Optimizing Sample Patches Selection of CNN to Improve the mIOU on Landslide Detection

Omid Ghorbanzadeh[a] and Thomas Blaschke[b]
Department of Geoinformatics - Z-GIS, University of Salzburg, Salzburg, Austria

Keywords: Convolutional Neural Network, RapidEye, mean Intersection Over Union, Training Data Set.

Abstract: Remarkable improvement has been made in object detection and image classification, mainly due to the availability of large-scale labelled data and also the progress of deep convolutional neural networks (CNNs). Thus, this amount of training data enables CNNs to learn data-driven image features. However, generating the efficient sample patches from the satellite images for training the CNNs remains a challenge. In this study, we use a CNN for the case of landslide detection based on the optical data from the Rapid Eye satellite. We separate the image into training and test areas of the highly landslide-prone Rasuwa district in Nepal. Thus, the sample patches were extracted from the training area of the Rapid Eye image. Although the approach of random sample patches is considered as the most common for feeding the CNNs, it is not the best solution for all object detection aims. We feed our structured CNN with the randomly selected sample patches as our first approach. For the second approach, the same CNN architecture is trained by the patches that selected based on only the central areas of any landslide. The trained CNNs based on both approaches were used to detection the landslides in an area where considered as our test zone. The detection results are compared against a precise inventory dataset of landslide polygons through a mean intersection-over-union (mIOU). The mIOU value of the first approach is 53.56%. However, that of the second one is 56.24%, which shows an approximately 3% improvement in the resulting accuracy of the landslide detection using the sample patches generated by the second approach. Rather, the current performance of CNNs in object detection domain they strongly depend on the quality of the training data and augmentation strategies.

1 INTRODUCTION

Landslide detection has been considered as one of the important active study domains in remote sensing today because of the adverse consequences of this natural hazard on the human habitation (Hong et al., 2017). It is essential regarding fast response actions after a destructive landslide. Although there are some new field surveying methods for landslide detection and mapping, e.g. laser rangefinder binoculars by applying a GPS receiver (Guzzetti et al., 2012), the problems of the accessing to such areas still remains a challenge . Therefore, remotely sensed imagery is the most accessible data providing critical information required for supporting humanitarian response (Lang et al., 2017). Analysis and classification of the remotely sensed imagery for extracting landslides have done in several studies. Previous researches have primarily focused on detecting the changes occurred on the environment due to the landslides based on the remotely sensed imagery and some knowledge-based methods or manually image processing methods (Amit and Aoki, 2017). Moreover, different machine learning techniques, e.g. MLP Neural Nets have been used for landslide detection (Mezaal et al., 2017; Bui et al., 2016). (Moosavi et al., 2014) proposed a landslide detection approach based on support vector machines to find whether the occurrence of the landslide.

Recently, convolutional neural networks (CNNs) have become the new hot topic in various image processing domains and object detection in particular (Zhang et al., 2018). CNNs are specific

[a] https://orcid.org/0000-0002-9664-8770
[b] https://orcid.org/0000-0002-1860-8458

Ghorbanzadeh, O. and Blaschke, T.
Optimizing Sample Patches Selection of CNN to Improve the mIOU on Landslide Detection.
DOI: 10.5220/0007675300330040
In *Proceedings of the 5th International Conference on Geographical Information Systems Theory, Applications and Management (GISTAM 2019)*, pages 33-40
ISBN: 978-989-758-371-1

kind of deep learning techniques based on artificial neural networks. CNNs can directly get the images as input data, to avoid the traditional approaches with pre-processing methods and feature extraction by the complex operations (Yu et al., 2017). They have achieved acceptable results in wide range of image analysis tasks in computer vision (Zhu et al., 2017; Ghorbanzadeh et al., 2018). There are several studies that used CNNs for image segmentation (Längkvist et al., 2016), scene classification (Qayyum et al., 2017), and object detection (Radovic et al., 2017). The large numbers of labelled images along with CNNs were used to some object detection aims, e.g. airplanes, vehicles, and some specific trees. The availability of massive amount of labelled images is considered as one of the main reasons for achieving fairly good results by CNNs. However, the way of using these data for training the CNNs is still a topic of discussion. Randomly sample patches selection is the common and traditional way to patch extraction for the CNNs, but not the best method for any application. A critical problem in object detection using CNNs is the selection way of sample patches. Because in some cases such as landslide detection results with poor quality when the sample patches are selected randomly. Therefore, the method of selection of the patches can be improved regarding the target object that should be detected. For example, the Genetic Algorithm was used to identify the best sample patches from all of the selected patches of tile-based texture synthesis by (Dong et al., 2005). In another study (Zhang et al., 2018) used the Moment bounding (MB) box for identifying the location of the optimal patches on objects in the urban land use classification. However, using the mentioned approaches for the case of landslide detection has some difficulties regarding the various shapes of landslides.

