

航空レーザ測量データの機械学習による地震被害建物の検出

千葉大学大学院融合理工学府 学生会員 ○籠嶋 彩音
千葉大学大学院工学研究院 正会員 劉 ウェン, 丸山 喜久

1. 研究背景と目的

2016年4月の熊本地震では、二度の大きな地震により建物やインフラ施設に甚大な被害が生じた。地震発生後の対応として、迅速に建物の被害状況を把握することは災害対応や復興・復旧支援の点で重要である。しかし、被害が広範囲である場合や二次災害が考えられる場合、直接足を運んで被害状況を把握することは困難である。罹災証明書は、被災者の生活再建に重要な役割を果たす。しかし、住家の被災度判定には膨大な時間を要するため、熊本地震の際には本震から1か月後の時点で申請件数の3割ほどしか交付されず、罹災証明書の発行の遅れが生じた。

そこで、対象地から離れていても迅速に広範囲を観測できるリモートセンシング技術の利用が有効である。本研究では、航空レーザ測量により本震の前後で取得した数値表層モデル(DSM)の機械学習によって地震被害建物の検出を試みる。ここで、DSMは建物や樹木を含む3次元の表層モデルを指す。

2. 分析対象データ

本研究は、熊本地震の被害の大きかった益城町を対象に行った。航空レーザ測量データには、Moya らが整備したデータを用いた(図1)。航空レーザ測量は前震後(2016年4月15日15:00~17:00)と、本震後(2016年4月23日10:00~12:00)に行われた。国土地理院の基盤地図情報²⁾の2015年10月もしくは2016年4月に更新された建物の外周線を用いて、建物の位置を特定した。建物の被災度には、罹災証明書の発行のために行われた被害認定調査の結果を用いる。地震動の強さは、松岡³⁾が推定した250mメッシュの地震動分布を用いた。

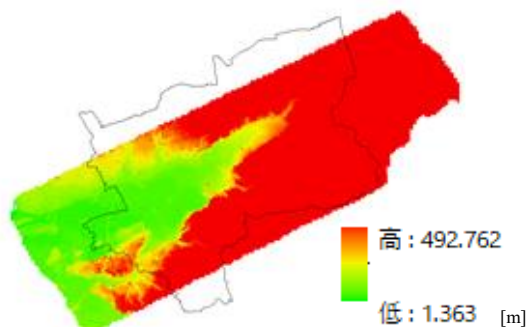


図2 Moya らの整備した航空レーザ測量データ

3. データセットと学習条件

(1) DSM データの使用方法

本研究では本震前後での建物の高さの変化に注目し、本震前後のDSMの差分のデータを用いた。図2のようなDSMデータを持ったラスターデータに対し、GIS上で50cmごとにポイントを生成する。建物の外周線の内部および外周線から1m以内の範囲にあるポイントをその建物のDSMデータとした。

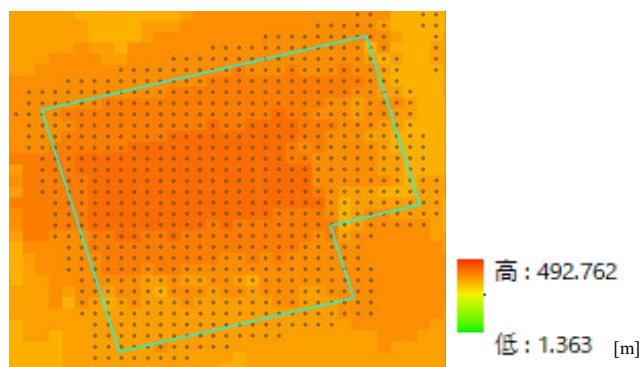


図1 ポイントに変換したDSMの例

(2) データセット

被害認定調査による建物の被災度とDSMの本震前後の差分が対応付けられたデータに対し、被害認定調査で全壊判定がなされた建物をクラス1、それ以外の判定がなされた建物をクラス0とした。建物の被災度ごとのデータ数の内訳を表1に示す。12,263件の全データを訓練データとテストデータに7:3の割合で分割し、機械学習を行った。説明変数には、本震前後のDSMの差分の平均・分散・最大値・最小値、地震動分布として最大地表速度(PGV)を用いている。

