

深層学習を用いた被災地衛星画像の被覆分類

Land cover classification of satellite images of disaster areas using deep learning

吉原 篤¹⁾
Atsushi Yoshihara
滝口 哲也²⁾
Tetsuya Takiguchi
有木 康雄³⁾
Yasuo Ariki

概要：地震や津波といった大規模災害が発生した場合、安全な避難・救援ルートの確保や復興対策を検討するため、広域の情報を迅速に把握する必要がある。このタスクを実現するため、近年ではリモートセンシングで取得した衛星画像を用いた被覆分類に関する研究が発展している。

これらの研究では、深層学習を用いて様々な分類手法が提案されている。中でも、Convolutional Neural Networks (CNN) を応用した研究が盛んに行われており、他の手法と比べて高い性能を示している。しかし、従来の CNN を用いた手法では、全結合層の構造上入力サイズが固定されてしまうという点や、プーリングの繰り返しにより位置情報が失われてしまうといった問題がある。

そこで、本論文では、上述の問題を解決する手法として、全層畳み込みエンコーダ・デコーダ型モデルによる画像のセグメンテーション手法を提案する。これを被災地衛星画像に適用することで、被覆分類精度向上および被害地域の識別を目的としている。これにより、広域の情報を把握することができ、政府や地方自治体がとるべき災害応急対策の支援に繋がると考えられる。

被災前衛星画像において、既存手法および提案手法を用いて分類精度比較を行い、衛星画像に対する提案する多クラスセグメンテーションモデルの有効性を示す。また、被災後の衛星画像に対し、事前学習モデルとして被災前衛星画像で学習したパラメータを用いるかどうかの2種類の方法で学習を行うことで、セグメンテーション精度に生じる差異を検証した。この結果から、事前学習モデルを用いることでセグメンテーション精度が向上することがわかった。

キーワード：リモートセンシング, 衛星画像, 被覆分類, 深層学習, CNN, Semantic Segmentation

1. はじめに

毎年、世界中で大規模な災害が発生し、大きな被害を受けている。なかでも日本列島は、複数のプレートの上に位置しているため、世界的にも地震が多い地域である。特に、2011年に発生した東日本大震災では、東北地方の太平洋沿岸地域が甚大な被害を受けた。このような大規模災害が発生した場合、広域の情報を迅速かつ正確に把握することが求められている。

このため、リモートセンシング技術および CNN を中心とした深層学習モデルを適用する手法が注目されている。しかし、従来の手法で用いられている一般的な CNN には、全結合層の構造上入力サイズが固定されてしまうという点や、プーリングの繰り返しにより位置情報が失われてしまうといった問題がある。

そこで、本研究では、上述の問題を解決する手法として、全層畳み込みエンコーダ・デコーダ型モデルを構築し、衛星画像に対して多クラスセグメンテーションを行う手法を提案する。これを衛星画像に適用することで、被覆分類精度の向上および被災地域の識別を目的とする。これにより、広域の情報を把握することができ、政府や地方自治体がとるべき災害応急対策の支援に繋がると考えられる。

2. 実験で用いた衛星画像

(1) 衛星画像

本研究では、Geoeye-1 という人工衛星に搭載されたマルチスペクトルセンサから得られる高解像度衛星画像を用いた。このセンサは R、G、B、NIR の 4 バンドを有している。画像全体のサイズは 10314×10312 pixels、空間解像度は 0.5 m/pixel である。研究対象領域は、2011 年に発生した東日本大震災で津波によって最も大きな被害を受けた地域の一つである宮城県石巻市とした。図 1 に宮城県石巻市上空の衛星画像を示す。

(2) Ground Truth

被災前衛星画像の Ground Truth として Esri ジャパンが提供する「ArcGIS データコレクション詳細地図」を使用する。このデータから、Facility area (施設領域)、Road area (道路領域)、Water area (水域) の領域を抽出し、Others を加えた 4 種類のクラスに被災前衛星画像のラベル割り当てを行った。

また、被災後衛星画像の Ground Truth として、復興支援調査アーカイブ¹⁾から取得した「建物被災状況」および「インフラ被害 (道路)」を使用する。復興支援アーカイブとは、国土交通省都市局の「東日本大震災津波被災市街地復興支援調査」の成果をアーカイブ化したものである。このデータは、津波で浸水した地域の全建物を対象として行われた現地調査の結果に基づいて作成されたものであり、建物一棟単位で地理座標付きのポリゴンと被災区分とを記録したものである。被災区分は、次の 8 つが設定されている。

