公開地理空間情報を利用したマルチチャンネル深層学習による土砂災害検出
－平成 30 年北海道胆振東部地震を事例に－

三好達也*・橋本雄一**

Landslide Detection by Multi-Channel Deep Learning Using Open Geospatial Data
— A Case Study of the 2018 Hokkaido Eastern Iburi Earthquake —

Tastuya MIYOSHI*, Yuichi HASHIMOTO**

Abstract: In Japan, earthquakes and heavy rain often cause landslides, requiring rapid assessment immediately after a disaster. We evaluated multi-channel deep learning that integrates aerial RGB imagery with open geospatial data from the 2018 Hokkaido Eastern Iburi Earthquake. YOLO11 models with elevation and slope inputs improved detection accuracy and supported wide-area screening, highlighting the benefit of incorporating topographic context. In contrast, the inclusion of land-use layers raised precision but reduced recall, underscoring that feature design and data representation remain key challenges for practical landslide detection.

Keywords:土砂災害（landslide），深層学習（deep learning），物体検出（object detection）

１．はじめに

近年，日本各地で地震や豪雨に伴う土砂災害が頻発している．こうした災害の発生直後には，被災箇所を迅速かつ広範囲に把握する手法が強く求められる．一般的に土砂災害の検出には，航空写真の目視判読や現地踏査による確認が行われているが，災害規模が広域に及ぶ場合には人的負担および時間的コストが非常に大きい．そこで近年，リモートセンシング技術の進歩と深層学習の発展に伴い，画像から土砂災害箇所を自動検出する研究が進められている．

１．１　先行研究と課題

これまで，リモートセンシング画像を用いた土砂災害検出に関する研究として，衛星のマルチスペクトル画像や SAR データを活用した手法が多数報告されている．しかし，これらのデータは特殊なセンサーによるものであり，画像の取得が容易ではない．

一方，入手しやすい可視光（RGB）画像のみを用いた土砂災害検出の試みも報告されている．例えば，叶井ら（2021）は災害直後に撮影された航空写真に深層学習モデル（Mask R-CNN）を適用し，効率的に土砂崩壊地を抽出する手法を提案した．また，河村ら（2018）は深層学習によって空中写真から土砂移動範囲を自動抽出する手法を提案している．

しかし，可視光画像のみを入力とする深層学習では利用可能な特徴量が限られるため，マルチスペクトル画像等を用いた場合と比較して検出精度の向上が難しい可能性がある．

１．２　本研究の目的

そこで本研究では，可視光航空写真（RGB3 チャンネル）に対して，第 4 チャンネル以降に基盤地図情報や国土数値情報といった一般公開されている地理空間データを付加するアプローチを提案する．再現性の高い環境構築を採用することにより，汎用性と検出精度向上の両立を図ることを目的とする．

２．深層学習を活用した研究手法と環境構築

２．１　２つの検出手法

画像から対象領域を検出する深層学習の手法は，物体検出（Object Detection）とセマンティックセグメンテーション（Semantic Segmentation）の大きく 2 つに分類される（図-1）．

それぞれの手法には特徴やメリットに違いがあるが（表-1），既往の土砂災害検出研究の多くは，セマンティックセグメンテーションを採用している例が多く，物体検出を採用している事例は少ない．本研究では，アノテーションの作成と発災時における検出速度を重視して物体検出を採用する．

