Reducing sampling bias of landslides by identifying unrecorded events from satellite images

Lijing Wang
Department of Geological Sciences
Stanford University
lijing52@stanford.edu

Zihan Wei
Department of Geophysics
Stanford University
zihanwei@stanford.edu

Abstract

Global records of landslides are subjected to incompleteness and sampling bias because of spatially different levels of monitoring and reporting. In this work, we use a convolutional neural network to detect landslides from globally available satellite images and compare the detection with existing records of landslides. We train the model using an open high-resolution landslide dataset containing labeled landslide and non-landslide satellite images from the Bijie region in China. We test the trained model on low-resolution satellite images of Oregon, USA and Rwenzori Mountains, Uganda taken by Sentinel-2, which provides global coverage. To achieve better generalization of our model, we conduct multiple operations including downsampling the original training set to match the lower resolution of the test images, data augmentation, and distribution matching for the input test images. We visualize the model's decision making process using Grad-CAM. By comparing the detected landslides with the existing record, we find that the landslides in Rwenzori Mountains, Uganda are much more under-reported than in Oregon, USA. This work can be used to detect potentially unrecorded landslides and understand the landslide sampling bias globally.

1. Introduction

The careful analysis of historical landslides is crucial for mitigating this natural hazard risk in the future. However, incomplete record of landslides can result in sampling bias, which may lead to problematic decision making [2] and incomplete understanding of the landslide impact at different regions. For example, from the global view, South America and Africa have high risk of landslides. However, according to the documented record, the Global Landslide Catalog, these two regions have fewer reported landslides but larger number of casualties than in North America. Many small landslides or even larger ones are left out in English publications or not recorded, but these events are happening and having a huge impact on people [17].

In this project, our goal is to reduce this sampling bias by detecting landslides from globally available satellite images of the earth surface and comparing the detection with existing records of landslides [13, 12]. Mapping landslides by exploiting the big data becomes an efficient and complete way compared to field mapping [28]. We train our convolutional neural networks given an open landslide dataset in China [9], including labeled landslides and non-landslide satellite images from the Bijie region. This model is tested on unlabeled satellite images (Sentinel-2) from Oregon, North America, and Rwenzori Mountains, Africa to help us understand the different sampling bias rates at different locations. The challenge is that the tested unlabeled satellite images (Sentinel-2) have 10x lower solution compared with the Bijie dataset. We address this issue by training the model using the Bijie dataset of which the resolution is reduced to the same as the Sentinel-2 dataset.

The input to our algorithm is a satellite image. We avoid other types of data with limited accessibility because we aim to develop a method useful for areas with limited data availability. We use a convolutional neural network with transfer learning to classify whether the image contain landslides, and visualize which part of the image contribute to the classification using class activation mapping [20]). We quantify the sampling bias of landslide record by comparing the number of detected landslides in the two test regions with the recorded landslides.

2. Related work

Recorded landslides provide the guidance for the government or local authorities to assess hazards and reduce risks. However, many landslides are unreported. This under-reported issue is associated with a global spatial sampling bias, such as in Rwenzori Mountains in Africa [8]. Researchers use additional resources to correct the bias: geological surveys, local media reports [22, 8], geophysical methods such as InSAR [24, 2]. However, to our knowledge, there is no quantitative metrics to evaluate the sampling bias in recorded global landslide inventories.

To understand and reduce the sampling bias, we need to identify unrecorded landslides from globally available datasets. Mapping landslides with globally and publicly available satellite images (e.g., Landsat, Sentinel-1 and Sentinel-2) are more efficient than manual field mapping.

Although visual interpretation is still very common [6], more automatic landslides mapping methods have emerged. Methods focusing on pixel-wise values in satellite images include thresholding [19, 25], change detection [15, 27], and supervised classification [14, 16, 21]. Deep convolutional neural networks (CNN) also become popular in this landslide detection task. [4, 26, 7] apply CNN with different depths to landslide detection. [9] take a step further using a pretrained ResNet model with an additional convolutional block attention module. They also use high-resolution (0.8m) TripleSat satellite images which are not public available yet. In spite of all the existing landslide detection methods using deep learning, high-quality landslide labels for different geology, weather conditions and land types are required.

