Progressive Label Refinement-Based Distribution Adaptation Framework for Landslide Detection

Hengwei Zhao[1,†], Junjue Wang[1,†], Yang Pan[1], Ailong Ma[1], Xinyu Wang[2] and Yanfei Zhong[1,*]

[1]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University 430074, China
[2]School of Remote Sensing and Information Engineering, Wuhan University 430074, China

Abstract

Efficient and accurate landslide detection is of great significance for an emergency response to geological disasters. However, detecting landslides from remote sensing images faces two challenges: small objects and class imbalance, and distribution inconsistency. In this paper, the progressive label refinement-based distribution adaptation for the landslide detection framework was proposed. The scale promotion, Lovasz loss, and online hard example mining strategy are adopted to alleviate the class imbalance, and the separated normalization and pseudo label refinement were proposed to encode the statistical inconsistency for reducing the distribution differences between the training and validation/testing data. The proposed framework has a significant potential for the large-scale global typical natural disaster monitoring rapidly from multi-sensor remote sensing imagery and ranking first place in the validation (F1-score=80.41%) and test leaderboard (F1-score=74.54%) in the LandSlide4Sense competition.

Keywords

Landslide detection, Small objects and class imbalance, Distribution inconsistency, Progressive label refinement

1. Introduction

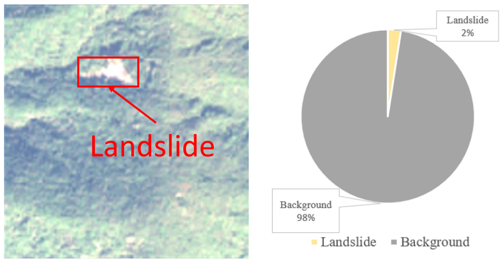
Landslide is a worldwide destructive natural phenomenon, usually following an earthquake or heavy rainfall, where thousands of small to medium-sized ground movements occur [1]. Landslides bring serious harm to society and the economy. Remote sensing technology offers the possibility of rapid and large-area land cover monitoring [2, 3], and the detection of globally distributed landslides from multi-source, multi-spectral remote sensing images using machine learning and computer vision algorithms facilitates rapid response and management of landslide-generated disasters.

In the early stage of the research, the methods for identifying landslides from remote sensing images were mostly semi-automatic two-stage methods: extracting discriminative features of landslides through expert knowledge firstly, and then using SVM or RF for classification [4]. Although using expert knowledge to construct discriminative features is transparent and flexible, it is laborious, time-consuming, and subjective [5]. Deep learning-based methods make fully automated one-stage landslide extraction possible, and these methods are well-reviewed in [5]. However, most of these methods were validated in small local regions, and the performance of these models applied directly to a new region in an emergency is unclear. To promote the development of the landslide detection field, the LandSlide4Sense competition was held and a large landslide dataset, which was collected from diverse geographical regions, is publicly available to help develop a new landslide detection algorithm [5]. Landslides detection from large remote sensing imagery will encounter the following two problems: **1) Small objects and class imbalance** and **2) Distribution inconsistency**.

As shown in Figure.1, small objects and class imbalance are the first challenges of landslide detection. In the real scene, the landslide will have some small branches or the landslide itself is relatively small in area, and as a result, there will be a serious imbalance in the number of pixels between the landslide and the background, as shown in the Figure.1(b) where the number of pixels in the background is 49 times that of the landslide in the training set. Small objects and class imbalance will lead to the problem of lower recall scores.

