

# Mask R-CNN による衛星画像を用いた地震時の土砂崩壊箇所の自動検出

Identification of the Locations of Landslides after an Earthquake based on Satellite Images using Mask R-CNN

23WM1332 原 恭也  
Kyoya Hara

指導教員 丸山 喜久, 劉 ウェン

## SYNOPSIS

Landslides with human casualties often occur in Japan. In areas where such major landslides have occurred, the Geospatial Information Authority of Japan (GSI) produces slope failure and sedimentation distribution maps. However, in preparing these maps, the interpretation of slope failure points is carried out manually by specialists who look at satellite images of the affected area, which causes problems such as a lack of human resources, the burden on workers and the difficulty in quickly identifying the extent of landslide damage. The aim of this study is to develop a method for automatic detection of landslide failure points using satellite images based on deep learning. As a result, we were able to construct a model which can detect landslides with an IoU of 0.641.

## 1. はじめに

我が国では人的被害を伴う土砂災害がしばしば発生している。2018年9月6日に発生した北海道胆振東部地震では、震源に近い厚真町を中心に土砂災害が多発した。土砂崩壊面積は推定約13.4 km<sup>2</sup>にわたり、死者43人、負傷者782人の甚大な被害をもたらした(2019年8月20日時点)<sup>1)</sup>。このような大規模な土砂災害があった地域では、国土地理院等が土砂災害の被害把握や分析および減災を目的として斜面崩壊・堆積分布図を作成する。しかし、本図の作成に当たって、斜面崩壊箇所の判読、抽出は、専門の作業員が被災地域の衛星画像から目視・手作業で行っている。そのため、将来的には人材不足の問題が懸念される。さらに、現時点での課題としては、作業員の負担が大きいこと、土砂災害範囲の迅速な把握が難しい場合もあることが挙げられる。

近年、深層学習(Deep Learning)が注目されている。深層学習は機械学習の一種で画像認識によく利用されており、精度・処理速度ともに従来の機械学習を上回る結果を記録している。深層学習を用いて土砂崩壊箇所を自動検出する手法を構築することによって、前述した斜面崩壊・堆積分布図の作成における3つの課題を解決できる。さらに、迅速な被害把握が可能となり、人命救助にも役立つと考えられる。そこで本研究では、地球観測衛星が撮影する土砂災害後の画像を用いて、土砂崩壊箇所の形状・範囲を自動検出する手法を深層学習により構築することを目的とする。

## 2. 本研究で用いる深層学習手法

本手法では、オープンソースのニューラルネットワークフレームワークであるTensorFlow<sup>2)</sup>で実装されたMask R-CNNを用いた物体検出およびセマンティックセグメンテーションを行う。物体検出はバウンディングボックスで行われ、これにより大まかな土砂崩壊箇所の位置を把握できる。また、セマンティックセグメンテーションは、画像上の全画素に対してクラス分類を行う手法であり、これにより土砂崩壊箇所の形状・範囲を検出することが可能である。

本研究で用いるMask R-CNNのネットワーク図を図-1に示す。Mask R-CNNのネットワークはBackbone部、RPN

(Region Proposal Network)部、Head部の3つに大別される。Backbone部は画像分類のタスクで用いるCNNと同様の機能を有しており、入力画像の特徴を抽出する役割を担う。RPN部は物体らしき領域の候補を抽出する役割を、Head部はその領域についてRoI Align<sup>3)</sup>で物体のクラスと位置を推定する役割をそれぞれ担っている。本研究では、Backbone部のネットワークにResnet101<sup>4)</sup>と呼ばれる101層のCNNを用いた。また、Mask R-CNNはMSCOCOデータセットにより事前学習が行われている。本研究では、土砂崩壊箇所の検出を目的に、事前学習されたモデル全体を再学習させるファインチューニングを行った。モデルのハイパーパラメータに関しては、基本的に初期値を用いており、学習率は0.0001、mrcnn\_mask\_lossを2.0と設定した。

