

TerraSAR-X 画像の機械学習による津波被災地の自動検出

Machine Learning Based Method for Detecting Tsunami Devastated Area Using TerraSAR-X Data

郷右近英臣¹・Joachim POST²・Enrico STEIN³・Sandro MARTINIS⁴
André TWELE⁵・Matthias MÜCK⁶・越村俊一⁷

Hideomi GOKON, Joachim POST, Enrico STEIN, Sandro MARTINIS,
André TWELE, Matthias MÜCK and Shunichi KOSHIMURA

A method for rapid detection of tsunami devastated areas using multi-temporal TerraSAR-X data is proposed. To develop the method, machine learning algorithm, a branch of artificial intelligence (AI), is applied. We focus on the multiple bounce reflection which is a specific feature on Synthetic Aperture Radar (SAR) data to estimate building devastated areas. Finally, classifiers which enable automated classifications of damage patterns intopredicted damage classes were built. The evaluation of the model was conducted through cross-validation and the best accuracy was obtained as 89.2 %.

1. 序論

2011年東北地方太平洋沖地震津波のような大規模地震津波災害の発生直後には、被災地との情報通信手段の途絶により、被害の全容把握が遅れ、災害救援活動に関わる全ての意志決定が後手にまわることがある(情報空白期)。この情報空白期において、被害を把握するための有効な手段は人工衛星や航空機等による緊急観測であり、特に昼夜・天候の条件に関わらず被災地を観測できる合成開口レーダ(Synthetic Aperture Radar: SAR)は、災害時の広域被害把握に有用である。

SAR画像による地震・津波被災地の建物被害把握に関する研究事例には多くの蓄積がある。例えば、リュウら(2012)は津波被災前後の高分解能合成開口レーダ衛星画像から得られる後方散乱係数の差分と相関係数からなる合成変化量、標高モデル、建物輪郭データを用いた統合解析による津波被災地の流失建物の被害推計を実施した。

本研究では、2011年東北地方太平洋沖地震津波の被災地において、被災前後の高分解能合成開口レーダ衛星画像(TerraSAR-X)を用いた新しい建物被害領域自動検出モデルを提案する。

2. 研究手法

(1) 解析に使用したデータ

本研究では、津波により甚大な被害を受けた宮城県名取市閉上を解析対象とする(図-1(A))。閉上地区の中でも、ほとんどの建物が流失した領域(図-1(A)(a))、流失建物と残存建物が混在する領域(図-1(A)(b))、残存建物のみが立地する領域(図-1(A)(c))の3領域を選択した。建物被害の抽出には、被災前の画像として2010年10月21日、被災後の画像として2011年3月13日に被災地を捉えた分解能約3mのStripMapモードのTerraSAR-X衛星画像(以下、TSX画像と記述する)を使用した(図-1(B),(C))。また、建物被害領域自動検出モデルの検証に使用するために、国土地理院が2011年3月12日に撮影した分解能約80cmの航空写真(国土地理院, 2011-04-01参照)とゼンリン住宅電子地図(Zmap-TOWN II)を使用した(株式会社ゼンリン, 2011-04-01参照)。

(2) 被害領域推定の原理

建物等の人工構造物はSAR画像上で、カージナル効果により高い後方散乱係数を示す。一方、津波により建物が流失し、基礎部のみが残った建物は、カージナル効果の減少により後方散乱係数が減少する。本研究ではその特性に着目し、被災前後のTSX画像上で、カージナル効果により後方散乱係数が高い値を示す領域を閾値処理(詳細は後述)により抽出し、その被災前後の面積変化

1	学生会員	東北大学大学院工学研究科
2	Ph.D	German Aerospace Center, Earth Observation Center, German Remote Sensing Data Center
3		German Aerospace Center, Earth Observation Center, German Remote Sensing Data Center
4	Ph.D	German Aerospace Center, Earth Observation
5		German Aerospace Center, Earth Observation Center, German Remote Sensing Data Center
6		German Aerospace Center, Earth Observation Center, German Remote Sensing Data Center
7	正会員 博(工)	東北大学教授災害科学国際研究所

