

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

○筒井駿吾, 平川翼, 山下隆義, 藤吉弘亘

○ Shungo TSUTSUI, Tsubasa HIRAKAWA,
Takayoshi YAMASHITA and Hironobu FUJIYOSHI

中部大学工学研究科, ts2hug30@mprg.cs.chubu.ac.jp, hirakawa@mprg.cs.chubu.ac.jp,
takayoshi@isc.chubu.ac.jp, fujiyoshi@isc.chubu.ac.jp

変化点検出とは, 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り, リモートセンシング技術の1つとして広く研究されている. 変化点検出は, 人的コストまたは時間的コスト削減のため, 高解像度マップの自動更新, 土地利用変化の推移など, 様々なアプリケーションに応用されている. しかし, 従来の変化点検出手法の多くは変化領域を検出するのみであるため, 画像内に複数の検出対象がある場合, 変化領域のオブジェクトクラスまでわからない. そこで本研究では, 変化領域のオブジェクトクラスを同時に推定する変化点検出手法を提案する. 提案手法は, セマンティックセグメンテーションモデルに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, 変化領域の検出精度向上を図る.

<キーワード> Change Detection, Semantic Segmentation, Satellite Image

1 はじめに

変化点検出は, コンピュータビジョンの分野において重要なタスクの1つであり, 異なる時刻に撮影された同一シーン画像2枚を使用する. このタスクは, 高解像度マップの自動更新や, 土地利用変化の推移など, 様々なアプリケーションに応用することができる. 土地の利用状況の変化を計測する場合, 様々な地域の調査を行うため, 広い範囲で撮影した画像が必要となる. 近年では, 車載カメラやドローンの開発により, 都市規模でのデータ収集が容易となった. これにより, 広範囲での調査を可能にしたが, 膨大なデータの中から必要な情報のみを選択するには人的コスト及び時間的コストがかかる.

変化点検出手法は, このようなコストを削減することができる. 従来の変化点検出手法 [1-6] では, 検出結果を2値ラベルまたは信頼値としてピクセル単位で推定する. しかし, これらの手法を前述のようなアプリケーションに適応する場合, 画像内に複数の検出対象が含まれるため, 変化領域のオブジェクトクラスがわからない. また, 高解像度マップを自動更新するためには, 変化したオブジェクトまたは領域を分類する必

要がある. 変化した領域を把握することで, 利用状況が変化した土地の特定を迅速に行うことができる. したがって, 変化した領域のセマンティックな情報を理解することは, 土地の利用状況を把握するアプリケーションにとって有益であると期待される.

そこで本研究では, 変化領域のオブジェクトクラスを推定する変化点検出手法を提案する. 提案手法では, オブジェクトクラスの推定を行うセグメンテーションモデルと変化領域の推定を行う変化点検出モデルの2つのモデルを使用する. 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, 変化領域の検出精度向上を図る.

2 関連研究

主な変化点検出手法アルゴリズムは, 対応するピクセル間の差分を求める処理と, 変化領域を決定する処理の2つの主要なステップで構成されている. 従来の変化点検出手法 [1] は, 2枚の画像間のRGB値の誤差より変化領域を推定しており, 視点の違いによる影響を抑えるため, SIFT [7] などを用いて対応付けを行う必要があった. 近年ではDeep Learningの影響を受け,