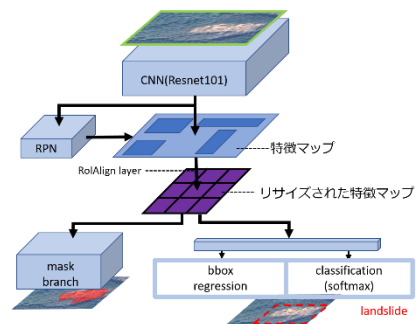


図-1 本研究で用いるMask R-CNNのネットワーク

## 3. 衛星画像のデータセット

本研究では、地球観測衛星の一つであるSPOT-7<sup>5)</sup>が北海道胆振東部地震発生後の2018年10月4日に撮影した衛星画像を用いる。撮影対象地域は北海道厚真町周辺であり、画像の大きさは縦14,396画素、横13,248画素である。この衛星画像は、赤・緑・青・近赤外線<sup>6)</sup>の4つの波長帯の電磁波を記録したマルチスペクトル画像であり、分解能は1.5mである。赤・緑・青の波長帯でカラー合成した衛星画像を図-2に示す。



図-2 本研究で用いる衛星画像（全体図）

さらに、土砂災害前の標高データを用いるため、国土地理院の数値標高モデル<sup>6)</sup>から衛星画像と同じ範囲の標高データを得た。数値標高モデルは地震前の2016年10月1日に更新されており、10mメッシュでデータが格納されている。本研究では、衛星画像の解像度1.5mに合わせるためリサンプリングを行った。

植生指標データを算出した。植生指標とは、植物による光の反射の特徴を生かし衛星データを使って簡易な計算式で植生の状況を把握することを目的として考案された指標で、植物の量や活力を表す<sup>7)</sup>。代表的な植生指標にはNDVI (Normalized Difference Vegetation Index: 正規化植生指標)があり、赤の波長帯のデータ  $R$  と近赤外線の波長帯のデータ  $IR$  から式(1)のように算出される。

$$NDVI = \frac{IR - R}{IR + R} \quad (1)$$

NDVIは-1から1までの値をとる。深層学習には、NDVIを式(2)のように整数化した植生指標データ ( $NDVI_{int}$ )を用いた。

$$NDVI_{int} = (NDVI + 1) \times 100 \quad (2)$$

植生指標データを取り入れる利点は、赤の波長帯のデータと近赤外線の波長帯のデータの2つの情報が1つの指標に集約される点である。これにより青、緑の波長帯のデータに植生指標データを加えることで、3チャンネルの画像に4つの波長帯のデータが含まれることとなる。

以上のデータを用いて、赤・緑・青の3チャンネルの衛星画像、赤・緑・青・近赤外線の4チャンネルの衛星画像、赤・緑・青・標高データの4チャンネルの衛星画像、緑・青・標高データ・植生指標データの4チャンネルの衛星画像の計4種類の衛星画像を作成した。

そして、Mask R-CNNを用いてセマンティックセグメンテーションを行うには、学習用画像と学習用画像ごとに画像内の検出対象とその他の部分をアノテーションしたマスク画像が必要となる。まず、学習用画像は、用意した4種類の衛星画像をそれぞれ正規化し、縦256画素、横256画素のメッシュで分割した。その後、雲で覆われている範囲や画素がない箇所が含まれる画像を除き、1,985枚の画像を得た。このように、4種類のデータセットを用意した。マスク画像の作成には、国土地理院が公開しているGeoJSON形式の斜面崩壊・堆積分布図のデータ<sup>8)</sup>と、株式会社地域環境計画が公開しているシェープファイル形式の土砂崩壊地、土砂堆積地マップデータ<sup>9)</sup>を用いた。この2つのデータをマージすることによって、見逃しが少ないと期待される土砂崩壊領域を作成した。