3. Dataset and Features

3.1. Satellite imagery datasets

Figure 1: Labeled landslide (left) and non-landslide (right) satellite images in Bijie region, China [9].

The high resolution (0.8m) training dataset is the open landslide dataset developed by [9], which contains 2775 satellite images (TripleSat) covering the Bijie region in China, with a block size 224 $\times$ 224. There are 771 images with landslides manually annotated by geologists and 2004 non-landslide images (Fig. 1). We split the Bijie dataset into training (60%), validation (20%), and test (20%) sets.

We then apply the trained model to two different test regions: Oregon, at the Pacific Northwest Coast of the United States and Rwenzori Mountains in Uganda, using Sentinel-2 satellite images with a spatial resolution of 10m and an area coverage of 108 km $\times$ 108 km [5]. Note that these images are different from the test set of the Bijie dataset. The test regions are prone to rainfall-triggered landslides. The satellite images we use are collected on 2019/03/17 (Oregon) and 2019/01/28 (Rwenzori). Note the Bijie dataset with high resolution (0.8m) for training the model are commercial data with extra costs. Therefore we use the open-source Sentinel-2 satellite data from the European Space Agency for the purpose of applying the model globally. We divide the satellite images into small blocks (18 $\times$ 18) covering 180m $\times$ 180m (Fig. 2). This block size covers similar area to the training images of Bijie region (224*0.8m = 179.2m). This block sizes results in 364816 images in each test region. The (potentially incomplete) locations and dates of recorded landslides in these two test regions are acquired from the NASA Global Landslide Catalog [13, 12].

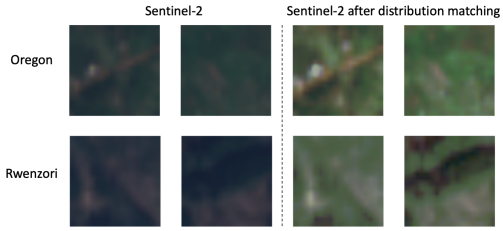

Figure 2: Sentinel-2 satellite images and corresponding images after distribution matching. Images are upsampled from 18 $\times$ 18 to 224 $\times$ 224 using bi-linear interpolation.

3.2. Image preprocessing: Normalization and Data augmentation

Every satellite image in the Bijie dataset contains RGB channels (0-255), which we normalize to 0-1. We perform different data augmentation methods to enlarge our training dataset and increase the generality of our model, including flipping (horizontal and vertical), rotation, and adjusting contrasts (Fig. 9). For each image in the Bijie dataset, we randomly apply augmentation operations and generate one new image so that we double the size of the training set. See Appendices for examples of augmented images.

4. Methods

We train our model using the Bijie dataset only because the Sentinel-2 dataset are unlabeled. We then test the trained model on both the Bijie dataset and the Sentinel-2 dataset. We use the deep learning framework Tensorflow Keras [1] to perform the data augmentation and train our model.

4.1. Baseline: a simple CNN model

We apply a simple CNN model as the baseline model, which consists of 3 identical convolution layers followed by 2 fully connected layers. Each convolution layer contains 32 filters with a size of 3 $\times$ 3 followed by a max-pooling layer. We use Adam algorithm for the optimization [11]. The loss function is the categorical crossentropy between the true label of the landslide classification ($y_0 = 1$ or $y_1 = 1$) and the predicted probability of non-landslide($\hat{y_0}$) / landslide($\hat{y_1}$): $L = -y_0 \log(\hat{y_0}) - y_1 \log(\hat{y_1})$.

4.2. Transfer learning with ResNet-50

To achieve a better performance, we apply transfer learning by fine-tuning the ResNet-50 model pre-trained with ImageNet [3]. We change the output dimension to 2 (2 classes), retrain the weights of the last two fully-connected layers, and freeze other parameters of the model. We use Adam algorithm for the optimization [11]. We use the same loss function proposed in the baseline method.

4.3. Visualizing activation maps using Grad-CAM

After training different models, we understand the decision making of our model by creating classification acti-

vation maps for different input images. Classification activation maps help us understand which parts of the image contributes more to the final classification decision. We use Gradient-weighted Class Activation Mapping (Grad-CAM) proposed by [20]. Grad-CAM uses the gradient of the final convolutional layer for different classes to provide weights for different feature maps. Then the activation map is a weighted combination of different feature maps with an additional ReLU layer. See Appendices for the mathematical details.