「1: 全壊 (流出), 2: 全壊, 3: 全壊 (1 階天井以上浸水), 4: 大規模半壊, 5: 半壊 (床上浸水), 6: 一部損壊 (床下浸水), 7: 被災なし, 99: 不明」

これら被災区分が衛星画像上でも区別できるか確認したところ、他の区分と明瞭に区別できたのは被災区分 1 のみでその他は見た目の違いが確認できなかった。そのため、本実験では被災区分 1 と判断された建物領域を Damaged facility area、その他の区分と判断された建物領域を Survived facility area として定義した。道路領域においても、「インフラ被害 (道路)」に格納されている領域を Damaged road area、その他を Survived road area と定義した。図 2 に、被災前衛星画像に対して作成した Ground Truth を示す。

赤色が Facility area、黄色が Road area、青色が Water area、黒色が Others を示している。また、図 3 に被害を受けた施設領域および道路領域を示す。



図 1. 衛星画像 (石巻市)

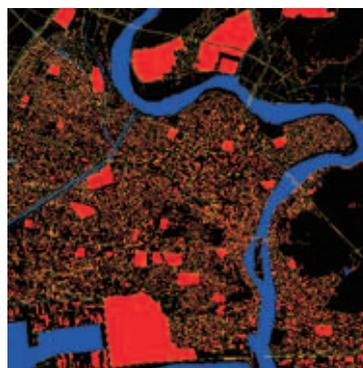


図 2. Ground Truth

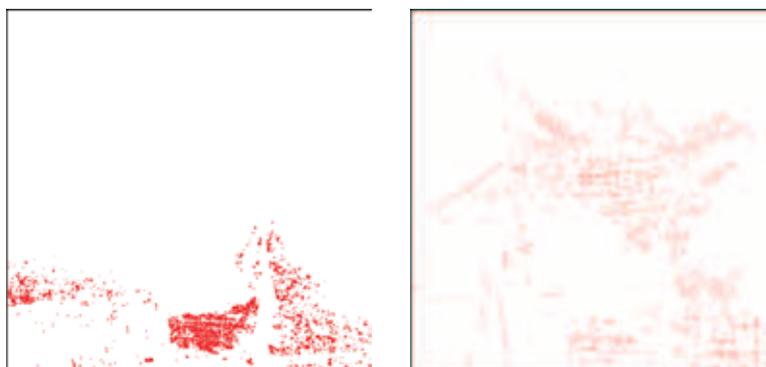


図 3. 被害施設領域 (被災区分 1) および被害道路領域

3. 提案手法

従来の CNN モデルでは、一般的に全結合層が接続されており、これにより入力画像サイズが制限されるという問題があった。これに対し、文献[1][2]では、全結合層を畳み込み層に変換することで、入力画像サイズと同じ大きさのセグメンテーションマップを出力するというモデルを構築し、任意の画像サイズの入力を許容する手法を提案している。

本研究では、文献[1][2]同様全結合層を畳み込み層に変換し、入力する衛星画像全体の多クラスセグメンテーションマップを出力するモデルを構築する。モデル構成は、セグメンテーションタスクにおいて state-of-the-art の精度を示している文献[3]の U-Net と呼ばれるモデルから着想を得て構築した。

このモデルの特徴は、低次元の特徴マップを上位層の特徴マップと結合させることで、畳み込みの繰り返しにより失われる位置情報を補う点にある。提案モデルは、文献[3]の U-Net 同様、エンコーダ・デコーダ型のモデル構造とした。

エンコーダ部は、3つの畳み込み層およびプーリング層のペアの繰り返しで構成し、各畳み込み処理後に、過学習抑制のため Batch Normalization および ReLU (Rectified Linear Unit) を適用した。

デコーダ部は、対応するエンコーダ出力との Concatenation、2つの畳み込み層および Deconvolution によるアップサンプリング処理のペアの繰り返しで構成した。

最終層は、多クラスセグメンテーションのため、特徴マップの深さはクラス数、出力サイズは入力と同じサイズとなるよう設計した。

U-Net との差異は、衛星画像を対象とした多クラスセグメンテーションモデルであること、特徴マップの深さを一定とし、その分層の深さを増やすことでより高次の特徴量を抽出できるようにした点、また、畳み込みにおけるパディング値を 1 とすることで、畳み込みの前後で特徴マップのサイズが小さくならないようにし、入力画像と同じサイズのセグメンテーション結果を出力するよう設計した。

