

1.はじめに

変化点検出は、異なる時刻に撮影された同一シーン画像から変化した領域を抽出する問題である。CNNを用いた変化点検出法[1][2]は、位置ずれや環境変化に頑健であるため、マップの自動更新や土地利用調査への利用が期待されている。土地利用調査に応用する場合、変化した領域の検出だけでなく、農地や森林など過去と現在の土地の属性情報が必要である。これにより、変化した場所の内訳を把握することができる。

そこで本研究では、変化領域とオブジェクトクラスを同時に推定し、変化領域に土地の属性情報を付与することができる変化点検出法を提案する。提案手法は、セマンティックセグメンテーションと変化点検出を1つのネットワークで実現する。また、セグメンテーションモデルより抽出した特徴マップを、変化点検出モデルの入力として使用することで、両タスクの精度向上を図る。

2.変化点検出の従来法

CNNを用いた変化点検出法は、特微量ベースと学習ベースの2つのアプローチに分類できる。

特微量ベースの変化点検出法: CNNより抽出した2枚の特徴マップの距離計算から変化領域を推定する。そのため、変化点検出のための学習データセットを必要としない。特微量ベースの代表的な手法であるCNMF-F[1]は、ImageNetで事前学習したVGG19を用いて、画像ペアから特徴マップ抽出し変化領域を検出する。検出にはユークリッド距離を使用し、変化領域を決定するために閾値処理を施す。変化領域のみを検出する手法であるため、変化領域に属性を付与することができない。

学習ベースの変化点検出法: 変化点検出のためのデータセットを用いて画像間の類似性(非類似性)を学習するため、前述のアプローチと比べて視点のずれに頑健な検出が可能である。学習ベースの代表的な手法であるFC-EF-Res[2]は、変化点検出タスクとセマンティックセグメンテーションタスクをそれぞれ別々のネットワークで処理し、その結果を統合する手法である。2つのネットワークはU-Net[3]をベースに構成され、残差ブロックを導入している。セグメンテーションネットワークから抽出した特徴マップの差分を、変化点検出ネットワークにスキップ接続することで、画像間の明示的な比較を行なながら学習する。しかし、タスクごとに独立したネットワークを用いるため、学習中にタスク間で情報を相互利用することができないという問題がある。

3.提案手法

本研究では、変化領域とオブジェクトクラスを同時に推定することができる変化点検出法を提案する。提案手法のネットワーク構造を図1に示す。ネットワークは、Encoder-Decoder構造を持つ2つのモデルから構成されており、変化点検出モデルにはセグメンテーションモデルの最終層から抽出した特徴マップを入力する。

3.1 セグメンテーションモデル

セグメンテーションモデルは、U-Netをベースとするネットワーク構造である。画像ペアのそれぞれを個別にネットワークへ入力する。なお、各ネットワークの重みは共有している。セグメンテーションモデルの損失関数は、従来のU-Netと同様、クロスエントロピー損失を使用する。

3.2 変化点検出モデル

変化点検出モデルは、2つのEncoderと1つのDecoderで構成したネットワーク構造となっており、それぞれ3層の畳み込み層を持つ。2つのEncoderにより、画像間で異なる特徴を獲得することができる。また、各Encoderにはセグメンテーションモデルより得られた特徴マップをそれぞれ入力し、Decoderには各Encoderより得られた特徴マップをチャンネル方向に連結し入力する。

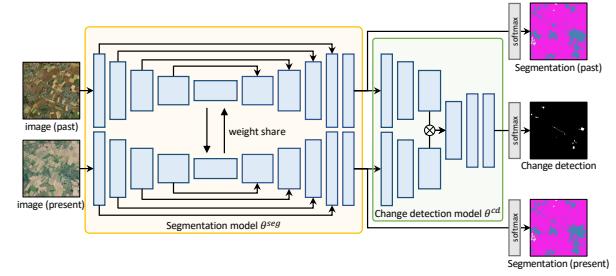


図1：変化点検出ネットワークの概要

3.3 段階的な学習と損失関数

提案手法は、セグメンテーションモデルと変化点検出モデルにより構成され、段階的な学習を行う。

Step1: まず、ベースとなるセグメンテーションモデル θ^{seg} を学習する。ピクセル総数を N 、クラス数を S 、正解ラベルを t 、ソフトマックス関数の出力を y とすると、セマンティックセグメンテーションタスクの損失関数 L_{Seg} は式(1)で表すことができる。

