

地すべり災害発生の危険性評価への ディープラーニング適用の試み

神山嬢子・杉本宏之・野坂隆幸・石田勇人

1. はじめに

将来起こりうる豪雨や地震への備えとして、事前防災対策を行うことが重要である。しかし、明瞭な地すべり活動が生じていない地すべりに対し、その危険性や対策の緊急性を判断することは難しい。これまで、様々な危険性評価手法が研究されているが、多くの災害データを収集することが難しく、十分なデータに基づく分析によって予測結果の妥当性を明らかにすることに課題がある。また、広域を対象に危険性評価を行う際、地形や地質等が地域により異なる中で、評価指標の種類や基準等を地域に応じて適切に設定することに課題がある。そして、この評価指標や基準等を技術者が地域ごとに分析し設定するには、その作業の効率化が望まれる。

ディープラーニング技術は、十分な数のデータがあれば、教師データにおける特徴を自動で抽出して指標を設定することが可能であり、先に述べた課題を解決できる可能性がある。そこで、全国で収集される地すべり災害データを用いて、地すべり災害発生の危険性を評価するディープラーニングモデル（以下「DLモデル」という。）の構築方法を検討した。本稿では、DLモデル構築に関する検討の概要とDLモデルの性能例を紹介する。なお本検討は、（国研）防災科学技術研究所、国土防災技術（株）、五大開発（株）、日本工営（株）と土木研究所地すべりチームの共同研究成果の一部である。本稿の詳細は、今後、共同研究報告書で公表予定である。

2. 地すべり災害発生の危険性評価DLモデルの概要

本検討では、既存の地すべり地形において地すべり災害発生の可能性を分類するDLモデルを構築した。DLモデルは、教師データとして災害が発生した地すべり地形（以下「発生」という。）と災害が発生していない地すべり地形（以下「非発生」という。）における、傾斜量や曲率といった地形量をもとに作成した画像を学習した（図-1(a)）。DLモ

デルによる「発生」と「非発生」の分類は、評価対象の地すべり地形の画像を学習済みのDLモデルに入力し、DLモデルから「発生」と「非発生」のもっともらしさを示す尤度を出力し、尤度が大きい方に分類した（図-1(b)）。

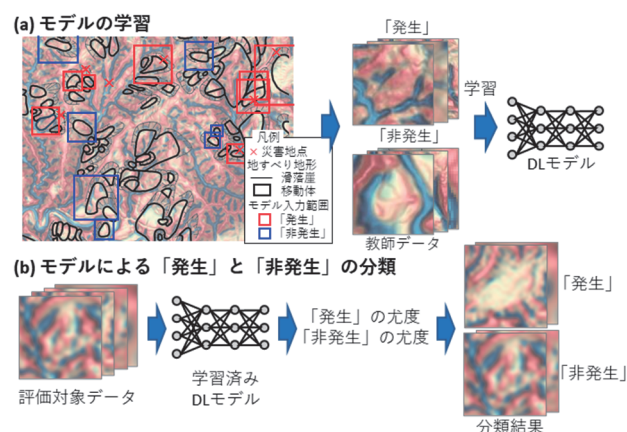


図-1 DLモデルの学習と分類のイメージ

3. DLモデル構築に関する検討方法

DLモデルの構築に用いる教師データを作成するため、地すべり地形から、地すべり災害データとの位置関係をもとに「発生」と「非発生」の事例を選定する（図-1(a)）。そして、「発生」と「非発生」に対してDLモデルに入力する範囲を設定し、地形量から作成した画像を切り出して教師データを作成した。教師データは、DLモデルの学習と分類性能の評価に用いるため、訓練・検証・テストデータの各データセットを作成した。また、DLモデルには、画像認識に用いられる代表的なモデル構成であるAlexNet¹⁾を用いた。

3.1 教師データ作成に用いるデータ

「発生」と「非発生」の選定に用いる地すべり災害データは、1989年～2019年に国土交通省へ「地すべり災害」として報告されたものを用いた。このデータには、発生日時や地名、地すべりの規模、緯度経度などの情報があるが、本検討では緯度経度のみ使用した。また、入力範囲の設定には、防災科学技術研究所の1:50,000地すべり地形分布図の移動体

ポリゴンを用いた。これは、過去に地すべりを生じた範囲を示す移動体を線で囲んだGISデータである。入力画像の作成には、国土地理院の数値標高モデル(5mDEM)を用いた。なお、教師データの作成は、新潟県、富山県、石川県、長野県、静岡県、徳島県、愛媛県、高知県を対象とした。