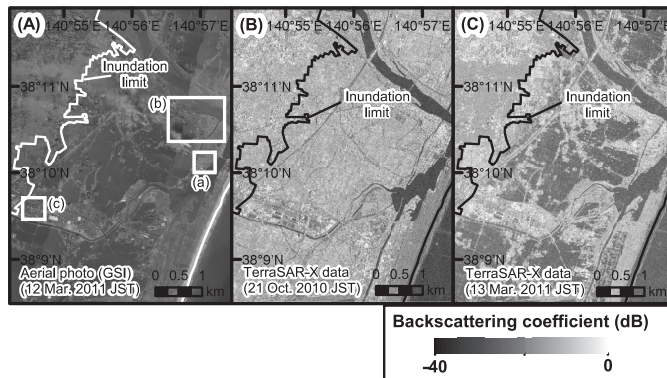


図-1 (A) 研究対象領域, (B) 被災前のTerraSAR-X衛星画像, (C) 被災後のTerraSAR-X衛星画像

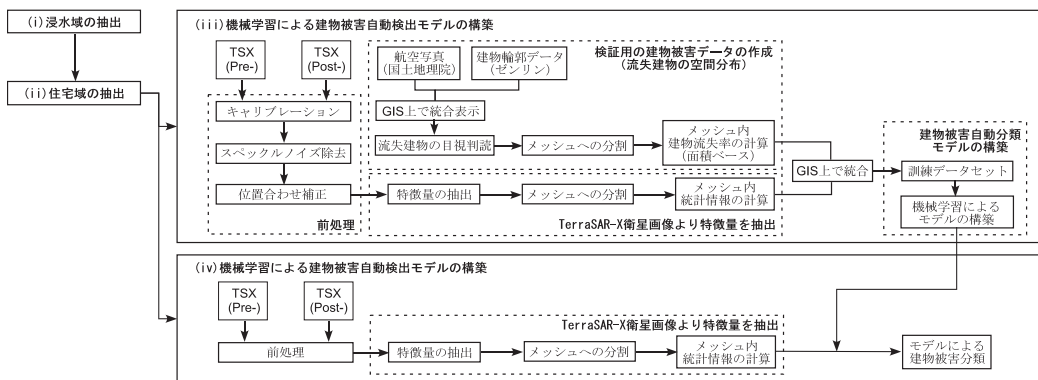


図-2 本研究の解析フロー

に着目することにより、建物被害領域の検出を行う。

(3) 浸水域と住宅域の抽出

TSX 画像による建物被害領域の検出は、まず「浸水域の抽出」、次に「浸水域内における住宅域の抽出」、最後に「推定住宅域内における建物被害分類」と、3段階の処理により行う。既往研究としては、浸水域の抽出についてMartinisら(2009)の手法、住宅域の抽出についてはEschら(2010)の手法に実績がある。いずれもTSX画像の画像処理に基づく手法であり、Martinisら(2009)によると、浸水域抽出の総合精度は95%程度、Eschら(2010)によると、住宅域抽出の総合精度は90%程度である。また、越村ら(2012)では光学衛星画像解析による浸水域抽出の手法も提案されており、本研究では、Martinisら(2009)や越村ら(2012)、Eschら(2010)の研究手法を適用して浸水域と住宅域が抽出できると想定し、建物被害域の抽出を目標として解析を進めていく。

(4) 機械学習

建物被害領域自動検出モデルを構築するために、機械学習を適用する。機械学習とは、人間が自然に行っている学習能力と同様の機能をコンピュータで実現しようとする技術・手法のことであり、(1) Trainingデータにより識別に必要な複数のパターンを学習しモデルを構築す

る段階と、(2) その構築したモデルに基づき新しい入力データを分類する段階に分けられる。

本研究では、被災前後のTSX画像と既存の機械学習アルゴリズムに基づき、建物流失率が大きい領域を自動検出するモデルを構築する。モデル構築には、オープンソースのデータマイニングツール WEKA (Ver.3.6.9) を使用した (Hallら, 2009)。

(5) 研究の手順

a) 前処理

本研究の解析フローを図-2に示す。まず具体的な解析に移る前に、株式会社バスコ(2011-04-30参照)の浸水域情報とゼンリン電子地図をGIS上で重ね合わせることで、津波により浸水した住宅域を抽出する。