$$L_{Seg} = -\frac{1}{N} \sum_i^N \sum_c^S w_c t_{i,c} \log y_{i,c} \quad (1)$$

ここで、 w_c は各タスクの重みである。衛星画像などの超高解像度画像で構成されるデータセットを使用する場合、クラス不均衡が生じる。ラベルが不均衡のまま学習すると、出現確率の高いクラスに対して早期に収束してしまい、出現確率の低いクラスがほとんど反映されない問題がある。そのため、式(2)に示すように、出現確率 p_c の逆数に対数を重みとする。

$$w_c = \log \frac{1}{p_c} \quad (2)$$

これにより、出現確率の高いクラスにオーバーフィッティングすることを防ぐ。

Step2: 次に、セグメンテーションモデル θ^{seg} の重みを固定し、変化点検出モデル θ^{cd} を学習する。式(2)で定義した重みを使用し、クラス数を D とすると、変化点検出タスクの損失関数 L_{Diff} は式(3)で表すことができる。

$$L_{Diff} = -\frac{1}{N} \sum_i^N \sum_c^D w_c t_{i,c} \log y_{i,c} \quad (3)$$

Step3: そして、2つのモデルの相互関係を保つため、 θ^{seg} と θ^{cd} を同時に学習する。相互関係を保つための損失関数 L_{Mut} を式(4)に示す。Step2と同様、 L_{Mut} の算出には θ^{cd} より抽出した特徴マップを使用する。

$$L_{Mut} = -\frac{1}{N} \sum_i^N \sum_c^D t_{i,c} \log y_{i,c} \quad (4)$$

Step4: 各iterationでStep1からStep3を繰り返す。

以上に示す段階的な学習により、タスク間で情報を相互利用し、両タスクの精度向上が期待できる。

3.4 学習データの入力方法

提案手法は、 $10,000 \times 10,000$ ピクセルの超高解像度画像を対象とし、学習時は 512×512 ピクセルにランダムクロップする。しかし、クラス不均衡が生じる場合、ランダムクロップでは出現確率の低いクラスが選択されず、学習が不安定となる。そこで、出現確率の低いクラスを多く含むパッチ画像を別で用意し、通常のパッチ画像とともに学習を行う。これにより、クラス不均衡による影響を抑えることができるため、学習の安定性につながる。

4.評価実験

提案手法の有効性を示すために、変化点検出及びセマンティックセグメンテーションの評価実験を行う。

表 2 : 変化点検出とセマンティックセグメンテーションの比較結果

	U-Net	Semantic segmentation					Change detection		
		Seg	Global Accuracy	Class Accuracy	mean IoU	Kappa	Global Accuracy	mean IoU	Kappa
	Diff	-	90.96	62.78	45.83	78.56	92.36	48.28	6.86
CNNF-F		-	-	-	-	-	99.25	54.56	17.72
FC-EF-Res		89.01	-	-	-	71.92	88.66	-	3.28
Proposed		91.46	63.21	53.48	79.24		98.30	-	25.49

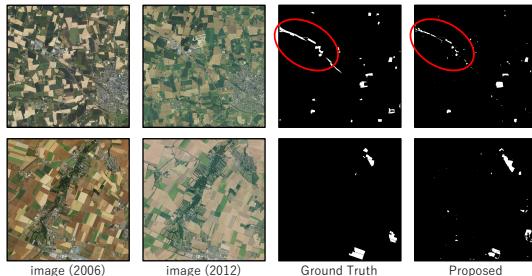


図 2 : 提案手法による変化点検出例

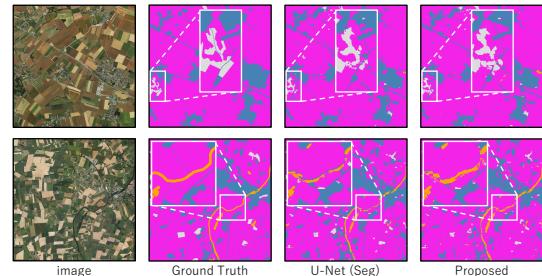


図 3 : 各手法のセグメンテーション例

4.1 データセット

本実験では、異なる時期に撮影された衛星画像ペアで構成される HRSCD データセット [2] を使用する。HRSCD データセットは、大規模な超高解像度の変化点検出用データセットである。RGB 画像ペアと、各ピクセルでの変化情報を付与した変化検出ラベル(変化領域なし、変化領域あり)、土地利用情報を付与したセグメンテーションラベル(情報なし、人工物、農業地帯、森林、湿地、水域)が含まれる。データ総数は 582 枚あり、学習用に 292 枚(146 組の画像ペア)、評価用に 290 枚(145 組の画像ペア)を使用する。表 1 に HRSCD データセットの各クラスにおける出現確率と重みを示す。表 1 より、出現確率が小さいほど、クラス重みが大きくなることがわかる。