図 4 に、提案する多クラスセグメンテーションモデルの概要図、図 5 に提案モデルの構成を示す。

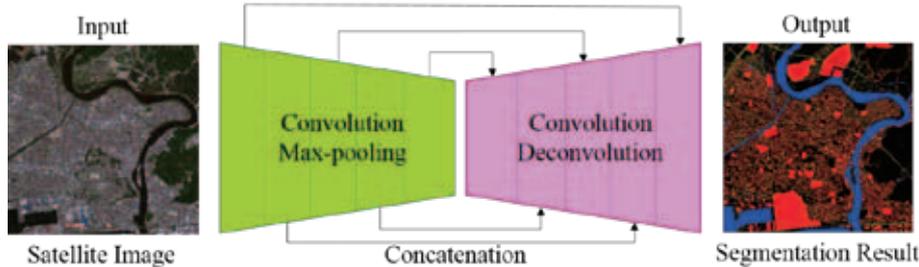


図 4. モデル概要図

Encoder (from top to bottom)					Decoder (from bottom to top)					
layer	kernel	stride	padding	output size		layer	kernel	stride	padding	output size
input				4*256*256		output				6*256*256
conv.	3*3	1*1	1	64*256*256		conv.	1*1	1*1	0	6*256*256
conv.	3*3	1*1	1	64*256*256	→	deconv.	3*3	2*2	1	64*256*256
conv.	3*3	1*1	1	64*256*256		conv.	3*3	1*1	1	64*128*128
max-pool.	2*2	2*2	0	64*128*128		conv.	3*3	1*1	1	96*128*128
conv.	3*3	1*1	1	64*128*128	→	concat.				128*128*128
conv.	3*3	1*1	1	64*128*128		deconv.	3*3	2*2	1	64*128*128
max-pool.	2*2	2*2	0	64*64*64		conv.	3*3	1*1	1	64*64*64
conv.	3*3	1*1	1	64*64*64	→	conv.	3*3	1*1	1	96*64*64
conv.	3*3	1*1	1	64*64*64		concat.				128*64*64
max-pool.	2*2	2*2	0	64*32*32		deconv.	3*3	2*2	1	64*64*64
conv.	3*3	1*1	1	64*32*32	→	conv.	3*3	1*1	1	64*32*32
conv.	3*3	1*1	1	64*32*32		conv.	3*3	1*1	1	96*32*32
conv.	3*3	1*1	1	64*32*32	→	concat.				128*32*32
max-pool.	2*2	2*2	0	64*16*16		deconv.	3*3	2*2	1	64*32*32
conv.	3*3	1*1	1	64*16*16	→	conv.	3*3	1*1	1	64*16*16
conv.	3*3	1*1	1	64*16*16		conv.	3*3	1*1	1	96*16*16
conv.	3*3	1*1	1	64*16*16	→	concat.				128*16*16
max-pool.	2*2	2*2	0	64*8*8		deconv.	3*3	2*2	1	64*16*16
						conv.	3*3	1*1	1	64*8*8
						conv.	3*3	1*1	1	64*8*8
						input				64*8*8

図 5. モデル構成

4. 被災前衛星画像に対する既存手法との分類精度比較

被災前の衛星画像に対し、既存手法 (U-Net) と提案モデルをそれぞれ適用し、セグメンテーション精度の比較を行った。

(1) 実験条件

学習用データと評価用データを区別するため図 6(a)のようにデータ分割を行った。学習時は、Training Area からランダムなパッチ領域を n 個選択し、ミニバッチ学習としてモデルへ入力する。バッチ数は実験的に得られた最適値として n=10 とした。評価も同様に行った。

