later renamed to Geotechnical Engineering Office), Civil Engineering and Development Department, Hong Kong Government)

Currently, the Geotechnical Engineering Office, Civil Engineering and Development Department (GEO, CEDD) maintains an Enhanced Natural Terrain Landslide Inventory (ENTLI). The ENTLI is an inventory of recent and relict natural terrain landslides in Hong Kong, updated every three years. Centrelines of landslides are identified and interpreted over aerial photographs, and are recorded into a geographical information system (GIS) (Dias et al., 2009). However, marking the centrelines alone is inadequate for analyzing the scale of landslides, such as areas and boundaries. Even though a polygon mapping is preferred for inventorying, manual interpretation on delineating boundaries of landslides is also subjective to operators' decisions and requires more labour and time.

1.2. Recent Research Studies on Landslide Mapping

Maps of landslide inventory, susceptibility, and hazard are three common types useful to the government and general public for landslide risks management (Highland & Bobrowsky, 2008). The focus of this work is to produce a landslide inventory map from post-disaster aerial photographs. A comprehensive and detailed landslide inventory is beneficial to understanding landslide mechanisms and processes, and developing regulative and preventive measures. Additional information can also be provided to landslide instances, including: geomorphic features of landslide (scarp, depletion zone, accumulation zone), depth of landslides (volume of material), geology (types of materials), activity status (active or relict), etc. Landslide inventory maps are created by comparing pre- and post-disaster images by manual effort or automatic methods. Recent research studies for automatic methods of landslide mapping can be classified into two types: pixel-based method and object-based method.

Pixel-based method. Pixel-based method is a single-shot classification, which classifies pixels at once. Most pixel-based methods are change detection on the spectral characteristics of pre- and post-landslide pixels, which means that they heavily rely on the availability of multitemporal images. However, this method cannot reject false features with similar spectral characteristics, and may produce salt-and-pepper false predictions. Yang and Chen (2010) proposed an image processing and vegetation change detection method on a pre-event Thematic Mapper image and post-event Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. Zhang et al. (2010) quantified a Normalized Difference Vegetation Index (NDVI) decline between Moderate Resolution Imaging Spectroradiometer (MODIS) and ASTER images, to be confirmed with terrain slope information. Mondini et al. (2011) applied change detection to very high resolution panchromatic and high resolution multispectral QuickBird images.

Object-based method. Object-based method is a two-shot classification, which first clusters similar pixels into objects by segmentation rules, then classifies the objects. Since objects are the basic element in this method, it helps to solve salt-and-pepper problem in pixel-based method. However, such segmentation rules are case-specific and lack of generalization for other regions in same dataset or other regions. Under-segmentation or over-segmentation may lead to inferior performance. Rau et al. (2013) utilized object-based image analysis (OBIA) with multilevel and hierarchical segmentation framework on high resolution multispectral imagery. Blaschke et al. (2014) used OBIA on NDVI, brightness and texture from SPOT-5 imageries in combination with a digital elevation model. Kurtz et al. (2014) proposed a top-down hierarchical approach based on OBIA to classify multimodal images of different resolutions at the same region.

1.3. Semantic Segmentation in Remote Sensing Images

Semantic segmentation refers to the task of grouping the regions of same class objects together in an image, which perfectly suits this work's objective of segmenting landslide areas and boundaries. Deep learning networks for semantic segmentation assign a semantic label to each pixel in the image, i.e., the prediction is a two-dimensional map of classes. These networks are fully convolutional, meaning that the size of output map is the same as the input image size and there is no specification on input image size. They are particularly useful for some remote sensing applications: land cover segmentation, land cover change detection, etc.

Land cover segmentation. Recent research studies had been conducted on urban land cover segmentation on multi-spectral, hyper-spectral, and SAR images. They aimed to segregate impervious urban surfaces, such as buildings and roads, from natural surfaces, such as vegetations and waterbodies. Kampffmeyer et al. (2016), Maggiori et al. (2017), and Volpi & Tuia (2016) performed segmentation on a dataset of ultrahigh-resolution NIR-R-G images (Figure 1.2). Chen et al. (2016), Hu et al. (2015), and Makantasis et al. (2015) performed segmentation on several hyperspectral imagery datasets. Duan et al. (2017), Zhang et al. (2017), and Zhou et al. (2016) performed segmentation on some datasets of SAR images, in particular polarimetric SAR.

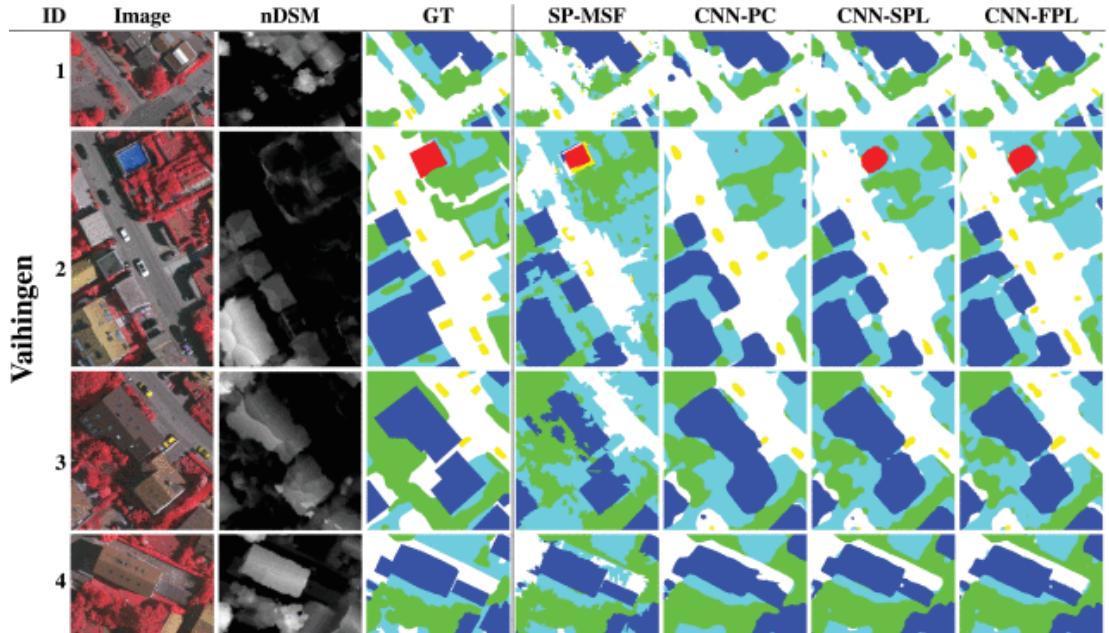


Figure 1.2. Land cover segmentation results on ultrahigh-resolution NIR-R-G image Vaihingen dataset by Volpi & Tuia (2016). This example shows the power of deep learning networks. nDSM: normalized digital surface model; GT: ground truth; SP-MSF: superpixels with multiscale features; CNN-PC: patch classification; CNN-SPL: semipatch labeling; CNN-FPL: full patch labeling.

Land cover change detection. Monitoring land cover change, both anthropogenic and natural, is important to urban planning and land resource management. Land cover change detection can be challenging because of colour variations and changing atmospheric conditions; however, deep learning is proven to be successful. Lyu et al. (2016) proposed a recurrent neural network to be convolved over bitemporal images, outputting a binary change or multi-class change map. Lyu et al. (2018) extended their application to long-term multitemporal images, to map urban area changes in four cities during 1984–2016 (Figure 1.3). Vakalopoulou et al. (2016) performed segmentation and change detection in multimodal data of different spatial, spectral, and temporal resolutions.

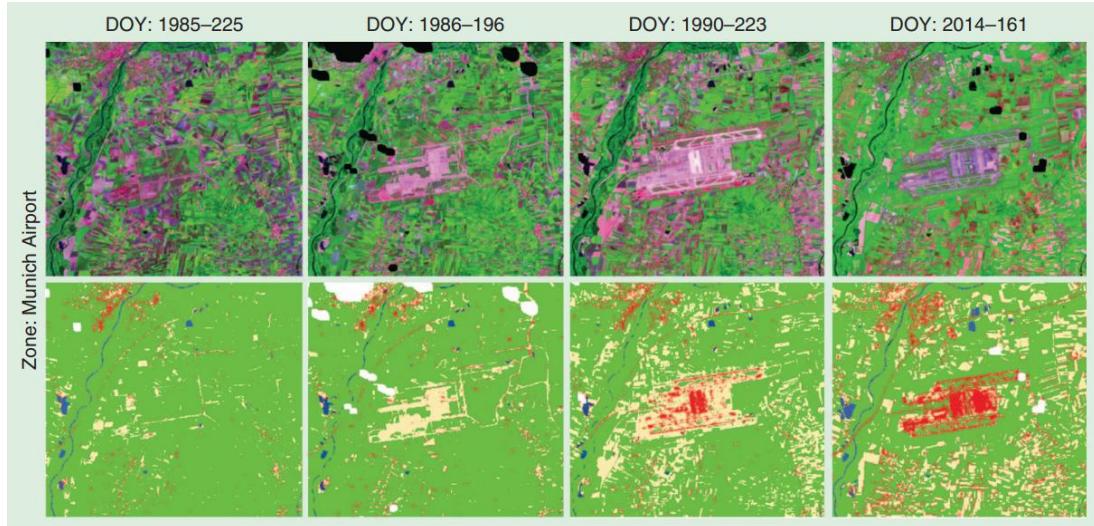


Figure 1.3. Segmentation results on Landsat images over Munich Airport during 1985–2014 by Lyu et al. (2018). This example shows long-term multitemporal land cover change detection can be successfully achieved by deep learning.

However, labelled datasets over natural terrain are often scarce while the urban area counterparts are readily available. Many applications have been conducted in urban settings, while researchers nowadays are extending the applications to natural terrains. With a landslide inventory map served as a ground truth label, this work presents a deep learning-powered method for segmenting landslide areas and delineating their boundaries over natural terrain.

1.4. Motivations and Objectives of This Thesis

In view of current limitations on the ENTLI in Hong Kong, there is a need to develop a framework for producing polygon mapping of landslide inventory. New innovations in deep learning frameworks and networks have achieved improved efficiency and accuracy in daily images. While they also showed promising results in land cover segmentation and land cover change detection, similar approaches can be applied in landslide inventory mapping (Shi et al., 2020; Su et al., 2021). An optimized network has great potential for generalizability of landslide features and transferability in many regions. Therefore, this thesis presents the following objectives:

- Develop an automated pipeline, which is easy to use, for landslide inventory mapping based on deep learning,
- Construct a deep convolutional neural network (DCNN) architecture for semantic segmentation on remote sensing images,
- Compare and evaluate performance of the proposed architecture with other existing DCNN architectures.

Through achieving these objectives, we are able to accurately map landslide boundaries in Lantau Island. This practical framework will promisingly be able to map landslides over the whole territory of Hong Kong.

1.5. Thesis roadmap

This thesis is organized as follows: we first describe our study area – Lantau Island, and our data – digital orthophoto (DOP), light detection and ranging (LiDAR), and ground truth landslide inventory map in Chapter 2; next, we derive a digital surface model (DSM) and infer geometrical properties of landslides in Chapter 3; we then describe our proposed automatic pipeline for mapping landslide, chosen DCNN architectures, use of a graph model – conditional random field (CRF), and post-processing filter by landslide geometries in Chapter 4; Chapter 5 presents our results; followed by a discussion of successful and failure cases, and limitation of the proposed method in Chapter 6; lastly, we conclude the thesis and lay out future directions in Chapter 7.

2. Study Area and Data

2.1. Lantau Island

Lantau Island is the largest island in Hong Kong. It is located in the Southwest of Hong Kong and mainly consists of mountainous terrain. The highest point of the island is Lantau Peak, with an elevation of 934 m. Other peaks include Sunset Peak (Tai Tung Shan) of 869 m, Lin Fa Shan of 766 m, Nei Lak Shan of 751 m, and Yi Tung Shan of 747 m. Half of the terrain in Lantau Island has a slope angle steeper than 25.1° . The lower and upper quartiles of the slope angles are 15.0° and 32.8° . Detailed maps of elevation and slope angles will be presented in Section 3.1.

The early development of Lantau Island can be traced back to the late 20th century with Tung Chung New Town. Other major infrastructure projects include the Hong Kong International Airport in 1998, Hong Kong Disneyland in 2005, and Ngong Ping 360 in 2006. Three major roads for commuting within Lantau Island include North Lantau Highway, Tung Chung Road, and South Lantau Road. The latter two are narrow and hilly, and are prone to landslide hazards. Figure 2.1 shows an overview of the peaks, major infrastructure locations, and roads of Lantau Island.

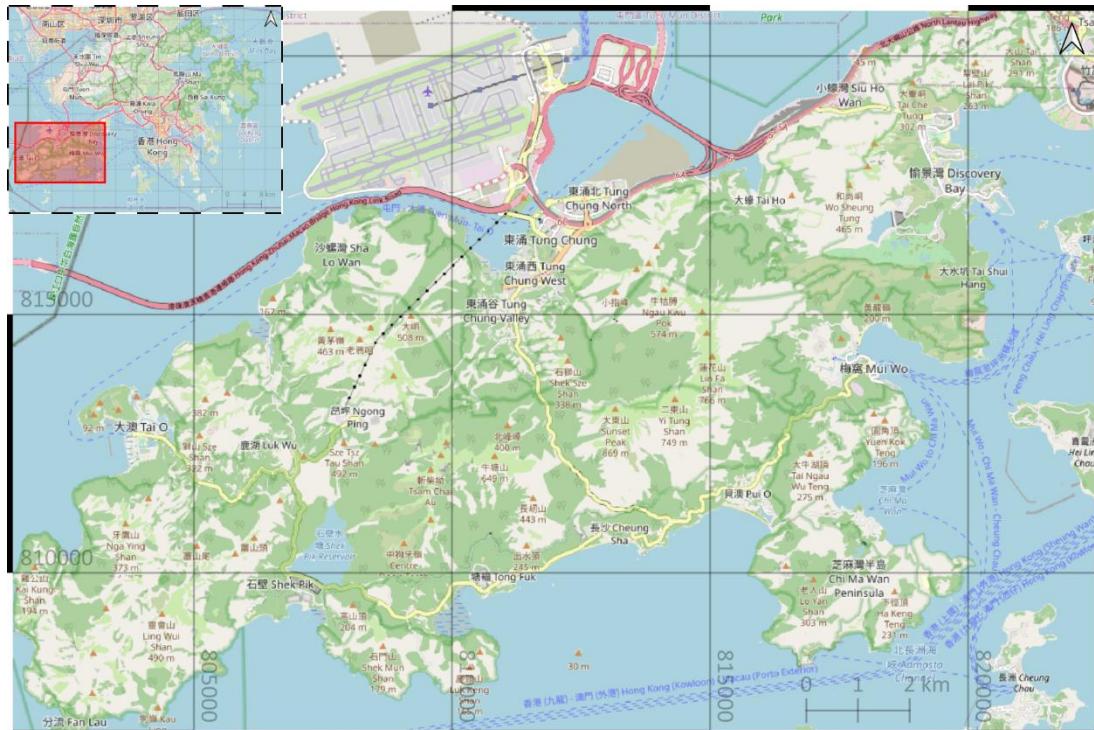


Figure 2.1. Overview of Lantau Island on OpenStreetMap.

2.2. Digital Orthophoto in 2008

The digital orthophoto (DOP) is a product of the Survey and Mapping Office, Lands Department (SMO, LandsD). It consists of three spectral bands: red, green, and blue (RGB). The original aerial photos were taken at a flying height of around 2400 m. The aerial photographs were then ortho-rectified to correct geometric distortions caused by topographic relief and camera tilt. The geometric rectification was only applied to the ground, elevated roads, and bridges, but not any buildings or trees. The resultant image map with coordinate reference in the Hong Kong 1980 Grid System (EPSG:2326) is the DOP.

This study uses the DOP in November 2008 with a very high spatial resolution of $0.5 \text{ m} \times 0.5 \text{ m}$. The DOP is in Enhanced Compression Wavelet Image (.ecw) format with associated geo-referencing in Earth Resource Mapper (.ers) format. The study extent is from $x_{\text{West}} = 801306 \text{ m}$ to $x_{\text{East}} = 822666 \text{ m}$, from $y_{\text{South}} = 806956 \text{ m}$ to $y_{\text{North}} = 820876 \text{ m}$. Figure 2.2 shows the DOP in 2008 covering the study extent.



Figure 2.2. Digital Orthophoto acquired in 2008 over the study extent (image source: SMO, LandsD).

2.3. Territory-wide Airborne Light Detection and Ranging Survey in 2010

The Geotechnical Engineering Office, Civil Engineering and Development Department (GEO, CEDD) conducted a territory-wide airborne light detection and ranging (LiDAR) survey from 1 Dec 2010 to 8 Jan 2011. The dataset contains the point cloud data in the LASer (.las) format and is open for public use. The product was geo-referenced to the Hong Kong 1980 Grid System (EPSG:2326) and the Hong Kong Principal Datum. Laser pulses were emitted and collected by the LiDAR equipment. With accurate positioning devices and measurement of the time difference between the emitted and received laser pulses, positions of ground and above ground objects can be detected.

2.4. Landslide Inventory Map in 2011

A landslide inventory map was provided by the Geotechnical Engineering Office, Civil Engineering and Development Department (GEO, CEDD). Government experts visually interpreted the DOP and produced a landslide polygon map in 2011. It represents landslide locations and boundaries over Lantau Island in 2008. In total, 1500 polygons are chosen to serve as ground truth data for the training of deep convolutional neural network (DCNN) models. Total area of the ground truth polygons is 726 938 m². Figure 2.3 shows an overview of landslide location distribution over Lantau Island. Figure 2.4 shows some example locations of detailed boundaries of landslides and some example locations of non-landslides.

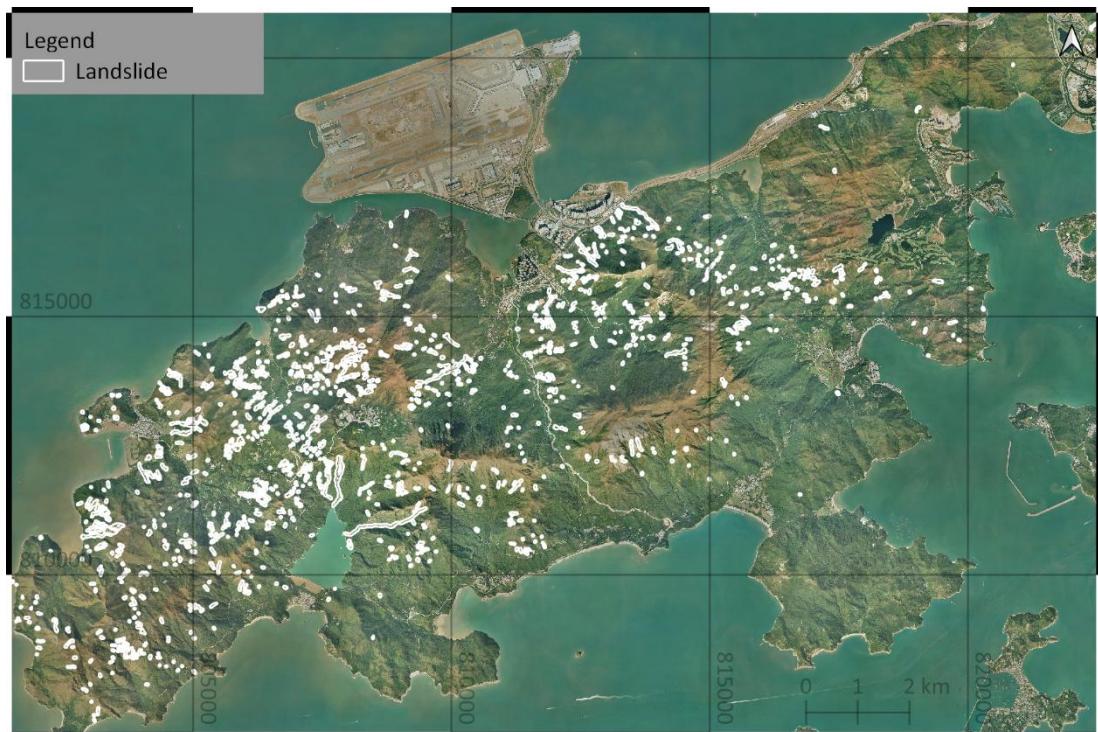


Figure 2.3. Overview of landslide location distribution over Lantau Island. The number of the ground truth polygons is 1500. The total area of the ground truth polygons is 726 938 m².