しかし、このデータと本研究で用いる衛星画像の土砂崩壊領域を見比べると、土砂崩壊箇所であるにも関わらず土

砂崩壊箇所とされていない範囲がいくつか存在した。この原因には、使用した空撮画像の撮影日の違いが挙げられる。国土地理院の斜面崩壊・堆積分布図および地域環境計画の土砂崩壊地、土砂堆積地マップを作成する際に用いた衛星画像の撮影年月日は2018年9月6~11日である。一方、本研究で用いている衛星画像の撮影年月日は2018年10月3日である。この間には台風24号、25号の影響による降雨<sup>10)</sup>が観測されており、土砂崩壊の状況が変わっている可能性がある。したがって、目視で土砂崩壊範囲のデータを修正した。この修正データをマスク画像化し、学習用画像データと同様に縦256画素、横256画素のメッシュで分割した。このようにして土砂崩壊箇所が白く塗りつぶされたマスク画像を学習用画像と同じ枚数だけ得た。マスク画像の一例を図-3に示す。そして、これらの学習用画像とマスク画像をペアとして6:2:2の割合で訓練、検証、テストデータに分割した。

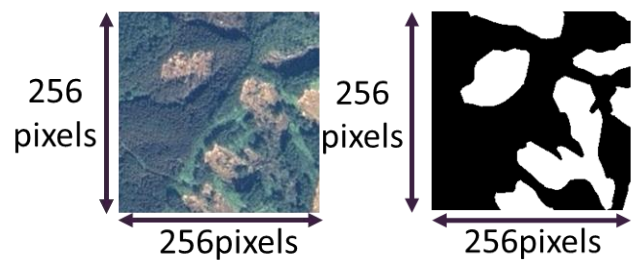


図-3 学習用画像（左）マスク画像（右）の一例

#### 4. 学習結果と精度評価

赤・緑・青の波長帯の3チャンネルの画像を学習に用いたモデルをモデル1、赤・緑・青・近赤外線の4チャンネルの画像を学習に用いたモデルをモデル2、赤・緑・青・標高データの4チャンネルの画像を学習に用いたモデルをモデル3、緑・青・標高データ・植生指標データの4チャンネルの画像を学習に用いたモデルをモデル4とする(表-1)。両者ともに100エポック時点の学習モデルを最良のモデルとし、テストデータによる汎化性能の評価を行った。

表-1 各モデルの学習データ内訳

モデル	データ内訳
1	赤・緑・青
2	赤・緑・青・近赤外
3	赤・緑・青・標高
4	緑・青・標高・植生指標

本研究ではセマンティックセグメンテーションの汎化性能を評価する上で、正解率、適合率、再現率、F値、IoUの5つの指標を導入する。各指標式を式(3)~(7)に示す。式中の $TP, FP, FN, TN$ は、表-2の混同行列のように定める。正解率  $Acc$  は、以下のように算出される。

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

適合率  $Pre$  は、モデルが土砂崩壊箇所であると検出した画素のうち、実際に土砂崩壊箇所であった画素の割合を示す。誤検出の割合を調べるときに使用する。

$$Pre = \frac{TP}{TP + FP} \quad (4)$$

再現率  $Rec$  は、実際の土砂崩壊箇所のうち、モデルが土砂

崩壊箇所であると正しく検出できた画素の割合を示す。検出見逃しの割合を調べるときに使用する。

$$Rec = \frac{TP}{TP + FN} \quad (5)$$

F 値  $F$  は、適合率と再現率の調和平均である。これらの 2 指標を総合的に見るときに使用する。

$$F = \frac{2 \cdot Pre \cdot Rec}{Pre + Rec} \quad (6)$$

IoU は、正解領域とモデルが検出した領域の積集合を和集合で除した、2つの領域がどの程度重なっているかを表す指標であり、セマンティックセグメンテーションにおいて総合的な精度を評価できる重要な指標である。

$$IoU = \frac{TP}{TP + FP + FN} \quad (7)$$

表-2 セマンティックセグメンテーションの混同行列 (単位: 画素)

