

機械学習による事後画像のみを用いた津波被災地の建物被害の抽出

東北大学 学生会員 ○千葉 周作
 東北大学災害科学国際研究所 正 会 員 ADRIANO Bruno
 東北大学大学院 学生会員 BAI Yanbing
 東北大学災害科学国際研究所 正 会 員 越村 俊一

1. 序論

合成開口レーダ(Synthetic Aperture Radar: SAR)を用いた地震・津波災害の被害把握は近年発達を遂げており、中でも機械学習を利用した研究が注目されている。2011年東北地方太平洋沖地震においても、被災前後のSAR画像から津波の浸水域分類モデルを自動構築し、適用が試みられた(郷右近ら,2013)。しかし、被災前後で土地利用状況が変化した場合には当該手法の適用が不適切となる。本研究では建物被害を対象に、被災後画像のみを用いて抽出を行うことで、SARによる災害被害把握の適用範囲を広げることが目的である。

2. 使用データ

本研究において、研究対象としているのは2011年3月11日に発生した2011年東北地方太平洋沖地震による津波被害である。調査対象域は宮城県仙台市若林区から亘理町にかけての太平洋沿岸部、解析に用いた建物数は15830棟である。SAR画像はTerraSAR-Xによって撮影(撮影日:2011年3月12日)されたものを、航空写真は国土地理院によって被災前後に撮影(撮影日:2010年2月10日, 2010年8月10日, 2011年3月14日)されたものを用いた。

3. 解析手法

(1) 前処理

まずSAR画像上のスペックルノイズ処理のために、Leeフィルタを施した。その後、GLCMフィルタによってSAR画像を属性・角度別に24種類に分類した。GLCMフィルタ処理によって、SAR画像の特徴量を多角的に観測できるようにした。そしてArcGISを用いて処理後の各SAR画像に建物被災データを合わせ、建物輪郭データ内の後方散乱係数を算出した。

(2) トレーニングデータ及びテストデータの作成

テストデータは使用データの項で述べた15830棟の建

表-1 建物被災区分(統合後)

建物被災区分(統合後)	
被災区分	使用被害建物数(棟)
1.流失	8106
2.半壊	6978
3.被災なし	746

物データを用いた、なお、本研究ではSAR画像による被災分類の明確化のため、建物被災分類を国土交通省都市局による7段階(流失, 全壊, 全壊2, 大規模半壊, 半壊, 一部損壊, 被災なし)から、表-1に示す通り3段階(流失, 半壊, 被災なし)に統合した。表-1では、元の被災区分における流失, 全壊, 全壊2を流失, 大規模半壊, 半壊, 一部損壊を半壊としている。

テストデータを作成した後、トレーニングデータを無作為抽出と目視判読、2種類の 방법으로作成した。無作為抽出では、各被災区分から5%ずつ(流失:410棟, 半壊:349棟, 被災なし:37棟)計791棟の建物被災データを抽出しトレーニングデータを作成した。目視判読では、調査対象域における被災前後の航空写真を見比べ、被災程度を判定しトレーニングデータを作成した。各被災区分における使用建物数は無作為抽出と同一である。

(3) Wekaを用いた機械学習による解析

本研究では、機械学習ツールとしてフリーソフトウェアのWekaを使用した。また、分類器はRandomForestである。RandomForestは分類設計の自由度が高く、対象数が膨大な場合でも高速で処理を行えるという利点を有する(河合ら,2010)。

