

機械学習と画像認識を組み合わせた斜面の広域危険度評価手法の開発

環境防災研究室 片平耀款
指導教員 大塚 悟

1. 研究背景

2001年に土砂災害防止法が制定されて以来、斜面災害のソフト対策が推進され、斜面災害の危険度評価が多数行われてきた。従来の斜面危険度評価は、数量化II類や重回帰分析等の多変量分析が主流であった。また、近年コンピュータ技術の急速な発展とともに、AIを用いた豪雨や地震による斜面崩壊に関する研究が徐々に行われてきている。これまで行われてきた上記の手法の多くは斜面の属性（標高、傾斜、地質等）をメッシュの要素毎に評価して、斜面の崩壊および非崩壊のイベントとの相関を解析したものである。しかし、斜面の崩壊規模を考えると、大規模の崩壊はマクロスケールの地形の影響を受けるのに対して、崩壊規模の小さい崩壊はミクロスケールの地形の影響を受ける。したがって、メッシュ単位で崩壊現象との関連を調査する従来の方法は斜面の崩壊特性を反映しない問題があり、出力される結果もモザイク絵画のように危険箇所と安全箇所が混在するような結果となるために、解析結果の工学的利用が限定される問題がある。そこで本研究では、様々なスケールの斜面の地形的要因を危険度評価に取り入れた予測を実施し、危険斜面を抽出するとともにその危険度を評価する解析手法の開発を行う。具体的にはAIを用いた2段階の解析システムを開発し、1段階目(Step1)ではスケールを変えた地形情報を用いる手法の提案を行い、その有効性を検証する。2段階目(Step2)では個別メッシュの危険度評価を用いた画像認識を実施し、崩壊ブロックを抽出する。解析結果に対するAI画像認識は従来にない手法であり、画像解析の新展開となる試みである。

2. 調査地域と資料概要

今回用いたデータは、平成23年7月に発生した新潟・福島豪雨において、崩壊の有無の調査が行われた図-1の灰色部の範囲である。この地域は、豪雨が発生する以前に地形調査が行われており、斜面崩壊の予測に適したデータである。また、1mメッシュで図の範囲内の領域をすべて扱うには、データ数が膨大であるため、図-1の範囲から対象地域を絞った太枠の範囲を今回の対象地域とした。調査地域での崩壊面積は、最小が3.5m²、最大が

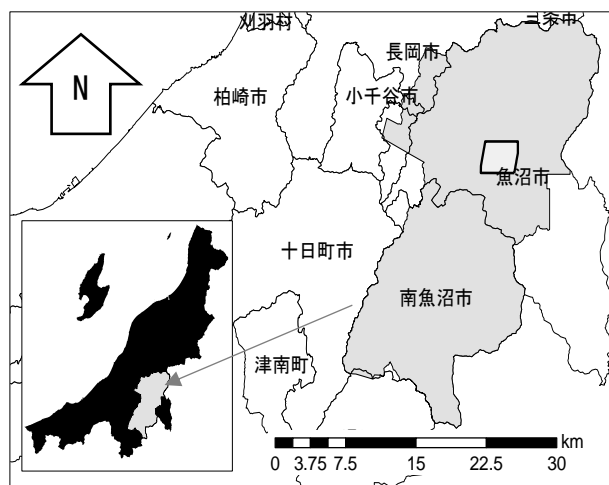


図-1. 調査地域及び対象範囲

24,319m²となっており、1mメッシュデータの適用が可能である。赤の対象範囲では、範囲が狭いため、降雨量に変化がみられず、AIの予測に使用しても意味がない。そのため、今回は降雨量を除いた定性的なデータでの予測を行うことにした。