本検討では、DLモデル構築に必要な教師データ数を確保するため、広範囲かつ長期間の地すべり災害データを用いた。また、地形データは同一分解能で広範囲に整備されているものを用いた。教師データに用いた地すべり移動体の多くが10ha未満であり、分解能5mの地形データを用いる場合、DLモデルに入力する画像上は災害前後の地形変化がそれほど大きく現れないことから、教師データには災害後に取得された地形データも使用した。また、地すべり発生に影響する因子として地質等があるが、因子別に十分な量の教師データを用意することが難しく、本検討では地質等を考慮していない。

3.2 「発生」と「非発生」の選定

「発生」は、地すべり災害データと移動体ポリゴンの位置関係をもとに選定した。地すべり地形(滑落崖と移動体)に災害地点が含まれるもの(図-2 A)や、これと上下に接し地すべり地形として上下の切り合いの関係にあるもの(図-2 B)、または同一の滑落崖を共有し隣接する関係にあるもの(図-2 C)を選定した。なお、地すべり災害データにおいて、どの地点を位置情報として記録するかは定まっておらず、位置精度が低いものが含まれている可能性がある。そのため、地すべり地形に災害地点が含まれない場合でも、実際は地すべり地形で発生している可能性も考えられることから、地すべり地形に災害地点が含まれると見なせる程度に近接しているもの(図-2 D)も「発生」とした。そして、「発生」以外のエリアにある移動体ポリゴンから「非発生」を選定した(図-1(a))。「非発生」の選定にあたっては、対象県ごとに「発生」と同程度の面積のものを同程度の数だけ選定し、「発生」と「非発生」の間に地域的な分布や規模の偏りが無いように留意した。なお、選定対象の移動体ポリゴンは、5mDEMから作成した微地形表現図を判読し、地すべり地形を視認可能な1ha以上のものを対象とした。

3.3 DLモデルに入力する範囲の設定

入力範囲は、地すべり地形(滑落崖と移動体)を概ね包含する範囲とした。地すべり地形分布図は

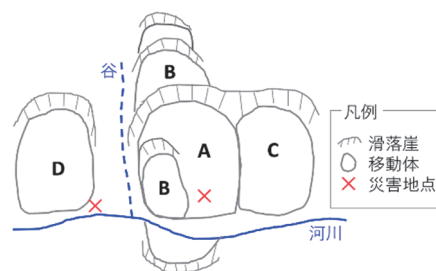


図-2 「発生」の選定イメージ

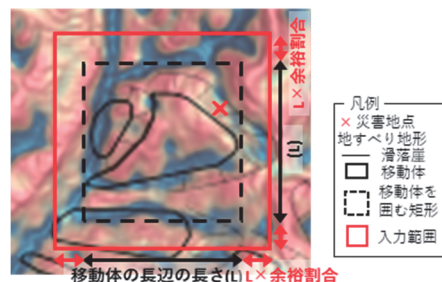


図-3 入力範囲の設定方法

データ上、滑落崖と移動体が紐付けられていないため、移動体を中心とする矩形に対し一定の余裕幅を設定し、滑落崖を概ね含む範囲の矩形を作成した(図-3)。予め余裕割合が異なる入力範囲の教師データを学習したDLモデルを構築し、その分類性能の評価結果から、再現率が最も高かった余裕割合20%をとることとした。

3.4 DLモデルに入力する画像データ

入力データには、傾斜量や曲率を色調の違いで表現したCS立体図²⁾を用いた。これは、技術者が地形判読を行う上で有効な微地形表現図である。CS立体図の作成には、G空間情報センターで公開されているGISソフトのCS立体図作成用プラグインから作成した傾斜と曲率のデータを用いた。地形表現方法の違いがDLモデルの分類性能に与える影響を把握するため、CS立体図の色設定をプラグインのデフォルト設定としたもの(以下「CS_{def}」という。)と、技術者が地すべり地形判読時に着目する凹状地形や緩勾配斜面を見やすくするために緑色で表現したもの(以下「CS_{ds}」という。)を用いた(図-4)。

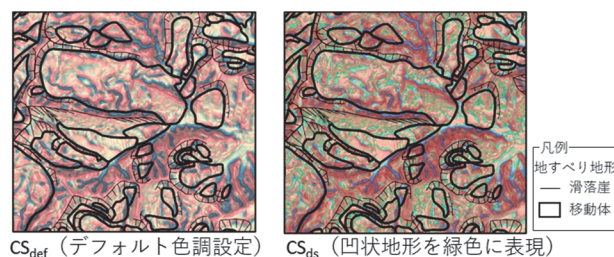


図-4 微地形表現図の例
(地すべり地形は入力データには含まない)

