LANDSLIDE DETECTION FROM SATELLITE IMAGERY USING A DEEP LEARNING TECHNIQUE

Yung-ming Tsai (1), Kuo-Chen Chang (1)

[1] National Taiwan Normal University, Department of Geography, No. 162 Heping E. Road, 106, Taipei City, Taiwan
Email: yungmingtsai@outlook.com

KEY WORDS: Landslide, Image Classification, Satellite Imagery, U-net

ABSTRACT:
Taiwan is subject to severe natural hazards like earthquakes and typhoon, which often cause landslides in mountainous area, claiming crops, property safety and even lives. Monitoring the occurrence of landslides using remote sensing is an annual task for government institutions. However, the task had been extremely labor-intensive and time consuming. In order to solve the problem, this study proposed a deep learning technique for automatic landslide classification from satellite imagery in order to get a more accurate and robust classification results.

The classification model is based on the U-Net convolutional neural network, implemented with the deep learning toolkit, CNTK. The model takes pairs of satellite imagery and ground truth label as the input and produces predicted classified labels as the output. The model is trained on pairs of FORMOSAT-2 imagery and ground truth labels. The ground truth is classified into 5 classes: vegetation, riverbed, landslides, water and miscellaneous. To best separate landslides from other impervious land cover like riverbed and farmlands, slope degree is added to satellite imagery to provide distinguish information for classification. Imagery from Kompsa-3 is used to test the sensor independency of the model.

The study produces a result of a robust classification model that is able to distinguish landslides from the satellite imagery. The landslides classification results are accurate for FORMOSAT-2 imagery. The model is also able to produce similar results on Kompsat-3 imagery when scaled to the image depth of 8-bit despite of the difference in spatial resolution. The model is reusable, and the process is fully automated. We expect the model will be useful on landslides monitoring and inventory mapping, which are elementary task for hazard mitigation and susceptibility mapping.

INTRODUCTION

Taiwan is a mountainous environment, with approximately 70 percent of the island being mountains. Given the limited flat areas, settlements and agricultural activities often occur in mountainous areas. The residents are exposed to the risk of hazardous landslides due to the frequency of earthquakes, typhoons and strong rainfalls. Landslide events often claim serious casualty and property loss. To monitor the landslide events for the sake of hazard mitigation, landslides inventory mapping is an annual routine for the authority.

Remote sensing techniques have been used widely for the purpose of landslide monitoring and inventory mapping. The task of landslide mapping can be done successfully with remote sensing data sources like aerial photographs and Very High Resolution (VHR) satellite imagery (Van Western, 2000). The mapping process can be done with manual interpretation. However, manual interpretation is both labor intensive and time consuming. Several techniques have been proposed for automatic or semi-automatic landslides image detection or classification. Research suggests that change detection analysis with optical satellite imagery can identify land cover changes caused by landslide events. (Hervás et al., 2003) Change detection analysis requires time sequential imagery that are geometrically rectified and radiometrically normalized, so multiple images can be compared pixel by pixel. Study also shows the integration of DTM data can help the extraction of landslides. (Chang, Wei & Chang, 2003)

Other studies have adopted image classification techniques to tackle the problems. Borghius, Chang & Lee (2007) compared the performance of automated image classification using maximum likelihood classifier with manual mapping. The results of unsupervised classifier using SPOT-5 imagery and slope filter shows a concordance of 63% with manual mapping results. The classifier shows less omission on small landslides than manual mapping, but larger commission to the classification of farmlands, riverbeds and road. A common approach currently for image classification is object-based image analysis (OBIA). Researchers have grown strong interest in OBIA, and the paradigm of pixel-based analysis has shifted to object-based analysis. (Blaschke, 2010) OBIA solves the "salt and pepper effect" of pixel-based classification. Furthermore, OBIA can extract more information from the imagery, providing extra spectrum statistical information like mean value or standard deviation, along with other object features like size, shape, texture or location. Object-based image classification usually involves with two processes, image segmentation and object classication. Aksoy & Ercanoglu (2012) uses object-based image analysis for landslide identification. Multiresolution segmentation is used as the image segmentation algorithm. Multiresolution segmentation (Baätz & Schäpe, 2000) is a bottom-up approach for segmentation that groups homogeneous pixels into a same object, and hence producing a more meaningful object segmentation. This algorithm is used in object-based image classification software like eCognition. For object classification, fuzzy classification is used in this study. Aside from spectral information, other variables like NDVI, slope, plan curvature, brightness, shape, texture and neighboring are also used for classification. The results are similar to the one mapped by experts. Classifiers like random forest are alternative options for classification. (Stumpf & Kerle, 2011)

