

深層学習を用いた画像解析及びクラスタリングによる土砂災害の素因特性の抽出

国土技術政策総合研究所 野呂智之、神山嬢子、鈴木大和、戸舘光

株式会社地圏総合コンサルタント ○中西宏彰、家田泰弘、片嶋啓介、大坪隆三、相澤明宏、潮見礼也

1. はじめに

土砂災害の発生には、降雨・地震等の誘因と地形・地質等の素因が影響する。誘因である外力の大きさはもちろんのこと、素因の違いによって発生状況が異なることが知られている。

土砂災害の崩壊規模等を予測することは、ハード・ソフト対策のいずれにおいても極めて重要である。特に、警戒避難等のソフト対策は複数の流域に跨った広範囲が対象となる。従来から様々な統計分析を用いた土砂災害の発生予測が試みられてきたが、複数の要素が複雑に入り組んだ素因からの的確に土砂災害発生に関わる素因特性を抽出することに統計分析の限界があった。

そこで、本研究は、新たなアプローチとして深層学習を活用し、土砂災害に関わる素因特性を分類した。

2. 解析手法

2.1 使用する深層学習（教師なし学習）の概要

本研究で使用する学習データ、画像解析およびクラスタリング手法を表-1 に整理する。

画像解析は、国内を包含する全ての第3次メッシュ約38万枚の学習データ全画像に対して、各画像データ自身の特徴をとらえる（新規学習）CNNモデルを利用した解析手法を用いた。CNNモデルは畳み込み層の異なる2種類のモデルを対象に試行学習を行って評価したモデルにより、全国約38万枚に展開し、素因特性の違いによるクラスタリングを行った。

表-1 解析手法の整理

学習データ	画像解析	クラスタリング
国内を包含する全ての第3次メッシュ約38万枚	CNN(畳み込みニューラルネットワーク) 2種類のモデル ①VGG16 ②LeNet	クラスタリングを行うための定番のアルゴリズムであるK-means

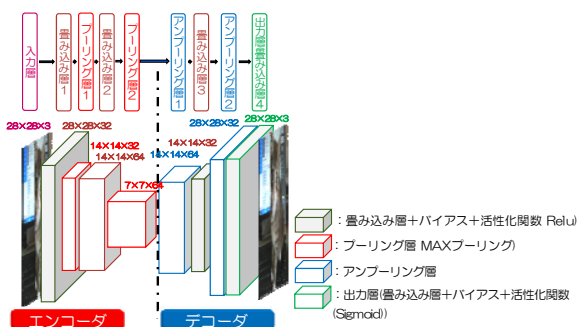


図-1 新規学習 CNN モデルの概要図 (LeNet)

2.2 学習データの作成

学習データは、数値標高モデル（10m メッシュ）からCS 立体図を作成し、その上に透過した地質図を重ね合わせた第3次メッシュごとのRGBの3チャンネルを持つRGB画像データとした。

またRGB画像データを作成するために必要な標高、曲率、傾斜、地質それぞれを4チャンネルとするマルチチャンネル画像データも作成した。（図-2）

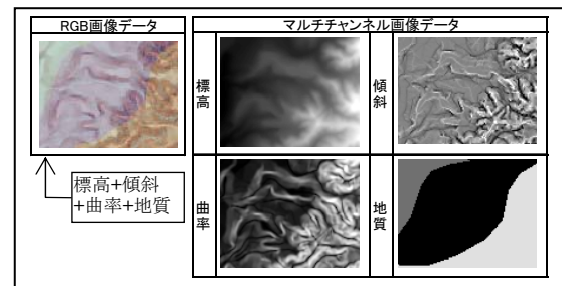


図-2 作成した学習データの一例

2.3 画像解析モデルと学習データの選定

過去に大規模な土砂災害が発生した46事例を対象に、RGB及びマルチチャンネルの画像データをそれぞれ8,731枚作成した。

試行検討ケースの条件を下記に整理する。

- ①モデル:VGG16、LeNet
- ②学習データの範囲:既往の災害46事例が発生した地域8,731メッシュ
- ③学習データ形式:RGB、マルチチャンネル画像
- ④フィルタサイズ:3×3
- ⑤学習回数:500回
- ⑥クラス数:10、20、30、40、50