As shown in Figure.2, distribution inconsistency is the second challenge of landslide detection. The mean and standard deviation values of the training, validation, and testing sets are counted band-by-band and displayed in the Figure.2, where the histogram is the mean and the error bars are the standard deviation. Because landslide remote sensing images are collected from diverse geographical regions, there are significant inconsistencies in

CDCEO 2022: 2nd Workshop on Complex Data Challenges in Earth Observation, July 25, 2022, Vienna, Austria
*Corresponding author.
† These authors contributed equally.
2019206190044@whu.edu.cn (H. Zhao); kingdrone@whu.edu.cn (J. Wang); panyang@whu.edu.cn (Y. Pan); maailong007@whu.edu.cn (A. Ma); wangxinyu@whu.edu.cn (X. Wang); zhongyanfei@whu.edu.cn (Y. Zhong)
https://github.com/Hengwei-Zhao96/ (H. Zhao); http://junjuewang.top/ (J. Wang)
0000-0001-5878-5152 (H. Zhao); 0000-0002-9500-3399 (J. Wang); 0000-0002-6190-4340 (Y. Pan); 0000-0003-3692-6473 (A. Ma); 0000-0002-0493-3954 (X. Wang); 0000-0001-9446-5850 (Y. Zhong)



(a) Sentinel-2 imagery with landslide (b) Ratio of the number of pixels

Figure 1: Small objects and class imbalance problem in landslide detection. The ratio is the number of landslide pixels to the number of background pixels in the LandSlide4Sense training set.

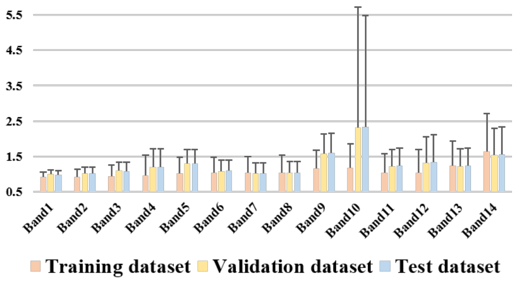


Figure 2: Distribution inconsistency between the training set and validation set. The histogram is the mean of the data and the error bars are the standard deviation of the data.

the means and standard deviation values of the training and validation/test sets, which pose a great challenge to the generality of the landslide detection algorithm.

This paper proposes a progressive label refinement based on a distribution adaptation landslide detection framework to overcome the above problems. The proposed framework achieves F1-score=74.54% in the test leaderboard of the LandSlide4Scence challenge.

2. Method

To address the two challenges above, this paper proposed the progressive label refinement framework for domain statistics adaptation, including data preprocessing, model ensemble, model training, model inference, and pseudo label refinement. The overview of the proposed algorithm is shown in Figure. 3.

2.1. Data Preprocessing

Because the small landslide areas account for few pixels and represent weak features, we take scale promotion to resize the original images (128 × 128 pixels) into 512 × 512 pixels. Besides, random flip, rotation, and color perturbation are adopted for data augmentation. As we stacked multi-spectral, DEM, and Slope data as inputs, the color perturbation is only applied to spectral data.

As the training and validation (testing) data have inconsistent statistics, the mean values of the pixel values of the data in the source and target domains are significantly different. Separated normalization is proposed to reduce the statistical difference between two domains, which takes different mean and standard deviation values to normalize the data in the source and target domains, respectively. The mean and standard deviation values were calculated from train and validation/test sets, respectively. Separate normalization is similar to cross-sensor normalization [3], but the domain-specific statistical normalization is performed in the input of the model.

2.2. Model Ensemble and Model Training

Several advanced networks are selected for the ensembling, including Swin-Transformer [6], EfficientNetV2 [7] and SegFormer [8]. The SegFormer adopts multilayer perceptron (MLP) for the decoder and the other networks adopt U-Decoder for resolution restoring. SegFormer utilizes self-attention operations to fit landslides of variant shapes as well as the MLP to enhance the difficult sample features.

To further increase the generalization capability of the model across different domains, the batch normalizations in the network are replaced with the cross-sensor normalizations [1] to automatically encode the statistical inconsistency during the training [3].

As for model optimization, Lovasz loss [9] and online hard example mining strategy were adopted to address the class imbalance problem, and Soft-cross entropy loss [10] was adopted to counteract the negative effects of noisy labels in the pseudo labels.

2.3. Model Inference and Progressive Label Refinement

In the inference phase of the model, the average of the probability values output by the above three models is taken as the final inference result.