表 1 : 各クラスにおける出現確率と重み

	Semantic Segmentation						Change Detection	
	情報なし	人工物	農業地帯	森林	湿地	水域	変化なし	変化あり
出現確率 [%]	17.70	11.51	61.84	8.36	0.02	0.58	99.23	0.77
クラス重み	0.0	2.16	0.48	2.48	8.54	5.15	0.01	4.87

4.2 実験概要

HRSCD データセットは、学習サンプルが少ないため、幾何変換、ノイズ付加等の Data Augmentation を行い、バリエーションを増加する。評価時は、超高解像度画像を 512×512 ピクセルの画像としてタイル状に切り出して入力する。定量的評価指標として、変化点検出、セマンティックセグメンテーション共に Global Accuracy, mean IoU, Kappa 係数を用いる。セマンティックセグメンテーションでは、さらに Class Accuracy を用いる。また、提案手法の性能比較として、U-Net および CNNF-F, FC-EF-Res を使用する。U-Net は、単一のネットワークとして変化点検出タスク U-Net (Diff) とセマンティックセグメンテーションタスク U-Net (Seg) それぞれの学習を行う。

4.3 各タスクにおける評価結果

各手法を使用した時の、変化点検出とセマンティックセグメンテーションの定量的評価を表 2 に、定性的評価を図 2 および図 3 に示す。U-Net (Seg) では、各画像に対するセグメンテーション結果を利用した変化点検出を行っており、画像間でクラスが異なる領域を変化領域とする。

変化点検出: 表 2 より、提案手法が従来手法と比較し、mean IoU と Kappa 係数において高精度であることが確認できる。一方、U-Net (Seg) ではセグメンテーション精度は高いものの、検出精度は低下した。これは、セグメンテーション結果においてクラス境界の誤識別が多くあるため、変化領域の誤検出を招いたと考えられる。また、図 2 より提案手法は道路などの細い領域の検出可能であることが確認で

きる。

セマンティックセグメンテーション : 表 2 より、提案手法は全ての評価指標で最も高精度であることが確認できる。特に U-Net (Seg) は、提案手法のセグメンテーションモデルと同一のネットワーク構造であるが、識別精度が最大約 0.7 ポイント向上した。これは、タスク間で情報を相互利用したからであると考えられる。また、図 3 より農業地帯などの領域の広いクラスと、検出が困難な川など領域が細いクラスにおいても高精度なセグメンテーションができるていることが分かる。

4.4 クラス不均衡に対する有効性の調査

人工物クラスを含むパッチ画像を追加することによる精度比較を行う。表 3 に、追加前と追加後それぞれの定量的評価を示す。出現確率の低いクラスを含むパッチ画像を追加することで学習が安定し、ほぼ全てのクラスで精度向上を確認した。また、出現確率の高い農業地帯の誤識別が減少したこと、森林クラスの識別精度が約 35.7 ポイント向上した。

表 3 : 入力方法変更による比較結果

	Semantic Segmentation					Change Detection	
	人工物	農業地帯	森林	湿地	水域	変化なし	変化あり
追加前	67.51	85.27	37.81	0.0	31.56	99.20	17.74
追加後	68.97	89.88	73.55	0.0	35.01	99.17	19.47

5.おわりに

本研究では、変化領域のオブジェクトクラスを推定するとのできる変化点検出法を提案した。2 つのモデルを使用することで、高精度な変化領域の検出を実現し、単一の U-Net 同様のセグメンテーション精度を達成した。今後は、出現確率の低いクラスの更なる精度向上を検討する。

参考文献

- [1] E. Amin, et al., “Convolutional neural network features based change detection in satellite images”, First International Workshop on Pattern Recognition, 2016.
- [2] R. C. Daudt, et al., “Multitask Learning for Large-scale Semantic Change Detection”, CVIU, 2019.
- [3] O. Ronneberger, et al., “U-Net: Convolutional Networks for Biomedical Image Segmentation”, MICCAI, 2015.

研究業績

- [1] 筒井駿吾 等,“セマンティックセグメンテーションによる超高解像度画像からの変化点検出”,動的画像処理実用化ワークショップ, 2021.

(他 1 件)