Initial Data Science Project Proposal

Justin Cai

13 September 2022

1 Overview

Landslides in the Front Range of Colorado are common and pose a significant threat to people and infrastructure. While past studies by the Colorado Geological Survey have documented historical landslides and mapped debris flow susceptible areas, there has yet to be a complete characterization of landslide susceptibility and at-risk infrastructure in this region. This proposal aims to map landslide susceptibility in Jefferson County, CO and identify buildings and other infrastructure located in landslide susceptible areas. This proposal involves deep learning and classification to extract building footprints from aerial imagery and to calculate landslide susceptibility based on LiDAR elevation and soil type data.

2 Related Work

The application of deep learning towards landslide susceptibility analysis has been shown to provide accurate and reliable results. Reference [1] uses convolutional neural network-based landslide susceptibility analysis models to balance the immense complexity of landslide-influencing factors with the need for an efficient and accurate method of studying landslides. Their model incorporated 16 different landslide influencing factors and provided information on which factors were most influential, classified land by landslide susceptibility, and achieved high prediction accuracy. Reference [2] provides a broader overview of typical GIS-based landslide susceptibility analyses, noting that a typical workflow consist of a landslide inventory, spatial data acquisition, landslide susceptibility map generation, and result accuracy assessment. These two references help provide a robust methodological basis for a landslide susceptibility analysis in Jefferson County.

3 Relevant Data

I intend to use Sentinel-2 aerial imagery, accessible from USGS EarthExplorer https://earthexplorer.usgs.gov/, as the aerial imagery upon which I will use deep learning to extract building footprints. Sentinel-2 contains R, G, B, and NIR bands at 10-meter resolution on a 10-day cycle, though for my analysis I plan on using the three visible light bands R, G, B. I will only need the most recent usable imagery of Jefferson County (i.e. no clouds, shadows, etc.). These building footprints will allow me to locate infrastructure in landslide-susceptible areas. For my digital elevation model, I intend to use LiDAR data, accessible from the USGS 3D Elevation Program at https://apps.nationalmap.gov/downloader/. LiDAR provides consistent, high-resolution elevation data that is necessary for an accurate assessment of landslide susceptibility. This data is a crucial factor in any landslide susceptibility model and will be integral for an accurate mapping of landslide susceptibility.

4 Analysis/Modeling

To extract building footprints from satellite imagery, I plan on using a deep learning classification model such as a neural network. I will train the model based on the same imagery using already established building footprints form aerial imagery from other sources or create the training data myself. Then, I will validate the model and assess its accuracy. For mapping landslide susceptibility, I also plan on using a neural network or a similar model that can incorporate many landslide-influencing factors and train and validate it similar to reference [1].

5 Results

My primary data product will take the form of a map of Jefferson County, with a raster overlay categorizing regions based on their landslide susceptibility as well as building footprint vector dataset. This map will visualize my model outputs and make it easy for any viewer to identify what regions and infrastructure are at highest susceptibility to landslide. I also plan on having additional figures such as a receiver operating characteristic curve, feature importance plot, and confusion matrix to provide information on the accuracy of my classification models and how they vary based on different thresholds and factors.

References

[1] Y. Wang et al. “Landslide Susceptibility Analysis based on Deep Learning”. In: *Journal of Geo-Information Science* 23.12 (2021), pp. 2244–2260. ISSN: 1560-8999. DOI: 10.12082/dqxxkx.2021.210057.

[2] W. Zhou et al. “GIS-Based Landslide Susceptibility Analyses: Case Studies at Different Scales”. In: *Natural Hazards Review* 22.3 (2021). ISSN: 1527-6988. DOI: 10.1061/(ASCE)NH.1527-6996.0000485.