4.4. Downsampling Bijie dataset for better generalization

Even though the Bijie dataset provides useful information on landslide classification, it is spatially limited and cannot be directly generalized to types of satellite data, especially on lower-solution data (Fig. 2). The goal for this work is to understand the sampling bias by classifying the low-resolution satellite images in other regions. Therefore, we downsample the Bijie dataset to the same resolution as the Sentinel-2 data, and re-train the ResNet-50 model. While we expect that the performance will drop, our goal is to achieve a satisfying performance even with low-resolution data.

We first shrink the images in the Bijie dataset to 18×18 such that the image resolution is the same as the Sentinel-2 data ($10m$), and then upsample the images to 224×224 (Fig. 3). We use bi-linear interpolation in these downsampling and upsampling. Similar to the training with the high-resolution data, we apply transfer learning by fine-tuning the ResNet-50 model pre-trained with ImageNet [3]. While we are aware that fine-tuning the ResNet-50 model pre-trained with the high-resolution data will result in better accuracy in the validation and test set of the Bijie dataset, we choose not to do so because the the model may not generalize well if it has already seen the high-resolution information in Bijie. After re-training the model, we input the Sentinel-2 images upsampled to 224×224 (Fig. 2) to detect potential landslides. We manually investigate a portion of the detected landslides in both study areas and confirm whether they are landslides based on knowledge.

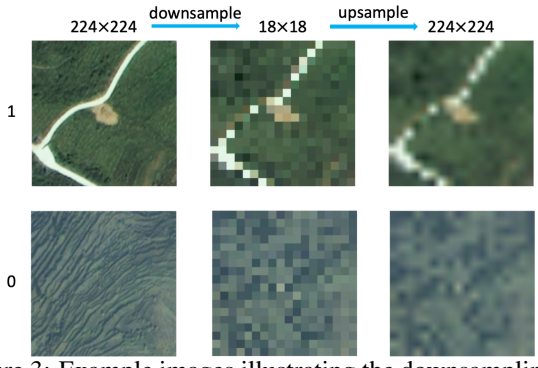

Figure 3: Example images illustrating the downsampling of Bijie dataset. The labels on the left refer to non-landslide (0) and landslide (1).

4.5. Distribution matching of different satellite images

Distributions of the value of satellite images of different satellites, land types or weather condition might be different. To generalize the trained model to Sentinel-2 datasets, we need to understand the difference in the distribution of the Bijie dataset and the Sentinel-2 dataset. We embed all satellite images in the two datasets using t-Distributed Stochastic Neighbor Embedding (t-SNE) [23]. We perform an additional Principal Component analysis (PCA) [10] before the t-SNE embedding, and use the first 100 principal components ($> 95\%$ of explained variance) as the input of t-SNE for efficiency. We use the machine learning package scikit-learn to perform PCA and t-SNE [18]. If satellite images from different datasets are embedded in different area of the lower dimensional space, we apply an distribution matching on the test Sentinel-2 dataset based on the distribution of RGB channels. After the distribution matching, the distribution of the Sentinel-2 dataset is the same as the distribution of the Bijie dataset. For example, if the Bijie dataset has 'brighter' images in general, after this distribution matching, the Sentinel-2 dataset becomes brighter. See Appendices for the details of distribution matching.

4.6. Quantifying the sampling bias of landslides in test regions

The Sentinel-2 images are unlabeled so that we cannot determine the correctness of all predictions in the test regions. Therefore, we first manually check the correctness of a subset of the predictions by our domain knowledge. We then compare the spatial distribution of the predictions against the land types of the test regions to check whether our model can handle different features in the satellite images such as cities/villages, waters and clouds. These above two steps demonstrate that the ability of our model in detecting landslides in the two test regions. We then quantify the sampling bias of landslides by comparing the number of detected landslides in the two test regions with the recorded landslides [13, 12]. We define a bias factor as $\Gamma = \frac{d_1/d_2}{r_1/r_2}$, where d_1, r_1, d_2, r_2 are the numbers of detected and recorded landslides in test region 1 and 2, respectively.