3.5 モデルの学習・評価

教師データのうち、訓練データを用いてDLモデルの学習を行った。作成した教師データ数は多くないため、教師データ数がDLモデルの分類性能に及ぼす影響を把握するために、作成したオリジナルの教師データを回転（90度、180度、270度）、左右反転、スライド（入力範囲を斜め四方にスライド）するデータ拡張処理を行った。また、検証データにおけるDLモデルの分類結果をもとに、DLモデルが正しく分類できるようパラメータを自動調整した。

DLモデルの分類性能は、テストデータに対する分類結果を示す混合行列（表-1）をもとに、正解率（式1）、再現率（式2）、適合率（式3）、F値（式4）を算出して評価した。また、Grad-CAM³⁾を用いて、入力画像のどの位置が分類結果に強く影響を与えたかを可視化したヒートマップを作成し、入力画像（CS立体図）と比較することで、DLモデルの分類の判断根拠の解釈を試みた。

表-1 混合行列

		ディープラーニングモデルの分類結果	
		発生	非発生
正解	発生	TP (True Positive)	FN (False Negative)
	非発生	FP (False Positive)	TN (True Negative)

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

$$\text{再現率} = \frac{TP}{TP + FN} \quad \dots (2)$$

$$\text{適合率} = \frac{TP}{TP + FP} \quad \dots (3)$$

$$\text{F値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}} \quad \dots (4)$$

DLモデルの汎化性能を確認するために交差検証を行った。交差検証では、データセット毎に訓練・検証データを8つに分割し、訓練と検証に用いるデータの組合せを変えて学習した8つのDLモデルを構築して性能を評価した。

4. DLモデルの性能と分類の判断根拠を解釈するための試み

構築したDLモデルの例として、データ拡張方法と入力データの地形表現方法の条件が異なるデータセット（表-2）を学習したDLモデルの性能評価結果を図-5に示す。交差検証により構築した8つのDL

モデルのうち、正解率が高い3つのDLモデルから出力される尤度の平均値をもとに分類した結果（正解率上位3モデル）、および正解率が最も高い1つのDLモデルで分類した結果を示す。テストデータに対する正解率は59%～69%となり、No.3の「正解率上位3モデル」が最も高い結果となった。No.1～No.3はデータ拡張処理による訓練データ数が異なるが、No.1からNo.3にかけて訓練データ数が多くなるほど正解率が高く、データ拡張による訓練データの増加が正解率向上に寄与していると考えられる。また、地形表現方法の異なるNo.2（CS_{def}）とNo.4（CS_{ds}）を比較すると、正解率、適合率、F値ともにNo.4が高い値となった。

No.2とNo.4のDLモデルにおいて、「発生」と正しく分類された画像とヒートマップを図-6に示す。なお、暖色系の色がDLモデルの分類に強く影響を与えた箇所である。No.2のヒートマップでは地すべり移動体付近に暖色系の色が見られないが、No.4では地すべり移動体付近に暖色系の色が着色され、地すべり移動体に着目した分類が行われているように見える。そのため、入力画像を技術者が地すべり地形判読時に着目する凹状地形などの微地形要素を視認しやすい色表現としたことで、DLモデルが地すべり地形に着目しやすくなった可能性がある。

表-2 データセット

No.	地形表現方法	データ拡張処理方法	訓練・検証データ数 発生/非発生		テストデータ数 発生/非発生
			オリジナル	データ拡張後	
1	CS _{def}	なし	501/506	—	120/120
2	CS _{def}	回転・左右反転	501/506	4,008/4,048	120/120
3	CS _{def}	回転・左右反転・スライド	501/506	19,688/19,984	120/120
4	CS _{ds}	回転・左右反転	501/506	4,008/4,048	120/120