OBIA have been used extensively in the task of image classification, yet there are some drawbacks for it. First, in supervised object-based classification, the selection of training sites is very crucial, and have direct impact on the results. Training selection will require human operations, so the process can not be fully automized. The involvement of human operation also adds uncertainty to the model, and the results of classification will be inconsistent when operated by different people. To produce consistent results, we will need a classifier that can be reused. Currently, neural network with multiple hidden layers structures, also known as deep learning, is the trend. Deep learning has been used widely on the task of classification for multiple applications. Most deep learning approaches for image object detection or classification are based on convolutional neural network (CNN). CNN can perform different tasks of image recognition. Image classification, like ImageNet (Krizhevsky, Sutskever & Hinton, 2012), classifies images into different categories. Object detection finds the location of objects in the image and produce the bounding box of the objects. Common object detection algorithms are Fast R-CNN (Girshick, 2015), Faster R-CNN (Ren et al., 2015) and YOLO. (Redmon et al., 2016) Semantic segmentation classifies each pixel of the image into different classes. U-net is a semantic segmentation deep learning framework, which is originally designed for biomedical image segmentation. (Ronneberger, Fischer & Brox, 2015) U-net is also adopted widely in satellite imagery analysis for different purposes like land cover mapping. (Iglovikov, Mushinskiy & Osin, 2017; Rakhlin, Davydow & Nikolenko, 2018) Zhang, Liu & Wang (2018) proposed a deep residual U-net framework for road extraction, combining ResNet and the original U-net. The algorithm is also implemented in the GeoAI platform from Microsoft for pixel level land cover classification. (Microsoft, 2018) Bai, Mas & Koshimura (2018) utilities the GeoAI platform to perform damage-mapping from satellite imagery. U-Net has been proofed to be able to do semantic classification with air photo or satellite imagery. The results are generally accurate.

This paper proposes a reusable image classifier using U-net deep learning framework for accurate and fast landslides mapping from satellite imagery.

DATA SOURCES

In this study, FORMOSAT-2 satellite imagery is used as the sources of training data. FORMOSAT-2 is an optical satellite, with four multispectral bands: red, green, blue and near infra-red. The spatial resolution of the panchromatic band is 2-meter, while that of the multispectral bands is 8-meter. The radiometric resolution of the imagery is 8-bits.

The data used in this study is pansharpened, meaning the imagery is multispectral imagery with red, green, blue and near infra-red bands, and the spatial resolution of the imagery is 2-meter. Pansharpened imagery can keep the details from panchromatic image while keeping the spectral information from multispectral bands. A total of 250 imagery is used as training data and 25 imagery as test sets. The size of each image is 1000 meters * 1000 meters. Each imagery contains 250000 pixels.

Each imagery is paired with a land cover map. A land cover map serve s as the ground truth labels of the data. It determines the real class of each pixel in the image. The land cover mapping is done manually from each image. Land cover is classified into 5 categories for this model: vegetation, water, riverbed, landslides and miscellaneous land use like farm or buildings. The land cover maps are presented in raster format, with a resolution of 2-meter in corresponding to the satellite imagery.