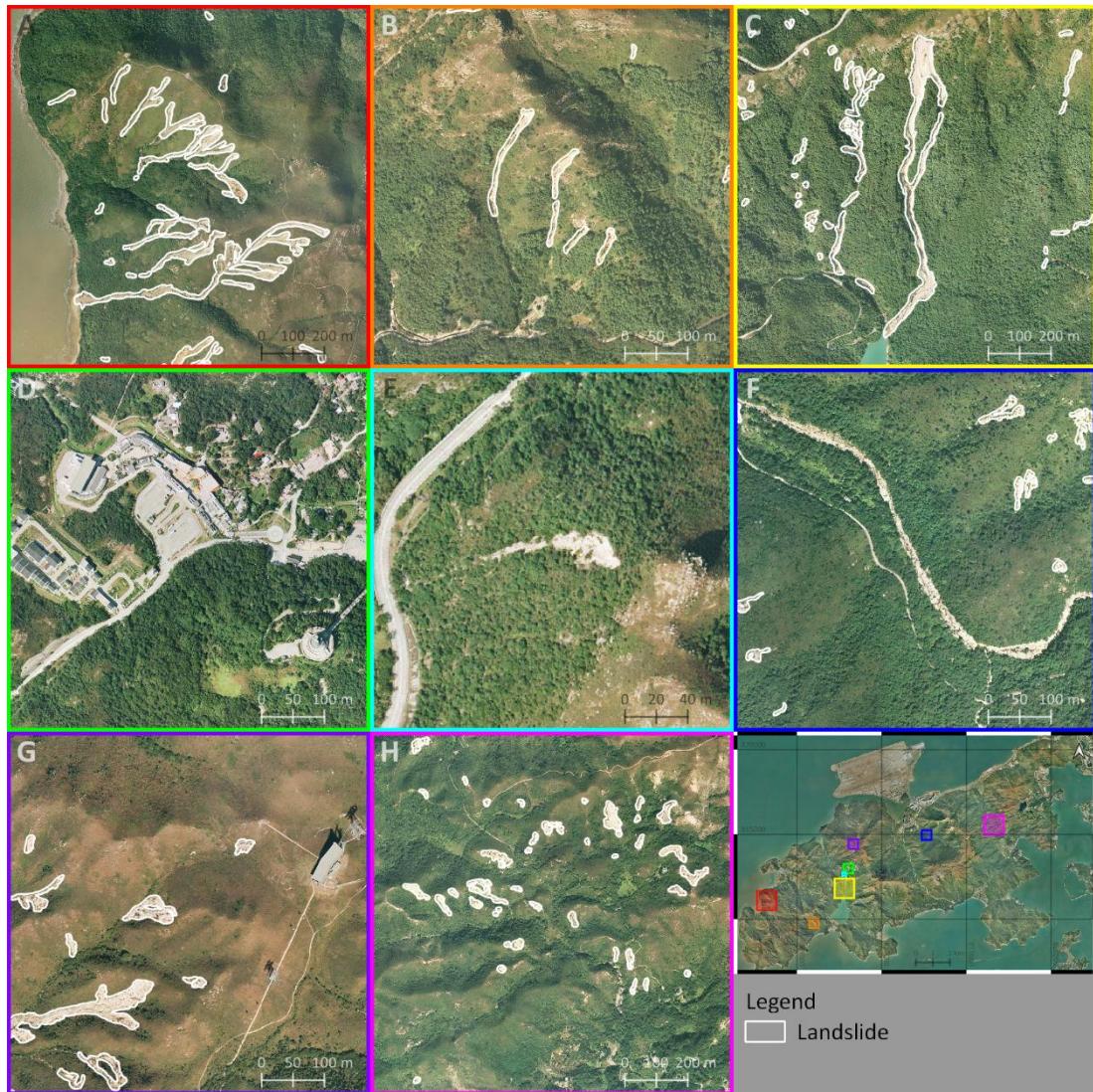


Figure 2.4. Example locations of detailed boundaries of landslides and example locations of non-landslides. Five examples of detailed boundaries of landslides: area A (Nga Ying Shan), area B (Keung Shan), area C (Shek Pik Reservoir), area G (Nei Lak Shan Angle Station), and area H (A Po Long and Wong Kung Tin). Three example locations of non-landslide objects: area D (Ngong Ping Market), area E (Sze Tsz Tau Shan), and area F (Wong Lung Stream, few small landslides within this scene).

3. Data Derivatives

Using the territory-wide LiDAR survey, we derived a digital surface model and geometric statistics of landslides. Such geometric properties are useful in distinguishing landslide objects against non-landslide objects in the outputs of our deep learning models, so that non-landslide objects can be rejected.

3.1. LiDAR-derived Digital Surface Model

The LAS LiDAR point clouds are imported into a raster map. The extent and resolution of the raster map are set to match the DOP. In each cell, the minimum of z-coordinates of points from any last return is selected as the raster cell value. Empty cells (cells with no LiDAR points) are filled with inverse distance weighting (IDW) interpolated values. The resultant raster map is a digital surface model (DSM), which includes buildings, vegetation, and ground surfaces. Figure 3.1 shows the surface elevation of the LiDAR DSM in 2010 of the study extent.

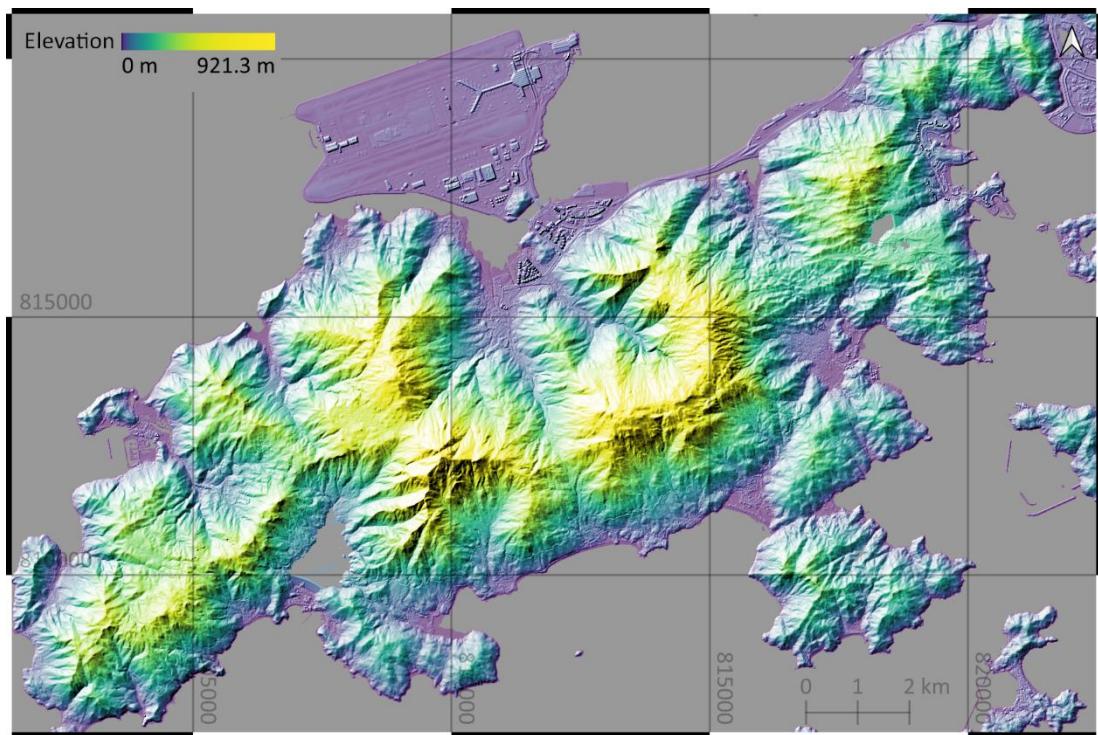


Figure 3.1. Light Detection and Ranging (LiDAR) digital surface model (DSM) in 2010 of the study extent.

From the DSM, raster maps of slopes, aspect angles, profile curvatures, and tangential curvatures can be generated. These topographic attributes, especially the curvatures, are important for landslide studies. Profile curvature measures the rate of change of slope along the gradient direction (or equivalently, perpendicular to contour). Negative profile curvatures are concave upward and indicate decelerated flow of soils over the surface. Tangential curvature measures the rate of change of slope perpendicular to gradient direction (or equivalently, tangential to contour). Negative tangential curvatures are concave upward and indicate convergent flow of soils over the surface. Figure 3.2 to 3.5 shows the slope, aspect, profile curvature, and tangential curvature raster maps.

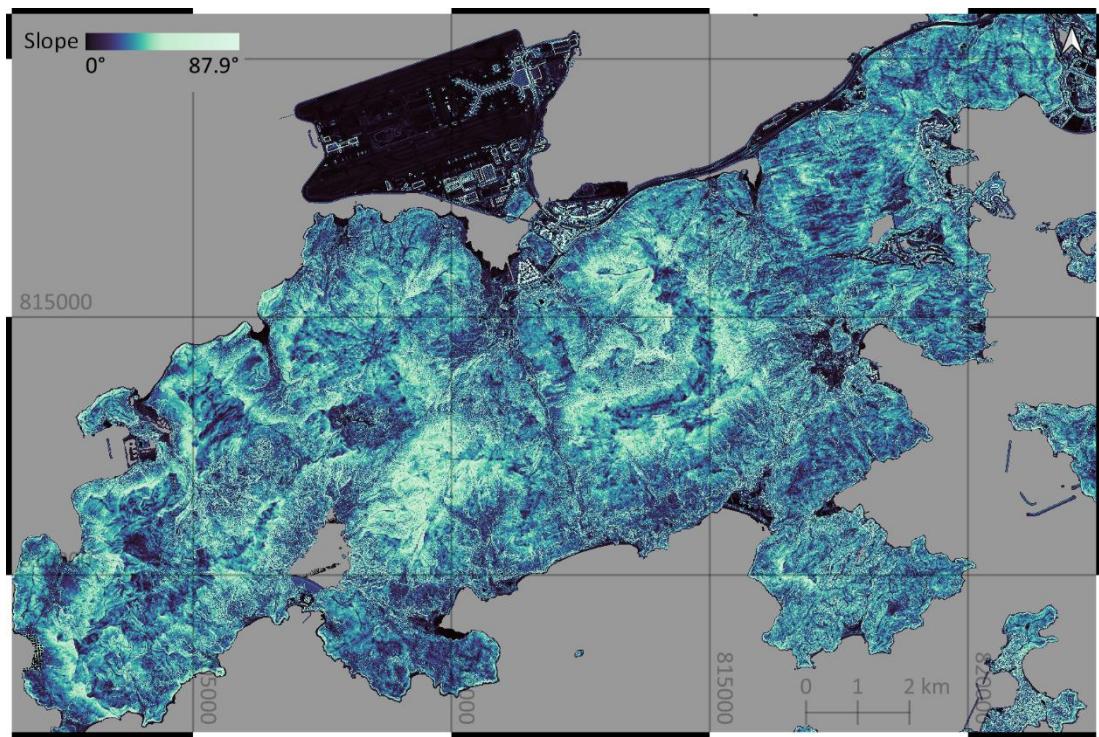


Figure 3.2. Similar to Figure 3.1 but for the slope angles.

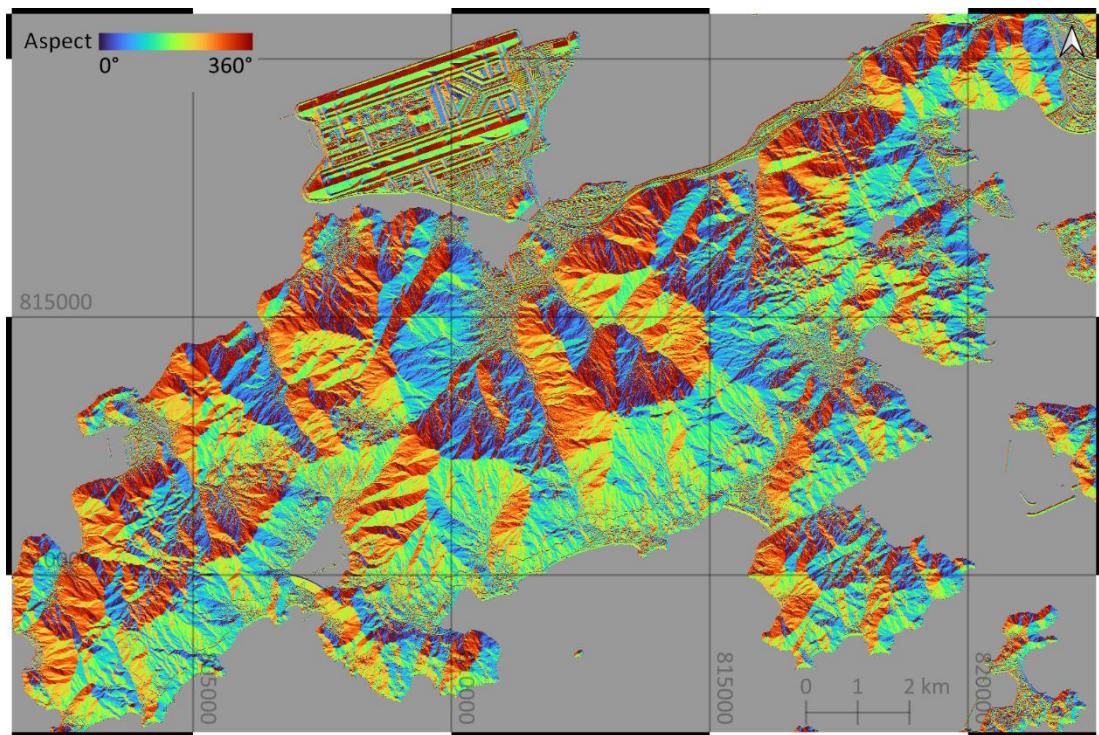


Figure 3.3. Similar to Figure 3.1 but for the aspect angles.

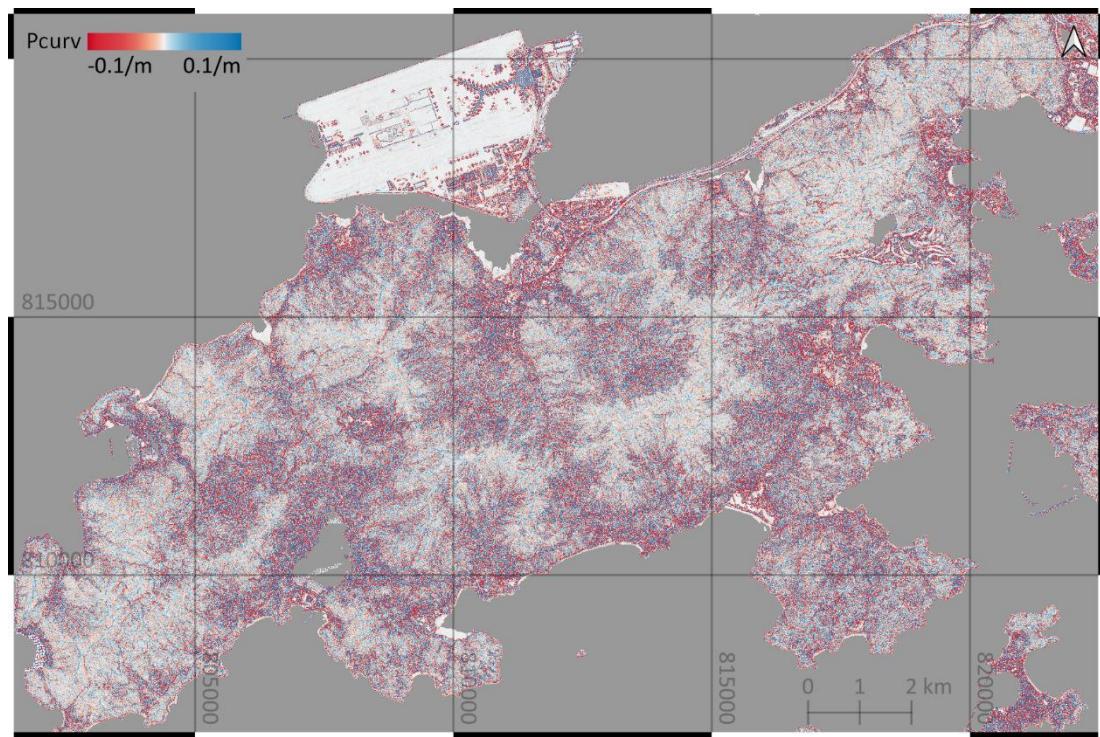


Figure 3.4. Similar to Figure 3.1 but for the profile curvatures (Pcurv).

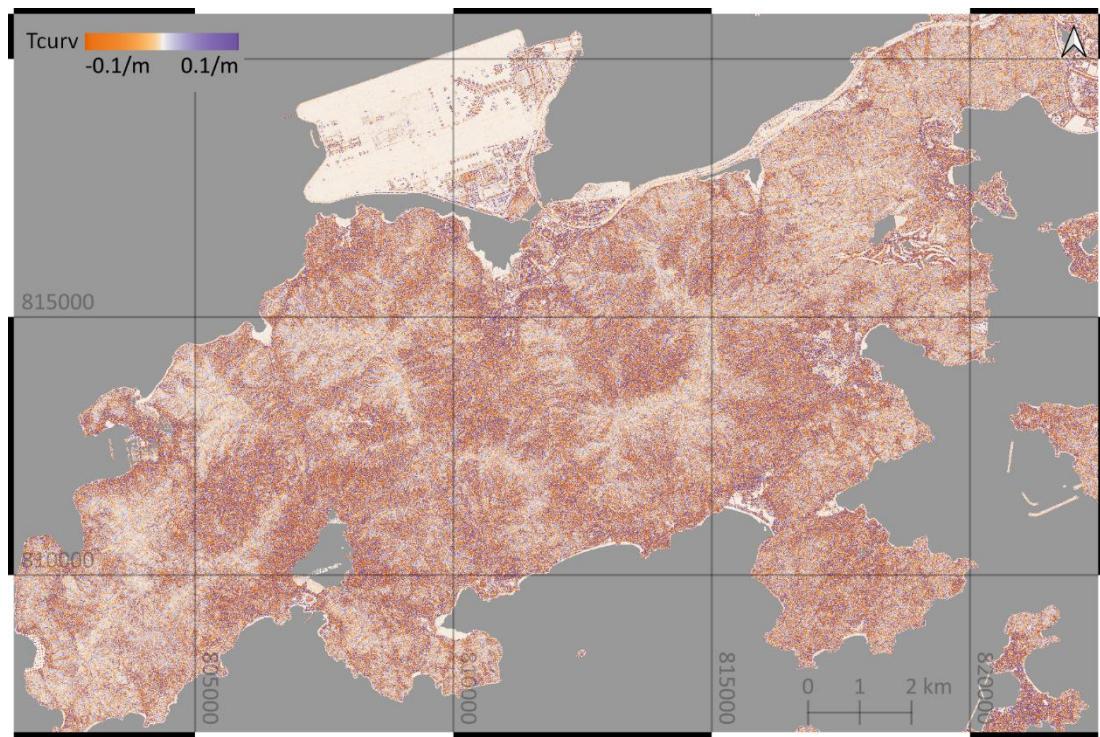


Figure 3.5. Similar to Figure 3.4 but for the tangential curvatures (Tcurv).

