

# 深層学習を用いた航空レーザフィルタリング 修正箇所自動検出

橋本 寛治・王 婭・鈴木 久美子・高橋 元気（国際航業株式会社）

## 1. はじめに

航空レーザフィルタリング業務では、地盤高の取得を目的としてレーザ計測点から樹木や建物といった地表面以外の点群を除去して地盤点群を得る。従来、除去のために点群処理ツールを用いて自動でフィルタリング処理を行っているが、自動フィルタリング結果に含まれる誤りを人手で修正（手動フィルタリング）する作業時間が長く、人的コストが高いといったことが課題となっている（図1）。そこで、本研究では手動フィルタリング作業量の低減に資するべく、深層学習モデルを用いた自動フィルタリング後の点群の修正箇所検出手法を開発した。また深層学習モデルとして複数のモデルを比較検討した。

## 2. 既存手法

既存手法では手動フィルタリングで自動フィルタリングの誤り箇所を修正している。修正する自動フィルタリングの誤りは、「フィルタリング漏れ」と「過剰フィルタリング」の2種類に分けられる。フィルタリング漏れは建物や植生を、地盤点群に含めてしまうケースである。過剰フィルタリングとは地盤と見なすべき地物を地盤点群から削除してしまうケースである。この2種類の誤りについて、2.1,2.2項に示す修正箇所を判読

により特定したのちに、人手で点群修正している。

### 2.1 フィルタリング漏れ箇所の判読

自動フィルタリングの誤りのうちフィルタリング漏れの箇所は、自動フィルタリング後の地盤点群から作成した段彩陰影図による判読を行い、修正箇所を決定している（図4）。修正箇所は建物、植生、橋梁など地盤以外の地物が対象である。またダム・堤防は地盤として扱うとし、修正の対象ではない。地物は段彩陰影図では周囲の地盤から凸形状と表現されるという特徴があるため判読することができる。

### 2.2 過剰フィルタリング箇所の判読

自動フィルタリングの誤りのうち過剰フィルタリング箇所は、段彩陰影図での地形表現に加えて、陰影点群図を利用して判読している。陰影点群図とは自動フィルタリング前の段彩陰影図と、自動フィルタリング後の地盤点の有無を示した点群図を重畳した図である（図2）。但しここで利用する段彩陰影図はグレースケール変換したものを使う。この陰影点群図では自動フィルタリング前の地物を背景の段彩陰影図で判読しながら、その地物地盤点の画素と重畳してい

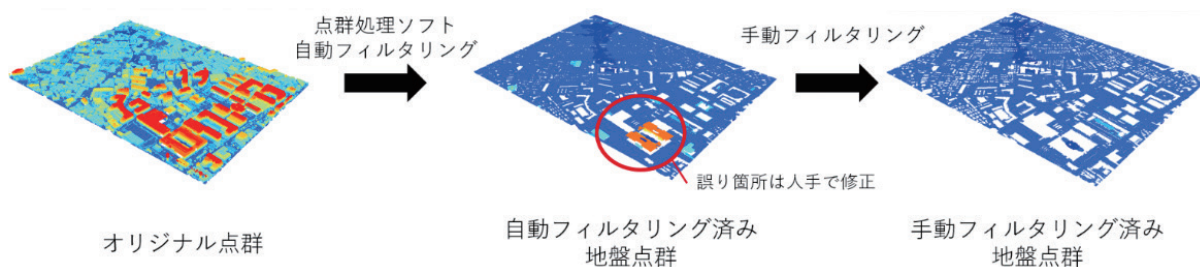


図1 既存のフィルタリング手法

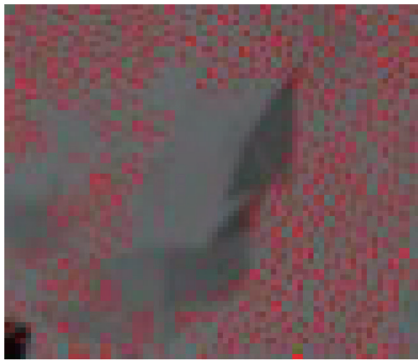


図2 陰影点群図

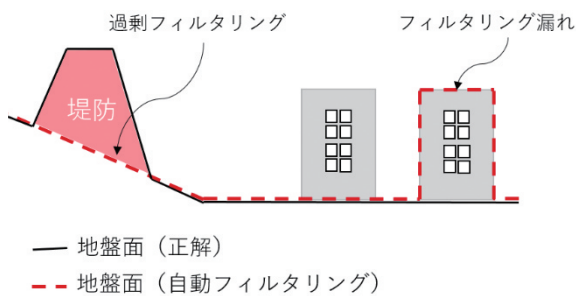


図3 修正箇所例

れば地盤点から削除されておらず、反対に地物に対応する地盤点の画素がなければ、自動フィルタリングにより地盤点から削除されていると判読できる。過剰フィルタリングの修正箇所は、堤防や砂防ダムの堤体、河川の護岸や擁壁といった地盤と見なす地物が削除されている領域が対象となる(図3)。