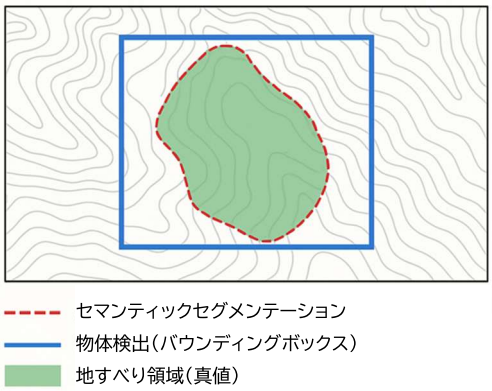


図 1　物体検出とセグメンテーションの違い

*　正会員　　株式会社ドーコン　TEL:011-801-1590
**　正会員　　北海道大学　大学院　文学研究院　人間科学部門　地域科学分野　地域科学研究室

表 1　2つの検出手法の特徴

手法	物体検出	セマンティック セグメンテーション
データ出力	矩形ボックス＋ラベル	ピクセル単位のマスク
想定用途	発災直後の位置把握を短時間で行う	面積の精密抽出や土量推計などの詳細解析
利点	推論が軽く高速でアノテーション作成が比較的	境界の再現性が高いく定量評価に向く
留意点	検出範囲は矩形近似となるため面積形状の精密性	学習データ整備の負荷が相対的に高い

２.２　使用するデータ

本研究では，国土地理院で作成され公開されている平成 30 年(2018 年)北海道胆振東部地震で撮影された航空写真（オルソ化済）と，判読された地すべりのポリゴンベクトルデータ（GeoJSON 形式）を用いて分析を行う．実際の航空写真データは，地理院地図で提供されており地理院タイル（レベル 18）をダウンロードしたものを使用する（図-2)．

拡張チャンネル用のラスターデータは基盤地図情報の数値標高モデル 10m メッシュ標高と国土数値情報の土地利用細分メッシュデータ（100m メッシュ）の平成 28 年度を使用する．

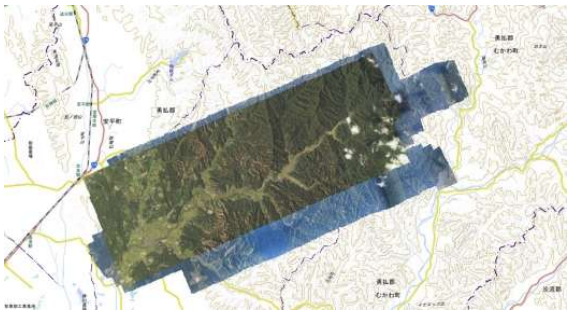


図 2　厚真川地区（9 月 6 日撮影）　出典：国土地理院

２.３　解析環境の構築

ここでは，深層学習を実行するための環境構築について述べる．本研究では，一般的に入手可能なゲーミング PC 程度のスペックで構成している（表-2)．この環境において，Windows 11 上に Python の仮想環境を構築し，データ処理・学習・推論を実行する．

表 2　本分析における開発環境

項目	名称	性能/役割
CPU	Ryzen 7 7800X3D	8C/16T
メモリ	DDR4-DIMM	32GB
保存領域	NVNE SSD	2 TB
VGA	GeForce RTX 4070	VRAM 12GB
OS	Windows11 Pro	
ドライバ	CUDA Runtime 12.6	GPUライブラリ
言語	Python 3.11.9	開発言語
フレームワーク	Pytorch 2.8.0+cu126	深層学習
	YOLO11/Ultralytics	物体検出
主張ライブラリ	numpy/pandas	数値データ処理
	OpenCV	画像処理
	GDAL/Rasterio	ラスタ処理
	GeoPandas/pyproj	ベクタ処理

３．マルチチャンネル拡張のアプローチ

３.１　チャンネル拡張と地理空間情報

本研究では，航空写真 RGB の 3 チャンネル（基準）から出発し，5 チャンネル（RGB＋標高＋傾斜），さらに 8 チャンネル（土地利用）へと入力情報を段階的に拡張する．標高や傾斜などの地形派生量は斜面安定性と強く関連し，外観が類似する裸地・造成地・採石場等との識別に有効な補助情報となることが知られている．加えて，土地利用や河川・道路といった周辺環境を表す情報を重ねることで，災害発生の地理的文脈をモデルに学習させ，検出の一貫性と汎化性の向上を狙う．

３.２　５チャンネル拡張（標高系データ）

先行研究で地形情報をレイヤに組み込み精度向上を目指した事例がある．そこでまず，標高データと傾斜角度データの２つ用意し，合計５チャンネルに拡張することを試みた．

本研究では，技術的制約としてすべての値を 8bit の整数ラスタに変換する必要があるため，標高データについてはメートル単位で格納が難しい．よって，全国の標高に対応できるように 4096mのレンジを基準とし 6m を 1 単位とする 8bit 標高ラスタに圧縮して作成した．