3.2. Statistics of Surface Geometry of Landslides in our Dataset

Combining the landslide map (Section 2.4) and LiDAR derived elevation, slope, profile curvature, and tangential curvature raster maps (Section 3.1), geometric properties of the ground-truth landslides can be determined. Figure 3.6 shows an example of surface geometry raster maps and landslide map on Nga Ying Shan. Negative profile curvatures, coloured in red, are concave. Negative tangential curvatures, coloured in orange, are concave upward.

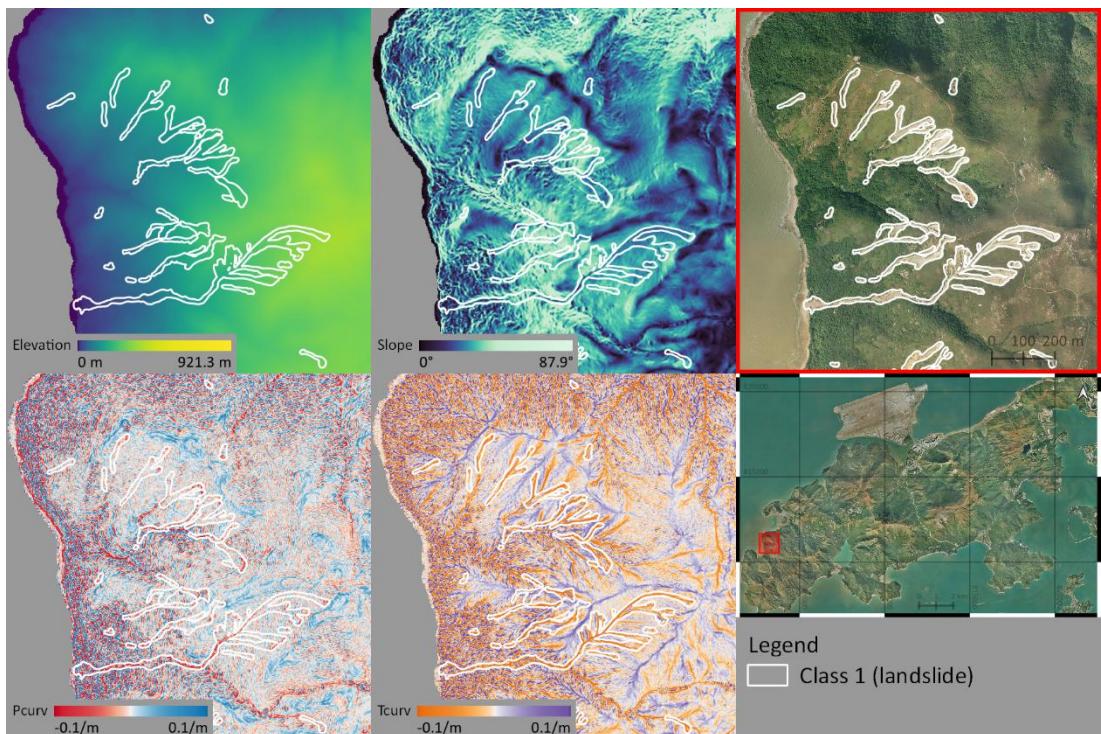


Figure 3.6. Elevation, slope, profile curvature, and tangential curvature raster maps of landslides in area A (Nga Ying Shan).

Extracting mean, standard deviation, maximum, and minimum values of elevation, slope, profile curvature, and tangential curvature in each landslide polygon forms statistical distributions of the surface geometry metrics. These statistics serve as empirical thresholds for removing predicted polygons with invalid surface geometry in Section 5.5. Figure 3.7 shows the histogram of these geometry matrices. Table 3.1 lists the statistics of some matrices of surface geometry of landslides and Table A.1 in the Appendix provides full matrices.

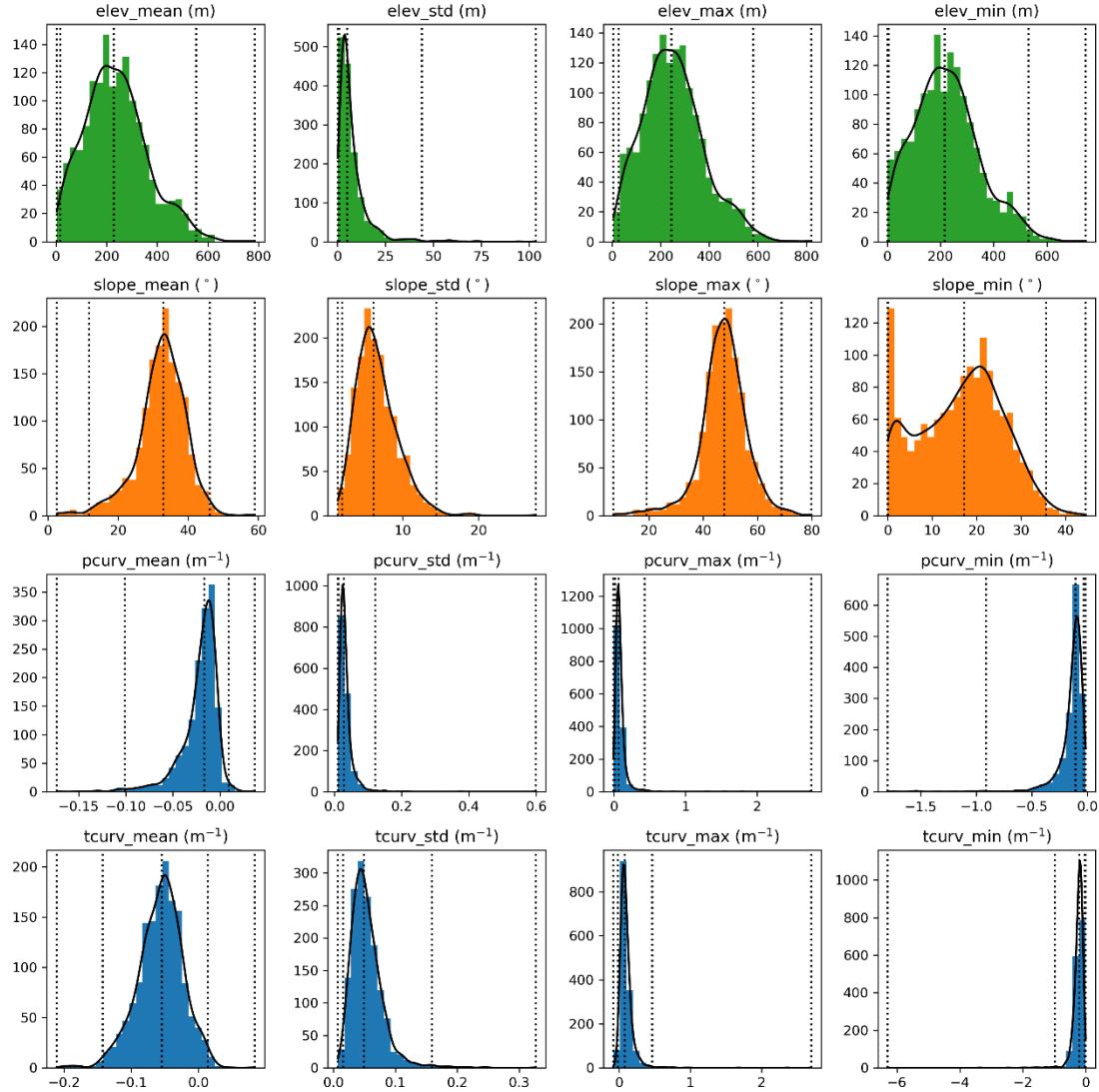


Figure 3.7. Distributions of Surface Geometry of Landslides. In each panel, vertical dotted lines represent minimum, 1st percentile, median, 99th percentile, and maximum values of the distribution.

Table 3.1. Statistics of Surface Geometry of Landslides.

Metric	Minimum	1st percentile	Median	99th percentile	Maximum
elev_std (m)	0.1	0.7	4.9	44.0	103.4
elev_max (m)	5.2	26.4	243.6	581.2	819.8
elev_min (m)	2.2	5.3	215.2	530.9	746.0
slope_mean ($^{\circ}$)	2.5	11.7	33.0	46.2	59.0
slope_std ($^{\circ}$)	1.5	2.1	6.2	14.4	27.5
slope_min ($^{\circ}$)	0	0.1	17.2	35.6	44.5
pcurv_mean (m^{-1})	-0.173	-0.101	-0.016	0.009	0.037
pcurv_std (m^{-1})	0.010	0.013	0.028	0.121	0.599
tcurv_mean (m^{-1})	-0.211	-0.143	-0.054	0.015	0.084
tcurv_std (m^{-1})	0.007	0.016	0.049	0.159	0.327

Remarks: Boldfaced values are empirical thresholds to be presented in Table 4.1. These empirical thresholds are useful for filtering out false positives (Section 4.5). Refer to Table A.1 in the Appendix for full table.

4. Methods

4.1. Overview of Automated Pipeline for Landslide Mapping

The automated landslide mapping pipeline proposed by this work is summarized in Figure 4.1 and briefly described as follows. Combining the digital orthophoto in 2008 and the landslide inventory map in 2011, we first split the large image into smaller image patches of dimension 480×480 pixels with corresponding label patches. Next, we perform data augmentation to obtain a synthesized training dataset (Section 4.2). We then train our DCNN models to obtain optimized models (Section 4.3). We then perform inference on the original DOP with conditional random fields post-processing to obtain a landslide probability map (Section 4.4). Combining with the LiDAR DSM, we further filter out small polygons and invalid landslide surface geometry to obtain a landslide polygon map (Section 4.5). Finally, we evaluate the performances of our models against the original ground truth landslide polygons. The next few subsections provide full details of the key steps.

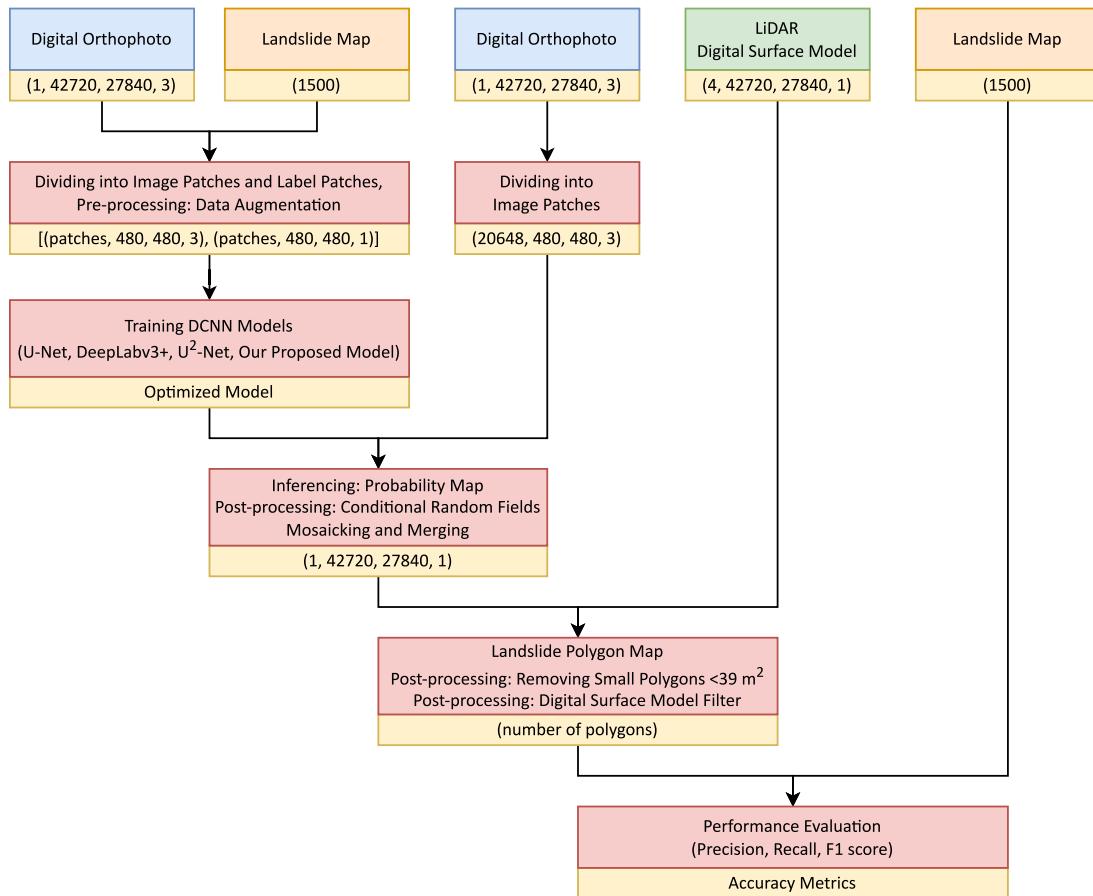


Figure 4.1. Overview of Automated Landslide Mapping Pipeline. Large boxes are data and processes. The small yellow boxes are data size and process outputs.

4.2. Pre-processing: Data Augmentation

Deep learning has a huge hunger for data, while our labeled landslide data is limited for training deep neural networks. Data augmentation helps to mitigate data scarcity issues and reduce overfitting. We utilize some conventional image augmentation strategies used in computer vision, including: rotation, flipping, Gaussian blur, and brightness, contrast, colour balance.



Figure 4.2. Data augmentation methods: original image (red box), rotation by 90° , 180° , 270° (magenta boxes), flipping horizontally, anti-diagonally, vertically, and diagonally (blue boxes), random Gaussian blur (yellow boxes), random brightness, contrast, colour balance (cyan boxes).

4.3. Training and Inference: Deep Convolutional Neural Network Models

This work uses four DCNN architectures for semantic segmentation tasks: 1) U-Net (Ronneberger et al., 2015), 2) DeepLabv3+ (Chen et al., 2018), 3) U²-Net (Qin et al., 2020), and 4) our proposed network. Semantic segmentation networks perform pixel-level class prediction, which perfectly suits our objective of segmenting landslide pixels and delineating landslide boundaries.

The network inputs are cropped RGB images of 480×480 pixels, and the outputs are landslide probability maps of 480×480 pixels. We use binary cross entropy (BCE) as the loss function. The batch size is 8. We train the network with Adam (adaptive moment) optimizer. The initial learning rate is 0.001, with a decaying factor of $\sqrt[3]{10} \approx 2.154435$ if the loss value has no improvement for three successive epochs. Training is early stopped if the loss value has no improvement for nine successive epochs. Regularization strategies are Gaussian noise at the input image layer and spatial dropout of rate 1/16 at the convolution layers.

U-Net (Ronneberger et al., 2015). The network contains a contracting path and an expanding path (Figure 4.3). The image is down-sampled via the contracting path and up-sampled via the expanding path. The contracting path acts as a feature extractor to capture image contexts while the expanding path with skipped connections acts as a feature localizer to assemble features of different levels. This network is proven capable of localizing objects accurately with large receptive context.

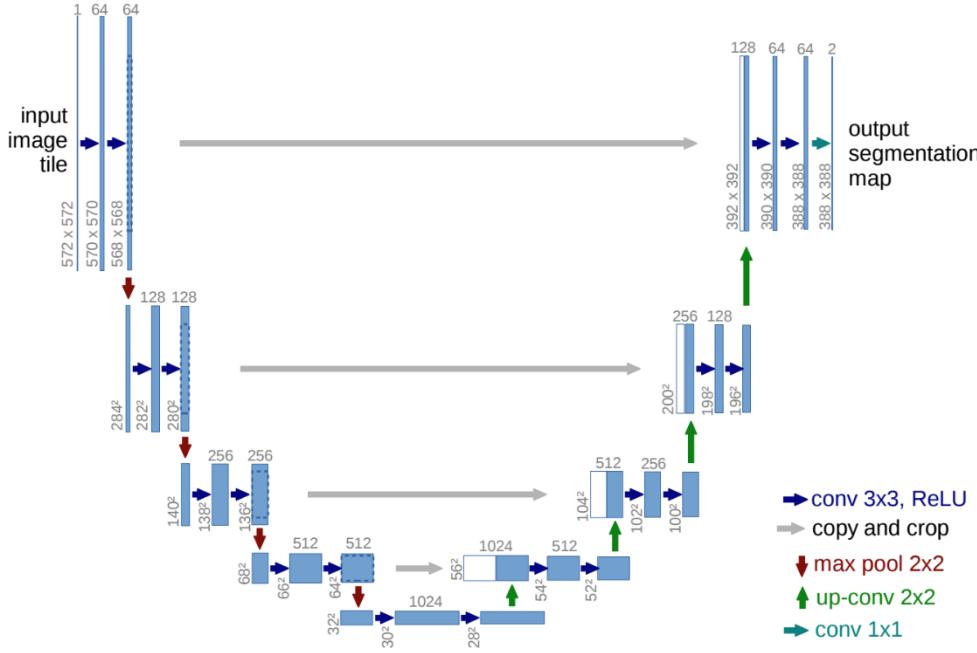


Figure 4.3. U-Net architecture (Ronneberger et al., 2015). The original feature map dimension in the vanilla version of U-Net is provided at the lower left corner of each box and the number of channels is provided on top of each box. In our adaptation, the input image and output segmentation map dimensions are 480×480 pixels, with a pooling factor of 2 and an up-sampling factor of 2.

DeepLabv3+ (Chen et al., 2018). The network combines an encoder-decoder structure as in U-Net and an atrous spatial pyramid pooling (ASPP) module (Figure 4.4). The ASPP module provides filters at multiple dilation rates and larger receptive fields compared to ordinary convolution for encoding multiscale contextual information. DeepLabv3+ extends the previous version by adding a decoder module for object boundary recovery. This version demonstrated faster and stronger segmentation performance.