		検出結果	
		崩壊	非崩壊
正解データ	崩壊	$TP$	$FN$
	非崩壊	$FP$	$TN$

物体検出の汎化性能を評価するために、適合率、再現率、F 値の 3 つの指標を導入する。各指標式は式(8)~(10)に示す。

式中の  $TP_{bbox}$ ,  $FP_{bbox}$ ,  $FN_{bbox}$  は、表-3 の混同行列のように定める。なお、本研究ではモデルが検出したバウンディングボックスと正解バウンディングボックスとの適合率、もしくは再現率のどちらかが 0.5 以上の場合を  $TP_{bbox}$ 、0.5 未満の場合は  $FP_{bbox}$  とした。正解バウンディングボックスと紐づく検出バウンディングボックスが存在しない場合を  $FN_{bbox}$  とした。なお、非崩壊領域を正しく検出しない割合 ( $TN$ ) は評価できない。

適合率  $Pre_{bbox}$  は、モデルが土砂崩壊箇所であると検出したバウンディングボックスのうち、実際に土砂崩壊箇所を示すバウンディングボックスの割合を示す。誤検出の割合を調べるときに使用する。

$$Pre_{bbox} = \frac{TP_{bbox}}{TP_{bbox} + FP_{bbox}} \quad (8)$$

再現率  $Rec_{bbox}$  は、実際の土砂崩壊箇所を示すバウンディングボックスのうち、モデルが土砂崩壊箇所であると正しく検出できたバウンディングボックスの割合を示す。検出見逃しの割合を調べるときに使用する。

$$Rec_{bbox} = \frac{TP_{bbox}}{TP_{bbox} + FN_{bbox}} \quad (9)$$

F 値  $F_{bbox}$  は、適合率と再現率の調和平均である。これらの 2 指標を総合的に見るときに使用する。

$$F_{bbox} = \frac{2 \cdot Pre_{bbox} \cdot Rec_{bbox}}{Pre_{bbox} + Rec_{bbox}} \quad (10)$$

表-3 バウンディングボックスの混同行列 (単位: 個)

		検出結果	
		崩壊 bbox	非崩壊領域
正解データ	崩壊 bbox	$TP_{bbox}$	$FN_{bbox}$
	非崩壊領域	$FP_{bbox}$	

モデル 1~4 のテストデータによる全画素についてのセマンティックセグメンテーションの精度指標を表-4 に、モデル 1~4 のテストデータによる全バウンディングボックスについての物体検出の精度指標を表-5 に示す。

表-4 テストデータによる精度指標 (セマンティックセグメンテーション)

指標	モデル 1	モデル 2	モデル 3	モデル 4
正解率	0.916	0.922	0.914	0.917
適合率	0.756	0.792	0.792	0.758
再現率	0.789	0.770	0.715	0.797
F 値	0.772	0.781	0.751	0.777
IoU	0.629	0.641	0.602	0.635

表-5 テストデータによる精度指標 (物体検出)

指標	モデル 1	モデル 2	モデル 3	モデル 4
適合率	0.659	0.793	0.778	0.653
再現率	0.849	0.845	0.817	0.920
F 値	0.742	0.818	0.797	0.764

まず、セマンティックセグメンテーションについて各モデルを比較する。モデル 1 とモデル 2 の精度の比較をみると、再現率のみモデル 1 の方が高いが、他の指標はモデル 2 の方が高い結果となった。特にセマンティックセグメンテーションにおいて重要な指標である IoU がモデル 2 の方が高い。このことから、近赤外線波長帯のデータはセマンティックセグメンテーションによる土砂崩壊領域の検出に有効だと考えられる。