5. Results

5.1. Evaluation

We use accuracy, precision, recall, and F1-scores to evaluate the classification performance in the Bijie dataset. We first train the model using the original Bijie dataset (224×224). Table 1 show results for the baseline model and the transfer learning model ResNet-50. The models are trained for 10 epochs. The hyperparameters we used are learning rate $= 0.001$, $\beta_1 = 0.9, \beta_2 = 0.999$, $\epsilon = 10^{-7}$, and batch size $= 32$. We do not apply data augmentation in training these two models. We do not perform hyperparameters tuning, therefore the validation and test datasets can both be seen as test datasets. Note in Table 1, the test set is

part of the Bijie dataset, not the Sentinel-2 satellite images for the test regions where we apply this model.

	Accuracy	Precision	Recall	F1-score
Baseline				
Training	0.9747	0.9438	0.9650	0.9543
Validation	0.9395	0.9037	0.8841	0.8938
Test	0.9194	0.8721	0.8304	0.8508
ResNet-50				
Training	0.9979	0.9950	0.9975	0.9963
Validation	0.9896	0.9854	0.9783	0.9818
Test	0.9471	0.9266	0.8783	0.9018
ResNet-50 with the downsampled Bijie dataset				
Training	0.9892	0.9695	0.9922	0.9807
Validation	0.9435	0.8941	0.9056	0.8998
Test	0.9097	0.8182	0.8710	0.8437

Table 1: Evaluation results of trained models.

We then train the ResNet-50 model using the downsampled Bijie dataset (fig. 3) with the same resolution as the Sentinel-2 dataset, with data augmentation applied. We use the same hyper parameters listed above. Due to the increased dataset size after augmentation, we increase the training epochs to 15. No apparent overfitting issue presents in our training. Table 1 shows the results of the model. Although the accuracy is lower than the ResNet-50 model trained with the original images, the performance is enough for our purpose of detecting potential landslides.

5.2. Visualizing activation map

We do sanity check on using activation maps (Fig. 4). For different classes, our trained models focus on the correct locations. The class-0 (non-landslide) activation map focuses on buildings, roads, and waters, while the class-1 (landslide) activation map focuses on exposed soils on mountains which are likely to be landslide regions. In images where both landslides and other objects present, the model can locate them separately. In Fig. 4E, the model focuses on the road and the landslide in the in the class-0 and class-1 activation maps, respectively. In Fig. 4F, the model focuses on the water (the brown area at the bottom right corner) and the landslide in the in the class-0 and class-1 activation maps, respectively.

5.3. Distribution matching of different satellite images

The training Bijie dataset and the test Sentinel-2 dataset are located at different area in the t-SNE space (Fig. 5a1). The outliers of the Oregon Sentinel-2 dataset at the left part of the t-SNE map are the ocean area. We apply distribution matching illustrated in Section 4.5. Fig. 5 b are distributions of the RGB channels for the Bijie dataset. Fig. 5 c1 and c2 are distributions of the Sentinel-2 dataset at Rwenzori mountains before and after distribution matching. The t-SNE embedding after distribution matching (Fig. 5 a2) shows that the datasets shares similar lower dimensional representations. We also show the satellite images of Sentinel-2 before and after distribution matching in Fig. 2.

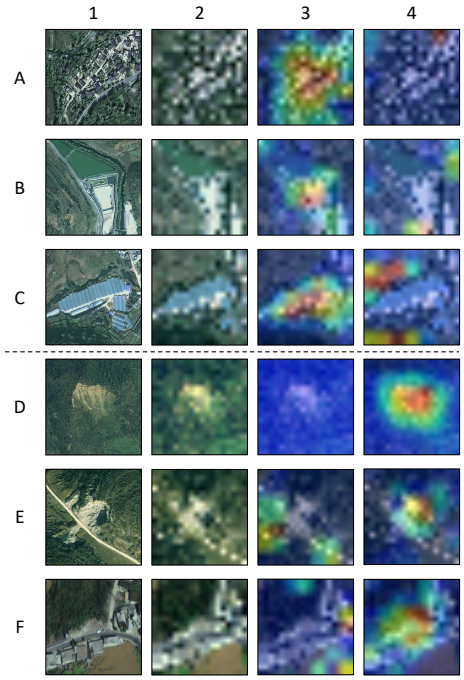

Figure 4: Activation maps for the ResNet-50 model trained with downsampled Bijie dataset. A-C and D-E are non-landslide and landslide images, respectively. Columns 1 and 2 are the original 224×224 and downsampled images, respectively. Columns 3 and 4 are the activation maps for non-landslide and landslide class, respectively.