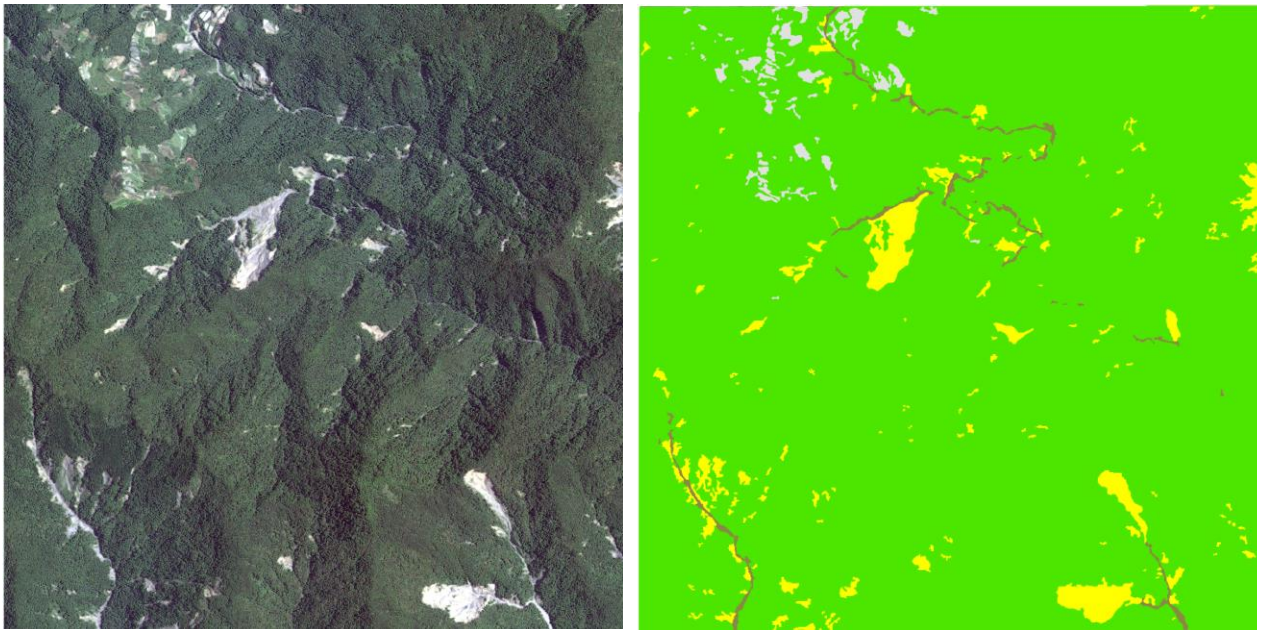

Figure 1 Samples for Formosat 2 imagery and the corresponding land cover mapping. (Green: vegetation, Yellow: landslide, Brown: riverbed, Gray: miscellaneous)

Topographic data are also used in the model. The source of the topographic data is the DTM retrieved from open data platform of the government. The DTM is produced by LiDAR, downscaled to the resolution of 20 Meters.

METHODOLOGY

This study proposes a U-Net classification model that takes satellite imagery as the input and produces outputs with 5 classes: vegetation, water, landslide, riverbed and miscellaneous land cover. The figure below is the flowchart of the study design.