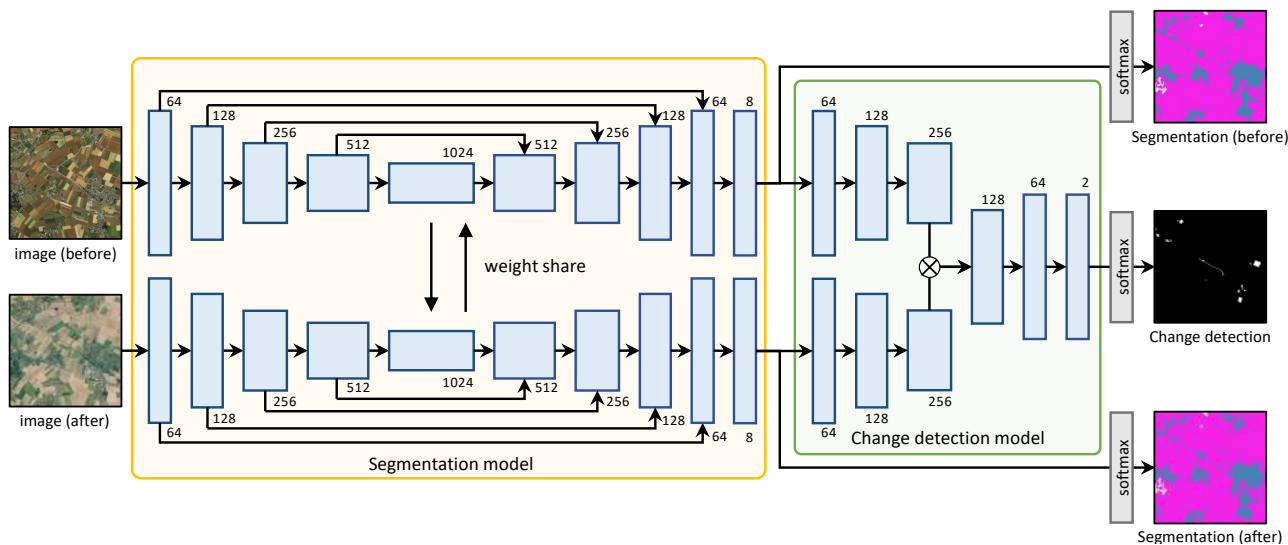


図 1 セグメンテーションを用いた変化点検出ネットワーク

CNN を用いた変化点検出手法が提案されており、画像の位置ずれに頑健な検出が可能となった。ここで、CNN を用いた変化点検出手法は大きく 2 つのアプローチに分類される。一つ目のアプローチは、ImageNet データセット [8] などの大規模データセットで学習した CNN に基づく変化点検出手法 [3] である。このアプローチは、CNN より抽出した 2 枚の特徴マップ間で距離計算を行い、変化領域を推定する。また、既存のデータセットを用いているため、変化点検出のための新しいデータセットを必要としない。しかし、推定した変化領域は連続値であるため、目的に応じて閾値の細かい調整が不可欠である。

2 つ目のアプローチは、2 枚の RGB 画像を入力した CNN から変化領域を直接推定する変化点検出手法 [4-6] である。このアプローチは、画像間の類似性 (非類似性) を学習するため、環境による変化や軽微な視点のずれに頑健な検出が可能である。しかし、学習を行うための大量のデータが必要となり、作成コストが大幅にかかる。

3 提案手法

本研究では、変化領域におけるオブジェクトクラスの推定を可能にするため、オブジェクトクラスを推定するセグメンテーションモデルと、変化領域を推定する変化点検出モデルの 2 つのモデルを用いた変化点検出ネットワークを提案する。提案手法の変化点検出ネットワークを図 1 に示す。図 1 のネットワーク構造は、Encoder-Decoder 構造となっており、2 つの異なるサ

イズのモデルから構成されている。セグメンテーションモデルでは RGB 画像を入力としており、変化点検出モデルで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とする。

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、セマンティックセグメンテーションネットワークの 1 つである U-Net [9] を使用する。U-Net は、Encoder 部分と Decoder 部分の対応した解像度の特徴マップを連結させることで、大域的な特徴と局所的な特徴を捉えることができる。これにより高精度なセマンティックセグメンテーションが可能となり、物体の境界を詳細に認識することができる。セグメンテーションモデルでは、画像ペアそれぞれを入力としており、画像に対する 2 つのネットワークの重みは共有している。また、従来の U-Net と同様に、セグメンテーションモデルの損失関数には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 を持つネットワーク構造となっており、それぞれ 3 層の畳み込み層で構成されている。2 つの Encoder を用いることで、画像間で異なる特徴を獲得することが可能である。図 1 より変化点検出モデルは、セグメンテーションモデルより得られた特徴マップを、それぞれの Encoder に入力する。これにより、end-to-end で学習することができ、セマンティックセグメンテーショ

表 1 各クラスにおける出現確率と重み

	Change Detection		Semantic Segmentation					
	変化領域なし	変化領域あり	情報なし	人工物	農業地帯	森林	湿地	水域
出現確率 [%]	99.232	0.768	17.696	11.508	61.835	8.364	0.02	0.578
クラス重み	0.008	4.869	0.0	2.162	0.481	2.481	8.541	5.154