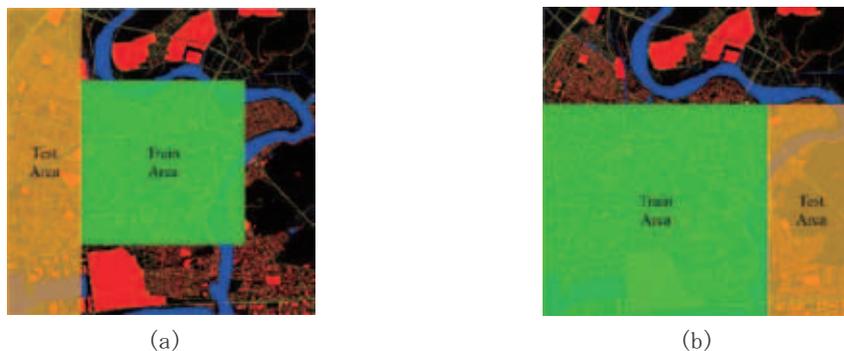


図 6. データ分割

(2) 実験結果

表 1 および表 2 に、U-Net を適用した場合の分類精度と提案モデルを適用した場合の分類精度を示す。また、図 7(a)に Ground Truth および各モデルを適用した際のセグメンテーション出力結果例を示す。上段が Ground Truth、中段が既存手法 (U-Net) の出力結果、下段が提案モデルの出力結果を示している。

表より、各クラスにおいて既存手法を上回る結果を得ることができたことわかる。特に Facility area および Road area において、既存手法よりも良い分類精度を得ることができた。

また、出力結果を見ると、提案モデルにおいてわずかではあるが既存手法よりも他クラス間の境界をうまく識別できていることがわかる。しかし、大まかなオブジェクトの位置は捉えることができているが、まだ境界識別の精度が不十分であることがわかる。これは、Ground Truth を詳細地図から作成した際、空間座標のずれが存在していたことが原因であると考えられる。

表 1. 既存手法 (U-Net)

	Facility area	Road area	Water area	Others
Recall	76.59	76.74	98.77	90.68
Precision	76.06	88.64	97.22	89.57
F-measure	76.32	82.26	97.99	90.12

表 2. 提案手法

	Facility area	Road area	Water area	Others
Recall	90.42	89.34	99.88	91.31
Precision	86.67	83.47	99.17	94.2
F-measure	88.51	86.31	99.52	92.73

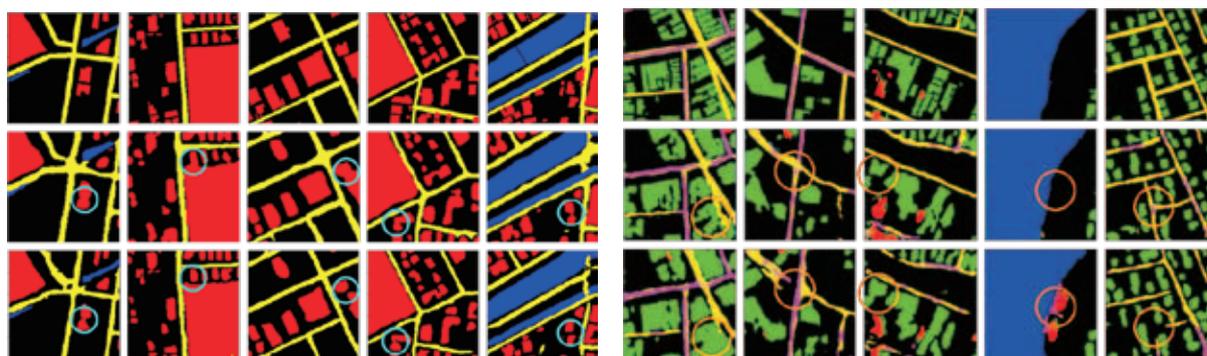


図 7. セグメンテーション結果

5. 被災後衛星画像への適用

被災後の衛星画像に対して提案モデルを適用し、セグメンテーション精度の評価を行った。ここで、被災前衛星画像で学習したネットワークパラメータを事前学習として用い、新たに被災後衛星画像から取得したデータで fine-tuning を行う学習方法 (以下、事前学習あり) と、事前学習を行わず、ランダムパラメータを用いて学習する方法 (以下、事前学習なし) の 2 種類の方法で学習を行い、セグメンテーション精度の差

異を検証した。

(1) 実験条件

前述の実験同様、学習用と評価用データを区別するため、図 6(b)のようにデータ分割を行った。ここでもミニバッチ学習を採用し、バッチ数は $n=10$ とした。

(2) 実験結果

表 3 に事前学習ありの場合の分類精度、表 4 に事前学習なしの場合の分類精度を示す。また、図 7(b)に Ground Truth およびそれぞれの方法で学習したモデルを用いてテストした際のセグメンテーション出力結果例を示す。上段が Ground Truth、中段が事前学習ありの出力結果、下段が事前学習なしの出力結果をそれぞれ示している。緑色が Survived Facility、黄色が Survived Road、青色が水域、赤色が Damaged Facility、桃色が Damaged Road、黒色が Others を示している。