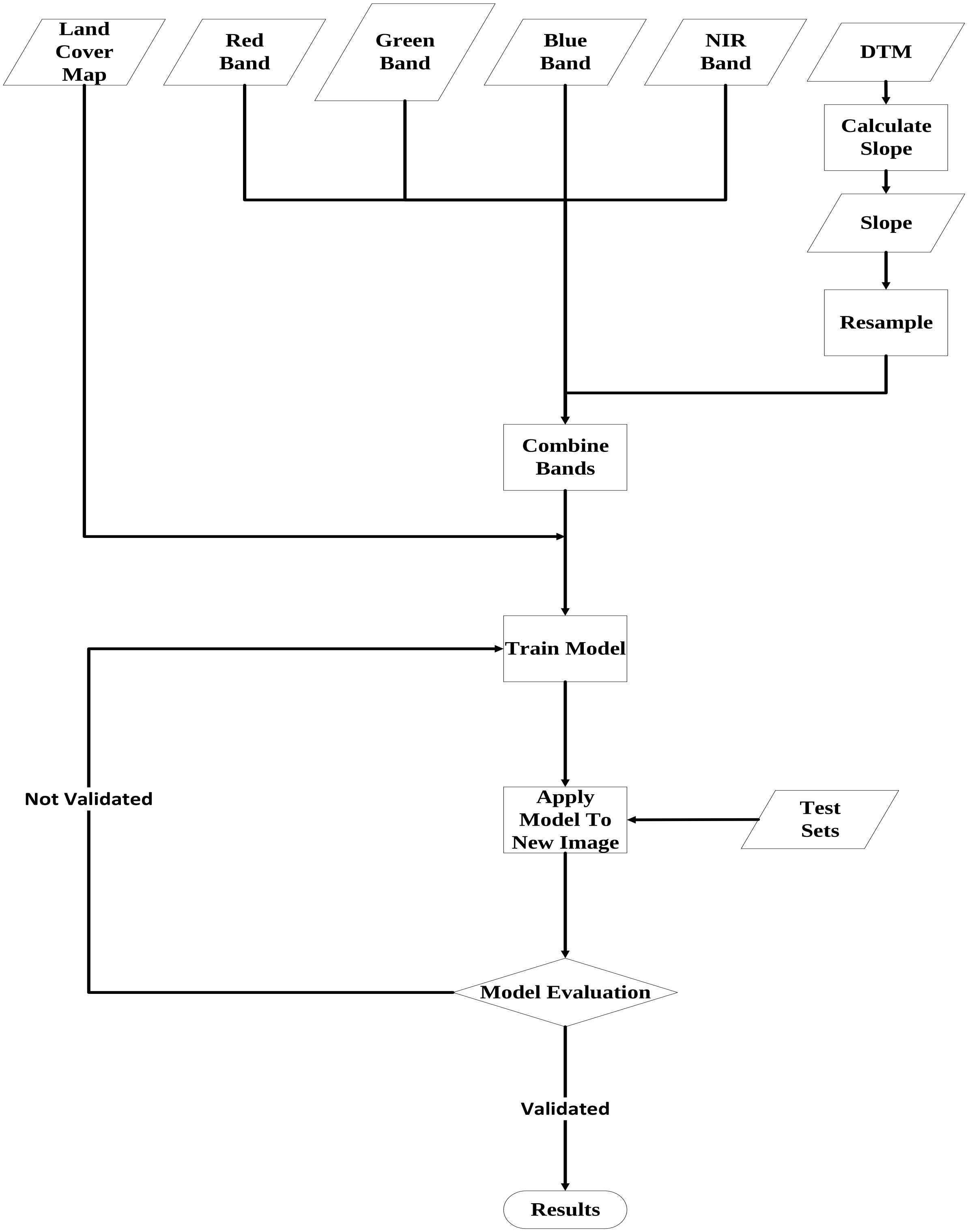

Figure 2 Flowchart for study design.

The model takes all four bands from the satellite imagery. The satellite imagery is orthorectified and pansharpened using PCI Geomatica software. In order to better distinguish landslides from riverbed and farmland, slope derived from 20x20 meters DTM is added as additional input. To match the resolution of satellite imagery, the slope layer is resampled and upscaled to 2-meter resolution. Satellite imagery and slope are combined, forming a 5-band imagery. Land cover maps are manually digitized from the corresponding satellite imagery and stored as GeoTiff with 2-meter resolution.

Table 1 Composition of training imagery samples.

Order	Band	Resolution
1	Red	2 meters
2	Green	2 meters
3	Blue	2 meters
4	Near Infra-red	2 meters
5	Slope	20 meters (upscaled to 2 meters)

The deep learning framework used for the study is the Deep Residual U-net framework developed by Microsoft in cooperated with ESRI for the GeoAI platform. The model is originally designed for land cover mapping using air photo with four bands: red, green, blue and near infra-red. For the model to take slope information as inputs, the input layer is altered for inputting 5 bands. The input for the model is a pair of satellite imagery and its corresponding land cover map, each with sizes at 1000 * 1000 meters. Batch normalization is adopted in model. The ReLU function is used as the activation function. Figure 3 below shows the framework of the deep learning model.