### 2.3 手動フィルタリング

自動フィルタリング後の点群の修正箇所は、2.1や2.2のように図面で判読し決定する(図4)。決定した修正領域に対して手動フィルタリングにて点群を修正することで、最終成果品として

いる。なお、いずれの場合もオルソ写真を参考情報として参照することもあるが、基本的には段彩陰影図等を利用している。

## 3. 提案手法

自動フィルタリングの誤りであるフィルタリング漏れでは地物が対象であるのに対して、過剰フィルタリングの対象は本来地盤であるものの削除された領域であり、誤りの箇所は異なる性質を示している。そこで修正が必要となる漏れ箇所と過剰箇所それぞれを別個に検出するセマンティックセグメンテーションモデルを構築した。セマンティックセグメンテーションモデルは画素単位で分類を行う深層学習モデルである。

モデルの教師データとして、手動フィルタリングの作業領域を示すマスク画像を利用した。マスク画像は手動フィルタリング前後の標高差分に対して閾値を設定することで作成した。標高差分が負の領域は漏れ箇所を推定するモデルの教師、正の領域は過剰箇所を推定するモデルの教師となる。

### 3.1 フィルタリング漏れ箇所検出モデル

既存手法の漏れ箇所の判読では、地形表現が特徴的で地物が判読しやすい段彩陰影図が利用されている。提案手法での漏れ箇所の検出においても地形の特徴が大きく影響すると考えられることから、自動フィルタリング後の点群から作成した段彩陰影図を入力データとするモデルを構築した。

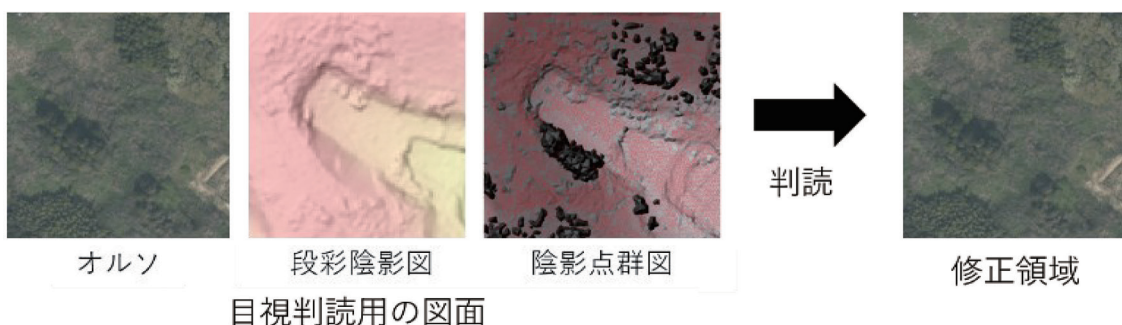


図4 修正領域の判読

### 3.2 過剰フィルタリング箇所検出モデル

過剰箇所の検出においても地形的特徴は重要であると考えられるため、段彩陰影図を入力データの1つとした。一方で、既存手法の判読に用いていた陰影点群図は、段彩陰影図と点群図という2つの図の重畳であり、特に点群図の画素値はカテゴリ値であり通常の画像とは異なる性質をもつ。CNN(畳み込みニューラルネットワーク)に用いられる畳み込み演算では、近接領域の画素値で演算する機構であり、カテゴリ値と連続値が混交して演算が行われる。そのため陰影点群図は、畳み込み演算を用いるCNNベースの深層学習にそのまま用いることは適さないと考えられる。

そこで、自動フィルタリングが作用した領域を示す陰影点群図の代替として、自動フィルタリング前後の標高差分を使用した。この標高差分は、自動フィルタリング前の点群の最低点から生成した標高と、自動フィルタリング後の地盤点から生成した標高の差分であり、自動フィルタリングによって地物が削除されていれば、標高差が大きくなる性質をもつ。また差分の情報は連続値であるため、CNNで扱いやすく入力データに適していると考えられる。

以上のことから、過剰フィルタリング箇所検出においては、段彩陰影図と自動フィルタリング前後の標高差分を入力データとするモデルを構築した。

## 4. 実験

### 4.1 実験概要

実験ではセマンティックセグメンテーションの手法の中で良く利用される、DeepLabV3+<sup>1)</sup>、PointRend<sup>2)</sup>、OCRNet<sup>3)</sup>、BiSeNetV2<sup>4)</sup>、SegFormer<sup>5)</sup>の5つのモデルの比較検証を行った。DeepLabV3+はマルチスケールの情報を統合する機構を持つモデルで、デファクトスタンダードなモデルである。PointRendは、CGのレ