次に，標高データから傾斜度を算出し，角度データとして 8bit の整数に格納した．なお，両データとも 10m メッシュ標高であることから，航空写真よりも解像度が低いため，航空写真の解像度に統一してマルチページの GeoTIFF を生成した．

３.３　８チャンネル拡張（土地利用データ）

国土数値情報からダウンロードした土地利用細目メッシュ（100m）の GML から地すべりに関係深いと考えられる「森林」「道路」「河川及び湖沼」のメッシュをそれぞれ，0 と 1 のバイナリの値を入力した 8bit ラスタ 3 に分割し，5 チャンネルの GeoTIFF に統合した．

表 3　各チャンネルデータ一覧

データ	ch	項目	格納値
航空写真	1	R	0-255
	2	G	0-255
	3	B	0-255
10mメッシュ標高	4	標高データ	0-255
	5	傾斜度	0-255
土地利用細分メッシュ	6	森林	0 or 1
	7	道路	0 or 1
	8	河川及び湖沼	0 or 1

３.４　教師データの作成

学習単位の画像は 1024×1024 ピクセル，オーバーラップ 256 ピクセルで元の航空写真から切り出しを行う．完全な無地・透明域が一定割合を超えるタイルは除外し，タイルごとのチャンネル統計値を確認して外れ値を抑制

し，合計 869 個のタイル画像を用意する．

物体検出用の教師データ（物体検出用ラベル）は，国土地理院の作成した地形判読済みの地すべりポリゴン（GeoJSON）をタイル画像単位でクリップし，外接矩形ポリゴンへ変換して YOLO 形式に整形する．結果，矩形ポリゴンがあるタイル画像は 396 枚制作し，ポリゴンが無い空のタイルを 473 枚となる．

これらのタイルを学習・検証を空間バイアスが偏らないように，乱数を用いて分類し，学習用 520 枚，検証用 349 枚のタイル画像を振り分ける．

４．YOLO11 を用いた学習と考察

本研究で用いた YOLO11 は，Ultralytics により公開されており，物体検出の高速性と高精度性を両立する特徴を持つ．YOLO（You Only Look Once）系列は，画像を一度の推論で分割し，各領域から同時にバウンディングボックスとクラスを推定する「ワンステージ検出器」であり，従来の R-CNN 系列に比べて大幅に高速であることが特長であり，従来モデルに比べて少ない計算量で高い検出性能を得られる．また，パラメータ数に応じて 5 種類のモデル規模が用意されており，利用環境や目的に応じた柔軟な選択が可能である．そこで，本研究ではこれら 5 種類を用いてそれぞれ学習を試み，規模による性能差を評価した（表-4）．

表 4 ３チャンネルによる学習スコア

モデル	パラメータ数	Precision（適合率）	Recall（再現率）	F1（F値）	mAP @0.5	mAP @0.5–0.95
n	2.6M	0.742	0.59	0.657	0.643	0.411
s	9.4M	0.816	0.625	0.708	0.704	0.509
m	20.1M	0.764	0.666	0.711	0.707	0.515
l	25.3M	0.729	0.642	0.683	0.682	0.47
x	56.9M	0.713	0.634	0.671	0.671	0.463

学習の結果，全モデルで安定して収束し，s モデルの Precision が 0.816，m モデルの Recall が 0.666 を示し，F1 値は m モデルが 0.711 で最良となった．なお，本研究では地すべりポリゴンを外接矩形に変換して教師データを作成しているため，mAP@0.5-0.95 の評価値は厳しくなりやすいが，それでも m モデルが最良値 0.515 を示した．このことは，矩形近似による不利な条件下でも中規模モデルがバランスの良い性能を示すことを意味する可能性があり，現実的なラベリング環境における有効性がある．

