

Sentinel 衛星画像データセットを利用した斜面崩壊地の自動抽出に関する検討

長岡技術科学大学大学院 ○市原 直人
長岡技術科学大学 力丸 厚, 高橋 一義

1. はじめに

リモートセンシング技術の発展により取得可能になったデータは、地震等の災害により発生した地表に生じる変化を抽出するための重要な情報源として扱われている。地表の変化抽出手法の一つとして教師あり学習が挙げられる。教師あり学習は、学習データを用いた分類器の学習能力を利用して変化を抽出することができる手法で、解析者が事前に特徴量を付与する必要がない点が利点の一つである。

しかし、現在利用されている教師あり学習手法は、対象とした単一の災害のみに対応したネットワークを作成し、地表面の変化を抽出していることが多い。

2. 研究目的

2004年の新潟県中越地震を対象とする著者らの先行研究で作成した学習データ¹⁾から、事前学習済みネットワークを転移学習させ、斜面崩壊を検出するネットワークを作成する。そして、作成したネットワークを、2018年の北海道胆振東部地震で斜面崩壊が多数発生した北海道厚真町周辺地域に適用し、異なる災害を学習データとしたネットワークによる斜面崩壊地の自動抽出がどの程度可能か検討する。

3. 方法

著者ら先行研究では、4つの事前学習済みネットワークを使用し、新潟県中越地震を対象とした斜面崩壊地判別ネットワークを作成した。本研究では先行研究で作成した学習データセットを流用し、新規に斜面崩壊地判別ネットワークを転移学習により作成する。

3-1. 斜面崩壊地判別ネットワークの作成

著者ら先行研究では、SqueezeNetを転移学習させた時の再現率が崩壊地80.2%、非崩壊地88.9%、正解率84.7%であった。先行研究の内容を参考に、崩壊地と非崩壊地の教師データセットをそれぞれ1000パッチとする。またパッチサイズを拡大し(75m四方)、パッチの75%以上を崩壊地が占めた場合、崩壊地とする。これらを設定した上でSqueezeNetを転移学習する。採用するネットワークは、ハイパーパラメータを調整し、最終検証精度が80%以上となったものとする。

3-2. 北海道胆振東部地震の画像データセット作成

発災前後データとして北海道厚真町周辺地域のSentinel-2衛星画像(発災前:2017年9月5日,発災後:2018年10月20日,画素:10m)を使用する。画像はGEE(Google Earth Engine)を利用して取得する。

Sentinel-2衛星画像は、TOA反射率を表す13の16bitスペクトルバンドを含んでいる。そのため、前処理として16bitデータから8bitデータへの形式変換と、コントラストを上げるための画像強調処理を行う。その後、判定用の画像として教師データセットと同様の発災後の赤と緑・発災前の緑の3つを組み合わせた合成画像を作成する。

崩壊地の正解データとして、2018年9月6日から2018年9月11日の期間で撮影された空中写真を元に、国土地理院が作成した斜面崩壊・堆積分布図を使用する³⁾。

3-3. 斜面崩壊の判別と結果の評価

判定用の画像を70m(7pixel)四方の小領域(パッチ領域)に分割し、判別する。判別結果の評価には正解データから作成される混合行列とそこから作成される精度、適合率、再現率の3つの指標で行う。混合行列は、判別結果と崩壊地の正解データとの一致度を表している(表-1)。評価指標は、混合行列を用いてそれぞれ式(1)、(2)、(3)の通りに導出する。

表-1 混合行列

判別結果	not land slide		land slide	
	not land slide	TN	FP	FN
正解	land slide	FN	TP	TP

$$\text{精度} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{適合率 (崩壊地)} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{再現率 (崩壊地)} = \frac{TP}{TP + FN} \quad (3)$$

4. 結果と考察

4.1. 斜面崩壊地判別ネットワーク

作成した教師データを用いて転移学習したネットワークの学習率は0.0001、ミニバッチサイズは32、エポック数は30とした。また最終検証精度は80.50%となった。

4.2. 北海道胆振東部地震での判定

判別用の画像を7×7pixelで分割し、2,874,144個の小領域を得た。判別は小領域ごとに実施した。その結果を図-1に示す。また正解データと判別結果を比較し、混合行列の作成と評価指標の算出した(表-2, 表-3)。表から、TNが全体の89%を占め、TNとFNを足し合わせると画像全体の約99%を非崩壊地と判別していることが分かる。

評価指標では、精度約 90%、適合率約 58%、再現率約 6%と、精度に比べて適合率と再現率が低く、正解データの示す崩壊地を判別することができていない。

4.3. 非崩壊地と誤判別した崩壊地

誤判別された領域について調べると、図-2 (a)に示すように、判定用の画像上では地形による影が生じている領域の一部が、正解データ上の崩壊地に含まれている。また、図-2 (b)のように、発災前画像に含まれる雲領域では地表面の確認できない。この場合、ネットワークは崩壊地と判定しない。これらが再現率の低さの一因と考えられる。