ンタスクにおける同一クラス間の誤識別の抑制できる。また、変化点検出タスクでは2クラス分類として検出を行なっているため、損失関数にはクロスエントロピー損失を使用する。

3.3 損失関数

ネットワークの学習は、end-to-endで行う。学習の損失関数は、セグメンテーションモデルを学習するための損失、変化点検出モデルを学習するための損失、2つのモデルの相互関係を保つための損失の3つの損失を用いて学習を行う。相互関係を保つための損失では、変化点検出結果を用いて算出する。

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。また、提案手法を用いて変化領域のオブジェクトクラスを推定する。

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 [10] 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のデータセットであり、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) が含まれている。データ総数は 291 組の画像ペアであり、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。

4.2 クラス重みの定義

衛星画像で構成されるデータセットを使用する場合、クラス不均衡が存在することがある。ラベルが不均等なまま学習を行うと、出現確率の高い損失が始めに収束され、出現クラスの低いクラスの損失がほとんど反映されないといった問題が存在する。そこで、不均衡

問題を解決するため、損失関数に付与する各クラスの重みを定義する。各クラスの重みを式 (1) に示す。

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

ここで、 w_c は各クラスの重みであり、 p_c は各クラスの出現確率である。出現確率の逆数に対数をとることで、出現確率の高いクラスを相対的に小さくし、出現確率の低いクラスの損失を大きくする。表 1 に各クラスにおける出現確率と重みを示す。表 1 より、セグメンテーションでは 6 割以上が農業地帯であり、変化点検出では 9 割以上が変化領域なしであることが確認できる。また、出現確率が小さいほど、クラス重みが大きくなることわかる。これにより、出現確率の高いクラスがすぐに収束して局所解に陥ることを防ぐ。

4.3 実験概要

HRSCD データセットの画像サイズは 10,000×10,000 であるが、学習時に使用する画像は、2,000×2,000 に分割したパッチ画像としてあらかじめ用意しておく。さらに、変化領域の大部分を占める人工物クラスの学習を促進するために、人工物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。人工物クラスは出現確率の低いクラスの1つであるため、学習の安定性にも繋がる。本実験では、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し、画像サイズを 512×512 に切り出す。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 係数を使用し、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には以下の手法を用いた。

FC-EF-Res [10]: この手法は、FC-Siam-diff [6] に基づいて構成されており、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それぞれのネットワークを統合したマルチタスクネットワークとなっている。こ

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

		Change detection			Semantic segmentation			
		Global Accuracy	mean IoU	Kappa	Global Accuracy	Class Accuracy	mean IoU	Kappa
U-Net	Seg	92.36	48.28	6.86	90.96	62.78	45.83	78.56
	Diff	99.25	54.56	17.72	-	-	-	-
FC-EF-Res		98.30	-	25.49	89.01	-	-	71.92
Proposed		99.17	63.50	32.20	91.46	63.21	53.48	79.24