In this study, ones we use the conventional approach of a random selection of sample patches. Then we selected the sample patches were located on the central part of any landslide. Most of the landslides have linear shape started from SCAR (area of initial failure) to the deposition area (Fan) that leads to a high ratio of length to width. Thus, we selected the patches of the central areas of the landslides to get those with the most area from landslides. Both approaches of randomly and central selection of sample patches were implemented on optical satellite imagery from the Rapid Eye sensor. We compare the results from the CNNs based on both approaches to illustrate the performance of each approach and its impact on landslide detection. For comparison, the resulting detected landslide the mean intersection-over-union (mIOU) accuracy assessment method was used.

2 STUDY AREA

The case study area lies in the southern part of the Rasuwa district in Nepal (see figure 1). The study area has an area of about 1544 km2. The land cover is mostly forest, followed by shrub land, grassland, agriculture, and villages. This district is located in the higher Himalayas and is one of the most landslide-prone areas along the Trishuli River. Some of the known landslides had adverse consequences on the built-up areas and have already caused casualties in settlement areas. Landslides have also destroyed the bridges and roads of the main transport corridor between this country and China.

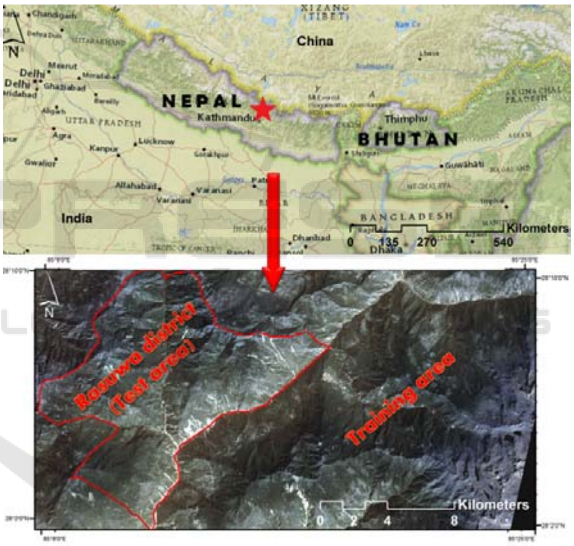

Figure 1: The geographic location of the study area.

3 METHODOLOGY

3.1 Overall Methodology

The Rapid Eye images were used to evaluate the performance of two approaches of randomly and central sample patches selection within a structured CNN for the detection of landslides. The workflow of the present study is as follow:

- Landslide inventory data set creation;
- Designing the training data set of the spectral information;

- Generating the random sample patches by considering a window size of 32×32 pixels;
- Generating the central sample patches by considering the same window size;
- Structuring CNN;
- Testing and validating the performances of each sample patches selection approaches using mIOU method.

The experimental results and related descriptions of this study are organized in the following sections. More explanations and discussions about the impact of using different approaches on the resulting landslide detection can be found in the conclusion section.

3.2 Landslide Inventory

Our landslide inventory data set was generated within an extensive field survey in the Rasuwa district in the higher Himalayas using a GPS device (Garmin Etre 20X). The resulting GPS polygons of landslides were then manually boosted using the satellite images. Therefore, our inventory data set was generated using the GPS data, correcting instances, and finally adding landslide areas visible in the image but not mapped in the field. The Geographic Information System ArcGIS 10.3 was used for the correction process.

3.3 Data

The data used for the present study is from RapidEye that is a constellation of five Earth-observing satellites with a height of 680 km, the swath width of 77 km and a 5-day revisit period.