4. 解析結果

(1) 無作為抽出によるトレーニングモデルを用いた解析結果

無作為抽出によってトレーニングデータを作成した場合の精度結果は表-2のようになった。流失区分について

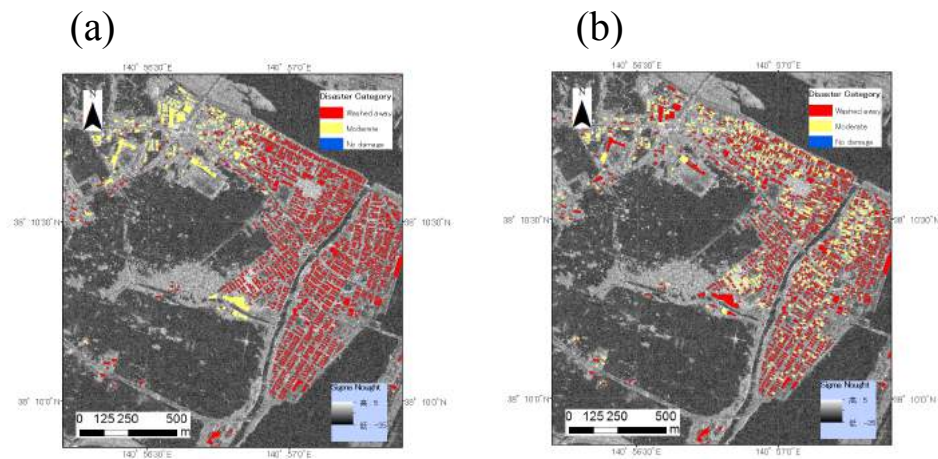


図-1 ダメージマッピング (a)GTD (b)RandomForest

表-2 無作為抽出によるトレーニングデータを用いた解析結果

		RandomForest			
		Washed away	Moderate	No damage	P.A.(%)
GTD	Washed away	5516	2588	2	68%
	Moderate	2613	4360	5	63%
	No damage	338	370	38	5%
	U.A.(%)	65%	60%	84%	
		Overall Accuracy=63%			

はP.A.が68%の精度で検出されている。また、半壊区分についてはP.A.が63%であり、流失区分と比較してやや精度が低く出た。被災なし区分については検出自体がほとんどなされていない結果となったがU.A.が84%を示し、検出分については高精度であった。解析結果は図-1のようにダメージマッピングされた。

(2) 目視判読によるトレーニングデータを用いた解析結果

目視判読によってトレーニングデータを作成した場合の精度結果は表-3のようになった。無作為抽出の場合に比べ、流失区分の検出数が減少し半壊区分の検出数が増加した。しかし、流失区分のU.A.が69%と高く、半壊区分のU.A.が56%と低くなった。被災なし区分については無作為抽出の場合とほとんど変化が見られなかった。

5. 考察

表-3の結果が表-2に比べ流失区分のU.A.が高く、半壊区分のU.A.が低く出た原因としては、目視判読で流失区分の建物を抽出した際に、元の被災区分における流失区分のものを集中的に選択しており、全壊、全壊2に該当する建物をほとんど選択していなかったことが挙げられる。

被災なし区分の建物がほとんど検出されなかった原因としては、第一に津波による軽度の建物被害が、壁面方向に集中しており、SAR衛星からは捉え辛いことが挙げられる。また、機械学習を用いた解析においてトレーニングデータに用いた建物数が少なかったことも検出数の過少化を引き起こしていると考えられる。しかし、トレー

表-3 目視判読によるトレーニングモデルを用いた解析結果

		RandomForest			
		Washed away	Moderate	No damage	P.A.(%)
GTD	Washed away	4600	3502	4	57%
	Moderate	1840	5135	3	74%
	No damage	235	473	38	5%
	U.A.(%)	69%	56%	84%	
		Overall Accuracy=62%			

ニングデータにおける被災なし区分の建物データの割合を大きくした場合、被害検出モデルの構築において過学習が発生し、著しく過剰に検出される事態につながる可能性がある。

6. 結論

本研究では、機械学習を用いることで2011年東北地方太平洋沖地震における建物の津波被害を約60%の精度で検出することができた。しかし、トレーニングデータにおける各被災区分の建物の割合により、検出結果に大きな差が生じる可能性も示唆された。従って、トレーニングデータの作成の際には、用いる各データの数及び割合について考慮が必要であると結論付けられる。

参 考 文 献

- 郷右近ら (2013): TerraSAR-X画像の機械学習による津波被災地の自動検出, 土木学会論文集B2(海岸工学) Vol. 69, No. 2, L1441-L1445
- 堺ら (2013): TerraSAR-X強度画像の変化に着目した津波被災地の建物被害程度の把握, 土木学会論文集B2(海岸工学) Vol. 69, No. 2, L1456-L1460
- リュウら (2012): 高解像度SAR画像を用いた東北地方太平洋沖地震における津波湛水域と建物被害の抽出, 日本地震工学会論文集 第12巻, 第6号 (特集号), pp.72-85
- Matsuoka, M. and Nojima, N. (2010): Building Damage Estimation by Integration of Seismic Intensity Information and Satellite L-band SAR Imagery, Remote Sensing 2 (9), pp.2111-2126
- 河合ら (2010): 不均衡データからのランダムフォレストを利用した高速高次特徴抽出, 映像情報メディア学会誌 Vol.64, No.12 pp.1951-1955