試行ケースを用いた検討の結果、画像解析モデルにはおおまかな特徴抽出が可能なLeNet、学習データにはコンピューターに識別させる深層学習に、より適していると考えられるマルチチャンネル画像データを採用した。（図-3）

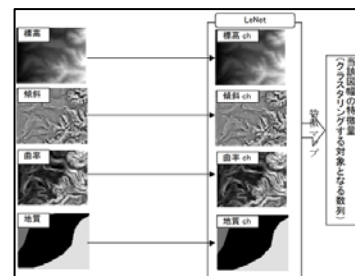


図-3 LeNet を使用した学習のイメージ

2.4 ハイパーパラメータ（フィルタサイズ）の決定

土砂災害の種類に応じた地形特徴を深層学習（画像解析）により抽出するためには、地形特徴の規模等の特性に応じたパラメータを設定する必要があると考え、模擬地形データ（図-4 参照）を使用しフィルタサイズの検討を実施した。規模の大きい表層崩壊

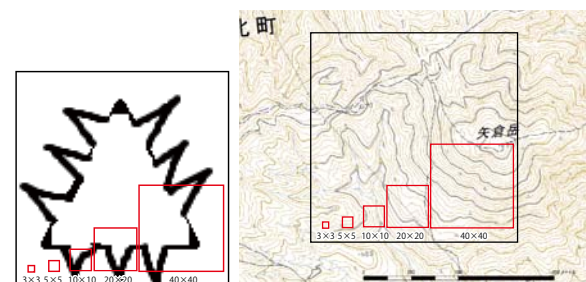


図-4 フィルタサイズのイメージ

や小規模な地すべりが捉えられる程度が妥当と考え、30～50m 程度の地形を認識できることを目標とした。
 検討結果より、30m～50m 程度の地形規模がうまくクラスタリングされている 5×5 を採用した。

3. 解析結果

3.1 全国を対象としたクラスタリング結果確認

日本全国を対象とした約 38 万枚の学習データについて、特徴を抽出し、クラスタ数を 10 から 50 まですべて 10 刻みの 5 区分でクラスタリングを実施した結果は以下となる。

- ・クラスタ数 10 の場合でも地質や地形量（起伏量、斜面形状、傾斜角）のばらつきはみられるため、同じ特徴を持ったクラスタが形成されていると考えられる。
- ・クラスタ数を増やすごとに地質や地形量（起伏量、斜面形状、傾斜角）に明瞭な差があるクラスタが増え、段々と分類が細分化されていく。

3.2 個別事例の分析によるクラスタ分類基準検討

平成 29 年 7 月九州北部豪雨時の崩壊発生領域を対象として確認した結果について、10 クラスタに区分した結果を図-5 に、50 クラスタに区分した結果を図-6 に示す。