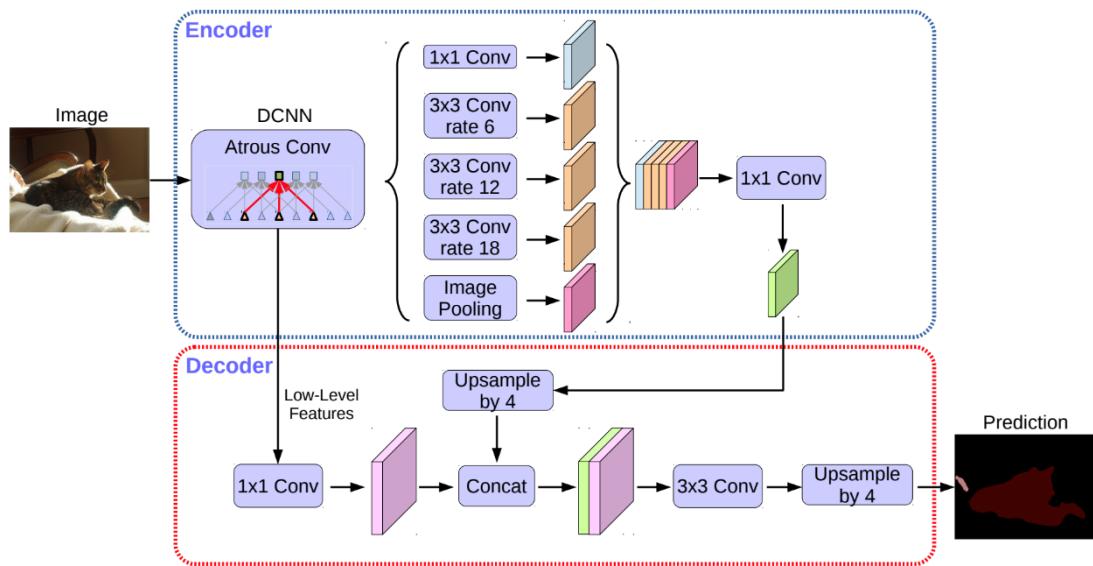


Figure 4.4. DeepLabv3+ architecture (Chen et al., 2018).

U²-Net (Qin et al., 2020). The network, on a global level, is a U-Net like structure. On a stage level, each encoder/decoder is replaced by a residual U-block module, which extracts intra-stage multi-scale features in the original resolution. This allows going deeper while maintaining high resolution. This network also spotlights an optional training strategy, deep supervision. It is used for supervision at each intermediate decoder, so that decoder outputs at all resolutions predict separately meaningful results.

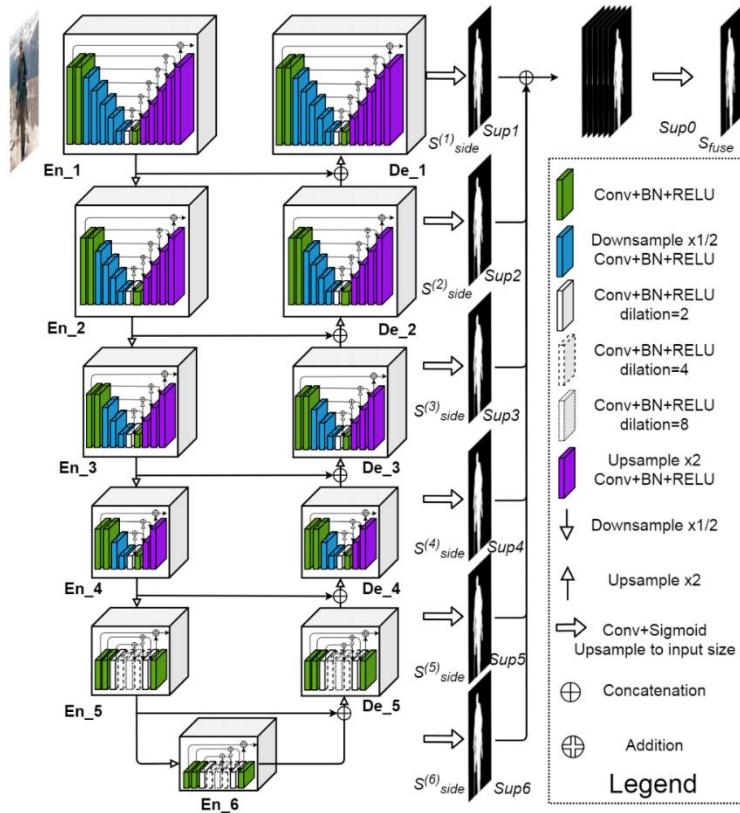


Figure 4.5. U²-Net architecture (Qin et al., 2020).

Our proposed network. Our proposed network (This work) combines the advantages from the previous three architectures. It employs EfficientNetB3, pretrained using ImageNet, as encoder backbone. It has a U-structure with skipped connections in the original and every downscaled resolution (as in U-Net and U²-Net). It is equipped with improved dilated spatial pyramid pooling (DSPP) modules at every skipped connection. Every convolutional sequence is equipped with a batch normalization layer (as in U²-Net) and exponential linear unit (ELU) activation function layer. Figure 4.6 presents our proposed architecture.

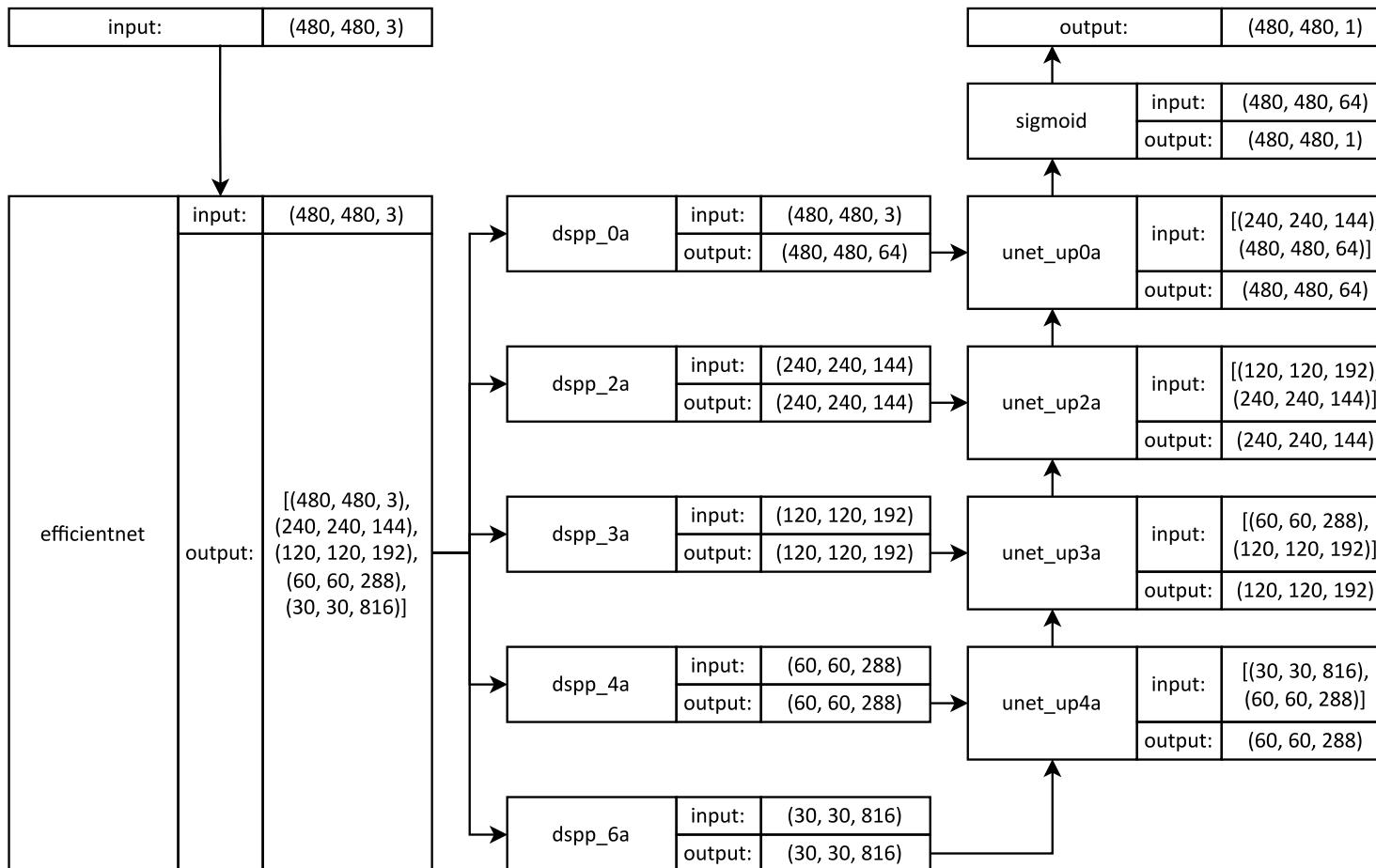


Figure 4.6. Architecture of our proposed network (This work).

The DSPP module is an improved version of the ASPP module in DeepLabv3+. The ASPP in DeepLabv3+ consists of one pointwise convolution layer, three standard convolution layers with dilation rates of 6, 12, and 18, and one global average pooling layer (Figure 4.4). Inspired by the depthwise separable convolution operation in Xception, our newly proposed DSPP allows more convolution kernels by increasing depthwise convolution channels and has fewer trainable parameters by splitting the operation. The DSPP in our proposed model consists of two stages: 1) depthwise convolutions with dilation and global average pooling, and 2) pointwise convolution.

Figure 4.7 shows the DSPP module architecture.

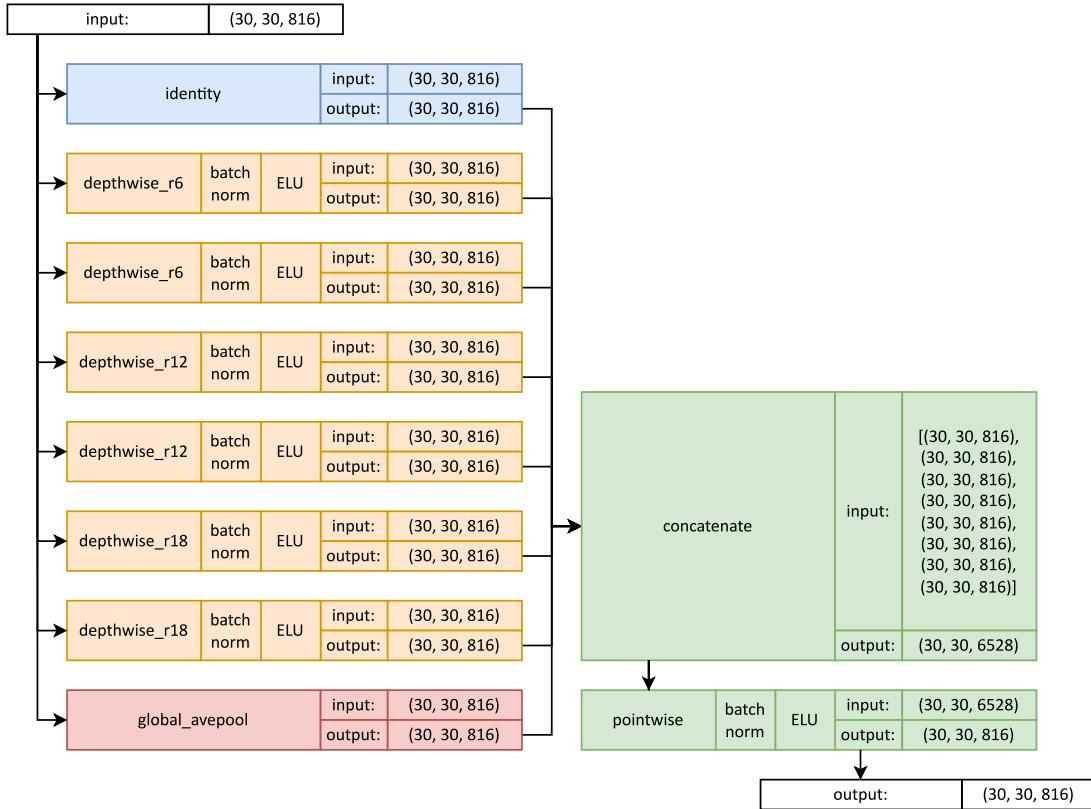


Figure 4.7. Architecture of our proposed dilated spatial pyramid pooling (DSPP) module. (30, 30, 816) are width, height, and number of channels of feature maps of DSPP block 6a.

4.4. Post-processing: Conditional Random Fields

Fully-connected Conditional Random Field (CRF) (Krähenbühl & Koltun, 2011) is employed for accurately recovering landslide boundaries. When a DCNN goes deeper, both the number of convolution layers and the number of pairs of max-pooling layers and up-sampling layers increase. However, a large number of convolution layers may produce an overly smoothed probability map due to a large receptive field. And a large number of up-sampling layers may produce checkerboard artifacts (Sugawara et al., 2019). Both limitations lead to inaccurate segmentation boundaries. The aim of adding a fully-connected CRF module is to improve landslide boundaries prediction from DCNN models.

The fully-connected CRF module uses the following Gibbs energy function (1),

$$E(x) = \sum_i \underbrace{\theta_i(x_i)}_{\text{unary potential}} + \sum_{ij} \underbrace{\theta_{ij}(x_i, x_j)}_{\text{pairwise potential}} \quad (1)$$

with unary potential function $\theta_i(x_i) = -\log P(x_i)$, where $P(x_i)$ is the landslide probability at pixel i predicted by a DCNN model. The pairwise potential function connects all image pixel pairs, i, j , and adopts the following expression (2),

$$\theta_{ij}(x_i, x_j) = \mu(x_i, x_j) \left[w_1 \underbrace{\exp\left(-\frac{|p_i - p_j|^2}{2\sigma_\alpha^2} - \frac{|I_i - I_j|^2}{2\sigma_\beta^2}\right)}_{\text{appearance kernel}} + w_2 \underbrace{\exp\left(-\frac{|p_i - p_j|^2}{2\sigma_\gamma^2}\right)}_{\text{smoothness kernel}} \right] \quad (2)$$

where $\mu(x_i, x_j) = 1$ if $x_i \neq x_j$ and zero otherwise. The Potts compatibility is used because it only penalizes pixel pairs of different classes. The appearance kernel is a colour-dependent term. It inputs pixel positions p_i, p_j and RGB vectors I_i, I_j , and forces pixels with similar colours and close position to have similar predicted classes. The smoothness kernel is a colour-independent term. It only inputs pixel positions p_i, p_j , and enforces spatial proximity, that is, forcing predicted boundary to be simpler and more compact. Hyperparameters $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ are set to be 10, 13, 3. Instead of computing the exact CRF distribution, the inference algorithm is successively updated every 5 iterations to minimize the Kullback-Leibler divergence of mean field approximation. The convergence on distribution outputs a probability map of landslides.

4.5. Post-processing: Digital Surface Model as Filter

Surface geometry properties of predicted landslide polygons can be extracted from the digital surface model. Mean, standard deviation, maximum, and minimum values of elevation, slope, profile curvature, and tangential curvature in each predicted landslide polygon is calculated. False positives, which usually fall outside the threshold, are then filtered and removed from the predictions.

Elevation thresholds. Polygons retained in the prediction are in between the minimum threshold “elev_max > 26.4 m” and the maximum threshold “elev_min < 746.0 m”. The first percentile of the maximum elevation is selected as the minimum threshold. This removal contains around 15 landslides instances of the least risks. The maximum of the minimum elevation is selected as the maximum threshold. This threshold ensures at most only one false removal instance of landslide of high risks. In addition, the first percentile of the elevation standard deviation (0.7 m) is selected as minimum standard deviation threshold “elev_std > 0.7 m”. This removal contains around 15 landslide instances of the least risks.

Slope thresholds. Polygons retained in the prediction are in between minimum threshold “slope_mean > 11.7°” and maximum threshold “slope_min < 44.5°”. The first percentile of mean slope is selected as the minimum threshold. This removal contains around 15 landslide instances of the least risks. The maximum of minimum slope is selected as the maximum threshold. This threshold ensures at most one false removal instance of landslide of high risks. In addition, the first percentile of slope standard deviation (2.1°) is selected as minimum standard deviation threshold “slope_std > 2.1°”. This removal contains around 15 landslide instances of the least risks.

Profile curvature and tangential curvature thresholds. Polygons retained in the prediction fall under maximum thresholds: “pcurv_mean < 0.009 m⁻¹” and “tcurv_mean < 0.015 m⁻¹”. Both the 99th percentile of mean profile curvature and the 99th percentile of mean tangential curvature are selected as the maximum thresholds. These removals contain around 15 landslide instances of the least risks for each case. In addition, the first percentile of profile curvature standard deviation (0.013 m⁻¹) and the first percentile of tangential curvature standard deviation (0.016 m⁻¹) are selected as minimum standard deviation thresholds: “pcurv_std > 0.013 m⁻¹” and “tcurv_std > 0.016 m⁻¹”. These removals contain around 15 landslide instances of the least risks for each case.

Table 4.1. Empirical thresholds for filtering false landslide geometry.

Metric	Threshold
elev_std (m)	> 0.7
elev_max (m)	> 26.4
elev_min (m)	< 746.0
slope_mean (°)	> 11.7
slope_std (°)	> 2.1
slope_min (°)	< 44.5
pcurv_mean (m ⁻¹)	< 0.009
pcurv_std (m ⁻¹)	> 0.013
tcurv_mean (m ⁻¹)	< 0.015
tcurv_std (m ⁻¹)	> 0.016

Remarks: These values are in bold in Table 3.1.

5. Results

5.1. Overview of Inferred Landslides from DCNN Models

This section evaluates models' ability to generalize landslide features and to recall landslide instances. Figure 5.1 shows an overview of landslide segmentation maps.

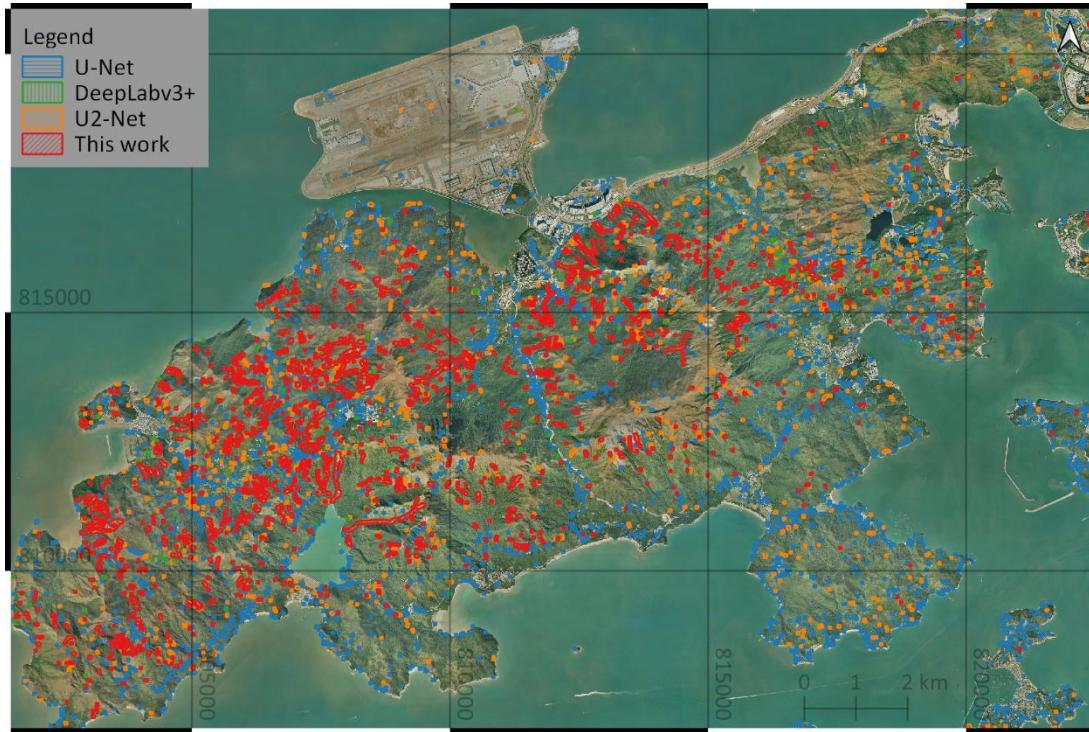


Figure 5.1. Overview of landslide segmentation maps by four DCNN models.