These five satellites deliver sun synchronous of 5 m spatial resolution images (Mahdianpari et al., 2018). Two RapidEye cloud-free satellite images were used for this study. We used multispectral bands (Red, Green, Blue, Red Edge, and Near Infrared) of RapidEye as following:

- Blue 440 – 510 nm;
- Green 520 – 590 nm;
- Red 630 – 685 nm;
- Red Edge 690 – 730 nm;
- Near-Infrared 760 – 850 nm.

Moreover, the normalized difference vegetation index (NDVI) as a widely used ratio was calculated from the near-infrared and the red spectral bands (Modzelewska et al., 2017). Therefore, we prepared a data set of the spectral information of RapidEye and the NDVI.

3.4 Convolution Neural Network (CNN)

CNNs have introduced state-of-the-art results for image processing and computer vision (Zhang et al., 2018). Multi-layer neural networks of a CNN can obtain the important feature representations of an image. Thus, these networks can distinguish the visual laws in the image without any expert-designed complex rule (Ding et al., 2016). CNNs have a basic architecture, where each so-called hidden layer normally contains convolutional and pooling layers, whereby the convolutional layers are considered as the main building block of any CNN (Ghorbanzadeh et al., 2018). The sample patches of the input image are convolving with a set of trainable kernels that scan across the entire input patch resulting in a group of feature maps. Therefore, the set feature maps result from the convolution of the filter, with its corresponding local region on the original sample patches of the input image.

Structuring a CNN with the architecture that results in the best performance vary regarding the application and still is an ongoing discussion in the deep learning field (Csillik et al., 2018). In this study, a seven-layer depth CNN was structured and trained separately with sample patches resulting from both random and central approaches. This layer depth was selected according to our sample patches size of 32×32 through cross-validation. By using two different sample patches and the same CNN, we could investigate the impacts of sample selection approaches on landslide detection. Our structured CNN was fed by the input sample patches with 32×32×6 units, where 32×32 is the size of one layer of sample patches and 6 is the number of image layers (Red, Green, Blue, Red Edge, and Near Infrared). The first convolution layer was implemented with a filter size of 5 continuing with further convolution layers with a smaller filter size of 3. A max-pooling layer of 2×2 was used immediately after any convolution layer except the last one. The architecture of the CNN is shown in figure 5.

3.5 Sample Patches Selection

In this section, the generation of two different the datasets based on random and central approaches as well as the problem of using the moment bounding (MB) box for our case is detailed. Generally, the scope of the datasets is to obtain a consistent set of patches with the aim of training the CNNs for any object detection or classification aims (Depeursinge et al., 2012). The random selection of the patches approach was used in several studies, and the randomly extracted patches were applied to train their network (Wei et al., 2014; Ghorbanzadeh et al., 2019). The moment bounding (MB) box is considered as a useful method for finding the position of the sample patches and also the size of the patches. However, for the object of the landslide, on the one hand, this method leads to defining a wide range of patch sizes and consequently much more computations. On the other hand, considering the specific shape of some landslides (see figure 2), selecting the patches based on the position that defined by MB box results in having much more non-landslide areas in the patch. It means the CNN would train by the patches that have less useful data for landslide detection. Using the MB box for CNN is fully described by (Zhang et al., 2018).

Figure 2: An illustration of different sizes and shapes of the landslides that resulted in different moment bounding (MB) boxes.

In this study, we used this approach for generating our first training data set. The CNN that trained with this approach was named as random-CNN. More than 3000 original samples were generated from the training area (see figure 1). However, approximately 2000 sample patches were manually extracted from the central areas of landslides. The lower number of central sample patches is because of avoiding much overlap of patches on the image. By selecting the patches from the central areas of the landslides, it is more likely to have more areas from the landslide polygon in the extracted patch than the non-landslide areas. Therefore, the central-CNN will train with the patches that have more data from the landslide areas. The difference of sample patches selection is illustrated in figure 2.