表1 建物の被災度ごとの内訳

被災度判定	件数
無被害	943
一部損壊	4879
半壊	2178
大規模半壊	714
全壊	3549
合計	12263

Key words 熊本地震, 地理情報システム(GIS), ランダムフォレスト, 被害認定調査, 航空レーザ測量

連絡先: 〒263-8522 千葉県千葉市稲毛区弥生町 1-33

千葉大学大学院融合理工学府地球環境科学専攻都市環境システムコース TEL 043-290-3555

(3) 機械学習の手法

機械学習のアルゴリズムにはランダムフォレストを用いた。ランダムフォレストとは、複数の決定木を統合させて行うアンサンブル学習である。また、汎化性能を向上させるため、訓練データに対して5分割のクロスバリデーションを行い、ハイパーパラメータの調整にはグリッドサーチを行った。

全壊のデータが少なくデータに偏りがあるため、学習を二回行いアンダーサンプリングした。本研究のアンダーサンプリングでは、一回目の学習で85%以上の確率でクラス0(その他)に分類されたデータを訓練データから除き、二回目の訓練データとした。本研究では、一回目の学習で構築した正答率もしくはF値が最大となるモデル1・2と、学習を二回行い正答率もしくはF値が最大となるモデル3・4の分類精度を比較する。

4. 学習結果

学習条件を変えたモデル1～4を用いてテストデータを分類した結果を表2、表3に示す。表2ではテストデータの判別結果を混同行列に表している。また、表3では学習結果から得られた精度指標を示している。

表2 モデル別のテストデータの混同行列

モデル1		予測結果	
学習1回・正答率最大		その他	全壊
実際の結果	その他	2319	275
	全壊	618	467
モデル2		予測結果	
学習1回・F値最大		その他	全壊
実際の結果	その他	1886	708
	全壊	323	762
モデル3		予測結果	
学習2回・正答率最大		その他	全壊
実際の結果	その他	2318	276
	全壊	641	444
モデル4		予測結果	
学習2回・F値最大		その他	全壊
実際の結果	その他	1936	658
	全壊	361	724

表3 モデル別の精度指標の比較

	モデル1	モデル2	モデル3	モデル4
正答率	0.76	0.72	0.75	0.72
適合率	0.63	0.52	0.62	0.52
再現率	0.43	0.70	0.41	0.67
F値	0.51	0.60	0.49	0.59

モデル1・3は、モデル2・4に比べその他の判別は比較的できてきているものの、全壊の誤判別は多かった。そのため、適合率は高く再現率は低い結果となった。全壊のデータの方が少ないため、全壊の建物の判別精度が良くないことが正答率に与える影響は小さい。そのため、その他の建物の判別精度が良くなったと考えられる。モデル2・4は、全壊の見逃しが少ないため再現率が高い。防災分野では、全壊の見逃しを減らすことは重要な課題である。モデル2は、適合率と再現率の調和平均であるF値の値が最も高く、釣り合いの取れたモデルであると考えている。モデル3・4は二回の学習を行ったモデルであるが、モデル1・2に劣る精度となった。アンダーサンプリングの効果よりも、学習データが減ってしまったことに要因があると考えられる。

5. 結論

本研究では、2016年熊本地震における益城町を対象として、航空レーザ測量データの機械学習による地震被害建物の検出のモデルを構築することを試みた。どのモデルも正答率は70%を超えていたが、再現率を保ちつつ、適合率とバランスをとることが困難であった。今後は、今回用いた特徴量のうち何が一番有効であったかを分析するとともに、新たな特徴量を追加することで予測精度向上を目指す。

謝辞:本研究は、首都圏を中心としたレジリエンス総合力向上プロジェクトの一環として行ったものである。記して、謝意を表する。

参考文献

- 1) L. Moya, F. Yamazaki, W. Liu and M. Yamada: Detection of collapsed buildings from lidar data due to the 2016 Kumamoto earthquake in Japan, Natural Hazards and Earth System Sciences, 18, pp.65-78, 2018.
- 2) 国土地理院: 基盤地図情報, <https://fgd.gsi.go.jp/download/menu.php>
- 3) 松岡昌志: 強震観測記録に基づく2016年熊本地震の地震動分布の推定, 地域安全学会梗概集, No.42, pp.23-26, 2018.