To further align the distributions of the two domains, the progressive label refinement is designed to improve the pseudo-labels. Based on the model prediction, the pseudo labels can be generated from the best models in the ith round, using the threshold of 0.7. As for the $i+1th$ round, the source samples come from the train set and the target samples are test images with pseudo labels. The

[1]https://github.com/Junjue-Wang/LoveCS

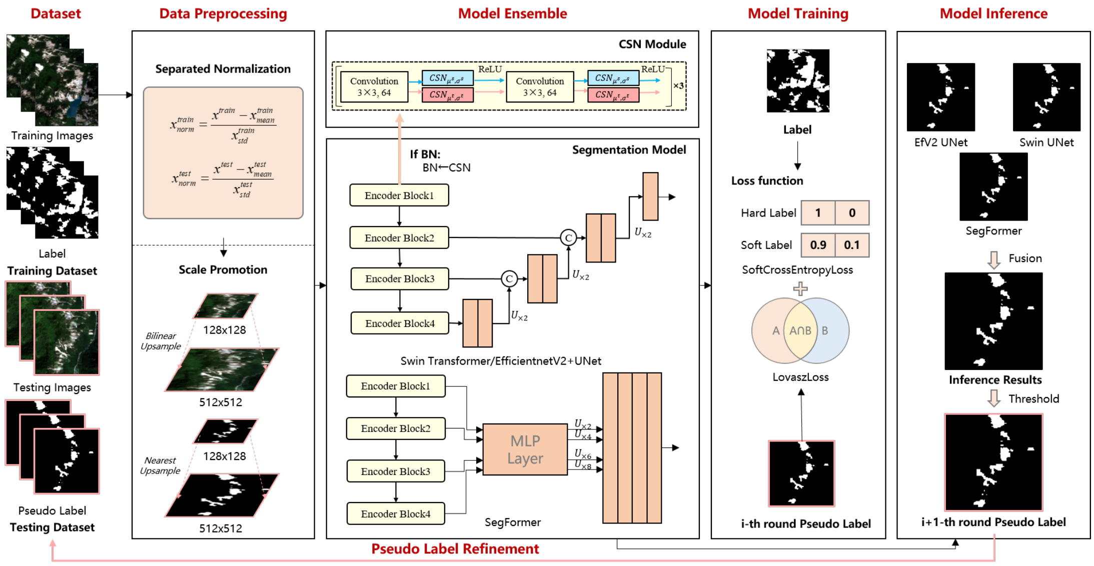


Figure 3: Progressive label refinement-based distribution adaptation for landslide detection.

pseudo-label generation and domain-adaptation training perform iteratively, progressively refining the test labels.

3. Challenge Results

The data used in LandSlide4Scence [5] are collected from diverse geographical regions, which consists of training, validation, and test sets containing 3799, 245, and 800 image patches, respectively. Each image patch is a composite of 14 bands that include: multi-spectral data from Sentinel-2 (B1-B12), slope data from ALOS PALSAR, and DEM from ALOS PALSAR. All bands in the competition dataset are resized to the resolution of about 10m per pixel. The image patches have the size of 128×128 pixels and are labeled pixel-wise. We set batch size as 16 and each model was trained for $20k$ steps.

The last round refinement results on the validation leaderboard are shown in Table 1. Compared with the baseline, separate normalization significantly improved the accuracy. The selected advanced networks were refined with several rounds and we ensemble them to obtain the highest F1-sorce=80.41%.

The results from the best models serve as a baseline and achieve F1-sorce=73.07% on the test leaderboard. Similar to the validation development, the label refinement was continuously performed on the test set. The test results in Table. 2 show that the performances of the models are progressively improved as the round increases. Round3 obtains the best result F1-sorce=74.54%.

Table 1
Results on the validation leaderboard. SN: Separated Normalization. SLO: Soft cross entropy loss+Lovasz loss+OHEM. SP: Scale promotion.