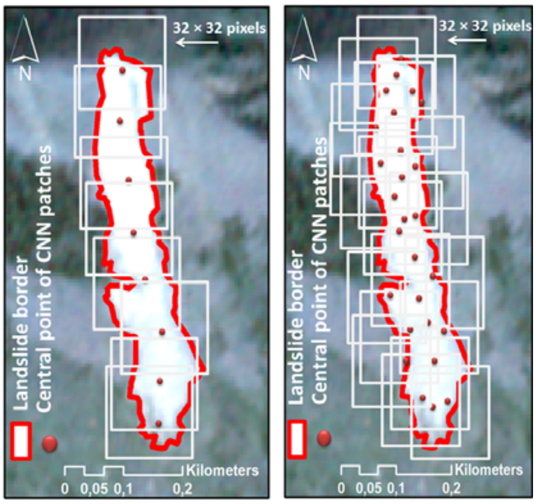

Figure 3: An example of the sample patches selection based on central (left) and random (right) approaches.

4 RESULTS

The same CNNs trained with different sample patches extracted from both random and central approaches were tested on the Rasuwa district where considered as our test area. For both CNNs, we used the same threshold of 95% and those detected landslides which were smaller than 70 pixels were removed. As described earlier, the main goal of this study is to investigate the impact of using different input sample patches of CNN on the accuracy of landslide detection. The sample patches extracting from both approaches are presented in figure 4.

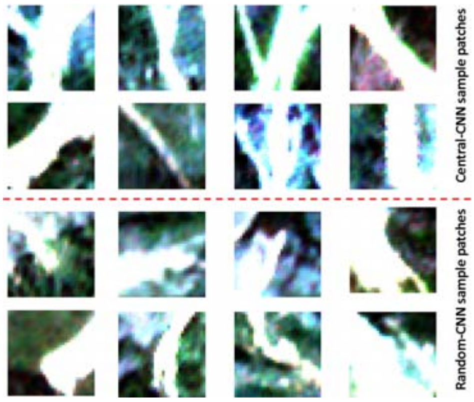

Figure 4: An illustration of convolution input sample patches extracting based on central (upper) and random (lower) approaches.