表より、事前学習モデルを用いることで、被災後衛星画像に対するセグメンテーション精度が向上することがわかった。しかし、被災前画像における精度と比較すると、全体的にかなり低い精度となってしまう。また、出力結果より、全体的に他クラス間の境界識別が十分に行えていないことや、被害領域ではない領域を被害領域と判定している、あるいはこの逆となる判定をしまっていることがわかる。これは、学習に用いた被害施設領域および被害道路領域のデータ数が不足していたことが原因であると考えられる。

表 3. 事前学習あり

	Survived facility	Survived road	Water	Damaged facility	Damaged road	Others
Recall	68.05	48.18	94.04	31.46	18.18	78.53
Precision	48.98	34.69	91.56	80.66	36.31	86.75
F-measure	56.96	40.38	92.79	45.26	24.23	82.43

表 4. 事前学習なし

	Survived facility	Survived road	Water	Damaged facility	Damaged road	Others
Recall	53.72	40.95	96.26	21.33	14.28	77.58
Precision	51.93	26.56	88.07	72.26	38.9	82.6
F-measure	52.81	32.21	91.99	32.93	20.9	80.01

6. まとめ

本論文では、被災前衛星画像に対して既存手法および提案モデルを適用し、セグメンテーション精度の比較を行った。この結果から、提案モデルにより既存の手法を上回る性能を示すことができた。しかし、出力結果から境界部分の識別がまだ正確ではないことが確認できたため、今後の課題として、学習パラメータやモデル構成についてさらなる調整が必要であると考えられる。

また、被災後の衛星画像に対して提案モデルを適用し、2 種類の方法で学習を行うことでセグメンテーション精度にどのような差異が生じるか検証を行った。この結果から、被災前衛星画像で学習したパラメータを用いた場合に精度がわずかに良くなることが分かった。しかし、出力結果から、被災領域の識別はまだ精度が良くない結果となった。これは被害データの不足や Ground Truth と衛星画像の間に存在する空間座標のずれが一因であると考えられる。そのため、今後の課題として、他クラス間の境界をさらにうまく識別できるよう、学習パラメータやモデル構成の調整および学習・テストデータの分割方法についての検討が必要である。また、他領域の被災地域のデータを用いることが必要である。

参考文献

- [1]: J. Long et al., “Fully convolutional networks for semantic segmentation,” in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431-3440.
- [2]: V. Badrinarayanan et al., “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” arXiv preprint arXiv:1511.00561, 2015.
- [3]: O. Ronneberger et al., “U-net: Convolutional networks for biomedical image segmentation,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234-241.

筆者： 1) 吉原篤，システム情報学研究科，学生； 2) 滝口哲也，都市安全研究センター，教授； 3) 有木康雄，都市安全研究センター，名誉教授

Land cover classification of satellite images of disaster areas using deep learning

Atsushi Yoshihara
Tetsuya Takiguchi
Yasuo Ariki

Abstract

In the event of a large-scale disaster such as an earthquake or tsunami, it is necessary to quickly grasp the wide area information in order to secure safe evacuation / rescue routes and consider recovery measures. In order to realize this task, research on coating classification using satellite images acquired by remote sensing has been developed recently.

In these studies, various classification methods have been proposed using deep learning. Among others, Convolutional Neural Networks (CNN) are actively conducted, showing higher performance than other methods. However, with the conventional method using CNN, there is a problem that the input size is fixed due to the structure of the entire coupling layer, and positional information is lost due to repetition of pooling.

Therefore, in this paper, we propose a segmentation method based on fully convolutional encoder-decoder model as a solution to the above problem. We aim to improve land cover classification accuracy of the satellite image and to distinguish between damaged areas and areas that are not affected by satellite imagery in the disaster area. By doing this, we can grasp information on a wide area and think that it will lead to the support of disaster emergency measures to be taken by the government and local governments.

We showed the effectiveness of the proposed model in multi-class segmentation task of satellite image by comparing the classification accuracy of the conventional method with that of the proposed method for the pre-disaster satellite image. Furthermore, we applied the proposed model to the post-disaster satellite image, and two methods were used to evaluate the difference in segmentation accuracy. From this result, it was found that segmentation accuracy improves by using learned parameters of pre-disaster satellite image.

©2018 Research Center for Urban Safety and Security, Kobe University, All rights reserved.