非崩壊地と誤判別した大きな要因は、ネットワークが学習した崩壊地の特徴と判別用画像内の崩壊地の特徴が異なっていたことが考えられる。ネットワークは、発災前後で赤バンドの値が増加し、緑バンドの値が減少する変化を学習していると考えられる。一方、判別用画像の崩壊地は、赤・緑バンドの値がどちらも増加しているため（画像内では白く明るく見える）、崩壊地と判別できなかったと考えられる。Sentinel-2 の画像データは、12bit の情報量を持つことから、輝度値の調整を丁寧に行うことで崩壊地の検出精度改善の余地がある。

5. まとめ

本研究では、新潟県中越地震で発生した斜面崩壊から学習した SqueezeNet を用いて北海道胆振東部地震を対象に、斜面崩壊の検出性能の検討を行った。作成したネットワークは再現率が低く、崩壊地の判別ができていない結果となった。今回、ある事象で発生した斜面崩壊の画像データセットから学習させた事前学習済みネットワークを別地域の斜面崩壊判定に転用できると示す結果とはならなかった。判定用画像の輝度の調整する余地がある場合、崩壊地の検出精度の改善が期待できる。

参考文献

- 1) 高田 竜司ら：事前学習済みネットワークの転移学習による斜面崩壊地検出結果の比較，日本写真測量学会，写真測量とリモートセンシング 60 (6)，350-353，2021
- 2) 関口 辰夫ら：新潟県中越地震における斜面崩壊の特徴と分布，日本地すべり学会誌，46 (3)，142-154，2006
- 3) 国土地理院：平成 30 年 (2018 年)北海道胆振東部地震に関する情報，6.斜面崩壊・堆積分布図，2018
<https://www.gsi.go.jp/BOUSAI/H30-hokkaidoiburi-east-earthquake-index.html#10>

表-2 混合行列

判別結果		not land slide	land slide
正解	not land slide	2,572,280	14,106
	land slide	268,103	19,655

表-3 判別精度の評価結果

精度 (%)	適合率 (%)	再現率 (%)
90.18	58.22	6.83

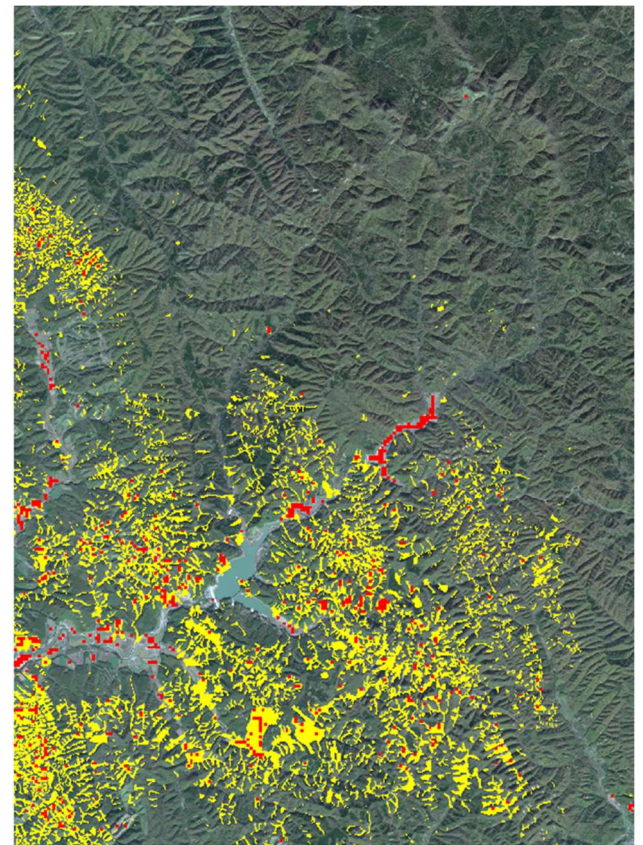


図-1 判別崩壊地 (赤)と実際の崩壊地 (黄)の分布

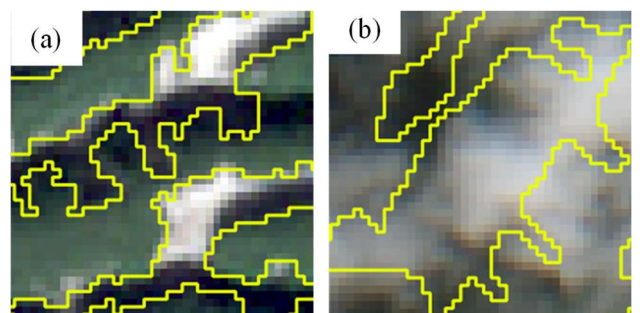


図-2 崩壊地と判別されなかった地形の一例 (a)崩壊地に影が色濃く出ている箇所，(b)崩壊地が雲によって隠れている箇所