In the training phase, we use 250 imagery as the input. The model was trained with 500 epochs with 2000 minibatches per epoch. RMSProp is used as the optimization algorithm for learning. The model is implemented with Python language interface and Microsoft Cognitive Toolkit (CNTK). CNTK enables multiple GPU parallel processing for faster training performance.

The output of the model can be evaluated using ESRI ArcGIS Pro. The model can be applied to a new imagery that is not seen by the model during training. The results of the evaluation will be a raster layer with 5 predicted classes from the classification. Producer accuracy and user accuracy of landslide class can then be calculated to evaluate the performance of the classification.

Figure 3 Deep Residual U-Net framework.

RESULTS AND DISCUSSION

The training progress in this study is done under a single workstation environment. The configuration of the workstation is as follows. The CPU is a hexa-core CPU with the base clock of 3.70 GHz. The capacity of RAM is 32 GB. The GPU is a NVIDIA GeForce GTX 1070 Ti with 8 GB of video memory.

Table 2 Specifications of workstation environment for model training.

Hardware	Specifications
CPU	Intel Core i7- 8700K
RAM	32 GB
GPU	NVIDIA GeForce GTX 1070 Ti
Operating System	Windows 10 x64

The model is trained for 500 epochs with 2000 minibatches per epoch. It takes approximately 4.8 hours to finish the training using GPU parallel processing. The training loss of each epoch is shown in Figure 4.

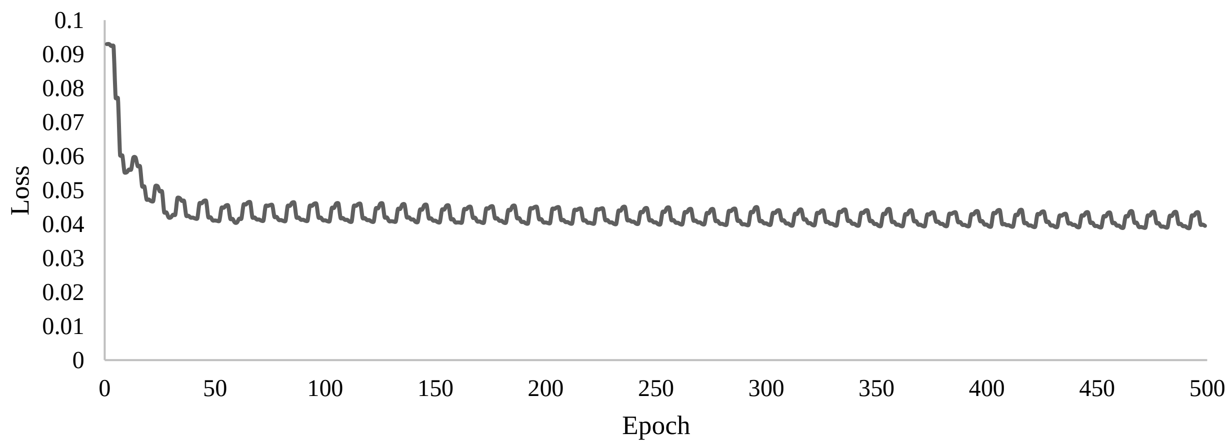

Figure 4 Training loss of each epochs.

The model is applied to a set of new imagery that the model has never seen before during training. The evaluation process is done under a client environment. The environment is a laptop configured with a hexa-core CPU with the base clock of 2.20 GHz, a 16 GB RAM and a Max-Q design NVIDIA GeForce GTX 1060 with 6 GB of video memory. It takes approximately 5 seconds to apply classification to an imagery with 2500 * 2500 pixels. The classification results are in Figure 6, and the confusion matrix calculated using 500 checkpoints is shown in Table 4. As shown in the results, the classifier performs well on distinguishing landslides. The producer accuracy and user accuracy of landslides mapping are approximately 0.926 and 0.862 respectively. The classifier successfully distinguished landslides from riverbeds, which are easily confused with each other when using traditional image classification methods. The classifier performs relatively poor on darker regions. Also, border effect is presented in the results. Pixels around the corner tends to be classified incorrectly. Tiling the imagery could solve the problem of border effect.