次に、被災前後のTSX画像にキャリブレーションを施し、画素値から後方散乱係数 (dB) への変換を行う。Gamma-Mapフィルタによりスペckルノイズを除去する。最後に、被災前後の画像の位置合わせ補正を行う。

次に被災前後のTSX画像上で、カージナル効果により高い後方散乱係数を示す領域を抽出する。本研究では簡単のため、この領域を C-areaと呼ぶことにする。まず、抽出に使用する閾値を決定するために、図-1 (B) の被災前のTSX画像全体より、後方散乱係数とその頻度分布の

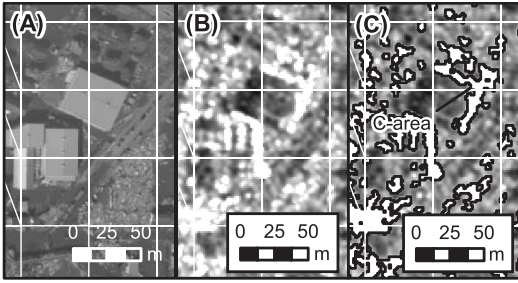


図-3 特徴量の抽出例。(A) 国土地理院航空写真、(B) 被災後 TSX 画像、(C) 抽出した C-area

関係を示すヒストグラムを作成する。そして、その累積分布関数の上位10%の地点を閾値と決定する。その結果、後方散乱係数が-4.1dBよりも大きくなる領域を C-area として抽出した (図-3)。

次に、検証に使用するために、流失建物の空間分布の解析を行う。まず、震災後に撮影された国土地理院の航空写真 (国土地理院, 2011-04-01 参照) とゼンリン電子地図を GIS 上で統合表示し、航空写真と建物輪郭データの位置合わせをする。そして、津波浸水域の情報 (たとえば株式会社パスコ, 2011-04-30 参照) に基づき、津波により浸水被害を受けた建物のみを抽出する。最後に、直下視の航空写真を用いて建物屋根の「有り・無し」に着目し、目視による流失建物の被害判読を行った (Gokon・Koshimura, 2012)。

b) C-areaの面積変化率と建物流失率の比較

被災前後の C-area の面積変化と建物流失率の関係を比較・検討する。まず対象領域を、平均的な建物一棟とそれに関係する C-area が十分納まる大きさの 50 m × 50 m の区画に分割し、区画毎に C-area の面積変化率 (C_r) と建物流失率 (P_r) を以下のように定義し、算出する (図-4)。

$$C_r = 1 - \frac{\text{被災前後のC-areaの重複面積}}{\text{被災前のC-areaの面積}} \quad \dots (1)$$

$$P_r = \frac{\text{各区画内の流失建物の総面積}}{\text{各区画内の被災前建物の総面積}} \quad \dots (2)$$

次に、これらの算出した値を区画毎に対応付け、その比較・検討を行う。そして、 P_r の値により表-1 のような4つの被害ランクを設け、“Damaged”, もしくは “Undestroyed” の分類を行う。最後に、各ランク毎に C_r の値と、 P_r の被害発生確率の関係をプロットし、対数正規分布を仮定した以下の式により回帰分析を行う。

$$P_d(C_r) = \int_{-\infty}^{C_r} \frac{1}{\sqrt{2\pi}\xi t} \exp\left(-\frac{(\ln t - \lambda)^2}{2\xi^2}\right) dt \quad \dots (3)$$

ここで、 λ , ξ はそれぞれ $\ln C_r$ の平均値と標準偏差であ

表-1 被害ランクの定義と回帰に使用した変数

Rank	Damaged	Undestroyed	λ	ξ	R^2
A	$P_r \geq 0.8$	$P_r < 0.8$	-0.443	0.345	0.95
B	$P_r \geq 0.6$	$P_r < 0.6$	-0.503	0.324	0.90
C	$P_r \geq 0.4$	$P_r < 0.4$	-0.543	0.297	0.86
D	$P_r \geq 0.2$	$P_r < 0.2$	-0.606	0.313	0.79

る。関数の構築に使用したパラメータと回帰式の標本値に対する当てはまりの良さを示す基準となる決定係数を表-1に、得られた回帰曲線を図-5に示す。表-1では決定係数が0.79から0.95に分布し、いずれも高い相関を示した。