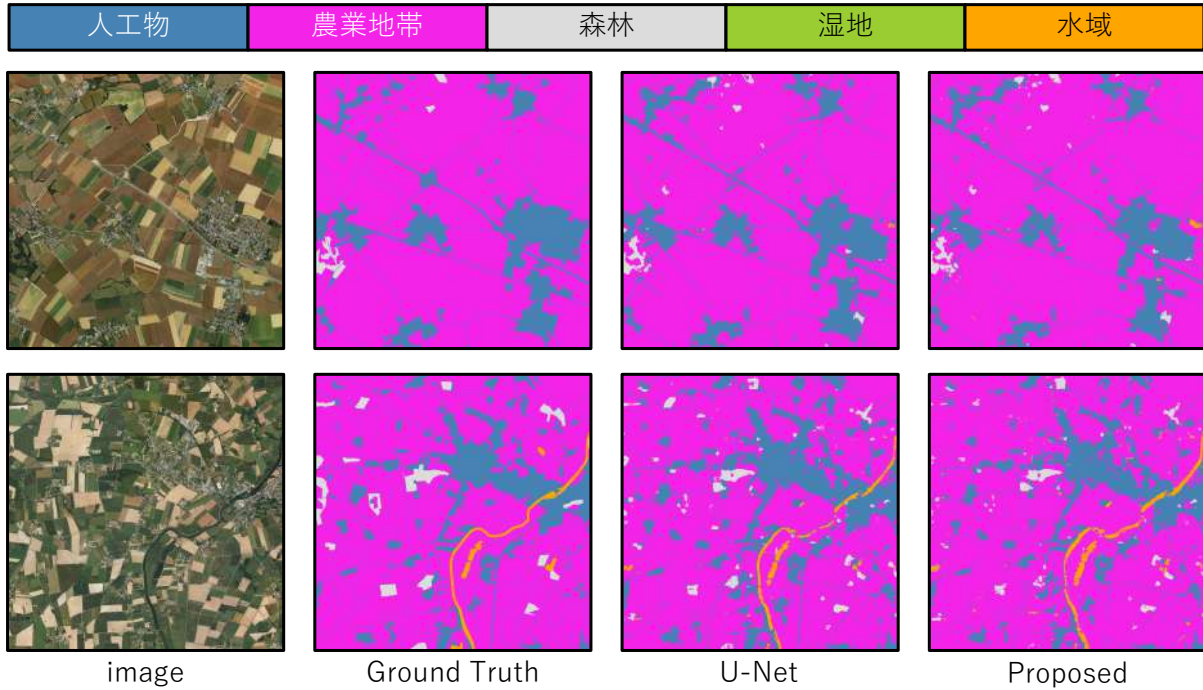


図 2 セグメンテーション比較結果

のネットワークは、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しており、画像間の明示的な比較を行いながら学習することができる。また、残差ブロック [11] を導入することにより、高精度な検出を可能にした。

U-Net: 単一のネットワークとして、変化点検出タスク及びセマンティックセグメンテーションタスクそれぞれの学習を行う。セマンティックセグメンテーションタスクのために学習した U-Net を U-Net (Seg) とし、変化点検出タスクのために学習した U-Net を U-Net (Diff) とする。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、画像ペア間でクラスが異なる領域を変化領域とする。また、U-Net (Diff) では 2 枚の画像を結合したものを 1 つのネットワークに入力し変化点検出結果を取得する。

2 枚の画像を同時に処理するため、学習の高速化と柔軟性を持つ。

4.4 各タスクにおける実験結果

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 3 および図 2 に示す。

セマンティックセグメンテーション: 表 2 より、提案手法が全ての評価指標で最も精度が高くなっていることが確認できる。中でも U-Net (Seg) は、提案手法のセグメンテーションモデルと同一ネットワーク構造となっているが、最大約 0.7 ポイントと識別精度が向上した。これは、同一ラベル間の誤識別が減り安定した学習が行えたためであると考えられる。また、図 2 より農業地帯などの出現確率の高いクラスのセグメンテーションの他、検出が困難な道路や川などの細か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ています。

