International Cartographic Association
Association Cartographique Internationale

Landslide identification in Hong Kong using generative adversarial networks based on inventory data

Xiaoyi Wu [a], Anran Zheng [b,*]

[a] *Stuart Weitzman School of Design, University of Pennsylvania – xiaoyiwu@upenn.edu,* [b] *Herbert Wertheim College of Engineering, University of Florida – anranzheng@ufl.edu*

Keywords: landslide identification, generative adversarial networks, machine learning, deep learning, GIS, Hong Kong.

Abstract:

Landslides are one of the most destructive geological hazards, which happen frequently at mountainous places with intensive rainfall events, disturbance of human activities and active geological activities. Consequently, accurate landslide prediction and susceptibility analysis become necessary and critical to mitigate financial loss and human life loss. In recent years, machine learning algorithms have been used in landslide susceptibility mapping but these models still need to be assessed and compared to enhance their performance on landslide susceptibility prediction. The scarcity of geological hazard data remains one of the challenges for these data-driven approaches. The imbalanced distribution of dataset usually leads to poor performance in landslide prediction.

To correct imbalanced landslide datasets used in mapping landslide susceptibility, this study applies a new proposed approach, which was based on the generative adversarial network (GAN), to landslide inventory data. GAN generates synthetic inventory data that makes up the shortage of original datasets. Based on fusing the original data and the synthetic data, prediction models were built, and the different types of landslide susceptibility were mapped. In this research, landslide inventory data of recent 10 years was identified in Hong Kong, China, one of the most frequent landslide prone regions. Based on previous studies and suggestions, multiple geo-environmental factors, including elevation, slope, aspect, curvature, vegetation density, land use and land cover (LULC), were integrated. The proposed model was compared with traditional machine learning and deep learning models that have been commonly used in landslide susceptibility mapping, including K-nearest neighbours (KNN), Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR). It is therefore necessary to explore how GAN improve and boost such models. Model performance was evaluated using the overall accuracy (OA), under the receiver operating characteristic curve (AUROC), confusion matrix and F1-score.

By adding new data generated by GAN, KNN, RF, DT, and LR models are trained with the best parameters respectively, and their AUROC values and accuracy are compared. The OA remained similar as the benchmark model, as OA or original KNN, RF, DT and LR model could achieve the OA values of (0.72, 0.77, 0.76, 0.73), while (0.71, 0.75, 0.70, 0.68) after adding synthetic data generated from GAN. However, the F1-score in minor categories were improved in models with GAN. The adversary generative model did not improve accuracies in landslide susceptibility mapping in Hong Kong Island with available datasets. Future studies could consider utilizing GAN as landslide prediction models by adding more significant conditioning factors.

Acknowledgements

We appreciate suggestions and encouragements by Prof Tao Wang to submit this work to EuroCarto 2022.