また，パラメータ数の大きな l および x モデルでは，計算コストに対してスコアの向上が限定的であった．これは学習タイル数が 869 枚と比較的少なく，大規模モデルの表現力を十分に活用できず最適化が進みにくかった可能性がある．大規模モデルは一般に豊富なデータ量を必要とすることが知られており，データ不足下では過学習や性能飽和を招きやすい．したがって，学習枚数が限られる場合は，計算資源と性能のバランスを考慮すると，s または m モデルが最も実務的に有用であると考えられる．

４.２ ５チャンネルの学習

学習条件は３チャンネル時と同一とし，入力のみを標高・傾斜を加えた５チャンネル（RGB+標高+傾斜度）に拡張した（表-5）．

結果をモデル別にみると，n モデルは５チャンネル化の効果が最も明瞭に現れた．Precision は 0.756，Recall は 0.648 と大きく伸び，mAP@0.5 は 0.687，mAP@0.5–0.95 は 0.472（+0.061）と大幅に改善した．その結果 0.698 へ上昇した．軽量モデルは高速に推論が可能なため，実用的な検出率への改善が期待できる内容となった．

s モデルは，適合率がさらに押し上げられ，Precision は 0.845 と顕著に改善し，Recall も 0.635 とわずかながら向上した．F1 は 0.725（+0.017）へと改善した．３チャンネル時には F1 は m モデルが最良であったが，５チャンネル化後は s モデルが僅差で最良となっている．

m モデルでは，mAP の伸びが大きい．mAP@0.5 は 0.728，mAP@0.5-0.95 は 0.546 と改善が大きく，Precision も 0.778 へと向上した一方で，Recall は 0.665 と横ばいであった．F1 は 0.717 と僅かながら上昇している．５チャンネル化によってボックスの位置合わせが安定し，検出品質の均質化が進んだと捉えられる．精度と再現のバランスを崩さずに定位精度を底上げできており，詳細確認や拾い漏れ抑制を重視するための採用候補となる．

l モデルでは，Precision は 0.752 とわずかな改善にとどまる一方，Recall は 0.627 と低下し，F1 も 0.684 と実質横ばいであった．高信頼な候補への絞り込みは強化されたが，その反面で拾い漏れがやや増えたといえる．

x モデルも同様の傾向を示し，Precision は 0.772 と向上した一方で，Recall は 0.622 とわずかに低下した．F1 は 0.689 と全体では改善しているが，軽量モデルのような一貫した底上げとはなっていない．大規模モデルでは確信度の高い検出に収束しやすく，その分だけ境界事例を取りこぼす（＝Recall 低下）傾向が表れたと解釈できる．

総評としては３チャンネル（RGB）から５チャンネル（RGB＋標高＋傾斜）への拡張により，全モデルで総じて精度指標が改善した．とくに mAP@0.5-0.95 が全モデルで上昇した点は重要であり，バウンディングボックスの位置合わせ精度が向上した．これは，標高・傾斜という地形文脈が，航空写真のみでは識別の難しい裸地・造成地・採石場等の見かけ上類似している地形を排除し，候補領域の絞り込みと中心の安定化に寄与したためと考えられる．

表 5 ５チャンネルによる学習スコア

モデル	Precision（適合率）	Recall（再現率）	F1（F値）	mAP @0.5	mAP @0.5–0.95
n	0.756	0.648	0.687	0.472	0.698
s	0.845	0.635	0.715	0.522	0.725
m	0.778	0.665	0.728	0.546	0.717
l	0.752	0.627	0.687	0.481	0.684
x	0.772	0.622	0.693	0.491	0.689

４.３ ８チャンネルの学習

学習条件は 3 および 5 チャンネル時と同一とし，入力のみを 8 チャンネルへ拡張した（表-6）．n モデルでは Precision が 0.787 と大幅に上昇した一方で，Recall は 0.323 に急落し，mAP@0.5 も 0.562 へ低下した．s モデルも同様に Precision は 0.876 に達したが，Recall は 0.285 まで低下し，mAP@0.5 は 0.588 にとどまった．