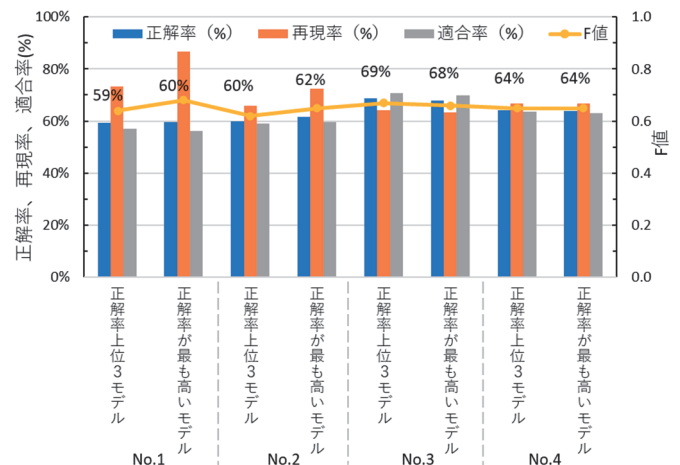
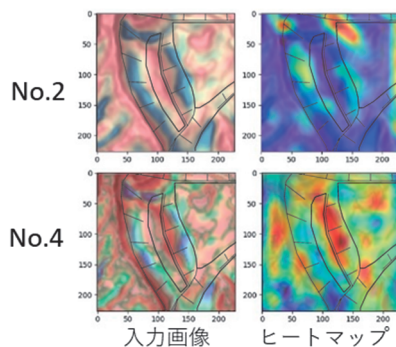
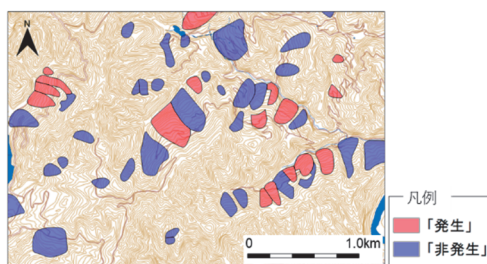


図-5 DLモデルの性能評価結果
（図中の数字は正解率を示す）

前述のDLモデル構築方法により、地すべり災害発生の危険性を評価するDLモデルのプロトタイプを構築することができた。これを用いて、ある地域における地すべり地形分布図の全ての移動体ポリゴンを対象として危険性評価を行った例を図・7に示す。赤く表示した移動体ポリゴンは、正解率約70%のDLモデルが「発生」と分類したものであり、地すべり災害発生の可能性が相対的に高いと解釈することができると思われる。



図・6 DLモデルが「発生」を正しく分類した画像の例



図・7 DLモデルの予測結果例

5. まとめと今後の課題

地すべり災害データと地すべり地形データをもとに「災害が発生した地すべり地形」と「災害が発生していない地すべり地形」を教師データとしてディープラーニングモデルに学習させることで、地

すべり災害発生の可能性の大小を評価可能なプロトタイプモデルを構築した。構築したDLモデルによる地すべり災害発生の危険性評価結果は、事前防災対策として対策の優先度や警戒避難体制の検討に活用できる可能性がある。一方、地質等の地すべりの発生に影響する因子を考慮できていない。そのため、地質等の影響を考慮したDLモデルを構築するための教師データの蓄積と併せて、教師データが少ない中でも一定の妥当性のある分類が可能なDLモデルの構築方法を検討することが課題である。

今後、DLモデルを用いた危険性評価手法を現場で実用化する上で、信頼性の向上を図ることが重要と考える。そのために、高い性能のDLモデルを構築するとともに、DLモデルの分類結果の妥当性の確認が重要と考える。精度を高めるためには、評価対象に応じた教師データの絞り込みやデータの加工、地すべり災害データの蓄積等が考えられる。また、技術者の地形判読に基づく危険性評価結果との比較や、DLモデルに入力するデータを変化させた場合の評価結果の感度分析などを通じてDLモデルの分類結果の妥当性を確認することが考えられる。

参考文献

- 1) Krizhevsky, A. et al.: ImageNet classification with deep convolutional neural networks, 25th International Conference on Neural Information Processing Systems, pp.1097-1105, 2012
- 2) 戸田堅一郎：曲率と傾斜による立体図法（CS立体図）を用いた地形判読、森林立地、Vol.56、No.2、pp.75~79、2014
- 3) Selvaraju R. R. et al.: Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization, 2017 IEEE International Conference on Computer Vision, pp.618-626, 2017

神山嬢子



土木研究所土砂管理研究
グループ地すべりチーム
主任研究員
KAMIYAMA Joko

杉本宏之



土木研究所土砂管理研究
グループ地すべりチーム
上席研究員
SUGIMOTO Hiroyuki

野坂隆幸



研究当時 土木研究所土砂
管理研究グループ地すべり
チーム 研究員、現 国土
交通省関東地方整備局富士
川砂防事務所 調査課長
NOSAKA Takayuki

石田勇人



土木研究所土砂管理研究
グループ地すべりチーム
交流研究員、博士（理
学）
Dr. ISHIDA Hayato