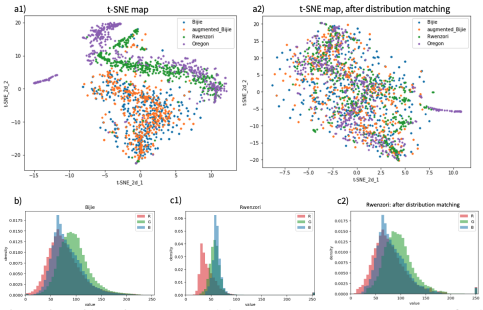

Figure 5: Distribution matching: a) t-SNE maps of all satellite images before and after distribution matching b) RGB channels histogram of Bijie dataset c) RGB channels histogram of Rwenzori mountains before and after distribution matching.

Images after distribution matching have higher brightness and more green tones.

5.4. Quantifying the landslide sampling bias

We test the model trained with the downsampled Bijie dataset on the Sentinel-2 images 1) near Eugene, Oregon, at the Pacific Northwest Coast of United States 2) at Rwenzori Mountains in Uganda (Fig. 2). As shown in Figs. 6, images with the highest probabilities correspond to mountains

Figure 6: Left: Test region 1 (Rwenzori). Right: Test region 2 (Oregon). Top (Bottom): randomly chosen 10 images with the predicted probability of containing landslide > 0.99 (< 0.01). The values above the images are the predicted probabilities.

with batches of exposed soil, while images with the lowest probabilities correspond to clouds and water. In Fig. 7, we compare the spatial distribution of the predicted landslide probabilities against the corresponding Sentinel-2 images. The regions with the most high probabilities are consistent with mountains, while the regions with less high probabilities are consistent with water, clouds, and cities/villages.

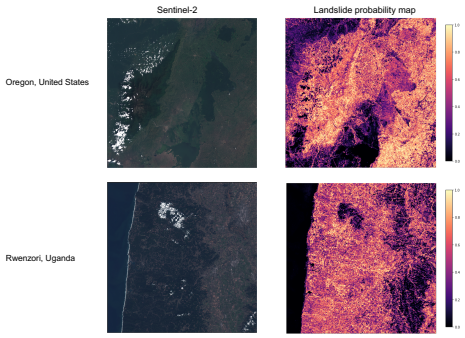

Figure 7: Comparison between the spatial distribution of the predictions against the corresponding Sentinel-2 images.

Fig. 8A plots the number of landslides detected in the two test regions with different probability thresholds applied to the predicted class probabilities. According to the NASA Global Landslide Catalog [13, 12], the numbers of recorded landslides are 38 for test region 1 (Rwenzori Mountains) and 189 for test region 2 (Oregon). Fig. 8B plots the events ratios (d_1/d_2 and r_1/r_2). The horizontal lines show the ratio in the record in different time periods. While the recorded events in test region 1 is much less than test region 1, our model detects more landslides in test region 1. The bias factor Γ is thus at the order of 10^1.

6. Discussion

With the high bias factor, our analysis suggests that the landslides in test region 1 (Rwenzori Mountains) are much more under-reported than test region 2 (Oregon). This under-reporting can be a result of a lower population density or a less developed monitoring system in Africa. Our work can help address this regional imbalance by detecting potential landslides from globally available satellite images. The detected landslides, or regions with high landslide probability, can be further inspected manually, which saves huge amount of human work in acquiring a more complete and less biased record of landslides.

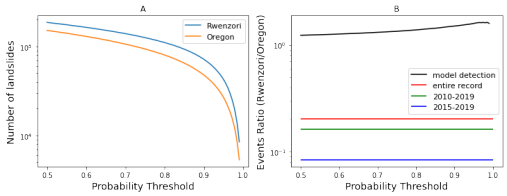

Figure 8: A: Number of landslides detected in the two test regions with different probability thresholds applied to the predicted class probabilities. B: detected events ratio (d_1/d_2) and recorded events ratio r_1/r_2.