Predictions over Large Landslides. Figures 5.2 to 5.4 show segmentation maps on some large- ($\text{area} > 346 \text{ m}^2$) and medium-scale ($142 \text{ m}^2 < \text{area} < 346 \text{ m}^2$) landslides in area A (Nga Ying Shan), area B (Keung Shan), and area C (Shek Pik Reservoir). All DCNN models have excellent performances in recalling large- to medium-scale landslides. For example, areas and boundaries of landslides in area A are perfectly delineated by all DCNN models compared to the ground truth. Next in area B, all DCNN models have similar recalling ability but with different false predictions on possibly relict landslides. Followed by area C, U-Net and U²-Net show some erosion on predicted boundaries, with numerous false predictions on trails and relict landslides, while the predictions by DeepLabv3+ and our proposed model are ideal.

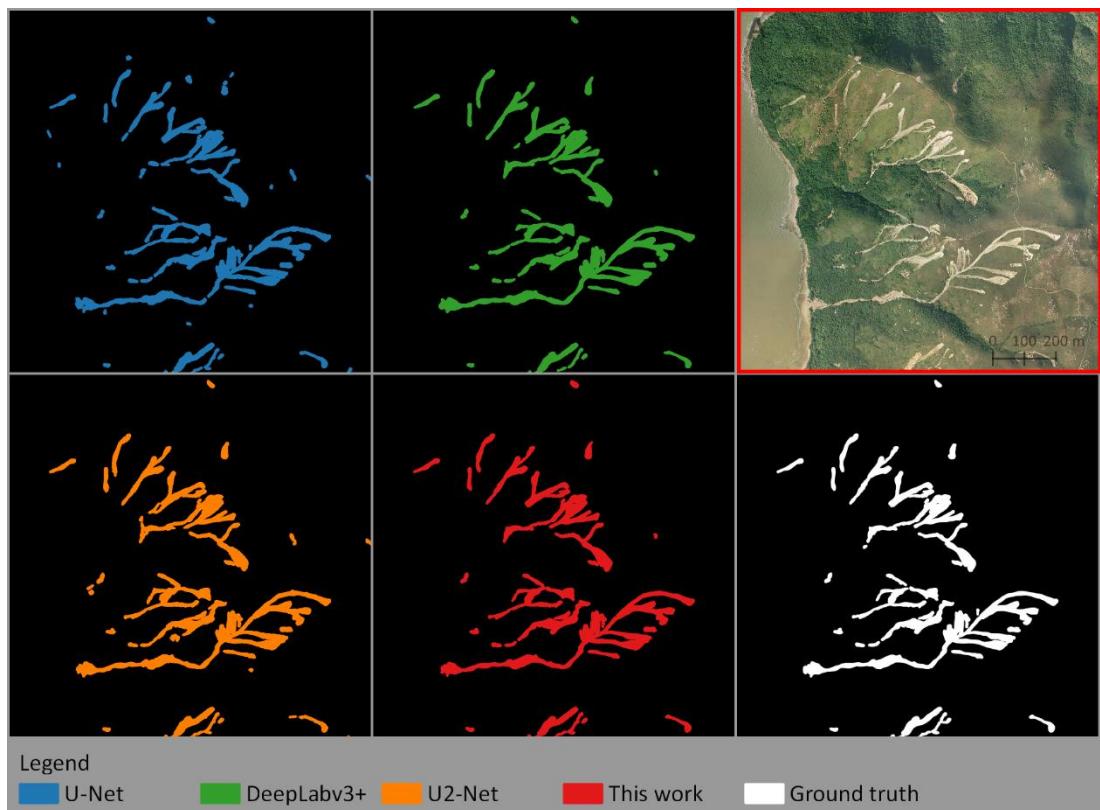


Figure 5.2. Landslide segmentation maps in area A (Nga Ying Shan). Segmentation map for U-Net in blue, DeepLabv3+ in green, U²-Net in orange, this work in red.

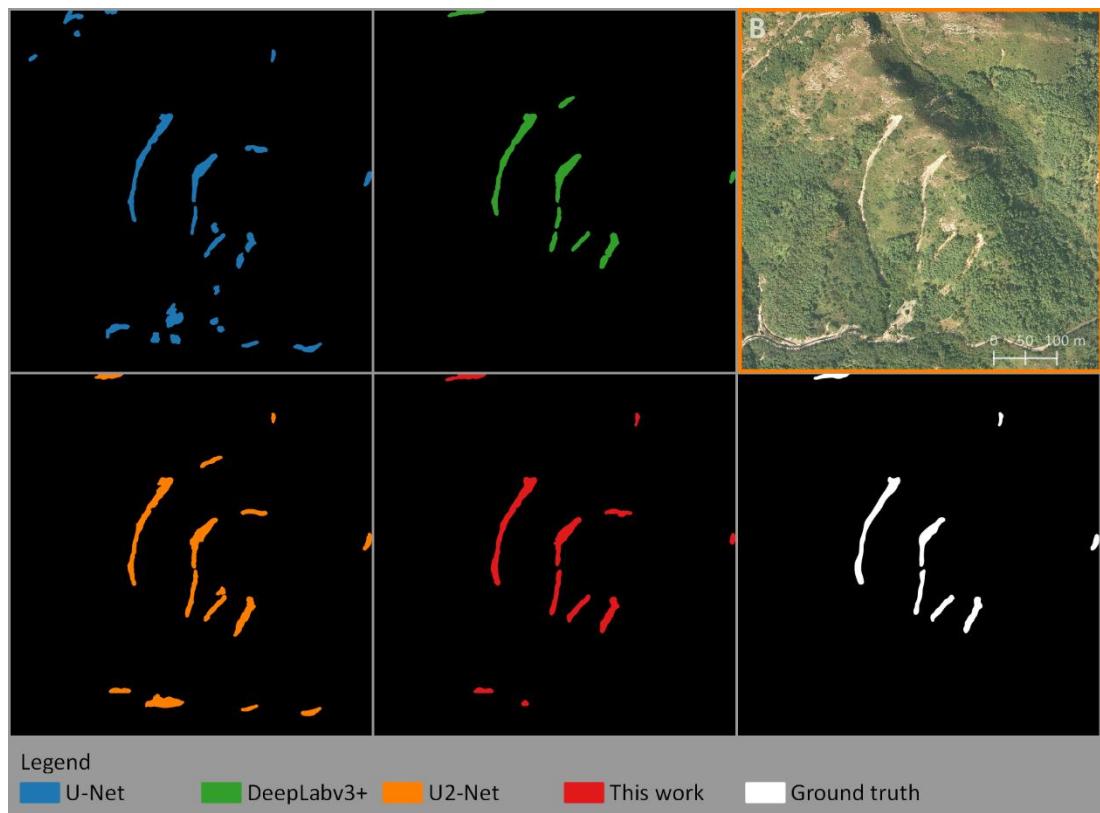


Figure 5.3. Similar to Figure 5.2 but for area B (Keung Shan).

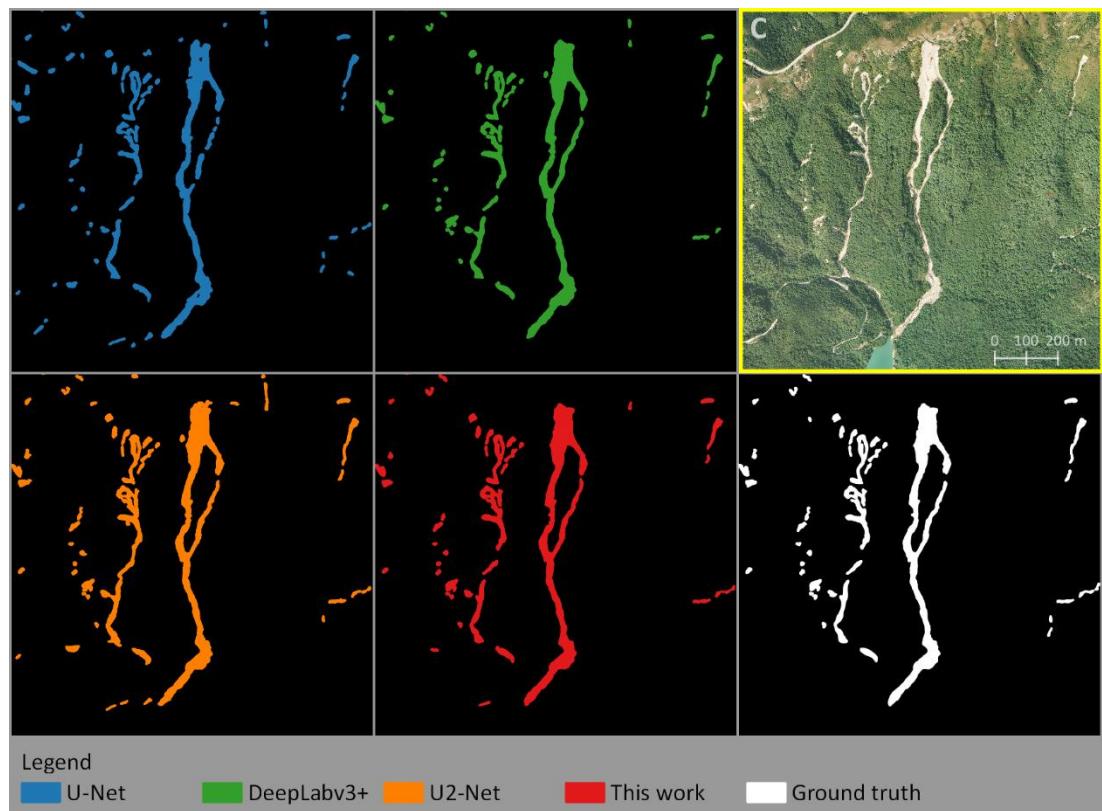


Figure 5.4. Similar to Figure 5.2 but for area C (Shek Pik Reservoir).

Predictions over Engineered Surfaces. Figure 5.5 shows segmentation maps on engineered surface in area D (Ngong Ping Market). U-Net shows many false predictions. On the other hand, DeepLabv3+ and our proposed model have only two and one false predictions, respectively. The prediction by U²-Net is perfect, completely no landslide in this area.

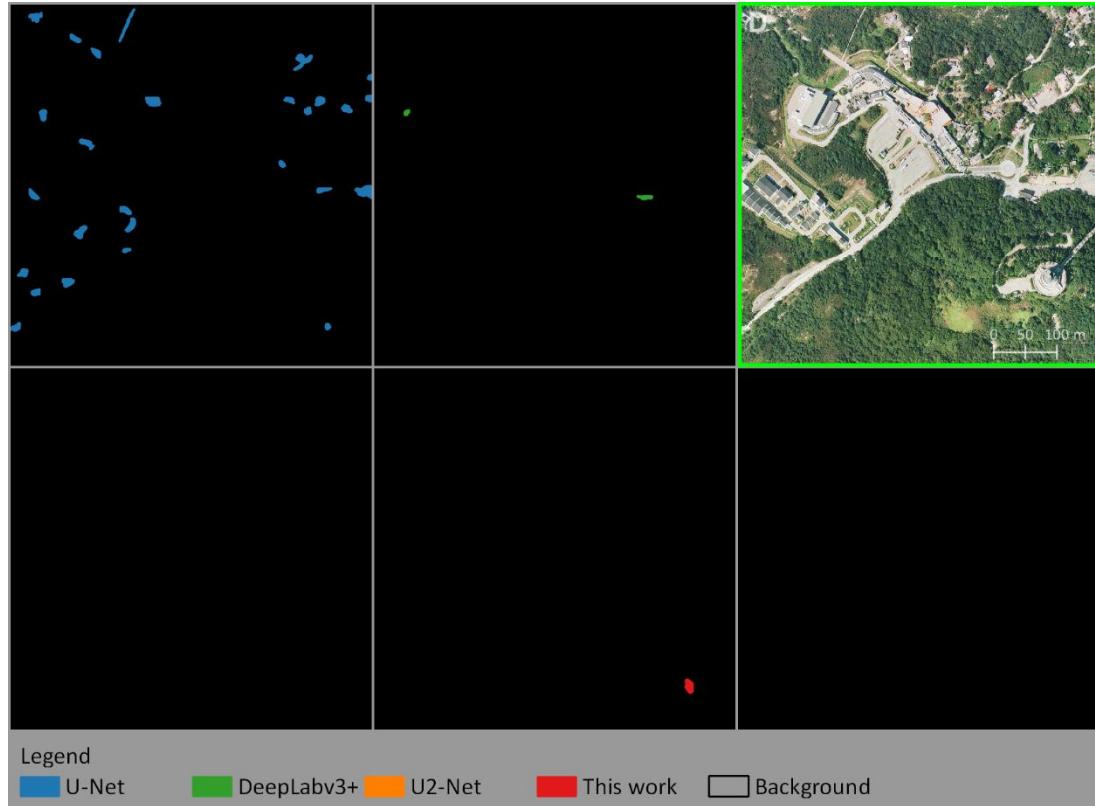


Figure 5.5. Similar to Figure 5.2 but for area D (Ngong Ping Market).

Predictions over Natural Surfaces. Figures 5.6 and 5.7 show segmentation maps on natural terrains in area E (Sze Tsz Tau Shan) and area F (Wong Lung Stream). Feature in area E, possibly bare soil or relict landslide, is not recorded in ground truth polygon produced by government expects and therefore is treated as background class. All the DCNN models, however, falsely segmented this feature as a landslide. Next in area F is a stream, thus the background class. The four DCNN models show different precision performance: U-Net falsely predicts the stream as a landslide, U²-Net and our proposed model segmented a minor part of stream as a landslide, and DeepLabv3+ has the most ideal performance in completely excluding the stream.

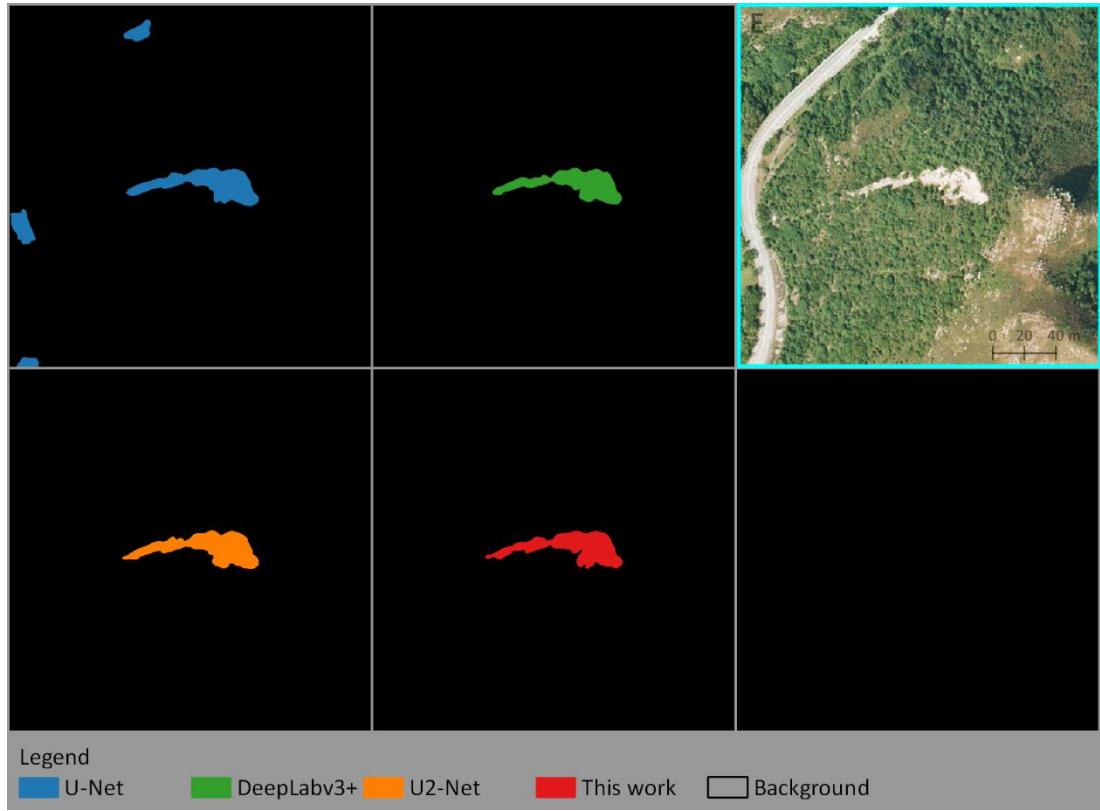


Figure 5.6. Similar to Figure 5.2 but for area E (Sze Tsz Tau Shan).

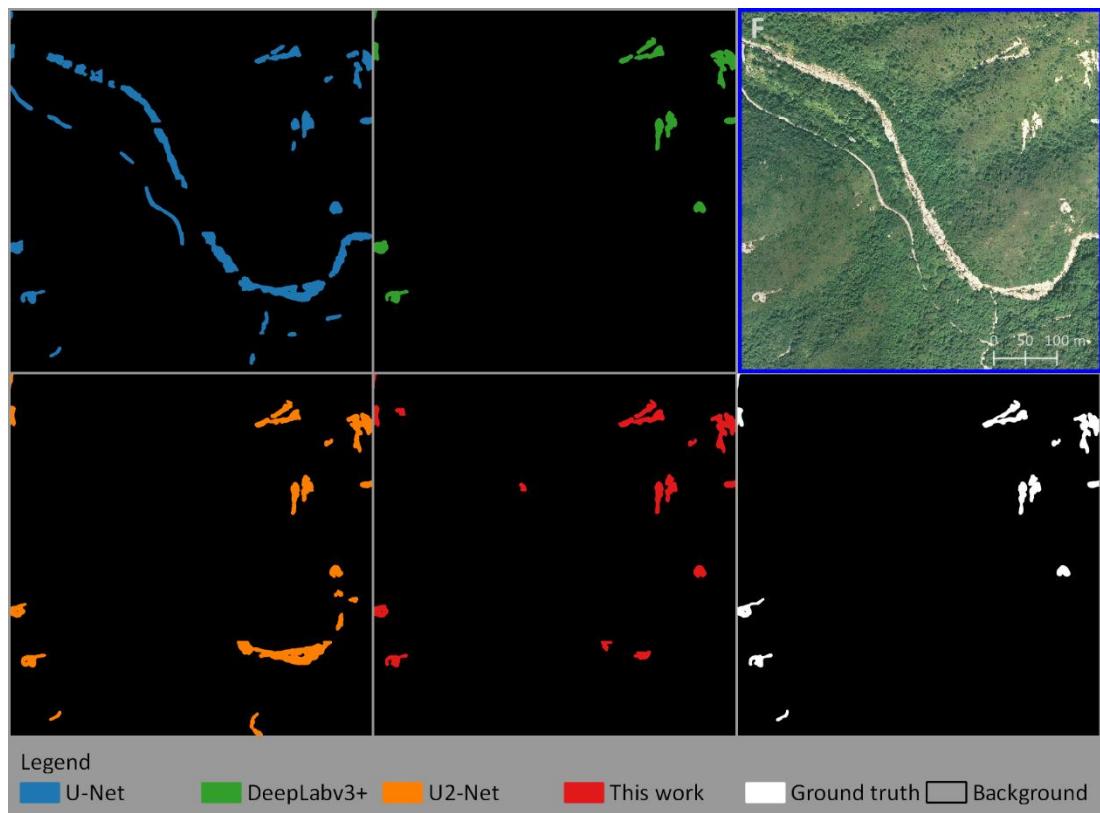


Figure 5.7. Similar to Figure 5.2 but for area F (Wong Lung Stream).