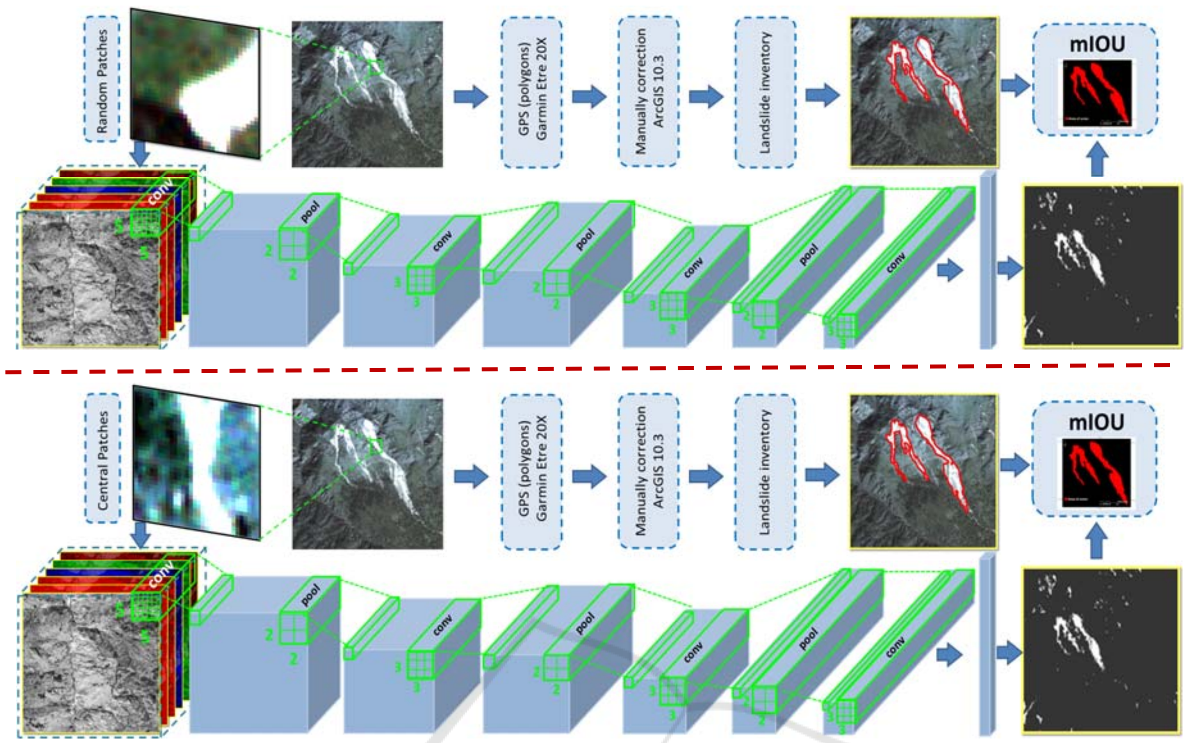

Figure 5: Flowchart of different sections of the methodology and accuracy assessment.

Two landslide maps were generated based on different sample patches selection approaches and the same CNN. Figure 6 shows the resulting landslide detected maps. Both approaches were implemented with five spectral layers from the RapidEye images (Red, Green, Blue, Red Edge, and Near Infrared) and the NDVI.

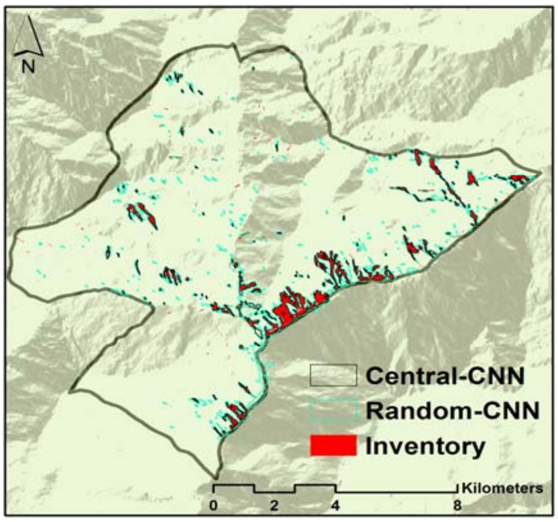

Figure 6: Landslide detection results using central and random-CNNs.

5 VALIDATION

5.1 Quantitative Results

In this section, we represent quantitative results of the resulting maps based on random-CNN and central-CNN. In this regard, the area and also the percentage of three classified pixels, namely, true positive (TP), false positive (FP), and false negative (FN) were assessed. These are the common measures that used in the remote sensing and the computer vision domains to validate the performance of the models. TP is referring to the pixels that were correctly detected as the target object. FP relates to pixels that were detected as the target object, but they are not. FN points to ground truths that are not detected as such by the applied model (Guirado et al., 2017). Regarding the calculation of these measures, a reliable inventory data set of the ground truths is required. The accuracy and details of the inventory data set can easily affect the final accuracy assessment results. Obtaining these measures make it possible to find any uncertainty among the location, and boundaries of the areas where the model detected as the landslide area. The areas and percentages of each measure and the approach were represented in table 1.

5.2 Mean Intersection Over Union (mIOU)

The mIOU is an accuracy assessment metric applied to measure the accuracy of the result of a predictor model based on ground truth. The mIOU is a known validation metric in computer vision domain, particularly for object detection studies (Liu et al., 2018). The mIOU is a general validation metric where any model that generates bounding polygons can be evaluated by using this metric based on an inventory dataset of ground truth polygons (see figure 7). It is defined as the mean of the following equation (1):