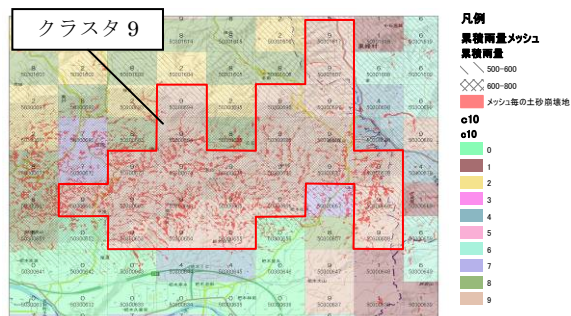


図-5 崩壊分布・累積雨量及びクラスタ区分（10 クラスタ）

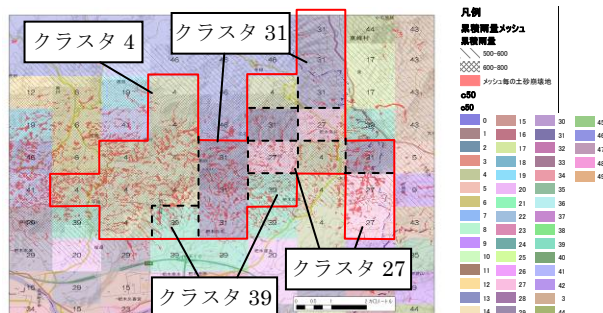


図-6 崩壊分布・累積雨量及びクラスタ区分（50 クラスタ）

10 クラスタに区分した場合、崩壊が発生している領域は、図-5 に示すとおりクラスタ 9 が卓越している様子がみられる。このクラスタは概ね深成岩の山地領域に該当する。この領域において崩壊が多いのは降雨量が多かったことによる影響も考えられ、実際に降雨領域の中心部近くはクラスタ番号によらず崩壊が多いことがわかる。しかしながら、図-7 に示すように豪雨領域の縁辺部となる累積雨量 600mm 以下の領域について着目すると、クラスタ 9 と比較して、クラスタ 2（黄）、クラスタ 8（深緑）は崩壊の分布が少ない傾向がみられる。

クラスタ数を 10, 30, 50 と増やしていくにしたがい、クラスタ 9 が細分化されていく推移を整理した結果を図-8 に示す。クラスタ数が 50 程度になると、微地形の区分や地質の混じり具合が反映されている。

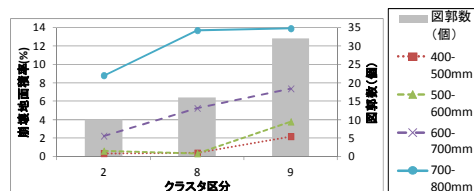


図-7 主要クラスタと崩壊面積率（10 クラスタ）

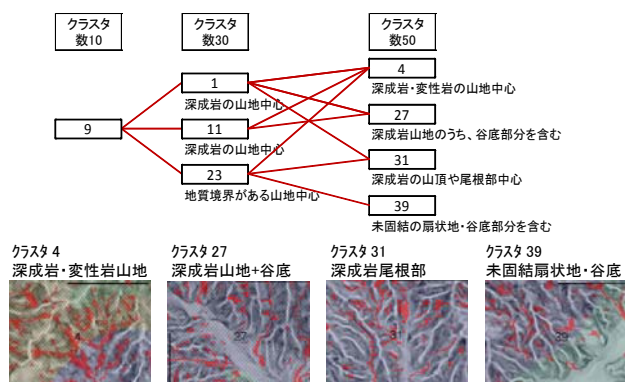


図-8 クラスタ数の増加に伴うクラスタ 9 の細分化

なお、上述までの「教師なし学習」の結果と比較・確認するために、実績崩壊分布を教師データとした「教師あり学習」（Deep Neural Network、Pix2Pix）を実施した結果、実績の崩壊と高い相関があるのはクラスタ 9 であることも確認した。

4. まとめと今後の展望

本研究では、日本全国の 3 次メッシュ単位の地形、地質をマルチチャンネル画像データとして学習し、クラスタ分類を行い、個別事例として平成 29 年 7 月九州北部豪雨時の崩壊発生領域を対象としクラスタ形成の推移の根拠を分析した。その結果下記の成果を得た。

- ・10 クラスタに区分した場合、地形や地質などの同じ特徴により区分することができた。
- ・50 クラスタに区分した場合、実績の崩壊発生領域における素因のうち微地形や地質の入り込み等の定性的な違いで区分することができた。

今後、クラスタ数の増加に伴い、別クラスタに分化した推移を精査、分析することで、土砂災害に関わる新しい視点（例えば、「天然ダムのできやすさ」「被害の発生しやすさ」等）の素因を抽出できる可能性がある。その際には、地形や地質分布の形状や複雑さなどに着目した結果の評価が有効であると考えられる。

謝辞

本検討を進めるにあたり、学習データの作成手法についてのご指導を長野県林業総合センターの戸田主任研究員、深層学習に関する技術のご指導はこだて未来大学の松原教授、深層学習に関する砂防的観点でのご指導を九州大学の久保田教授に頂きました。ご指導頂いた皆様へ感謝の気持ちと御礼を申し上げたく、謝辞にかえさせていただきます。

参考文献

- ・ゼロから作る Deep Learning 斎藤康毅
- ・岩崎ら(2017) Deep Learning での地図タイル活用の検討, 2017 年度人工知能学会全国大会論文集
- ・戸田堅一郎(2014)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 56(2)