加えて，m モデルでも Precision が 0.903 と高水準である一方，Recall は 0.319 に低下し，mAP@0.5 は 0.617，mAP@0.5-0.95 は 0.541 と 5 チャンネル時からほぼ横ばい～小幅減となった．l モデルでは Precision が 0.912 まで上昇しつつ，Recall は 0.304 に低下したが，mAP@0.5-0.95 は 0.522 と 5 チャンネル時よりわずかに改善した．x モデルでは，さらに同様の傾向が続き Precision が 0.931，Recall は 0.279 まで低下した．なお，全体として F1 は 0.43～0.47 の範囲に下降し，Precision 偏重／Recall 低下という共通傾向が n～x の各モデルで確認された．

この結果は，土地利用データが 100m メッシュの粗い区分情報（バイナリ化）であるため森林・水域等の表現が過度に単純化され，過剰な絞り込みに学習が誘導された可能性がある．

表 6 ８チャンネルによる学習スコア

モデル	Precision（適合率）	Recall（再現率）	F1（F値）	mAP @0.5	mAP @0.5–0.95
n	0.787	0.323	0.458	0.562	0.412
s	0.876	0.285	0.430	0.588	0.491
m	0.903	0.319	0.471	0.617	0.541
l	0.912	0.304	0.456	0.611	0.522
x	0.931	0.279	0.429	0.609	0.524

５．まとめ

本研究では，平成 30 年北海道胆振東部地震の航空写真に，公開地理空間情報を重ねたマルチチャンネル入力で土砂災害検出を評価した．3 チャンネル（RGB）に対して 5 チャンネル（RGB＋標高＋傾斜）を導入すると，全モデルで精度が底上げされ，とくに Precision と mAP@0.5-0.95 が一貫して改善した．これは，斜面の急峻さなど地形文脈（標高・傾斜）が本質的な説明力を持つためである．加えて，日本の基盤地図情報の 10mDEM は全国が整備済みであり，入手容易性と品質の両面で実運用に適する．

一方，8 チャンネル（土地利用の追加）では，多くのモデルで Precision は上がるものの Recall が低下する傾向がみられた．ただし，これは 8 チャンネルという枠組み自体の限界というより，100m メッシュ・二値分類という特徴量設計に起因する可能性が高い．先行研究でも，土地被覆そのものに加え，河川・道路までの距離や植生指数等の連続量を適切な空間スケールで与えると説明力が高まることが示されており，距離ラスタ化や高解像度データの導入，スケール整合を図ることで，Recall の回復と総合精度の向上は十分に見込めると考えられる．

総じて，４チャンネル以上の拡張は実用的な改良である．しかし，チャンネル拡張は“データの良し悪し”が精度に直結することから，入力データの種類・空間解像度・符号化・特徴量化（距離・連続量）等，データ設計が重要である．

６．今後の課題

本研究で有効性が確認されたチャンネル拡張をさらに発展させるためには，入力データの多様化が必要である．特に，曲率，相対標高，地形位置指数，地形起伏指数，集水量や流向分布といった地形派生量を追加することで，崩壊地の識別能力を高められる可能性がある．

土地利用については，今後は距離ラスタ化や高解像度の土地被覆データ，あるいは他の地物データの導入により，特徴量設計を精緻化・多様化することが課題である．

さらに，学習データ数の拡充も重要である．未利用の航空写真データを追加し，学習素材を増やすことで，モデルの汎化性能や再現性を一層高めることが期待される．

参考文献

叶井和樹・久保栞・山根達郎・全邦釘（2021）MaskR-CNN による航空写真からの土砂崩壊地自動検出手法.「AI・データサイエンス論文集」, 2 巻 J2 号 p.223-231

河村圭・中村優志・若月強・佐村俊和（2018）深層学習を用いた空中写真からの土砂移動範囲自動検出に関する研究.「土木学会論文集 F3（土木情報学）」, 74 巻 2 号 p.I_132-I_143

平成 30 年(2018 年)北海道胆振東部地震に関する情報「国土地理院」, https://www.gsi.go.jp/BOUSAI/H30-hokkaidoiburi-east-earthquake-index.html

基盤地図情報ダウンロードサービス「国土地理院」, https://service.gsi.go.jp/kiban/

国土数値情報ダウンロードサイト「国土交通省」, https://nlftp.mlit.go.jp/ksj/