While our model can detect potential landslides from the low-resolution Sentinel-2 images, it also misclassifies some trails and dirt roads (e.g., the second and fifth image in the first row of Fig. 6), which have similar color as the exposed soil in landslides. This can potentially lead to much more detected landslides than the actual landslides. It is also likely that large landslides cover larger area than the coverage of a single image (180m $\times$ 180m), so that a single landslide is detected multiple times.

7. Conclusion and Future work

Being trained with the downsampled Bijie dataset, our ResNet-50 model can detect potential landslides from low-resolution Sentinel-2 images and handle different land and object types in satellite images. We do distribution matching for the test images to accommodate images of different satellites from the training dataset. By visualizing the activation maps, we demonstrate that our model focuses on reasonable objects when making predictions. By comparing the recorded landslides in test regions with the detected landslides, we find that the landslides in test region 1 (Rwenzori Mountains) are much more under-reported than test region 2 (Oregon), suggesting that sampling bias exists in current records of landslides.

In this study, we use satellite images taken at a single time point. In the future, using series of satellite images taken at different time with labeled landslides can greatly improve the ability of detecting landslides because we can acquire the changes in the landscape, which can be caused by landslides.

Contributions

This project is a collaborative work, where Lijing Wang is only taking CS231N and Zihan Wei is only taking CS230. Considering that this project greatly match the scope of both

courses, we will collaborate on this project without explicitly distinguishing the work for a specific course. Lijing Wang conceptualized the project, collected and augmented the data, developed distribution matching, and applied the model to the test regions. Zihan Wei trained the models and quantified the sampling bias in the test regions. Both authors contributed to data preprocessing, visualizing activation maps using Grad-CAM, and writing the report.

Appendices

Data augmentation

Fig. 9 shows examples of augmented images.

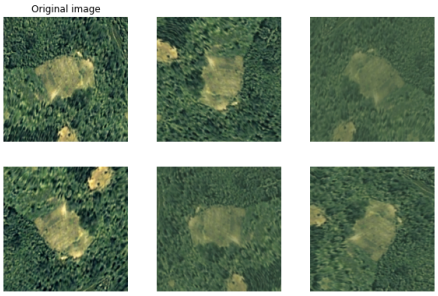

Figure 9: Data augmentation

Grad-CAM

We denote the k feature maps of the final layer as A_k, the score of class c as y^c, $c = 0, 1$. Then the activation map CAM^c for each class c is:

$$\text{CAM}^c = \textit{ReLU}(\sum_k \alpha_k^c * A_k)$$

The weights α_k^c are the average gradient of the score with respect to the feature map:

$$\alpha_k^c = \frac{1}{N} \sum_i \sum_j \frac{\partial y^c}{\partial A_k^{ij}}$$

N is the size of the feature map (column size $\times$ row size). We are only interested in the positive impact of the feature map. Therefore the additional ReLU layer is applied to remove any negative impact.

Distribution Matching

We learn the distribution of RGB channels from the training dataset, denoted as the cumulative distribution function (CDF): $F_R(r), F_G(g), F_B(b)$, and the distribution of RGB channels from the test dataset, denoted as the CDF: $F_{R'}(r'), F_{G'}(g'), F_{B'}(b')$. Then we replace every value (r', g', b') in the test dataset with the value (r, g, b) in the training dataset, where $F_{R'}(r') = F_R(r), F_{G'}(g') = F_G(g), F_{B'}(b') = F_B(b)$. This means that for all red channel values, the probability of any red channel values in the test set is less than r', $F_{R'}(r') = P_{R'}(R' \leq r')$, is the same as the probability of any red channel values in the training set is less than r, $F_R(r) = P_R(R \leq r)$. Then

References

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. 2

[2] L. Chen, P. Liu, W. Liao, K. Zhang, F. He, and L. Zeng. Regional-scale detection of unrecorded landslides in mountainous terrains by using interferometric stacking technique. *IOP Conference Series: Earth and Environmental Science*, 570:042052, nov 2020. 1