Table 3 Specifications of client environment for model evaluation.

Hardware	Specifications
CPU	Intel Core i7- 8750H
RAM	16 GB
GPU	NVIDIA GeForce GTX 1060 Max-Q
Operating System	Windows 10 x64

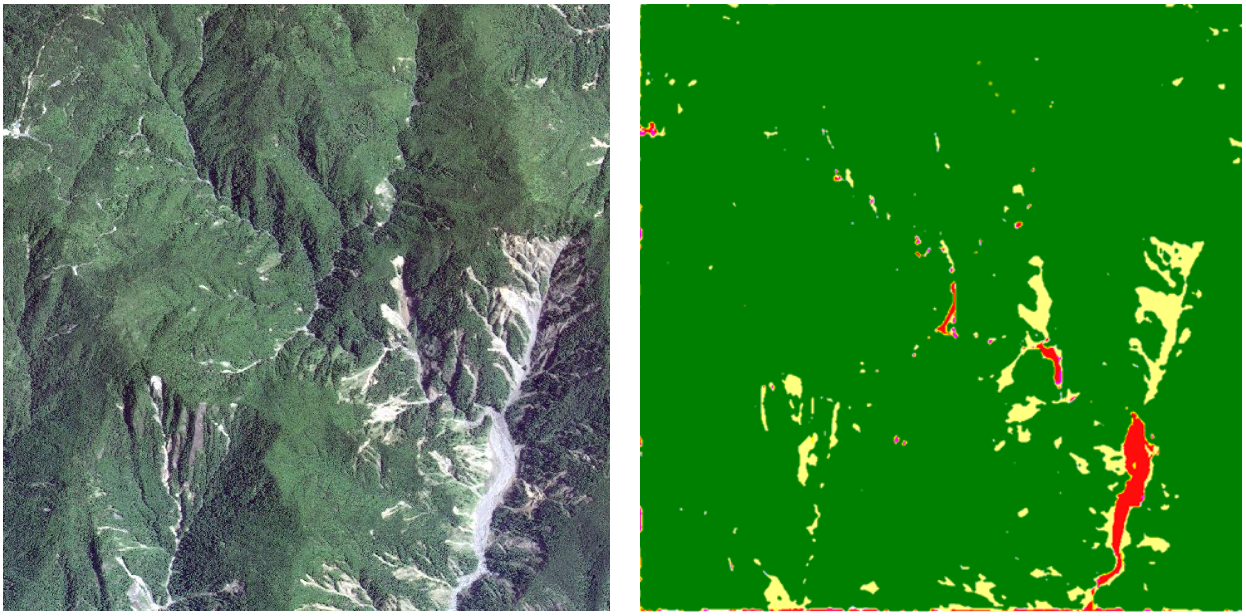

Figure 5 Classification results of FORMOSAT-2. (Dark Green: vegetation, Light Green: landslide, Red: riverbed, Magenta: Miscellaneous)

Table 4 Confusion matrix of FORMOSAT-2 classification based on 500 checkpoints

		Ground Truth Class	
		Landslides	Non-Landslides
Classified Class	Landslides	50	4
	Non- Landslides	8	438

To test the sensor independency of the model, the model is applied to data from other sources. Imagery from Kompsat-3 is used for the test. Compared to FORMOSAT-2, Kompsat-3 has higher ground resolution and radiometric resolution. Its ground resolution is 0.7 meter for panchromatic bands and 2.8 meter for multispectral bands consisting of red, green, blue and near infra-red bands. The radiometric resolution of the imagery is 14-bit. The imagery used is a pansharpened imagery.

When using the 14-bit Kompsat-3 imagery as the input, the model fails to classify the imagery and produces meaningless classification results. However, the model is able to classify the imagery when the DN value is scaled to the range of 8-bit unsigned integer. We use the adaptive color balancing function of PCI Geomatica software to scale the imagery to the image depth of 8-bit unsigned integer. The model performance of classification on the Kompsat-3 imagery is similar to that on the FORMOSAT-2 imagery despite the difference in spatial resolution, with the producer accuracy of 0.833 and the user accuracy of 0.929. Most commission error is caused by shadows, which the model is more incapable of dealing with.