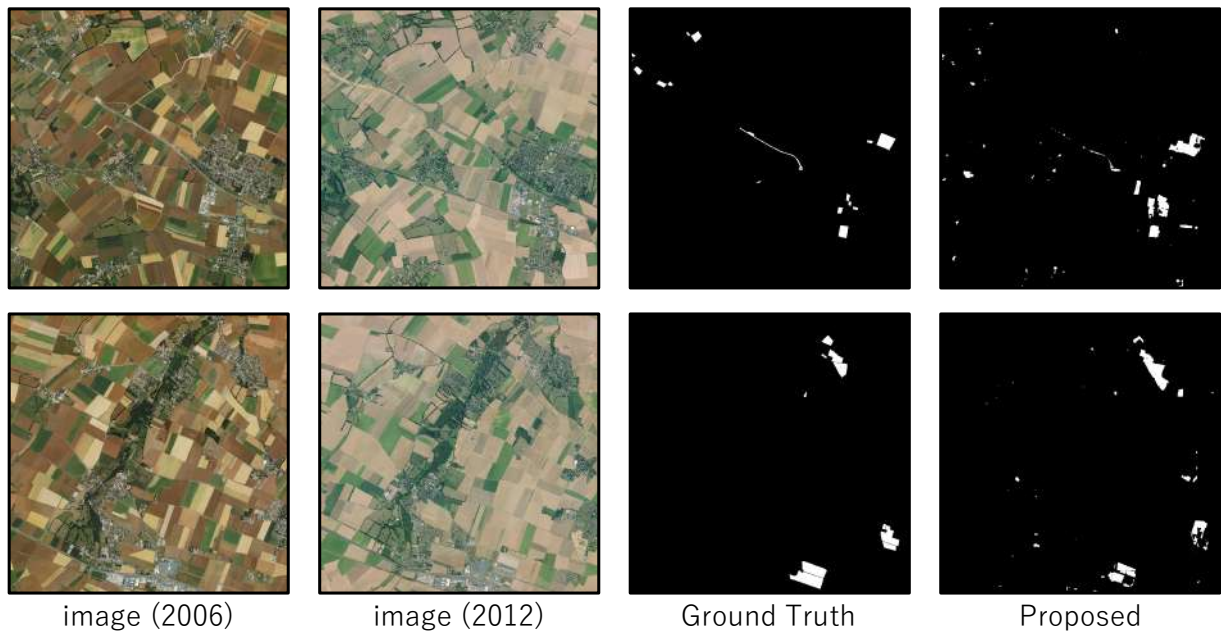


図 3 変化点検出の結果例

表 3 各クラスの出現確率と推定精度

	人工物	農業地帯	森林	湿地	水域
出現確率 [%]	86.975	2.207	0.056	0.652	10.110
IoU [%]	99.29	0.59	0.0	0.0	0.0

が分かる。

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 係数において精度が向上してい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，クラス境界での曖昧さが誤検出に繋がったと考えられる．また，図 3 より道路などの細い領域を検出できていることが確認できる．

4.5 変化領域におけるクラス推定

提案手法では，セグメンテーション結果と変化点検出結果を組み合わせることで，変化領域におけるオブジェクトクラスの推定を行うことができる．図 4 に変化領域におけるオブジェクトクラスの推定結果を示す．図 4 より変化領域を検出できているため，オブジェクトクラスの推定が可能であることが確認できる．また，表 3 に，変化領域における各クラスの出現確率と推定精度を示す．推定精度は，検出した領域が実際に変化領域である時の各クラスの識別精度を表す．表 3 より，土地の利用状況の変化した 8 割以上が人工物クラスへ

の変化であるため，人工物以外のクラスでは検出が困難となっている．

5 おわりに

本研究では，変化領域におけるオブジェクトクラスの推定を可能にするため，セマンティックセグメンテーションを用いた変化点検出手法を提案した．2つのモデルを使用することで，変化点検出およびセマンティックセグメンテーションにおいて従来手法と比較し精度が向上した．また，2つの結果を使用することで変化領域におけるオブジェクトクラスが推定できることを確認した．今後は，セグメンテーションモデルにおいて，U-Net をベースネットワークに使用したが，他の Encoder-Decoder 構造を持つネットワークに適用できる可能性があるため，調査する予定である．

6 謝辞

本研究の一部は，内閣府総合科学技術・イノベーション会議の戦略的イノベーション創造プログラム (SIP) 「国家レジリエンス (防災・減災) の強化」 (管理法人：防災科研) によって実施されました．

参考文献

- [1] A. Taneja, *et al.*, “Image based detection of geometric changes in urban environments”, In ICCV, (2011)