Predictions over Small Landslides. Figures 5.8 and 5.9 show small landslide instances over low to no vegetation cover in area G (Nei Lak Shan Angle Station) and area H (A Po Long and Wong Kung Tin). Although all the DCNN models predict large- to medium-scale landslides well, they show some extents of erosion when predicting small landslides. For example, in area G, DeepLabv3+ and our proposed model cannot fully recover landslide boundaries of two landslide instances in the middle of the image. Finally in area H, DCNN models show a tradeoff between reducing false predictions and recall landslides. U-Net and U²-Net can recall all small landslides, but give numerous false predictions over bare soils and hiking trials. On the other hand, DeepLabv3+ and our proposed model can reduce false predictions to an acceptable level, but they show one or two missing the ground truths.

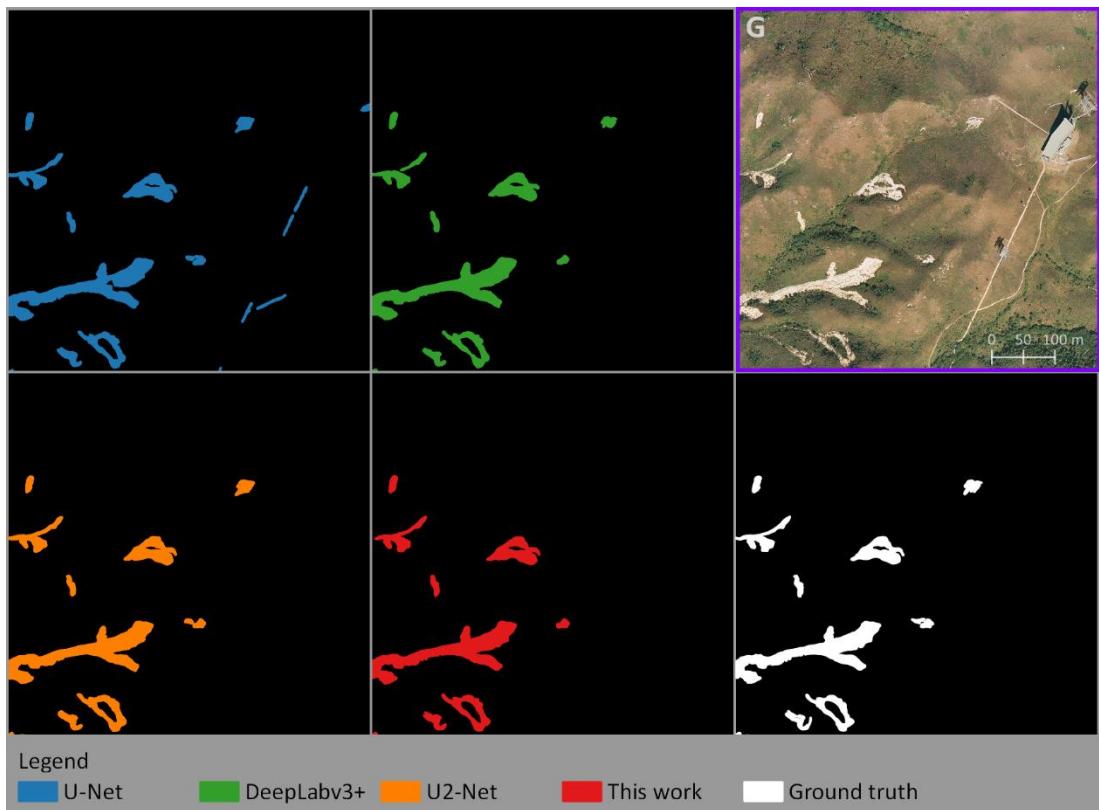


Figure 5.8. Similar to Figure 5.2 but for area G (Nei Lak Shan Angle Station).

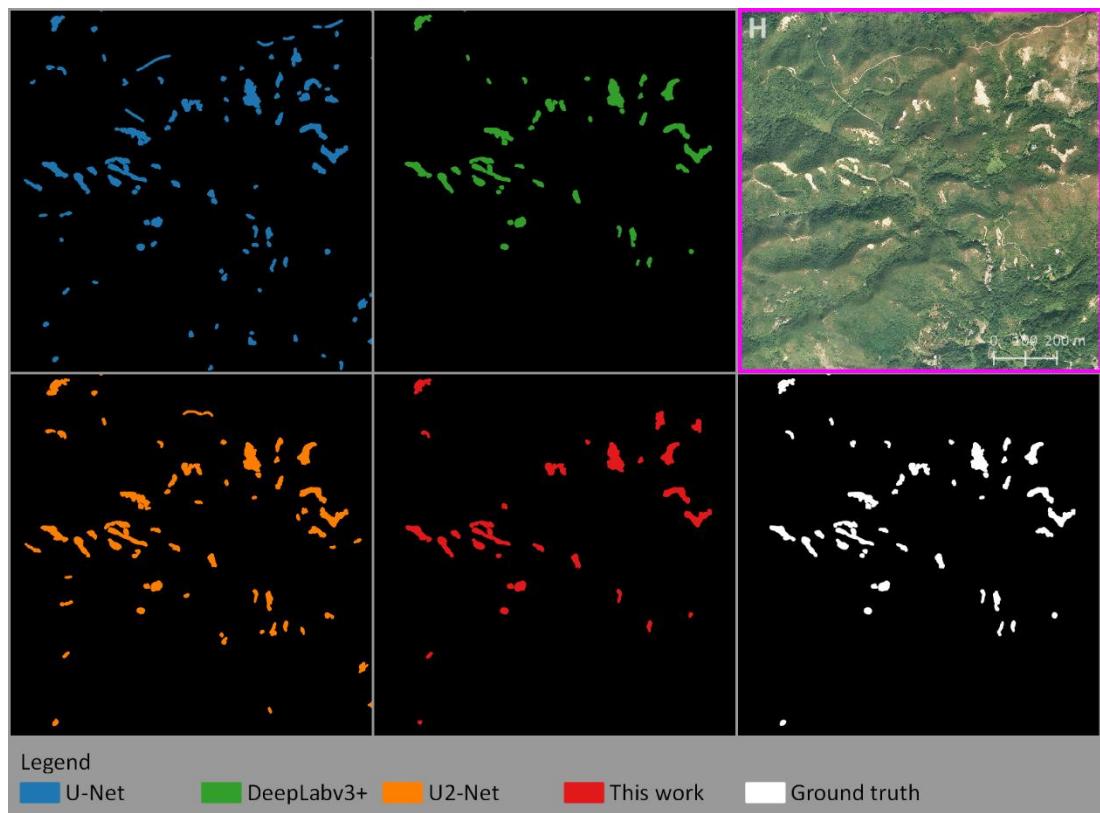


Figure 5.9. Similar to Figure 5.2 but for area H (A Po Long and Wong Kung Tin).

5.2. Accuracy Metrics of Inferred Landslides from DCNN Models

This section evaluates the DCNN models' accuracies in two ways: polygon-based (by count) and pixel-based (by area). Following the conventional similarity measure for object detection, this study uses Intersection over Union (IoU) to evaluate each polygon in the model predictions. True positive is an inferred landslide polygon with IoU between 0.5 and 1, i.e., the model successfully predicts a landslide polygon with good overlapping with the ground truth. False positive is an inferred landslide polygon with IoU smaller than 0.5, i.e., the model predicts a landslide polygon with poor or no overlapping with the ground truth. False negative is a landslide instance in the ground truth, in which the model fails to predict. Tables 5.1 and 5.2 summarize the validation metrics by count of four DCNN models, before and after applying the DSM filter post-processing. Another way to evaluate semantic segmentation model performances is by area. Table 5.3 summarizes validation metrics by area of four DCNN models.

Table 5.1. Summary of accuracy metrics by count of four DCNN models before DSM filter post-processing.

Model	# True Pos	# False Pos	# False Neg	Precision	Recall	F1 score
U-Net	1207	3883	293	0.237	0.805	0.366
DeepLabv3+	1180	387	320	0.753	0.787	0.769
U ² -Net	1271	1258	229	0.502	0.847	0.631
This work	1367	1293	133	0.514	0.911	0.657

Remarks: The number of ground truth polygons is 1500, which equals the sum of the true positive and false negative counts.

Table 5.2. Summary of accuracy metrics by count of four DCNN models after DSM filter post-processing.

Model	# True Pos	# False Pos	# False Neg	Precision	Recall	F1 score
U-Net	1136	1692	364	0.402	0.757	0.525
DeepLabv3+	1130	321	370	0.779	0.753	0.766
U ² -Net	1217	770	283	0.612	0.811	0.699
This work	1303	445	197	0.745	0.869	0.802

Table 5.3. Summary of accuracy metrics by area of four DCNN models

Model	Precision (by area)	Recall (by area)	F1 score (by area)
U-Net	0.554	0.833	0.665
DeepLabv3+	0.876	0.868	0.872
U ² -Net	0.709	0.911	0.797
This work	0.874	0.912	0.893

Before DSM filter post-processing, DeepLabv3+ has the best F1 score among four DCNN models, but it has the worst recall. After the post-processing, U-Net, U²-Net and our proposed model have an increase in F1 score, while the boost in our proposed model is significant (from 0.657 to 0.802). The post-processing is effective in removing many false positives in DCNN models, while slightly increasing false negatives. In summary, our proposed model has the best performance on F1 score both by count and by area. Our proposed model also has the best recall, and the second best precision slightly less than that of DeepLabv3+.

5.3. Comparison with Other Research Works

Recent research works by Shi et al. (2020) and Su et al. (2021) also carried out automatic landslide mapping using deep learning. Their study area was also over Lantau Island, and their data included 2008 Digital Orthophoto, manually interpreted landslide polygons, and a digital terrain model. This section directly compares the results from the aforementioned two research works against this study. Table 5.4. summarizes their results and compares with our results.

Table 5.4. Summary of results by area from other research works and this work.

Work	Method/Model	Precision	Recall	F1 score
Shi et al. (2020)	CNN designed by authors	0.545	0.884	0.675
	CDCNN + Post-processing	0.812	0.841	0.826
Su et al. (2021)	LanDCNN-BAI+DTM	0.723	0.841	0.777
	U-Net-BAI+DTM	0.700	0.821	0.756
This work	Our Proposed Model + Post-processing	0.874	0.912	0.893

In summary, our proposed model has the best performance by area in terms of precision, recall, and F1 score among recent research works on the same dataset. The next Chapter will focus on discussing the best performed model: our proposed model.

6. Discussion

6.1. Mapping Performance Breakdown by Size Groups

Collecting large- (areas greater than 346 m^2), medium- (areas from 142 m^2 to 346 m^2), and small-scale landslides (areas less than 142 m^2) into three sub-groups, this section evaluates the polygon-based performance breakdown over these three size groups. Class boundaries are chosen to be minimum, 33rd percentile, 67th percentile, and maximum of areas of the ground truth polygons, so that each bin represents 500 instances of ground truth landslide (which equals the sum of true positive and false negative). Corresponding accuracy metrics for each size group can be calculated. Figure 6.9 shows the histogram of prediction evaluation breakdown over three size groups and their corresponding accuracy metrics by count.

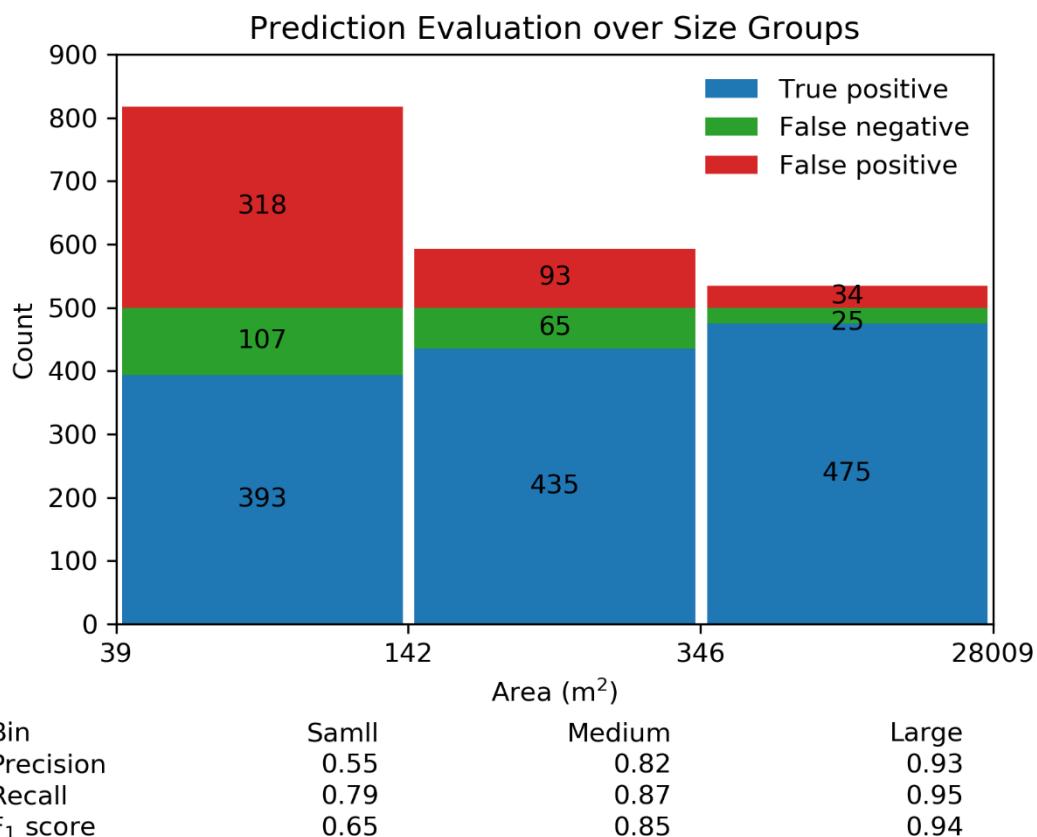


Figure 6.1. Evaluation breakdown by size groups and their accuracy metrics.

Mapping performance over large-scale landslides is excellent with an F1 score of 0.94, and over medium-scale landslides is great with an F1 score of 0.85. Large- and medium-scale landslides usually have distinguishing colours against the background and apparent landslide features, and occupy large numbers of pixels. Therefore, it is an easier task for our proposed model to segmentate these landslide areas and delineate corresponding boundaries.

However over small-scale landslides, the mapping performance is inferior with an F1 score of only 0.65. The low accuracy is attributed to the large number of false positives. Small-scale landslides may be obstructed by surrounding terrain or vegetation, have incomplete or obscure landslide features, and only occupy limited numbers of pixels, (examples refer to Section 6.3). Moreover, some landslide instances may be missing in the ground truth polygons and thus be evaluated as false positives. Therefore, it is a more difficult task for our proposed model to distinguish whether these small and yellowish objects are landslides or not.

6.2. Success Mode: True Positive over Large Landslides and True Negative over Engineered Surfaces

True Positives over Large Landslides. Our proposed model has an excellent performance in segmenting large- and medium-scale landslides, with areas greater than approximately 142 m^2 , which accounts for around 1000 landslide polygons in the ground truth. Figures 6.1 to 6.3 show landslide segmentation maps on large scale landslide examples and their corresponding elevation, slope, profile curvature, and tangential curvature maps. In area A (Nga Ying Shan), the predictions on large scale landslides are almost perfect, with mean IoU greater than 0.9. The only problem is with two separated polygons in the ground truth combined into one prediction polygon (see the pink polygon marked in Figure 6.1). This only affects validation by count but does not affect that by area. In area B (Keung Shan), the prediction on medium scale landslide is very good, with mean IoU greater than 0.85, with one false positive on exposed boulders (see the pink polygon marked in Figure 6.2). In area C, the prediction on large scale landslide is ideal, with mean IoU greater than 0.85, but four small landslides are missing.

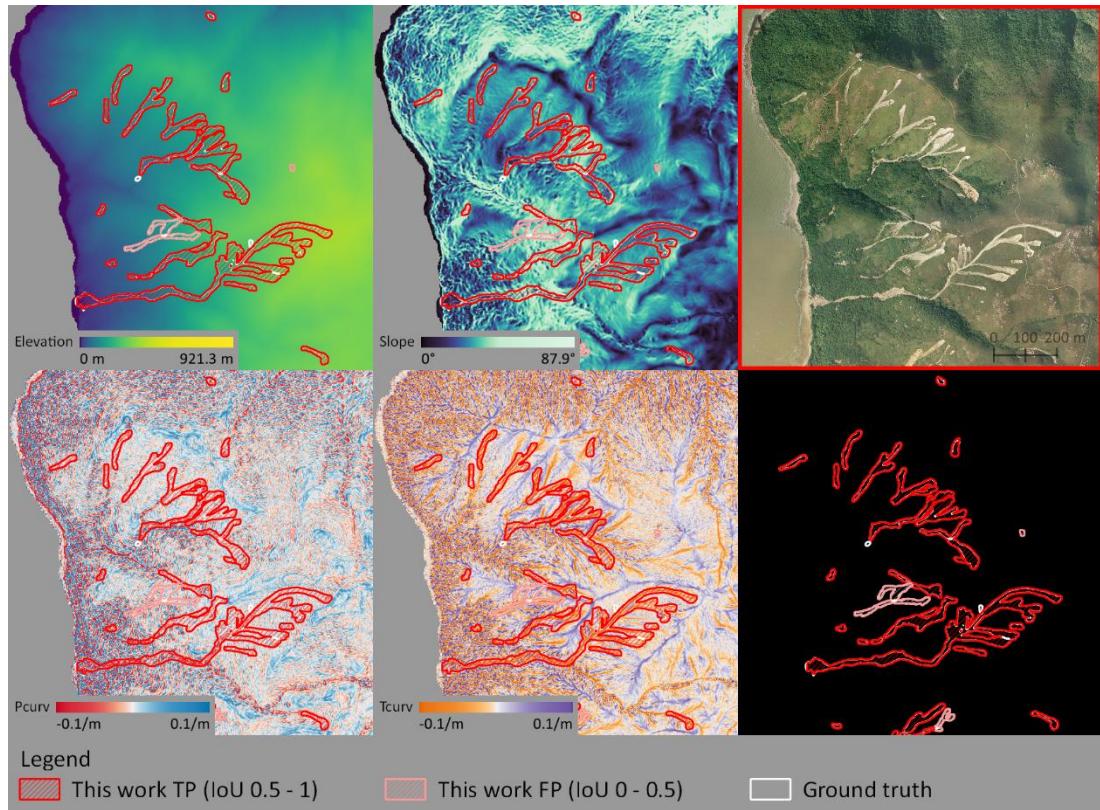


Figure 6.2. Landslide segmentation map by This work and elevation, slope, profile curvature, tangential curvature maps in area A (Nga Ying Shan). True positives (IoU between 0.5 to 1) by this work are in red, false positives (IoU < 0.5) by this work are in pink, the ground truth polygons are in white (many are not visible because of good match with the deep-learning-predicted polygons).