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. 2, 3

[4] A. Ding, Q. Zhang, X. Zhou, and B. Dai. Automatic recognition of landslide based on cnn and texture change detection. In *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pages 444–448. IEEE, 2016. 2

[5] European Space Agency. Sentinel-2, 2015. 2

[6] F. Fiorucci, M. Cardinali, R. Carlà, M. Rossi, A. Mondini, L. Santurri, F. Ardizzone, and F. Guzzetti. Seasonal landslide mapping and estimation of landslide mobilization rates using aerial and satellite images. *Geomorphology*, 129(1-2):59–70, 2011. 2

[7] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11(2):196, 2019. 2

[8] L. Jacobs, O. Dewitte, J. Poesen, D. Delvaux, W. Thiery, and M. Kervyn. The rwenzori mountains, a landslide-prone region? *Landslides*, 13(3):519–536, Jun 2016. 1

[9] S. Ji, D. Yu, C. Shen, W. Li, and Q. Xu. Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks. *Landslides*, 17(6):1337–1352, Jun 2020. 1, 2

[10] I. T. Jolliffe and J. Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016. 3

[11] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 2

[12] D. Kirschbaum, T. Stanley, and Y. Zhou. Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, 249:4–15, 2015. Geohazard Databases: Concepts, Development, Applications. 1, 2, 3, 5

[13] D. B. Kirschbaum, R. Adler, Y. Hong, S. Hill, and A. Lerner-Lam. A global landslide catalog for hazard applications: method, results, and limitations. *Natural Hazards*, 52(3):561–575, Mar 2010. 1, 2, 3, 5

[14] S. Lee. Application and verification of fuzzy algebraic operators to landslide susceptibility mapping. *Environmental Geology*, 52(4):615–623, 2007. 2

[15] Z. Li, W. Shi, S. W. Myint, P. Lu, and Q. Wang. Semi-automated landslide inventory mapping from bitemporal

aerial photographs using change detection and level set method. *Remote Sensing of Environment*, 175:215–230, 2016. 2

[16] M. Marjanović, M. Kovačević, B. Bajat, and V. Voženílek. Landslide susceptibility assessment using svm machine learning algorithm. *Engineering Geology*, 123(3):225–234, 2011. 2

[17] NASA. Overlooked landslides, 2017. 1

[18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 3

[19] P. L. Rosin and J. Hervás. Remote sensing image thresholding methods for determining landslide activity. *International Journal of Remote Sensing*, 26(6):1075–1092, 2005. 2

[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626, 2017. 1, 3

[21] S. Siyahghalati, A. K. Saraf, B. Pradhan, M. N. Jebur, and M. S. Tehrany. Rule-based semi-automated approach for the detection of landslides induced by 18 september 2011 sikkim, himalaya, earthquake using irs liss3 satellite images. *Geomatics, Natural Hazards and Risk*, 7(1):326–344, 2016. 2

[22] F. E. Taylor, B. D. Malamud, K. Freeborough, and D. Demeritt. Enriching great britain's national landslide database by searching newspaper archives. *Geomorphology*, 249:52–68, 2015. 1

[23] L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. 3

[24] C. Yonezawa, M. Watanabe, and G. Saito. Polarimetric decomposition analysis of alos palsar observation data before and after a landslide event. *Remote Sensing*, 4(8):2314–2328, 2012. 1

[25] B. Yu and F. Chen. A new technique for landslide mapping from a large-scale remote sensed image: A case study of central nepal. *Computers & Geosciences*, 100:115–124, 2017. 2

[26] H. Yu, Y. Ma, L. Wang, Y. Zhai, and X. Wang. A landslide intelligent detection method based on cnn and rsg_r. In *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 40–44. IEEE, 2017. 2

[27] W. Zhao, A. Li, X. Nan, Z. Zhang, and G. Lei. Postearthquake landslides mapping from landsat-8 data for the 2015 nepal earthquake using a pixel-based change detection method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5):1758–1768, 2017. 2

[28] C. Zhong, Y. Liu, P. Gao, W. Chen, H. Li, Y. Hou, T. Nuremanguli, and H. Ma. Landslide mapping with remote sensing: challenges and opportunities. *International Journal of Remote Sensing*, 41(4):1555–1581, 2020. 1