3. 予測方法

Step1では、応答変数を災害時の崩壊・非崩壊の事象とし、予測子変数を1.標高、2.傾斜角、3.曲率、4.断面曲率、5.平面曲率、6.起伏量、7.累積流量の7項目とした。また、スケールの異なる地形情報データでは、1~7に加えて、1mメッシュのデータと同項目の周囲5m四方の情報を含むデータを7項目、周囲10m四方の情報を含むデータを7項目の計21項目を入力変数とした。上記のデータについては、ArcGIS(ESRI社)によって、標高から導き出すことができる。また、今回用いたDEMデータは、国土交通省北陸地方整備局湯沢砂防事務所が2010年7月に作成したものをを用いている。ArcGISにて数値化したデータを、MATLABの機械学習ツールを用いて予測を行った。今回予測に用いたのは機械学習の分類手法である決定木、ブースティング木、バギング木、単純ベイズ、KNNの五種類の方法である。今回のデータの入力方法は、数値を直接代入する方式をとった。調査地域では、河川や池、湖などの調査不能な範囲を除いて、24,913,734サンプルのデータを得ることが出来た。そのうち崩壊データは、98,589サンプルであった。崩壊データ数のうち90,000サンプルを学習デー

タとし、残りの 8,589 サンプルをテストデータとした。過学習を起こさないように、非崩壊データの学習データをランダムに 90,000 サンプル抽出し、崩壊の学習データの個数と同数にし、それ以外のデータはこちらもテストデータにした。学習データのうち 25% を検証データとして、ホールドアウト検証を行い、データが過学習を起こしていないことを確認した。機械学習では、崩壊と非崩壊の 2 値分類の結果を用いて、有効性について判断した。

Step2 では、上記の方法で広域情報を入力した機械学習による危険度評価の結果について、ArcGIS を用いて 0 から 1 まで 20 段階に分けて色付けを行った画像を用いて、画像認識を行った。機械学習の結果画像については、上記で最も精度が高いバギング木の結果を利用した。処理の手法については、オブジェクト抽出に特化した YOLO、既存の画像を学習してそこから新たな画像を作り上げる敵対的生成ネットワーク GAN、学習した結果と同じように画像全体に色を塗る Semantic Segmentation の 3 つの手法を用いた。GAN では入力する画像のサイズは 256pixel×256pixel であり、入力枚数は 190 枚とした。Semantic Segmentation においては、画像サイズは同じで入力枚数は 160 枚とし、ピクセル数が少ないほど、重みを大きくする方法を用いてバランスを調整している。YOLO においては、四角形で崩壊・非崩壊を設定して、160 枚の画像を入力している。

4. 解析結果

表-1. 手法ごとの解析結果

(%)	決定木		バギング木		KNN		ブースティング木		単純ベイズ	
	従来	今回	従来	今回	従来	今回	従来	今回	従来	今回
正解率	59.3	64.4	59.8	76.9	58.1	63.1	58.8	63.9	54.3	57.0
精度	61.4	64.0	60.3	76.4	58.7	62.4	61.4	62.6	53.6	55.9
真陽性率	50.0	66.0	57.3	77.7	54.4	66.1	47.5	69.4	64.6	66.5
真陰性率	68.6	62.8	62.3	76.0	61.7	60.1	70.1	58.5	44.1	47.6
偽陰性率	50.0	34.0	42.7	22.3	45.6	33.9	52.5	30.6	35.4	33.5
偽陽性率	31.4	37.2	37.7	24.0	38.3	39.9	29.9	41.5	55.9	52.4