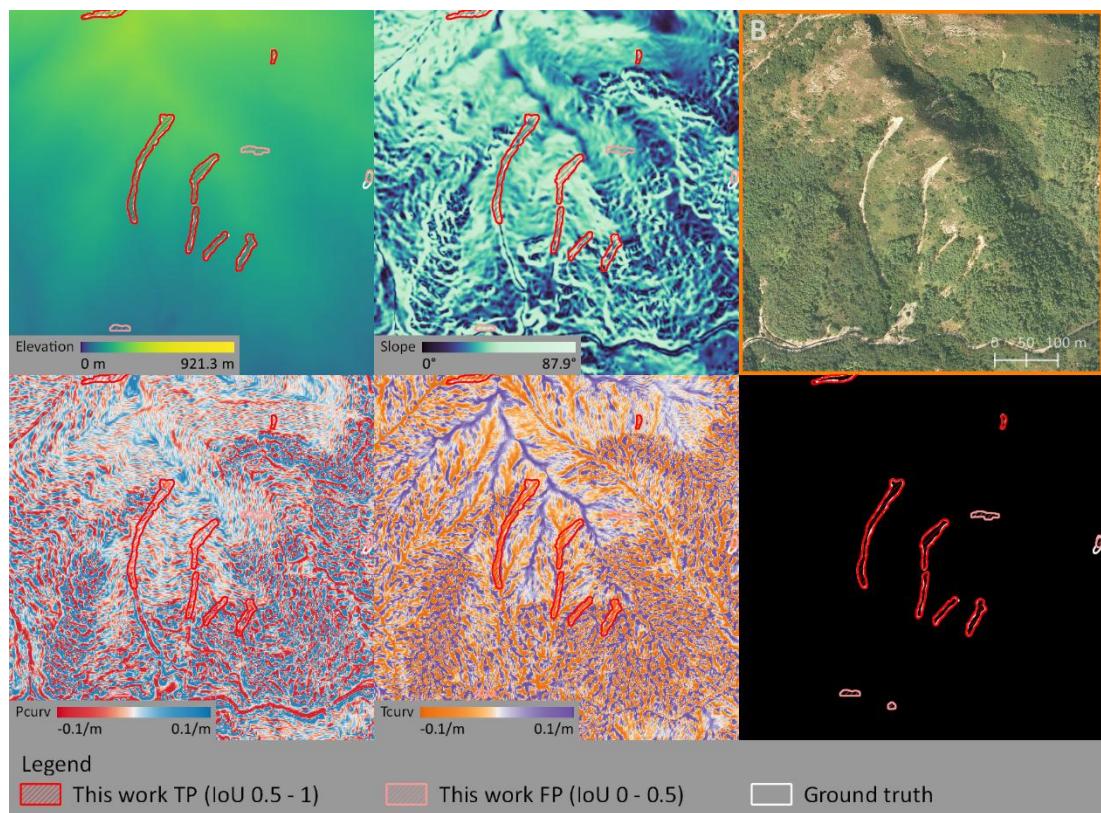


Figure 6.3. Similar to Figure 6.1 but for area B (Keung Shan).

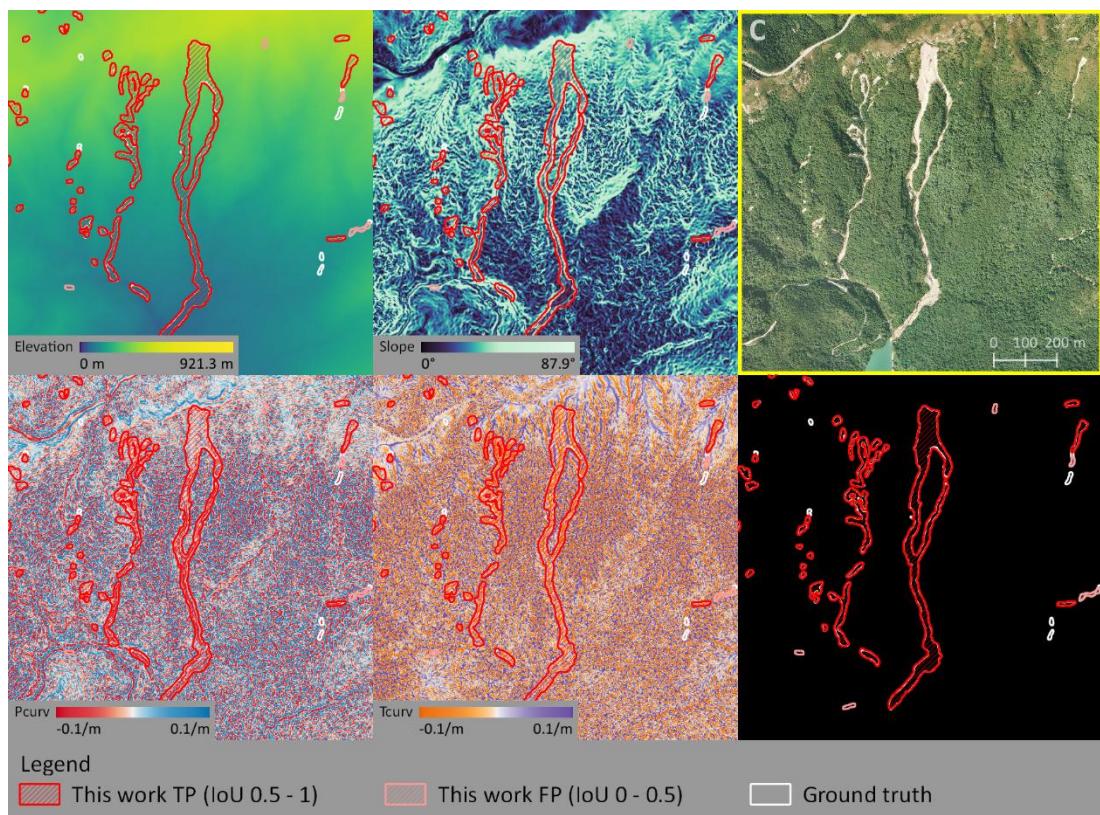


Figure 6.4. Similar to Figure 6.1 but for area C (Shek Pik Reservoir)

True Negatives over Engineered Surfaces. Our proposed model can successfully reject engineered surfaces, such as buildings, roads, and bridges. Figure 6.4 shows landslide segmentation maps over Ngong Ping Market (area D) and its corresponding elevation, slope, profile curvature, tangential curvature maps. In area D, the prediction on engineered surfaces are almost perfect, with all true negatives over Ngong Ping Market, except for one false positive on grave objects (pink polygon).

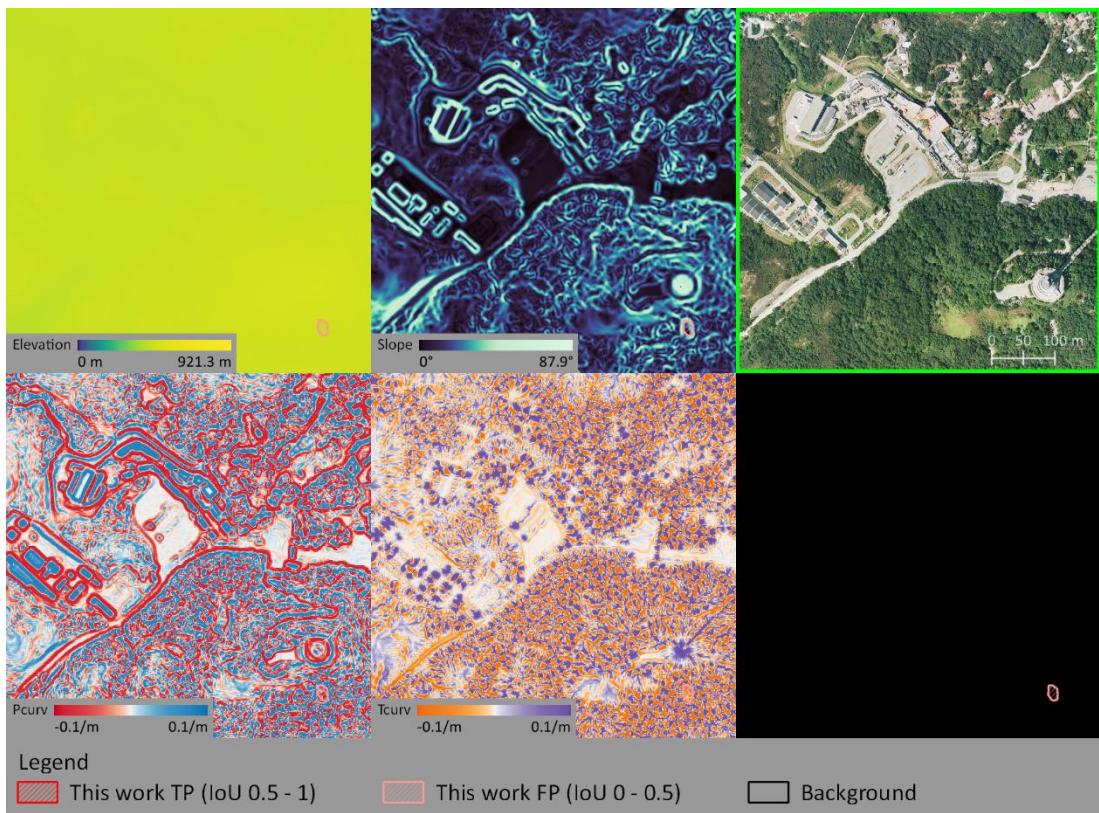


Figure 6.5. Similar to Figure 6.1 but for area D (Ngong Ping Market).

6.3. Failure Mode: False Positive over Natural Surfaces and False Negative over Small Landslides

False Positives over Natural Surfaces. Our proposed model still fails to reject some potentially confusing natural surfaces. Some bare soils, exposed boulders, and especially streams, have similar appearance in colours, textures, and patterns. A large portion of them have been rejected by our DSM-based filter post-processing, while a small portion remains as false positives. Figure 6.5 shows landslide segmentation map on bare soils as an example and Figure 6.6 shows landslide segmentation maps on streams. The predicted polygon on bare soils in area E (Sze Tsz Tau Shan) falls within the threshold of post-processing, and thus remains as a false prediction. Similarly, the predicted polygon on stream in area F (Wong Lung Stream) cannot be removed by the DSM filter post-processing.

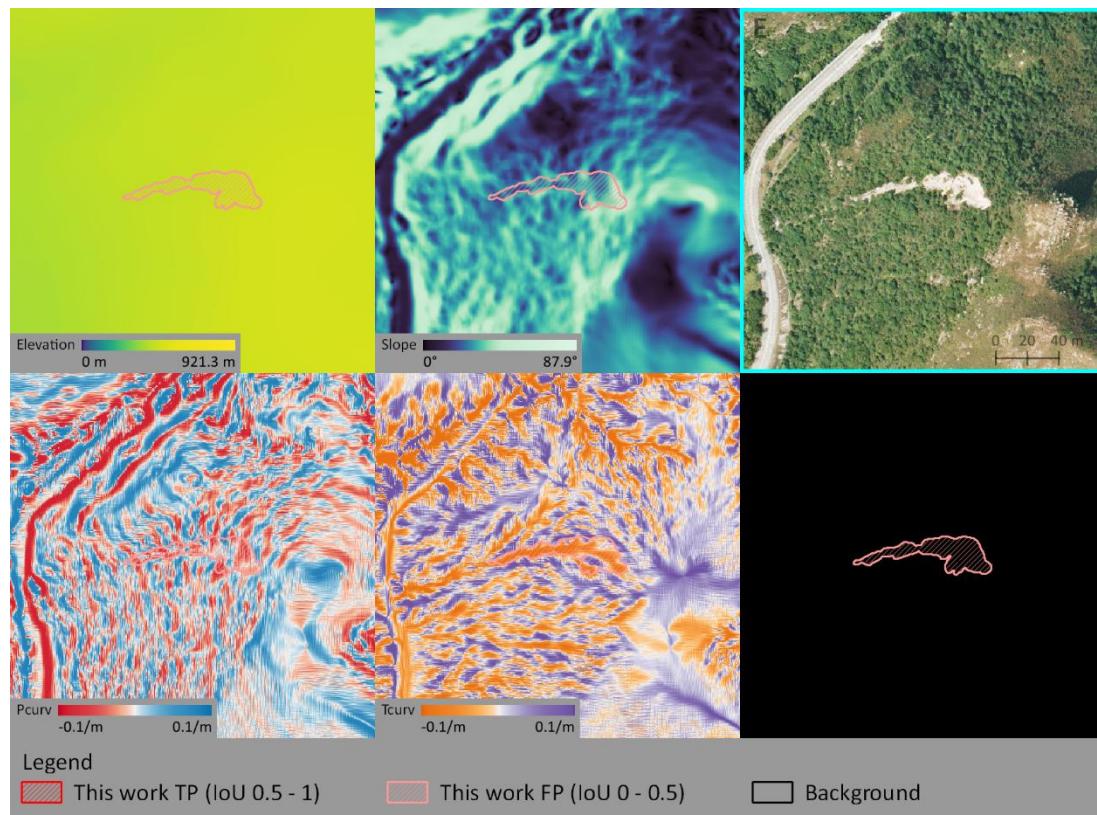


Figure 6.6. Similar to Figure 6.1 but for area E (Sze Tsz Tau Shan).

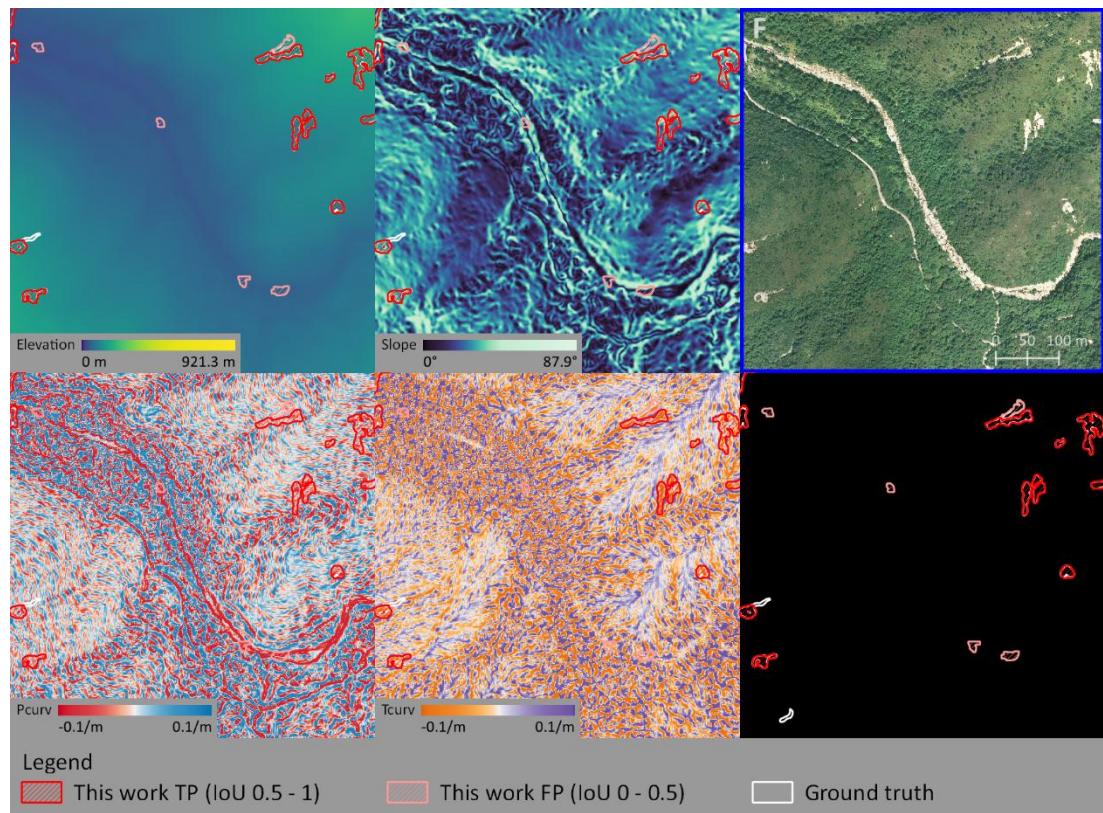


Figure 6.7. Similar to Figure 6.1 but for area F (Wong Lung Stream).

False Negatives over Small Landslides. Our proposed model has inferior performance on segmenting small-scale landslides, with area less than approximately 142 m², which accounts for around 500 landslide polygons in the ground truth. These landslides only occupy a few tens to hundreds of pixels in the image. Some of their landslide features such as their crown, flow zone, and accumulation area, are not visible. Our model therefore frequently misses small-scale landslide instances. Figures 6.7 and 6.8 show landslide segmentation maps on small scale landslides. In area G, a small-scale landslide is only partly segmented, and another is completely missed (see the white polygon marked in Figure 6.7). In area H, around 10 small-scale landslide instances are completely missed (see the group of white polygons marked in Figure 6.8).

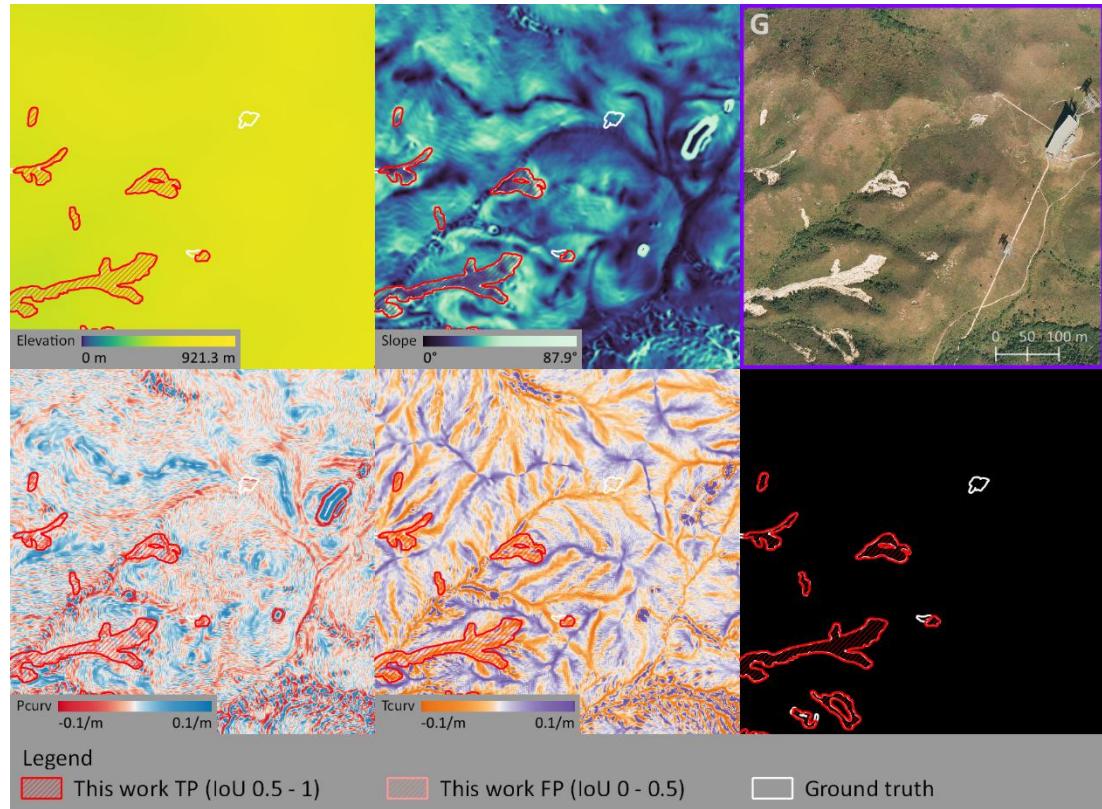


Figure 6.8. Similar to Figure 6.1 but for area G (Nei Lak Shan Angle Station).