ンダリング手法に着想を得たモデルで、低解像度の出力をBilinear補間でアップサンプリングする際に、確信度の低いピクセルは再度クラス予測することで精度を高める。OCRNetでは対象物領域に着目した自己注意機構を導入している。BiSeNetV2は、全体を高解像度のまま細部の情報を維持するネットワークと、高度なコンテキストを抽出するネットワークの2つから成り、最後に情報の統合を行う。これらのモデルは画像を扱う深層学習モデルで基本技術となっているCNNでの特徴量抽出をベースとしたモデルである。一方でSegFormerは他と異なりTransformerをベースとしたモデルであり、非常に大きなデータセットで学習した場合、性能が良いとされる。入力画像はパッチと呼ばれる小領域に分割されるが、パッチ同士の関係性についても学習を行うため、受容野は画像全体となると解釈される。

この特色の異なる5つのモデルを用いて、漏れ箇所及び過剰箇所をそれぞれ検出するモデルを構築して比較検討を行った。

### 4.2 学習

対象地域は、自動フィルタリングで誤りが起こりやすい大型建造物や河川護岸などの地物が含まれる東京都の都市部とし、航空レーザ点群から作成した標高画像や段彩陰影図を入力データとした。入力データは図郭単位で作成され、1図郭を3500×5000ピクセルの画像データとした。学習データセットはtrain/val/testの3群に分けて構築した(図5)。全図郭で192図郭あり、train群は144図郭、val群、test群は24図郭であった。本研究では1つのモデルにつき、20万回の学習を行った。学習にはtrain群のデータを使用し、1000回の学習ごとにval群で評価を行った。そのうち最良の結果を示した学習回数時点での重みを使い、test群で最終的な評価を行った。1回の学習に際して20枚の画像を同時に読み込み、3500×5000ピクセルの画像から600×



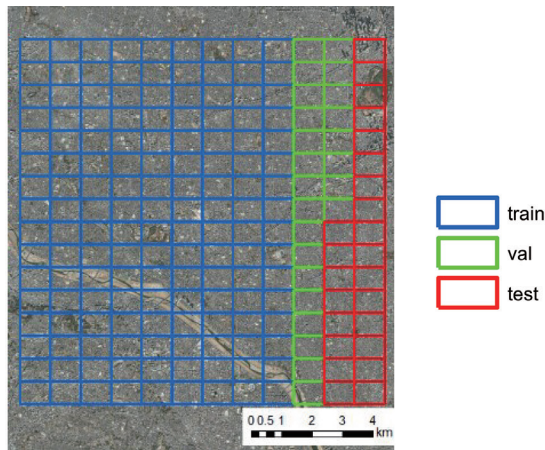


図5 対象地域

600の領域をランダムに取り出して用いた。

## 5. 実験結果

### 5.1 定量評価

表1、表2に各実験の定量評価を示す。漏れ検出モデルで最良の結果を示すOCRNetでは約68.8%の修正箇所を、次点のPointRendでは約67.3%の修正箇所を検出した。DeepLabV3+との差はわずかであった。

過剰箇所検出モデルでは、PointRendが約53.1%と最良の結果を示し、DeepLabV3+を上回った。OCRNetは約37.7%と、漏れ検出モデルのような性能は得られなかった。また5つのモデルの中で、特にSegFormerの性能は低かった。一般にTransformer系統のモデルでは、データ量が十分に大きいときには高い性能を示すが、データ量が限定的な場合には既存のCNNよりも劣ることが知られていることから、今回の学習データ量が不十分だった可能性がある。BiSeNetV2も過剰箇所の検出に対応できていない。BiSeNetV2は高速な推論を目的と

表1 漏れ検出モデルの性能比較

| 対象 | モデル         | IoU   | Recall |
|----|-------------|-------|--------|
| 漏れ | DeepLabV3 + | 49.38 | 66.96  |
| 漏れ | PointRend   | 50.63 | 67.29  |
| 漏れ | OCRNet      | 48.10 | 68.80  |
| 漏れ | BiSeNetV2   | 40.53 | 60.06  |
| 漏れ | SegFormer   | 36.82 | 45.53  |

して開発されたモデルであるため、細部に対しての検出が不十分であったと考えられる。

また両方の検出対象で安定的に性能を発揮したのはPointRendとDeepLabV3+であった。

### 5.2 定性評価

次に定性評価を示す。図6上段に漏れ検出モデル、下段に過剰検出モデルの検出結果を示す。検出結果は白い領域としてモデルごとに表示し、重畳して赤枠で示されているのが漏れ領域の、青枠で示されているのが過剰領域のグラウンドトゥールズである。図6上段では大型建築物が対象となっており、全てのモデルでおおよその領域は検出している。一方で、検出漏れや誤検出となった領域も認められる。DeepLabV3+やPointRendの結果では北側の箇所を除いて良好な結果を示している。OCRNetでは境界で良好に検出するも、誤り検出も大きい。またBiSeNetV2やSegFormerでは建物西側の細部への対応できなかった。高解像度の情報がネットワークでの情報の集約過程で消失したため、推論結果に反映されなかったと考えられる。