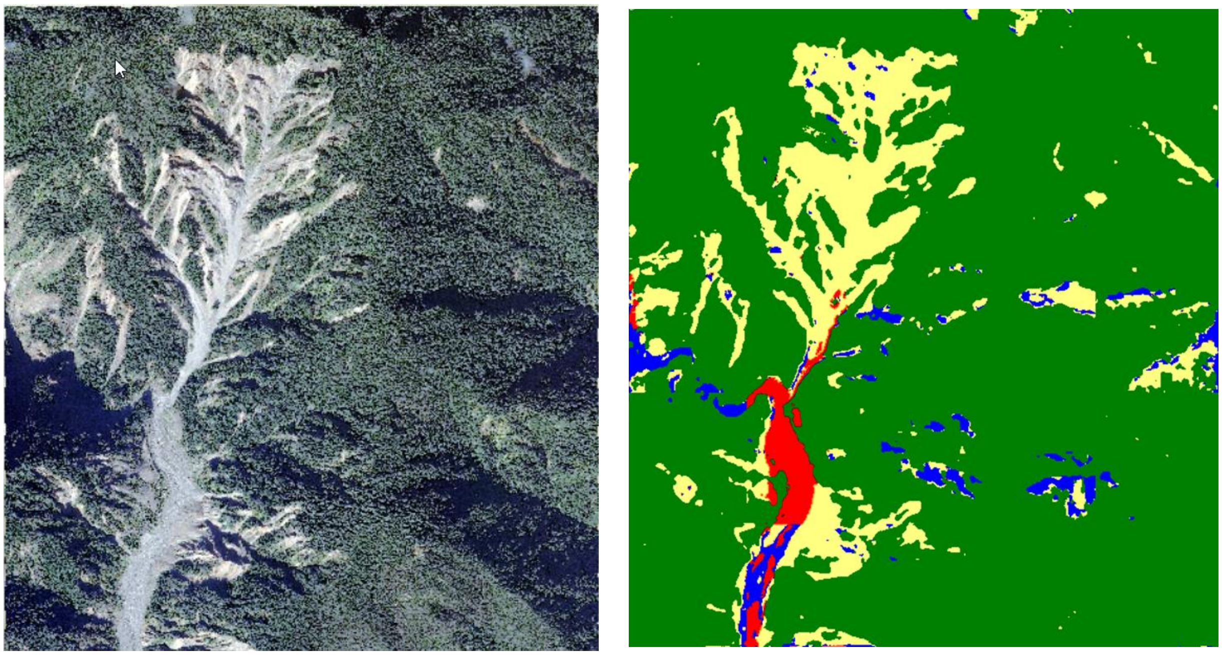

Figure 6 Classification results of Kompsat-3. (Dark Green: vegetation, Light Green: landslide, Red: riverbed, Blue: Water, Magenta: Miscellaneous)

Table 5 Confusion matrix of Kompsat-3 classification based on 500 checkpoints

		Ground Truth Class	
		Landslides	Non-Landslides
Classified Class	Landslides	65	13
	Non- Landslides	5	417

Based on the results, we can claim that U-Net classifier is capable of preforming the task of landslides mapping. There are several advantages of adopting U-Net deep learning framework to the classification workflow. First, the classification workflow is fast. The workflow can be completely automatic. Without the need for manually providing training sites and parameters, the classification results are consistent every time given the exact same inputs and models. Second, the model is more tolerant towards incorrect ground truth. In traditional supervised classification, selecting training sites is crucial for the classification results. Bad selection will cause poor classification. For the deep learning model, small error from the ground truth is more acceptable as long as there are enough samples to train. Third, the classification model is accessible. The model only requires a few data to perform. Multispectral satellite imagery with a combination red, green, blue and near infra-red bands, which is common for most optical sensors. Slope can be calculated from DTM, which can be easily accessed from open data platform. No need for extra topographic or lithographic information. The classifier produces good results. Comparing to pixel-based supervised classification, the deep learning framework produces less “salt and pepper” effect. It can also identify the border between landslides and riverbeds. Finally, the model can be applied to imagery from different sensors. The model produces consistent results on imagery with different spatial resolution if they have the same image depth.