3. 機械学習による建物被害領域自動検出モデルの構築

(1) 建物被害自動検出モデルの構築

様々な分野において機械学習アルゴリズムは多数提案・利用されているが、本研究では試行錯誤を行い、中でも特に高い分類精度が得られた Logistic function法 (Le Cessie・Van Houwelingen, 1992), SMO法 (Platt, 1999), Decision tree法 (Quinlan, 1993) を採用する。Logistic function法は線形識別関数の一つであり、線形関数をロジスティック・シグモイド関数により非線形に変換できる。SMO法 (Sequential Minimal Optimization) は2クラス問題の線形識別関数法である Support Vector Machine を拡張し、学習モデル構築と分類の計算負荷を下げたものである。Decision tree法では、Quinlan (1993) により提案された C4.5 という手法に基づき、パラメータの最尤推定を実施する。

全ての解析区画内の P_r の平均値は約0.6である。そこで本研究では、試験的にこの値を被害分類の閾値とし、 P_r が0.6を越える場合を “Damaged”, それ以下を “Undestroyed” と定義し、TSX 画像により算出したパラメータに基づき建物被害の自動分類モデルを構築する。入力するパラメータとして、被災前後の TSX 画像より得られる標準偏差, 平均値, 最大値など複数の変数を検討した。具体的な手順としては、各候補となる変数と “Damaged”, “Undestroyed” の関係を示す頻度分布図を作成し、変数の変化に伴い “Damaged”, “Undestroyed” の頻度分布を良好に分ける変数を試行錯誤的に検討した。そしてそれらの検討に基づき、“ C_r ” と 「被災後の C-area の解析区画内面積」 の二変数を、入力パラメータとして採用することにした。最後に、以上の入力パラメータ (説明変数) と被害に関する情報 (目的変数) を 288 のメッシュデータに属性として付与し、これらのメッシュデータを教師データとして Logistic function法, SMO法, Decision tree法に基づき機械に学習させることにより、各メッシュ毎に算出した “ C_r ” と 「被災後の C-area の解析区画内面積」 を説明変数と