$$IOU = (\text{Area of Overlap}) / (\text{Area of Union}) \quad (1)$$

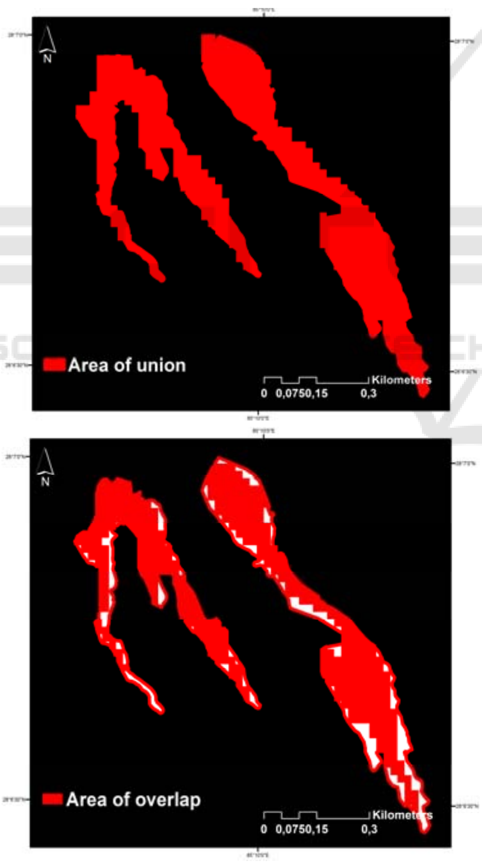

Figure 7: An illustration area of union and area of overlap.

The resulting mIOU value for both landslide maps generated by random-CNN and central-CNN were calculated and represented in table 1. According to the mIOU values, random-CNN yielded a landslide detection result with the mIOU value of 53.56. However, using the central-CNN improved the mIOU value to 56.24.

Table 1: The area and percentage of each measure along with the mIOU.

Model	TP (ha) TP (%)	FP (ha) FP (%)	FN (ha) FN (%)	mIOU (%)
Random-CNN	309.065 53.56 %	35.079 6.07 %	232.835 40.35 %	53.56
Central-CNN	186.839 56.24 %	81.092 24.42 %	64.227 19.34 %	56.24

6 DISCUSSION

In this study, we illustrated the importance of the quality of CNN training sample patches on the final result in the case of landslide detection. For the same model used, different training strategies will significantly influence the results. In this study, we generated two different training data sets. First, we randomly selected the sample patches from the landslides that occurred in the area where we considered as the training area. Second, we chose sample patches from the central area of the same landslides in the training area. Using the second approach improved the value of the mIOU metric. It means the landslides detected by the central-CNN have more overlap with those of indicated by the inventory map. However, it is not as simple as to generally compare, for instance, the TP value of the random-CNN is much more than that of central-CNN.

Moreover, random-CNN could not detect only 6 % of all landslides in the test area. Whereas, this is more than 24 % for the central-CNN. Therefore, the second approach was not successful to detect a quarter of the landslides, which is a significant portion. The better achievement of the central-CNN in the mIOU is because of it's lower FN value compare to that of random-CNN. Therefore, the second approach showed a better performance to differentiate between landslide and non-landslide areas.

7 CONCLUSIONS

The growing availability of remotely sensed imagery opens many options for updating any classification and object detection through the deep learning models. Generating of the appropriate training data sets for these models is still a challenging task due to the variety of the applications, scale of working and target classes or objects. CNN training data sets are traditionally generated by random sample patches from the whole image or region of interest. However, in parallel to the improvements in the methodology and training processes, several attempts have been made to improve the quality of training data sets generating approaches. In this study, we observed that selecting the CNN sample patches from only the central part of objects such as landslides is helpful to increase the final accuracy of the results. Although we used fewer sample patches for the central-CNN, we got a better result regarding mIOU. Thus, we can conclude the quality of the training data set for CNNs is as important as their quantity. For our future study, we aim to develop an object-based CNN method for the CNN sample patches generation. We also want to evaluate the multiple window sizes for the selection patches from the landslides of different sizes.

ACKNOWLEDGEMENTS

This research is partly funded by the Austrian Science Fund (FWF) through the GIScience Doctoral College (DK W 1237-N23). Special thanks are owed to Sansar Raj Meena, Department of Geoinformatics, University of Salzburg, Austria.

REFERENCES

Amit, S. N. K. B., Aoki, Y. Disaster detection from aerial imagery with convolutional neural network. *Knowledge Creation and Intelligent Computing (IES-KCIC), 2017 International Electronics Symposium on, 2017*. IEEE, 239-245.

Bui, D. T., Tuan, T. A., Klempe, H., Pradhan, B., Revhaug, I. 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides,* 13(2)**,** pp 361-378.

Csillik, O., Cherbini, J., Johnson, R., Lyons, A., Kelly, M. 2018. Identification of Citrus Trees from Unmanned Aerial Vehicle Imagery Using Convolutional Neural Networks. *Drones,* 2(4)**,** pp 39.