数値の精度： □ 良 □ 悪

表-2. 混同行列

		予測のクラス	
		崩壊(1)	非崩壊(0)
実際のクラス	崩壊(1)	TP	FN
	非崩壊(0)	FP	TN

$$\text{正解率} = (TP+TN) / (TP+TN+FP+FN) * 100$$

$$\text{精度} = TP / (TP+FP) * 100, \text{ 真陽性率} = TP / (TP+FN) * 100$$

$$\text{真陰性率} = TN / (TN+FP) * 100, \text{ 偽陰性率} = FN / (TP+FN) * 100$$

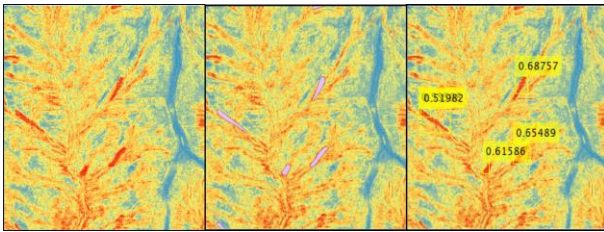
$$\text{偽陽性率} = FP / (TN+FP) * 100$$

解析結果を表-1 に示す。Step1 の機械学習による予測結果では、正解率及び精度等において全ての手法で、広域情報を入力した結果が、従来の手法より数値が向上したことが確認された。決定木・KNN・ブースティング木においては、真陰性率・偽陽性率の数値が低下しているが、偽陽性率は実際には崩壊していないものを崩壊したとみなすということで、安全側に働いたとも考えられるため、極端に悪くなったとも言い難い。また、実際の崩壊箇所を当てた確率である真陽性率では、すべての手法において数値の向上がみられる。これは、1m メッシュのみの微地形の特徴だけでは予測困難だった規模の大きな崩壊を、スケールの異なる地形情報を含んだ解析により、予測できたことを示していると考えられる。

次に、Step2 の機械学習による危険度評価の画像を用いた画像認識についての結果を図-2、図-3、図-4 に示す。機械学習では危険と判断された箇所である黄色やオレンジの部分と、本当に危険な箇所である赤色の部分を YOLO、Semantic Segmentation、GAN の 3 つすべての手法において判別できている。崩壊の形には特徴がほぼないことから、画像認識により AI が色の特徴を認識したことで、崩壊箇所と非崩壊所の分類が可能になったと考えられる。機械学習の結果では、実際は崩壊していない箇所を崩壊とする偽陽性率の数が非常に多く、2 値分類では崩壊箇所の数が非常に多くなっていった。しかしながら、画像認識を行うことで、周囲の情報をまとめる畳み込み層等の影響により、偽陽性個所の大部分について正しい予測ができた。これにより、崩壊箇所のみを表示する崩壊ブロックの抽出が可能になった。

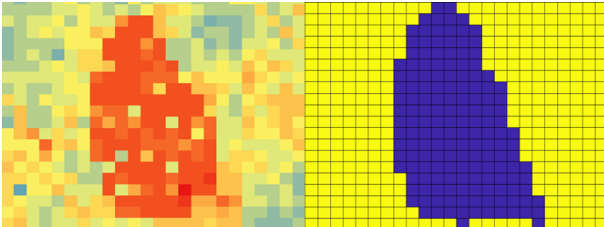
5. まとめ

解析結果より、機械学習による予測には、スケールの異なる地形情報の入力有効であることが確認された。また、スケールの異なる地形情報を入力した機械学習による危険度評価の結果を用いた画像認識により崩壊ブロックの抽出が可能であることが確認された。これより、本研究の結果から機械学習と画像認識を組み合わせることで、斜面崩壊の崩壊ブロックを抽出可能であるといえる。抽出した崩壊ブロックの危険度評価については、実施できていないため、今後も継続的な研究が必要である。



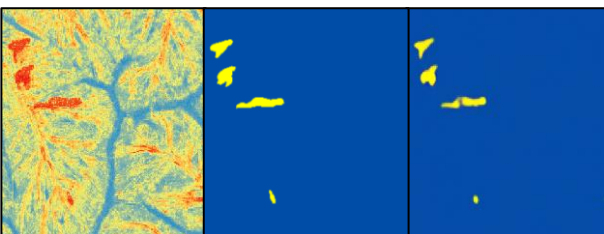
(A) 入力画像 (B) 正答画像 (C) 結果画像

図-2. YOLO



(A) 入力画像 (B) 結果画像

図-3. Semantic Segmentation



(A) 入力画像 (B) 正答画像 (C) 結果画像

図-4. GAN

6. 今後の展望

本研究の手法については、まず学習データとテストデータが同じ範囲の場合について手法の開発を行った。上記の結果より、同じ範囲については一定の成果が出せたといえる。そのため、今後は他地域へ適応できるような予測を行っていく予定である。既に、機械学習においては他の地域への適応に向けた取り組みを行っている。画像認識においても、赤色立体図を用いて露岩帯の抽出について他地域への適応を行っている研究¹があり、同様の手法を予測にも適用することが出来るか検討を行っていく予定である。

7. 参考文献

(1) 高山 陶子ら(2018)：ディープラーニングと赤色立体図を用いた微地形自動判読手法の検討，平成 30 年度砂防学会研究発表会概要集，p697~698

8. 謝辞

データを提供頂いた国土交通省北陸地方整備局および技術的な助言を頂いた北海道大学磯部公一先生に謝意を表します。