モデル 1 とモデル 3 の比較をみると適合率はモデル 3 の方が高く、その他の指標はモデル 1 の方が高い。さらに、全精度指標においてモデル 3 はモデル 2 以下となった。このことから、標高データは適合率を向上させる上では有用であるが、赤・緑・青の波長帯のデータに追加するだけでは IoU の向上には寄与せず、近赤外線波長帯のデータが標高データよりも有効であることが確認できた。

同じ標高データを用いているモデル 3 とモデル 4 を比較すると、適合率を除く全ての指標においてモデル 4 の方が高い結果となった。このことから、モデル 3 の赤の波長帯のデータの代わりに植生指標データを追加することで、誤検出は多少増えるが、総合的にはセマンティックセグメンテーションの精度が向上することが確認できた。

次に、物体検出について各モデルを比較する。モデル 1 とモデル 2 の精度の比較をみると、再現率はほぼ同等、適合率は 0.1 以上、F 値も 0.05 以上、モデル 2 の方が高い結果となった。このことから、近赤外線波長帯のデータはセマンティックセグメンテーションだけでなく物体検出による土砂崩壊箇所の検出においても有効と考えられる。

モデル 1 とモデル 3 を比較すると、適合率、F 値はモデル 3 の方が高い。このことから、標高データを追加することで物体検出の精度は総合的に上がることが確認できた。しかしモデル 2 と比較すると全ての指標においてモデル 2 の方が精度が高く、よってセマンティックセグメンテーションの場合と同様に近赤外線波長帯のデータが標高データよりも有効であった。

モデル 3 とモデル 4 を比較すると、再現率以外はモデル 3 の方が高い結果となった。モデル 4 は再現率が 0.92 と非常に高いが、適合率は 0.653 と低く、過剰にバウンディングボックスを生成してしまう傾向がある。よって、モデル

3の赤の波長帯のデータの代わりに植生指標データを追加することで、検出の見逃しは減少するが、物体検出の総合的な精度は低下することが確認できた。

最後にモデル1~4について総合的に評価する。セマンティックセグメンテーションについては適合率・IoUに関してモデル2が最も高く、誤検出が少ないかつ総合的にも一番精度が高いモデルといえる。再現率に関してはモデル4が最も高く、検出の見逃しが一番少ないモデルといえる。また、物体検出についても適合率・F値に関してモデル2が最も高く、誤検出が少なく、総合的にも一番精度が高いモデルといえる。再現率に関してはモデル4が最も高く、検出の見逃しが一番少ないモデルといえる。

そこで、実際の土砂災害時における本モデルの運用時においては、発災直後、すなわち迅速かつ確実に安全な場所、また危険な場所を判定する必要がある際には検出の見逃しが少ないモデル4を使用し、発災からある程度時間が経った後、すなわちより精度が高い土砂崩壊箇所の検出が求められる場合にはモデル2を使用するといった使い分けが考えられる。

これまでの結果から、赤・緑・青の波長帯のデータを用いたモデル1よりも、モデル1に近赤外線波長帯のデータを追加したモデル2が総合的に最も精度が高いことがわかった。そこで近赤外線波長帯のデータが検出にどのような影響を及ぼしているかを考察するため、テスト画像の検出結果の精度を一枚ずつ個別に比較し、どのようなケースでモデル1よりもモデル2の検出結果が優れているのかを検証する。397枚のテストデータのうち、モデル2の方がIoUが高いテスト画像は169枚、モデル1の方がIoUが高いテスト画像は134枚であった。そして、モデル2の方がIoUが高い169枚のテスト画像の中から抽出した画像の検出結果の比較を図-4に示す。正解領域は画像左の2つの土砂崩壊箇所であり、モデル1では正解領域は概ね検出できているが、影となっている領域で4か所の誤検出をしている。しかし、モデル2では、正解領域である画像左の2つの土砂崩壊箇所のみを検出し、誤検出はない。モデル1とモデル2の比較では、このような傾向が多くテスト画像で見られた。このことから、近赤外線波長帯のデータを追加することは、影が影響する誤検出を防ぐ効果があり、モデル2の方がより精度が高くなることがわかった。