Depeursinge, A., Vargas, A., Platon, A., Geissbuhler, A., Poletti, P.-A., Müller, H. 2012. Building a reference multimedia database for interstitial lung diseases. *Computerized Medical Imaging and Graphics,* 36(3)**,** pp 227-238.

Ding, A., Zhang, Q., Zhou, X., Dai, B. Automatic recognition of landslide based on CNN and texture change detection. *Chinese Association of Automation (YAC), Youth Academic Annual Conference of, 2016. IEEE*, 444-448.

Dong, W., Sun, S., Paul, J.-C. Optimal sample patches selection for tile-based texture synthesis. *Computer Aided Design and Computer Graphics, 2005. Ninth International Conference on, 2005. IEEE*, 6 pp.

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., Aryal, J. 2019. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sensing,* 11(2)**,** pp 196.

Ghorbanzadeh, O., Tiede, D., Dabiri, Z., Sudmanns, M., Lang, S. 2018. Dwelling Extraction in Refugee Camps Using CNN-First Experiences and Lessons Learnt. *International Archives of the Photogrammetry, Remote Sensing, Spatial Information Sciences,* 42(1)**,** pp.

Guirado, E., Tabik, S., Alcaraz-Segura, D., Cabello, J., Herrera, F. 2017. Deep-Learning Convolutional Neural Networks for scattered shrub detection with Google Earth Imagery. *arXiv preprint arXiv:1706.00917*.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T. 2012. Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews,* 112(1-2)**,** pp 42-66.

Hong, H., Chen, W., Xu, C., Youssef, A. M., Pradhan, B., Tien Bui, D. 2017. Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy. *Geocarto international,* 32(2)**,** pp 139-154.

Lang, S., Schoepfer, E., Zeil, P., Riedler, B. Earth observation for humanitarian assistance. GI Forum–J Geogr Inf Sci, 2017. 157-165.

Längkvist, M., Alirezaie, M., Kiselev, A., Loutfi, A. Interactive learning with convolutional neural networks for image labeling. *International Joint Conference on Artificial Intelligence (IJCAI)*, New York, USA, 9-15th July, 2016, 2016.

Liu, B., Dixit, M., Kwitt, R., Vasconcelos, N. Feature Space Transfer for Data Augmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018. 9090-9098.

Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F., Zhang, Y. 2018. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sensing,* 10(7)**,** pp 1119.

Mezaal, M. R., Pradhan, B., Sameen, M. I., Mohd Shafri, H. Z., Yusoff, Z. M. 2017. Optimized neural architecture for automatic landslide detection from high-resolution airborne laser scanning data. *Applied Sciences,* 7(7)**,** pp 730.

Modzelewska, A., Stereńczak, K., Mierczyk, M., Maciuk, S., Bałazy, R., Zawiła-Niedźwiecki, T. 2017. Sensitivity of vegetation indices in relation to parameters of Norway spruce stands. *Folia Forestalia Polonica,* 59(2)**,** pp 85-98.

Moosavi, V., Talebi, A., Shirmohammadi, B. 2014. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by Taguchi method. *Geomorphology,* 204(646-656.

Qayyum, A., Malik, A. S., Saad, N. M., Iqbal, M., Faris Abdullah, M., Rasheed, W., Rashid Abdullah, T. A., Bin Jafaar, M. Y. 2017. Scene classification for aerial images based on CNN using sparse coding technique. *International journal of remote sensing,* 38(8-10)**,** pp 2662-2685.

Radovic, M., Adarkwa, O., Wang, Q. 2017. Object recognition in aerial images using convolutional neural networks. *Journal of Imaging,* 3(2)**,** pp 21.

Wei, Y., Xia, W., Huang, J., Ni, B., Dong, J., Zhao, Y., Yan, S. 2014. CNN: single-label to multi-label. *arXiv preprint arXiv:1406.5726*.

Yu, H., Ma, Y., Wang, L., Zhai, Y., Wang, X. A landslide intelligent detection method based on CNN and RSG_R. *Mechatronics and Automation (ICMA)*, 2017 IEEE International Conference on, 2017. IEEE, 40-44.

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P. M. 2018. An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sensing of Environment,* 216(57-70.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F. 2017. Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine,* 5(4)**,** pp 8-36.