図6下段では、周囲の地盤より低い箇所を通る鉄道用地の擁壁部分での検出例を示す。DeepLabV3+やPointRendでは検出漏れの領域が多くなっているものの、検出自体は可能であった。一方でOCRNet、BiSeNetV2やSegFormerでは殆ど検出ができなかった。これらのモデルでは細部の情報を消失してしまったため、検出に問題が生じたと考えられる。こうした細部領域が多く、モデル間でのスコアの差を広げた要因であった。

表2 過剰検出モデルの性能比較

| 対象 | モデル         | IoU   | Recall |
|----|-------------|-------|--------|
| 過剰 | DeepLabV3 + | 33.99 | 51.48  |
| 過剰 | PointRend   | 34.48 | 53.10  |
| 過剰 | OCRNet      | 29.59 | 37.74  |
| 過剰 | BiSeNetV2   | 22.97 | 26.77  |
| 過剰 | SegFormer   | 25.55 | 33.69  |

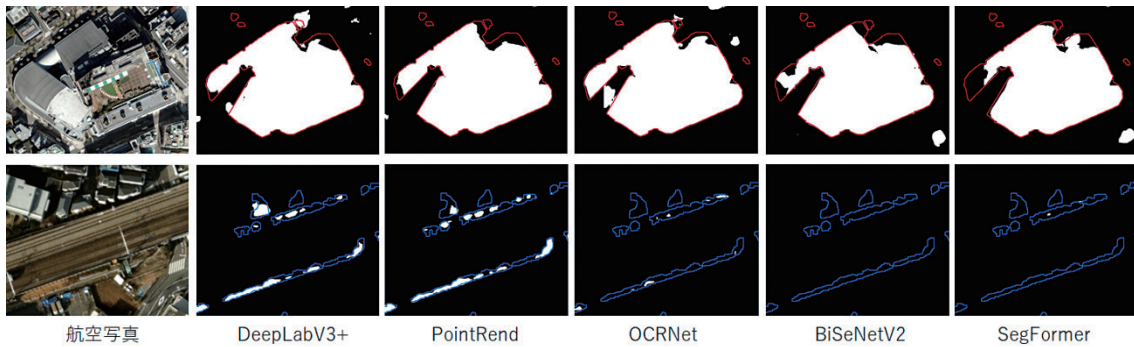


図6 定性評価結果

## 6. まとめ

航空レーザフィルタリングにおいて、自動フィルタリング処理時に発生する2種類の誤り（漏れ・過剰）をそれぞれ検出するモデルを開発した。複数のモデルを比較検証した結果、PointRendが安定的に高精度で頑健なモデルであった。

漏れ検出においては約7割の修正箇所を検出し、特に大型建造物の検出で良好な結果を示した。過剰箇所検出モデルにおいては、難易度の高い鉄道用地の擁壁部分過剰フィルタリング箇所の検出が可能であった。航空レーザフィルタリング業務においては、検出結果を利用したスクリーニングを行うことにより、判読作業のコストが低減する見込みであり、一定程度の目標を達したといえる。また検出結果を用いた点群の自動修正を行うことにより、手動フィルタリング作業の低減が期待できる。

今後は、教師データ作成基準の不一致からくる誤検出やモデルの出力結果に含まれるノイズ低減、水域でのブレイクラインへの対応について取り組み予定である。

### ■参考文献

- 1) Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, Hartwig Adam, Proceedings of the European Conference on Computer Vision (ECCV), pp. 801-818, 2018
- 2) Yuan, Y., Chen, X., Wang, J. (2020).

Object-Contextual Representations for Semantic Segmentation, Proceedings of the European Conference on Computer Vision (ECCV), pp. 801-818, 2018

- 3) A. Kirillov, Y. Wu, K. He and R. Girshick, PointRend: Image Segmentation As Rendering, Proceedings of the Computer Vision and Pattern Recognition (CVPR), pp. 9796-9805, 2020
- 4) Yu, C., Gao, C., Wang, J. et al. BiSeNet V2: Bilateral Network with Guided Aggregation for Real-Time Semantic Segmentation. Int J Comput Vis 129, 3051-3068, 2021
- 5) Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo, SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. Proceedings of the Neural Information Processing Systems, 2021

### ■執筆者

橋本 寛治 (はしもと かんじ)

国際航業株式会社

(共著者) 所属は筆頭著者に同じ

王 姪 (おう あ)

鈴木 久美子 (すずき くみこ)

高橋 元気 (たかはし げんき)