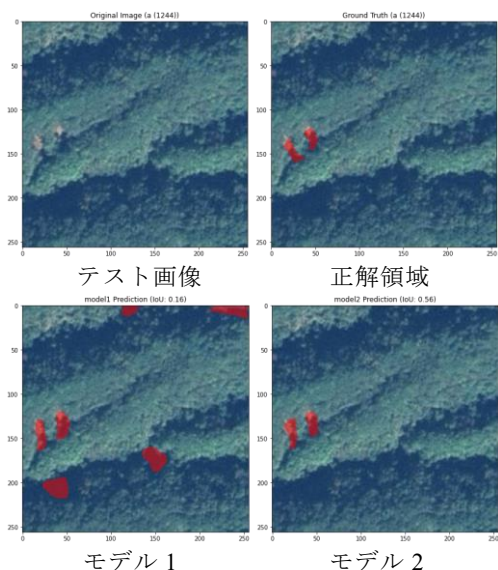


図-4 モデル1とモデル2の検出結果

## 5. 結論

本研究では、2018年北海道胆振東部地震を対象とし、衛星画像から土砂崩壊箇所を検出する手法に関する検討を行った。青・緑・赤・近赤外線の波長帯のデータから構成される4チャンネルの衛星画像を学習に用いたモデル2は、セマンティックセグメンテーションの精度で正解率0.922、適合率0.792、再現率0.770、F値0.781、IoU0.641、物体検出の精度で適合率0.793、再現率0.845、F値0.818と良好な精度を示した。さらに、青・緑の波長帯のデータと、標高データ・植生指標データから構成される4チャンネルの画像を学習に用いたモデル4は、再現率が高く、検出の見逃しが少ないモデルとなった。今後は、本モデルの汎用性を検証する目的で、2024年能登半島地震など他地域で起きた災害後の衛星画像を用いた精度の検討を予定している。

## 参考文献

- 1) 消防庁：平成30年北海道胆振東部地震による被害及び消防機関等の対応状況（第35報），  
<https://www.fdma.go.jp/disaster/info/items/190820hokkai-doujisinn35.pdf>
- 2) Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, Google Brain: "TensorFlow: A System for Large-Scale Machine Learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pp. 265–283, 2016
- 3) He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn, Proceedings of the IEEE international conference on computer vision, pp. 2961-2969, 2017.
- 4) He, K., Zhang, X., Ren, S., & Sun, J: Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.
- 5) 一般財団法人リモート・センシング技術センター：SPOT6,7，<https://www.restec.or.jp/satellite/spot-6-7>
- 6) 基盤地図情報ダウンロードサービス，  
<https://fgd.gsi.go.jp/download/menu.php>
- 7) 国土地理院：植生指標データについて，  
<https://www.gsi.go.jp/kankyochiri/ndvi.html>
- 8) 国土地理院：斜面崩壊・堆積分布図 GeoJSON，  
<https://www.gsi.go.jp/common/000204728.zip>
- 9) (株)地域環境計画：北海道胆振東部地域 土砂崩壊地、土砂堆積地マップ，<https://www.chiikan.co.jp/iburi/>
- 10) 小山内信智，海堀正博，山田孝，笠井美青，林真一郎，桂真也，古市剛久，柳井清治，竹林洋史，藤浪武史，村上泰啓，伊波友生，佐藤創，中田康隆，阿部友幸，大野宏之，武士俊也，田中利昌，小野田敏，本間宏樹，柳井一希，宮崎知与，上野順也，早川智也，須貝昂平：平成30年北海道胆振東部地震による土砂災害，砂防学会誌，Vol. 71, No. 5, pp. 54-65, 2019.