CONCLUSION

This study proposes a deep learning model for landslides detection from satellite imagery. Deep Residual U-Net framework is used as a classifier. Formosat-2 satellite imagery is used as the training data. The model takes ground truth label and satellite imagery with 5 bands as inputs. The 5 bands are red, green, blue, near infra-red and slope respectively. The output is classification results of 5 classes: vegetation, water, landslides, riverbed and miscellaneous land cover.

A total of 250 imagery is used as inputs. The model is trained for 500 epochs, with 2000 minibatches per epoch. The timespan of training is around 4.8 hours under a single GPU workstation environment. The classifier successfully identifies landslides, with the producer accuracy and user accuracy of landslides mapping being approximately 0.926 and 0.862 respectively when tested on FORMOSAT-2 imagery. The model produces similar results on Kompsat-3 imagery scaled to the image depth of 8-bit despite of the difference in spatial resolution, with the producer accuracy and user accuracy of landslides mapping being approximately 0.833 and 0.929 respectively.

The model improves the workflow of landslides inventory mapping. Without human operation, the process requires less labor and expertise, and the results are much more consistent each time. The model is more tolerant to incorrect ground truth, so that outdated land cover maps will not affect the overall performance of the classification. The model is very accessible, with less data required. All the model need is satellite imagery with the common four bands combination, and slope data derived from DTM. The model is also sensor independent if the imagery is scaled to the same image depth. The model can be easily adopted into workflow without the concern of niche data sources that are difficult to acquire.

The model will be useful on landslides monitoring and inventory mapping, which are elementary task for hazard mitigation and susceptibility mapping. For future work, we will seek to adapt the model to more land cover types and hazardous landscapes.

REFERENCES

Aksoy, B., & Ercanoglu, M. (2012). Landslide identification and classification by object-based image analysis and fuzzy logic: An example from the Azdavay region (Kastamonu, Turkey). Computers & Geosciences, 38(1), 87-98.

Baätz, M., Schäpe, A., Strobl, J., Blaschke, T., & Griesebner, G. (2000). Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. Angewandte Geographische Informations-Verarbeitung XII, 12-23.

Bai, Y., Mas, E., & Koshimura, S. (2018). Towards operational satellite-based damage-mapping using U-net convolutional network: a case study of 2011 Tohoku Earthquake-Tsunami. *Remote Sensing*, *10*(10), 1626.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, *65*(1), 2-16.

Borghuis, A. M., Chang, K., & Lee, H. Y. (2007). Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. International Journal of Remote Sensing, 28(8), 1843-1856.

Cheng, K. S., Wei, C., & Chang, S. C. (2004). Locating landslides using multi-temporal satellite images. Advances in Space Research, 33(3), 296-301.

Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

Hervás, J., Barredo, J. I., Rosin, P. L., Pasuto, A., Mantovani, F., & Silvano, S. (2003). Monitoring landslides from optical remotely sensed imagery: the case history of Tessina landslide, Italy. Geomorphology, 54(1-2), 63-75.

Iglovikov, V., Mushinskiy, S., & Osin, V. (2017). Satellite imagery feature detection using deep convolutional neural network: A kaggle competition. *arXiv preprint arXiv:1706.06169*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Rakhlin, A., Davydow, A., & Nikolenko, S. I. (2018, June). Land Cover Classification From Satellite Imagery With U-Net and Lovasz-Softmax Loss. In *CVPR Workshops* (pp. 262-266).

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91-99).

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.

Stumpf, A., & Kerle, N. (2011). Object-oriented mapping of landslides using Random Forests. *Remote sensing of environment*, *115*(10), 2564-2577.

Van Westen, C. J. (2000). Remote sensing for natural disaster management. International archives of photogrammetry and remote sensing, 33(B7/4; PART 7), 1609-1617.

Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, *15*(5), 749-753.