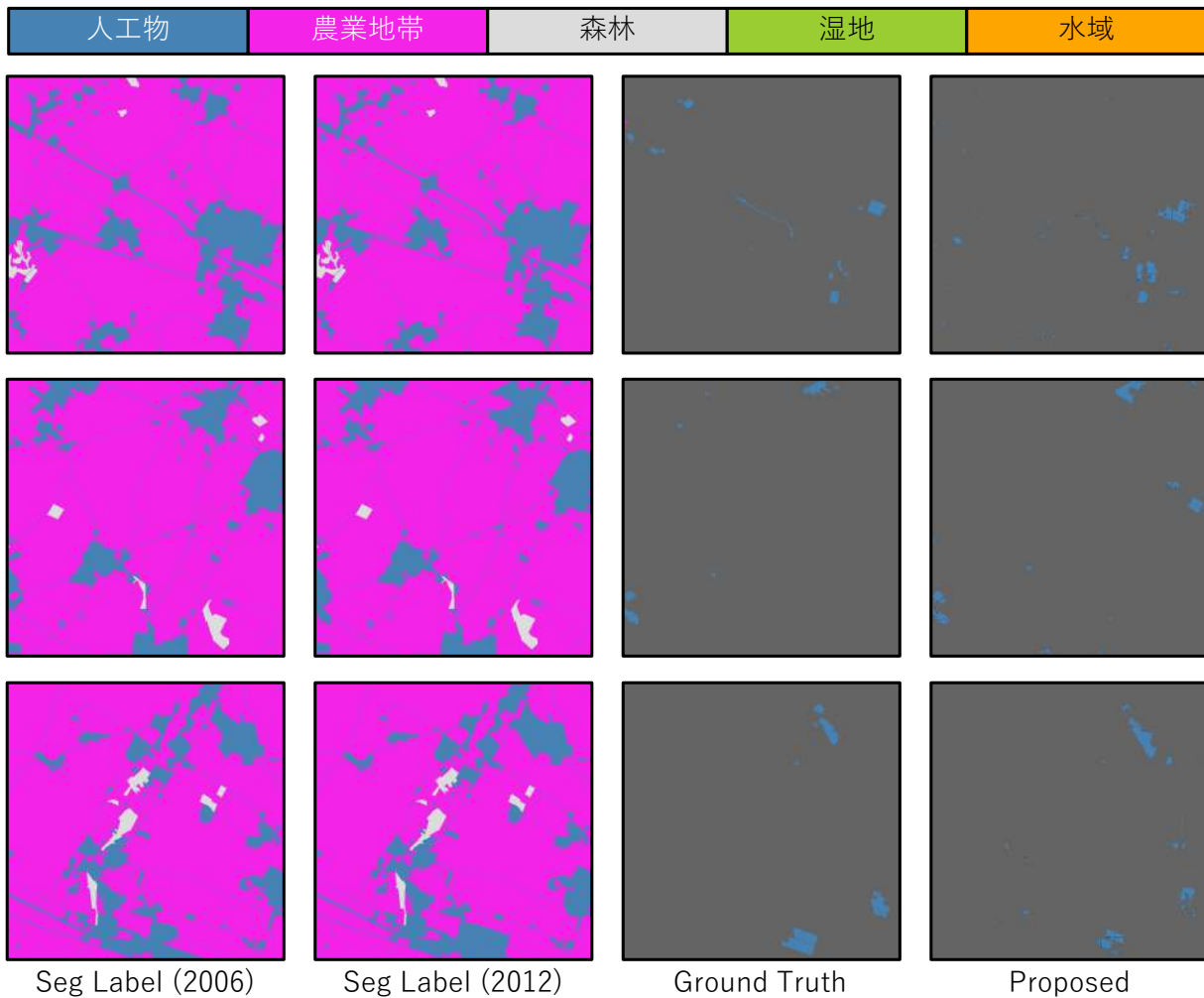


図 4 変化領域におけるオブジェクトクラスの推定結果

- [2] S. Stent, *et al.*, “Detecting change for multi-view, long-term surface inspection.”, In *BMVC*, (2015)
- [3] K. Sakurada, *et al.*, “Change Detection from a Street Image Pair using CNN Features and Superpixel Segmentation”, In *BMVC*, (2015)
- [4] S. Zagoruyko, *et al.*, “Learning to compare image patches via convolutional neural networks”, In *CVPR*, (2015)
- [5] S. Khan, *et al.*, “Learning deep structured network for weakly supervised change detection”, In *IJCAI*, (2017)
- [6] R. C. Daudt, *et al.*, “Fully convolutional siamese networks for change detection”, In *ICIP*, (2018)
- [7] D. G. Lowe, “Distinctive image features from scale-invariant keypoints”, In *IJCV*, (2004)
- [8] J. Deng, *et al.*, “ImageNet: A large-scale hierarchical image database”, In *CVPR*, (2009)
- [9] O. Ronneberger, *et al.*, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, In *MICCAI*, (2015)
- [10] R. C. Daudt, *et al.*, “Multitask Learning for Large-scale Semantic Change Detection”, In *CVIU*, (2019)
- [11] K. He, *et al.*, “Deep residual learning for image recognition”, In *CVPR*, (2016)