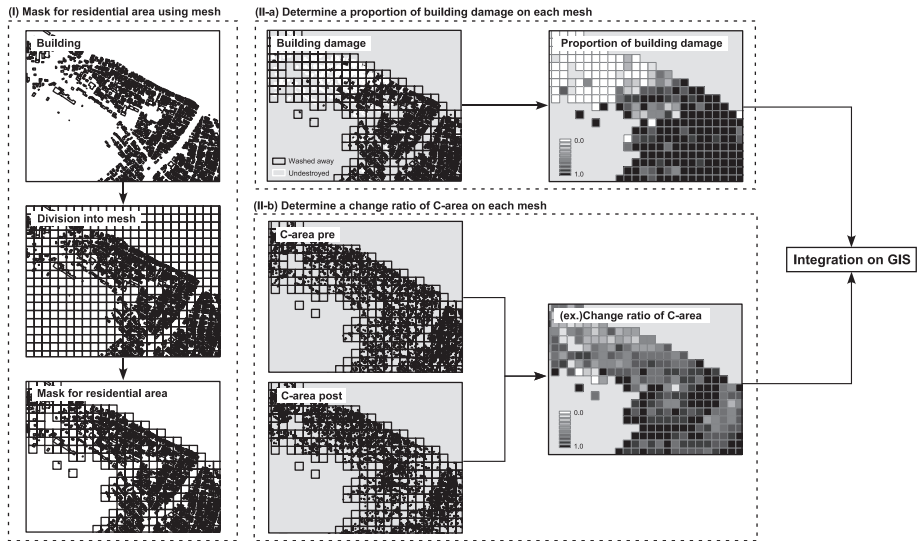


図-4 各区画内の建物流失率・C-areaの面積変化率の算出

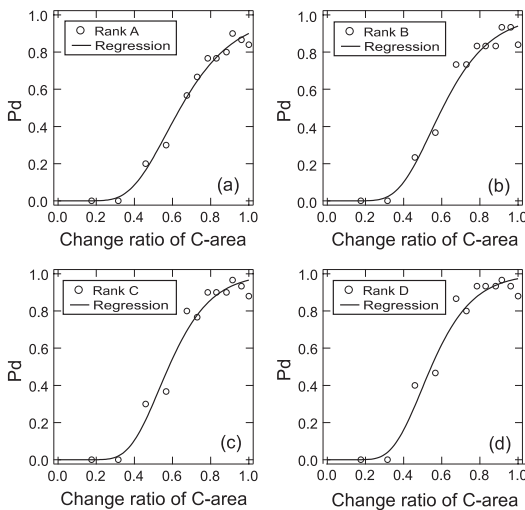


図-5 多重反射領域の面積変化率 ((a) RankA, (b) RankB, (c) RankC, (d) RankD)

し, “Damaged”, “Undestroyed”を目的変数とする建物被害自動分類モデルを構築した。

(2) 構築した建物被害領域自動検出モデルの検証

モデルの分類精度の推定は, 交差検証法により実施した(表-2)。その結果, いずれのモデルにおいても90%近い総合精度を得られ, 特にDecision tree法に基づくモデルでは, Kappa Statisticが0.77と高い値を得た。Kappa Statisticとはモデルによる分類結果と実際の被害分類の一致度を示す指標であり, 0から1の値を取る。Landis・Koch(1977)によるとKappa Statisticが0.6以上であれば十分高い(Substantial)と判断される。本研究の対象領域(図-1(A)(a)-(c))における分類精度は, それぞれ

表-2 建物被害分類モデルの精度検証

Constructed Model	Damaged (P.A.%)	Damaged (U.A.%)	Survived (P.A.%)	Survived (U.A.%)	Overall (%)
Logistic	94.3	87.7	79.8	90.1	88.5
SMO	95.4	87.8	79.8	91.9	89.2
Decision tree	97.7	86.3	76.3	95.6	89.2

(a) : 91.6%, (b) : 87.6%, (c) : 100%であり, 流失建物と残存建物が混在する領域では若干分類精度が落ちる事がわかった。その一方, 殆どの建物が流失した場所や, 多くの建物が残存した場所では, 本モデルにより良好に建物被害の分類が行えることがわかった。

次に, 検証用データでは“Damaged”, モデルによる分類結果が“Undestroyed”と誤分類された場合には, False Positive Error (F.P.), 検証用データでは“Undestroyed”, モデルによる分類結果が“Damaged”と誤分類された場合には, False Negative Error (F.N.)とし, 最も高い分類精度を示した Decision tree法による誤分類の要因について, 整理・考察を行った。その結果, 本分類におけるF.P.の総件数は27, F.N.の総件数は4であった。F.P.は, 建物の流失以外の要因により後方散乱係数が減少した場合に生じる誤分類であり, 被災前後の光学衛星画像とTSX画像を比較しながら以下のように考察結果をまとめた。(1) 建物以外の構造物や船舶, 車, 塀, 畑に張られているビニール等の流失により, 後方散乱係数が減少, (2) 建物周りや建物間の瓦礫堆積による建物側面のカージナル効果の減少により, 後方散乱係数が減少, (3) 建物と, カージナル効果が発生した領域が解析区画の境界により切れたために誤分類が発生。その一方, 建物が流

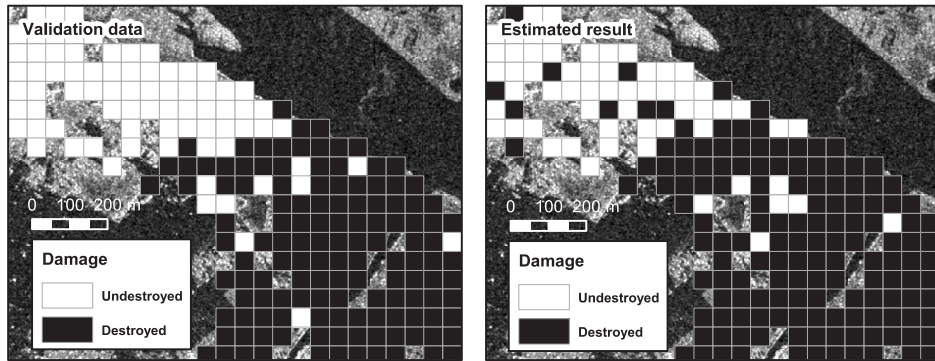


図-6 建物被害領域の検出結果 (左：検証用データ，右：分類結果)