Refinement	Strategy SN	SLO	SP	Model	F1(%)
			x1.	ResNet	63.52
	✓		x1.	ResNet	72.71
	✓	✓	x1.	ResNet	73.20
Round1	✓	✓	x1.	ResNet	75.29
			x2.5	ResNet	75.98
			x2.5	EfV2.	76.24
			x2.5	SegFormer	76.57
			x2.5	Swin.	76.78
Round2	✓	✓	x4	SegFormer	78.38
				Swin.	77.98
Round3	✓	✓	x4	EfV2.	79.37
				SegFormer	79.91
Round4	✓	✓	x4	EfV2.	80.06
				HRNet-OCR	80.17
				SegFormer	80.11
				Swin.	80.34
				Ensemble	80.41

4. Conclusion

By analyzing the Landslide4Sense dataset, we conclude two challenges in landslide detection including (1) small objects and class imbalance; (2) distribution inconsistency. Hence, the progressive label refinement-based dis-

Table 2
Results on the test leaderboard

Rfinement	Model	F1(%)
Baseline	Ensemble	73.07
Round1	Ensemble	73.62
Round2	Ensemble	74.03
Round3	Ensemble	74.54

tribution adaptation for landslide detection framework was proposed. Through multiple rounds of pseudo-label optimization and separately normalization, the performance of the model continues to improve. Our solution ranked first place on the Landslide4Sense challenge. In the future, we will extend the framework into multi-temporal images for landslide monitoring.

Acknowledgments

Thanks to the Institute of Advanced Research in Artificial Intelligence for organizing this competition. This work was supported by National Natural Science Foundation of China under Grant No.42071350, No.42171336 and No.42101327, in part by the Fundamental Research Funds for the Central Universities under Grant 2042021kf0070, and LIESMARS Special Research Funding.

References

[1] P. Lima, S. Steger, T. Glade, N. Tilch, L. Schwarz, A. Kociu, Landslide susceptibility mapping at national scale: A first attempt for austria, in: M. Mikos, B. Tiwari, Y. Yin, K. Sassa (Eds.), Advancing Culture of Living with Landslides, Springer International Publishing, Cham, 2017, pp. 943–951.

[2] H. Zhao, Y. Zhong, X. Wang, X. Hu, C. Luo, M. Boitt, R. Piiroinen, L. Zhang, J. Heiskanen, P. Pellikka, Mapping the distribution of invasive tree species using deep one-class classification in the tropical montane landscape of kenya, ISPRS Journal of Photogrammetry and Remote Sensing 187 (2022) 328–344. URL: https://www.sciencedirect.com/science/article/pii/S0924271622000715. doi:https://doi.org/10.1016/j.isprsjprs.2022.03.005.

[3] J. Wang, A. Ma, Y. Zhong, Z. Zheng, L. Zhang, Cross-sensor domain adaptation for high spatial resolution urban land-cover mapping: From airborne to spaceborne imagery, Remote Sensing of Environment 277 (2022) 113058. URL: https://www.sciencedirect.com/science/article/pii/S0034425722001729. doi:https://doi.org/10.1016/j.rse.2022.113058.

[4] S. Tavakkoli Piralilou, H. Shahabi, B. Jarihani, O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, J. Aryal, Landslide detection using multi-scale image segmentation and different machine learning models in the higher himalayas, Remote Sensing 11 (2019). URL: https://www.mdpi.com/2072-4292/11/21/2575. doi:10.3390/rs11212575.

[5] O. Ghorbanzadeh, Y. Xu, P. Ghamisi, M. Kopp, D. Kreil, Landslide4Sense: Reference Benchmark Data and Deep Learning Models for Landslide Detection, arXiv e-prints (2022) arXiv:2206.00515. arXiv:2206.00515.

[6] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10012–10022.

[7] M. Tan, Q. Le, Efficientnetv2: Smaller models and faster training, in: International Conference on Machine Learning, PMLR, 2021, pp. 10096–10106.

[8] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, P. Luo, Segformer: Simple and efficient design for semantic segmentation with transformers, Advances in Neural Information Processing Systems 34 (2021).

[9] M. Berman, A. R. Triki, M. B. Blaschko, The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4413–4421.

[10] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the Inception Architecture for Computer Vision, arXiv e-prints (2015) arXiv:1512.00567. arXiv:1512.00567.