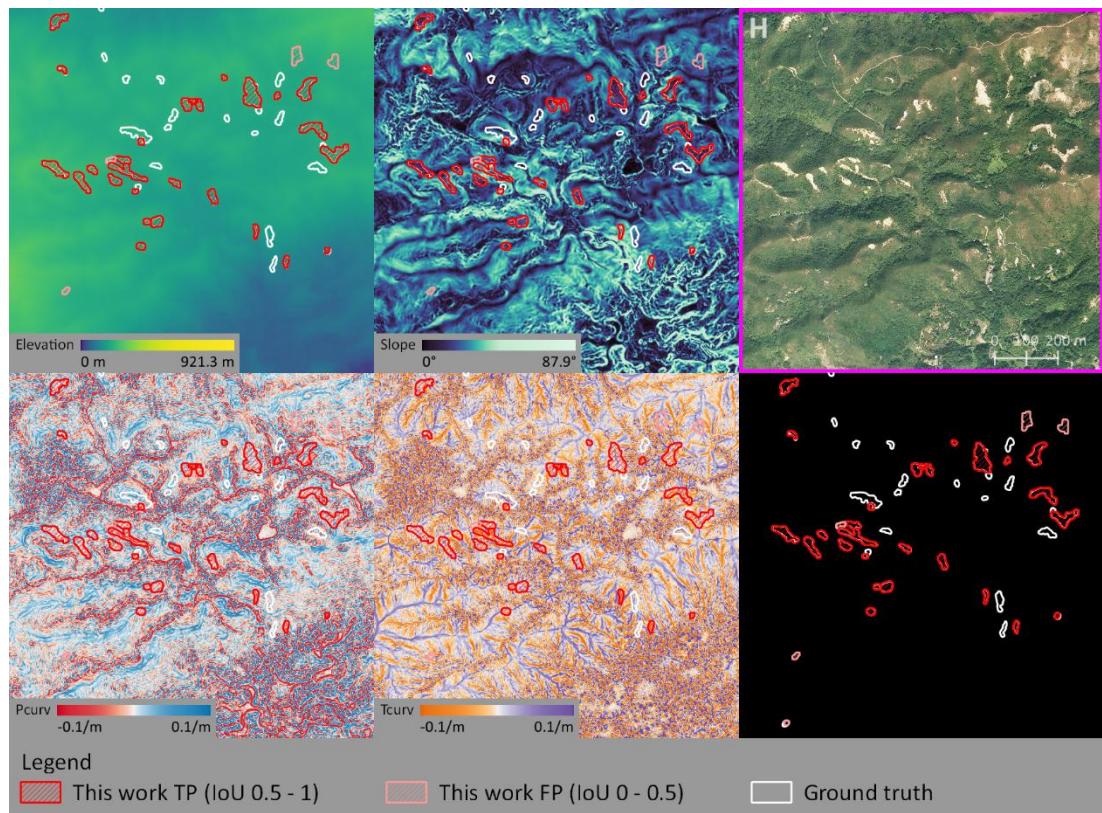


Figure 6.9. Similar to Figure 6.1 but for area H (A Po Long and Wong Kung Tin).

6.4. Choices of DCNN Architectures and Hyperparameters

Among the four DCNNs used in this study, U-Net is the simplest architecture and serves as the baseline for performance comparison. DeepLabv3+ uses an atrous spatial pyramid pooling (ASPP) modules for encoding multiscale contextual information at intermediate and final feature maps. U²-Net uses residual U-block modules for going deeper while maintaining high resolution. Our proposed network combines the advantages from the previous three architectures.

Precision and spatial pyramid pooling. In terms of precision by count, our network (0.745) outperforms U-Net (0.237) and U²-Net (0.502), and is comparable to DeepLabv3+ (0.753). In terms of precision by area, this work (0.874) also outperforms U-Net (0.554) and U²-Net (0.709), and is almost identical to DeepLabv3+ (0.876). Outstanding performances in this work and DeepLabv3+ are attributed to the use of the spatial pyramid pooling (SPP) modules. SPP can enlarge effective receptive fields at encoder stages; in other words, it can gather more contextual information compared to nondilated convolution operations. Large receptive fields are essential to reject land surfaces similar to landslides on the images, leading to higher precision.

Recall and skipped connection. In terms of recall by count, this work (0.869) outperforms all three other models: U-Net (0.805), DeepLabv3+ (0.787), and U²-Net (0.847). In terms of recall by area, this work (0.912) also outperforms U-Net (0.833) and DeepLabv3+ (0.868), and is almost identical to U²-Net (0.911). Outstanding performance in this work are attributed to extensive use of skipped connections at all resolutions. While DeepLabv3+ has only two connections (one skipped and one ordinary) linking the encoder and the decoder, the other two models and this work use skipped connections extensively at all resolutions. This configuration acts as “information highway” for simple features, such as, edges, etc., to skip through many layers of convolution and arrive at the decoder. Skipped connections are essential for retaining simple features for accurate recall of landslide pixels.

Therefore, this work has achieved very high F1 score by count (0.802) and F1 score by area (0.893), which is comparable to human performance.

Hyperparameters in conditional random fields. The CRF module has five hyperparameters: w_1 and $\sigma_\alpha, \sigma_\beta$ in the appearance kernel plus w_2 and σ_γ in the smoothness kernel. Following the default values for daily image segmentation, $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ are set to be 10, 13, 3, respectively. On the other hand, w_1, w_2 are set to be 1 instead of their default values of 3. This is because landslide objects in remote sensing images are usually thin and narrow. Having strong weighted regularizing kernels can erode the predicted segmentation boundaries. Lowering w_1, w_2 in the binary potential keeps the importance in prior information (segmentation map before process) in the unary potential term.

6.5. Choices of Surface Geometry Thresholds in Post-processing

The surface geometry thresholds are empirically determined from statistics and may not represent physical processes of landslides. Selecting the minimum or maximum value of a metric as a threshold ensures at most one landslide instance would be falsely removed. Selecting the 1st or 99th percentile of a metric as threshold may falsely remove around 15 landslide instances. The post-processing in total removes 912 instances, in which 848 are true removals of non-landslide instances and 64 are false removals of landslide instances (accounting for 64 out of 197 false negatives). Table 6.1. summarizes landslide instances removal by surface geometry thresholds.

Table 6.1. Summary of landslide instance removals by numbers by applying filters based on surface geometry thresholds.

Metric	Threshold	Total Removal	True Removal of Non-landslides	False Removal of Landslides
elev_std (m)	> 0.7	442	426	16
elev_max (m)	> 26.4	547	533	14
elev_min (m)	< 746.0	3	3	0
slope_mean (°)	> 11.7	316	301	15
slope_std (°)	> 2.1	211	195	16
slope_min (°)	< 44.5	13	12	1
pcurv_mean (m^{-1})	< 0.009	178	163	15
pcurv_std (m^{-1})	> 0.013	177	160	17
tcurv_mean (m^{-1})	< 0.015	163	147	16
tcurv_std (m^{-1})	> 0.016	339	327	12
All Thresholds		912	848	64

Minimum thresholds. We impose two minimum thresholds: “elev_max > 26.4 m”, and “slope_mean > 11.7°”. The threshold on minimum “maximum elevation” effectively removes false predictions on beaches. The threshold on minimum “mean slope” effectively removes false predictions on flat and leveled bare soil surfaces. However, these thresholds in total falsely remove 14 and 15 landslide instances, respectively.

Maximum thresholds. We impose four maximum thresholds: “elev_min < 746.0 m”, “slope_min < 44.5°”, “pcurv_mean < 0.009 m⁻¹”, and “tcurv_mean < 0.015 m⁻¹”. The thresholds on maximum “minimum elevation” and “minimum slope” effectively remove exposed boulders. However, these thresholds in total falsely remove 0, 1, 15, and 16 landslide instances, respectively.

Minimum thresholds on standard deviation. We impose four minimum thresholds on standard deviations: “elev_std > 0.7 m”, “slope_std > 2.1°”, “pcurv_std > 0.013 m⁻¹”, and “tcurv_std > 0.016 m⁻¹”. These thresholds ensure the predicted polygons are not on flat and leveled, or on flat and sloped natural surfaces. However, these thresholds in total falsely remove 16, 16, 17, and 12 landslide instances, respectively.

Transferability. The use of elevation and slope thresholds are empirical and only applicable to Lantau Island. When transferring the post-processing to another region, the elevation and slope thresholds are no longer suitable. On the other hand, the thresholds based on profile curvature and tangential curvature limit predictions to be negative or slightly positive, and remove large positive curvatures. Positive profile or tangential curvatures mean bulging outwards, which is impossible for landslide scars. Therefore, the profile curvature and tangential curvature thresholds represent general landslide surface geometries and are transferable to post-process in another region.

Using DSM in post-processing vs. as one input channel. Using DSM in post-processing is advantageous in interpretability, for instance: applying surface geometry thresholds represents selecting possible landslide scars geometry and removing false predictions. However, such post-processing method requires prior knowledge of the distribution of landslide attributes, such as, elevation and slope. Using DSM as one input channel is advantageous in predictive power, for instance: DCNN can learn about simple and complex spatial features on the DSM. However, transferring to another spatial domain may require to retrain DCNN parameters, otherwise it may have inferior prediction performance.

7. Conclusions and Future Work

This work demonstrated an automatic routine of landslide inventory mapping. With a high-resolution (0.5 m by 0.5 m) digital orthophoto and a previous manually compiled landslide inventory map, we first pre-process these data with image splitting and data augmentation for our training dataset. We then train our deep convolutional neural network (DCNN) models to obtain optimized models. The optimized models are put into inference on the original orthophoto with conditional random fields post-processing to obtain a landslide probability map. It is then converted into a landslide polygon map as a resultant product.

Our work designed a new DCNN architecture, with an improved dilated spatial pyramid pooling module, for remote sensing image segmentation. Among other existing DCNN architectures, our proposed model achieved state-of-the-art landslide object detection accuracy (F1 score by count: 0.802) and landslide segmentation accuracy (F1 score by area: 0.893). The model can be used standalone without digital surface model post-processing or can incorporate the post-processing filter for improved accuracy.

Our automatic pipeline provides an efficient and accurate network for mapping landslide inventory over large areas. Our method has great potential for generalizability of landslide features and transferability. Our pipeline can be applied to the whole territory of Hong Kong to obtain a territory-wide landslide polygon mapping. Provided with yearly or immediate post-disaster images, our pipeline can serve as an efficient tool for a more frequent update of landslide inventory, towards routine and effective monitoring of this kind of natural disasters to which Hong Kong is prone.

A possible future work direction is to incorporate multi-year aerial images for other objectives. With the existing Enhanced Natural Terrain Landslide Inventory (ENTLI), a well-trained model from this work can be tested using aerial images for another year, e.g., 2019, and the 2020 LiDAR data for Hong Kong, for evaluating testing accuracy. Another possible task is producing landslide susceptibility maps using pre-landslide images and post-landslide labels to train our model.

Appendix

Table A.0.1. Full Statistics of Surface Geometry of Landslides.

	Minimum	1st percentile	Median	99th percentile	Maximum
elev_mean (m)	2.9	16.3	229.1	553.8	784.8
elev_std (m)	0.1	0.7	4.9	44.0	103.4
elev_max (m)	5.2	26.4	243.6	581.2	819.8
elev_min (m)	2.2	5.3	215.2	530.9	746.0
slope_mean ($^{\circ}$)	2.5	11.7	33.0	46.2	59.0
slope_std ($^{\circ}$)	1.5	2.1	6.2	14.4	27.5
slope_max ($^{\circ}$)	6.9	19.2	47.8	68.8	79.8
slope_min ($^{\circ}$)	0	0.1	17.2	35.6	44.5
pcurv_mean (m^{-1})	-0.173	-0.101	-0.016	0.009	0.037
pcurv_std (m^{-1})	0.010	0.013	0.028	0.121	0.599
pcurv_max (m^{-1})	-0.009	0.015	0.064	0.427	2.756
pcurv_min (m^{-1})	-1.795	-0.909	-0.108	-0.035	-0.015
tcurv_mean (m^{-1})	-0.211	-0.143	-0.054	0.015	0.084
tcurv_std (m^{-1})	0.007	0.016	0.049	0.159	0.327
tcurv_max (m^{-1})	-0.079	-0.012	0.079	0.466	2.695
tcurv_min (m^{-1})	-6.287	-0.973	-0.203	-0.030	0.000

Remarks: Emboldened values are empirical thresholds to be presented in Table 4.1.

Bibliography

- Blaschke, T., Feizizadeh, B., & Hölbling, D. (2014). Object-based image analysis and digital terrain analysis for locating landslides in the Urmia Lake Basin, Iran. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(12), 4806-4817. <https://doi.org/10.1109/JSTARS.2014.2350036>
- Chatwin, S. C., Howes, D. E., Schwab, J. W., and Swanston, D. N. (1994). *A guide for management of landslide-prone terrain in the Pacific Northwest* 2d edition: Research Branch, Ministry of Forests, Province of British Columbia, Victoria, British Columbia, Crown Publications.
- Chen, Y., Jiang, H., Li, C., Jia, X., & Ghamisi, P. (2016). Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10), 6232-6251. <https://doi.org/10.1109/TGRS.2016.2584107>
- Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).
- Cruden, D. M., & Varnes, D. J. (1996). Landslides: investigation and mitigation. Chapter 3-Landslide types and processes. *Transportation research board special report*, (247).
- Dias, A., Hart, J. R., & Fung, E. K. S. (2009, April). The enhanced natural terrain landslide inventory. In *Natural Hillsides: Study and Risk Mitigation Measures. Proceedings of the Hong Kong Institution of Civil Engineers, Geotechnical Division, Annual Conference* (Vol. 17, pp. 71-78).
- Duan, Y., Liu, F., Jiao, L., Zhao, P., & Zhang, L. (2017). SAR image segmentation based on convolutional-wavelet neural network and Markov random field. *Pattern Recognition*, 64, 255-267. <https://doi.org/10.1016/j.patcog.2016.11.015>
- Highland, L., & Bobrowsky, P. T. (2008). The landslide handbook: a guide to understanding landslides (p. 129). Reston: US Geological Survey.

- Hu, W., Huang, Y., Wei, L., Zhang, F., & Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors, 2015*. <https://doi.org/10.1155/2015/258619>
- Kampffmeyer, M., Salberg, A. B., & Jenssen, R. (2016, June). Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 680-688). IEEE. <https://doi.org/10.1109/CVPRW.2016.90>
- Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected crfs with gaussian edge potentials. *Advances in neural information processing systems, 24*.
- Kurtz, C., Stumpf, A., Malet, J. P., Gançarski, P., Puissant, A., & Passat, N. (2014). Hierarchical extraction of landslides from multiresolution remotely sensed optical images. *ISPRS Journal of Photogrammetry and Remote Sensing, 87*, 122-136. <https://doi.org/10.1016/j.isprsjprs.2013.11.003>
- Lyu, H., Lu, H., & Mou, L. (2016). Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sensing, 8*(6), 506. <https://doi.org/10.3390/rs8060506>
- Lyu, H., Lu, H., Mou, L., Li, W., Wright, J., Li, X., ... & Gong, P. (2018). Long-term annual mapping of four cities on different continents by applying a deep information learning method to landsat data. *Remote Sensing, 10*(3), 471. <https://doi.org/10.3390/rs10030471>
- Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). High-resolution aerial image labeling with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing, 55*(12), 7092-7103. <https://doi.org/10.1109/TGRS.2017.2740362>
- Makantasis, K., Karantzalos, K., Doulamis, A., & Doulamis, N. (2015, July). Deep supervised learning for hyperspectral data classification through convolutional neural networks. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 4959-4962). IEEE. <https://doi.org/10.1109/IGARSS.2015.7326945>
- Mondini, A. C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., & Ardizzone, F. (2011). Semi-automatic recognition and mapping of rainfall induced shallow

landslides using optical satellite images. *Remote sensing of environment*, 115(7), 1743-1757. <https://doi.org/10.1016/j.rse.2011.03.006>

Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognition*, 106, 107404. <https://doi.org/10.1016/j.patcog.2020.107404>

Rau, J. Y., Jhan, J. P., & Rau, R. J. (2013). Semiautomatic object-oriented landslide recognition scheme from multisensor optical imagery and DEM. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2), 1336-1349. <https://doi.org/10.1109/TGRS.2013.2250293>

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

Shi, W., Zhang, M., Ke, H., Fang, X., Zhan, Z., & Chen, S. (2020). Landslide recognition by deep convolutional neural network and change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 59(6), 4654-4672. <https://doi.org/10.1109/TGRS.2020.3015826>

Su, Z., Chow, J. K., Tan, P. S., Wu, J., Ho, Y. K., & Wang, Y. H. (2021). Deep convolutional neural network-based pixel-wise landslide inventory mapping. *Landslides*, 18(4), 1421-1443. <https://doi.org/10.1007/s10346-020-01557-6>

Sugawara, Y., Shiota, S., & Kiya, H. (2019). Checkerboard artifacts free convolutional neural networks. *APSIPA Transactions on Signal and Information Processing*, 8. <https://doi.org/10.1017/AT SIP.2019.2>

Vakalopoulou, M., Platias, C., Papadomanolaki, M., Paragios, N., & Karantzalos, K. (2016, July). Simultaneous registration, segmentation and change detection from multisensor, multitemporal satellite image pairs. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1827-1830). IEEE. <https://doi.org/10.1109/IGARSS.2016.7729469>

Volpi, M., & Tuia, D. (2016). Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 881-893. <https://doi.org/10.1109/TGRS.2016.2616585>

- Yang, X., & Chen, L. (2010). Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides. *International Journal of Applied Earth Observation and Geoinformation*, 12(6), 487-495. <https://doi.org/10.1016/j.jag.2010.05.006>
- Zhang, W., Lin, J., Peng, J., & Lu, Q. (2010). Estimating Wenchuan Earthquake induced landslides based on remote sensing. *International Journal of Remote Sensing*, 31(13), 3495-3508. <https://doi.org/10.1080/01431161003727630>
- Zhang, Z., Wang, H., Xu, F., & Jin, Y. Q. (2017). Complex-valued convolutional neural network and its application in polarimetric SAR image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 7177-7188. <https://doi.org/10.1109/TGRS.2017.2743222>
- Zhou, Y., Wang, H., Xu, F., & Jin, Y. Q. (2016). Polarimetric SAR image classification using deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 13(12), 1935-1939. <https://doi.org/10.1109/LGRS.2016.2618840>