失したにも関わらず後方散乱係数が減少しなかった事による誤分類F.N.の要因としては、建物が流失したものの、他の建物が同じ場所に流されて来たため、後方散乱係数が減少しなかったという例が確認された。

4. 結論

本研究では、被災前後の2時期のTerraSAR-X衛星画像と機械学習による新しい建物被害領域自動検出モデルを提案した。得られた結論を以下に列挙する。

被災前後の SAR 画像上の後方散乱係数の高い領域の面積変化率と、建物流失率の関係を回帰式により分析した結果、両者に高い相関が確認され、後方散乱係数の高い領域の面積変化率が津波被災地の建物被害を推計する上で有用な指標である事がわかった

SAR 画像上の後方散乱係数の高い領域の「被災前後の面積変化率」と「被災後の面積」を入力データとする、機械学習による建物被害自動検出モデルを構築し、その活用について提案した。

交差検証法によりモデルの評価を行った結果、いずれにおいても高い分類精度が確認され、本手法が建物被害領域の即時探索に有効な手法であることがわかった。本研究で得られたモデルにより、実際に被害領域の自動検出を行い、その誤分類の要因について考察を行った。

謝辞：TerraSAR-X 画像と津波浸水域データは株式会社パスコとの共同研究の一環として提供を受けた。本研究の一部は東北大学災害科学国際研究所プロジェクト経費、科学研究費補助金（課題番号：25242035, 24241059）、JSPS 特別研究員奨励費（課題番号：24・5839）の補助を受けて実施された。ここに記して謝意を表す。

参考文献

- 株式会社ゼンリン：住宅地図データベース (Zmap-TOWN II) (オンライン), <http://www.zenrin.co.jp/product/gis/zmap/zmaptown.html>, 参照 2011-04-01
- 株式会社パスコ：平成 23 年 (2011 年) 東日本大震災に関する情報 (オンライン), http://www.pasco.co.jp/disaster_info/110-311/, 参照 2011-04-30
- 国土地理院：平成 23 年 (2011 年) 東北地方太平洋沖地震正射画像データ (オルソ画像) (オンライン), <http://saigai.gsi.go.jp/-h23taiheiyo-zort/index.html>, 参照 2011-04-01
- 越村俊一・郷右近英臣・福岡巧巳・林里美 (2012)：リモートセンシングによる 2011 年東北地方太平洋沖地震津波の広域被害把握, 日本地震工学会論文集, 第 12 巻, 第 6 号 (特集号), pp. 50 -62.
- リュウ・ウェン・山崎文雄・郷右近英臣・越村俊一 (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浸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.
- Esch, T., M. Thiel, A. Schenk, A. Roth, A. Müller, and S. Dech (2010)：Delineation of urban footprints from TerraSAR-X data by analyzing speckle characteristics and intensity information. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(2), pp. 905 -916.
- Gokon, H. and S. Koshimura (2012)：Mapping of building damage of the 2011 Tohoku earthquake tsunami in Miyagi Prefecture, *Coastal Engineering Journal*, Vol. 54, No. 1, 125006-1-12.
- Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reute-mann, Ian H. Witten (2009)：The WEKA data mining software: An Update; *ACM SIGKDD Explorations Newsletter*, Vol. 11, Issue 1, pp. 10 -18.
- Landis, J. R. and G. G. Koch (1977)：The measurement of observer agreement for categorical data, *biometrics*, Vol.33, No. 1, pp. 159 -174.
- Le Cessie, S. and J. C. Van Houwelingen (1992)：Ridge estimators in logistic regression. *Applied statistics*, pp. 191 -201.
- Martinis, S., A. Twele and S. Voigt (2009)：Towards operational near real-time .ood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data, *Nat. Hazards Earth Syst. Sci.*, 9, pp. 303 -314.
- Platt, J. C. (1999)：12 fast training of support vector machines using sequential minimal optimization, *Advances in kernel methods*, pp. 185 -208.
- Quinlan, J. R. (1993)：C4.5:Programs for Machine Learning. San Mateo, Calif.: Morgan kaufmann, 270 p.