

# セマンティックセグメンテーションやセンサーフュージョンを利用した橋梁の構造情報の推定のための点群処理手法の開発

板倉 健太<sup>1</sup>・林 拓哉<sup>1</sup>・上脇 優人<sup>1</sup>・全 邦釘<sup>2</sup>

<sup>1</sup> ImVisionLabs 株式会社 (〒113-8485 東京都文京区本郷 7-3-1)

東京大学南研究棟アントレプレナーラボ)

<sup>2</sup> 東京大学大学院工学系研究科総合研究機構 (〒113-8485 東京都文京区本郷 7-3-1)

E-mail: chun@g.ecc.u-tokyo.ac.jp (Corresponding Author)

本研究では、地上型レーザースキャナから得た点群データと画像データを組み合わせてセンサーフュージョンを実施し、画像でのセグメンテーション結果を点群にマッピングした。DeepLabv3+を利用して、画像のセマンティックセグメンテーションを行った。Segment Anything Model を利用して地覆と舗装面との接続線の情報を更新した後に、カメラの外部パラメータや内部パラメータを利用して、セグメンテーション情報を点群に格納した。この情報を利用することで、有効幅員などの橋梁の構造情報を計測できるようになった。画像データの詳細な情報と点群データの3次元情報を活用することで、データの詳細性と構造情報を両立した解析が可能となり、かつ大容量の点群データも効率良く処理することができた。

**Key Words:** Deep Learning, LiDAR, Point Cloud, Segment Anything Model, Sensor Fusion

## 1. はじめに

3次元点群とは、3次元空間内の点の集合であり、これらの点は通常、特定の物体や環境の表面を表現する。それぞれの点はXYZ座標を有しており、色情報なども保持することができる。近年、LiDAR (Light Detection and Ranging) や SfM-MVS (Structure-from-Motion and Multi-view-Stereo) といった技術を利用し、高密度な3次元点群データが取得されることが増えてきた。LiDAR点群を取得する方法は、点群を取得する場所や方法によって分類される。地上に設置された Terrestrial Laser Scanner (地上型レーザースキャナ) を使用する方法、自動車などの移動体に搭載された MMS (Mobile Mapping System) を用いて移動しながら取得する方法、さらに UAV (ドローン) やヘリコプター、航空機などの上空から取得する方法などが存在する。これらの方法は、それぞれ異なる用途や環境に応じて利用される。

特に地上型レーザースキャナは、地表面や構造物の高精度なデータを取得するのに非常に有効である。地上型レーザースキャナは、固定された位置から全方位にわたって詳細なスキャンを行うため、細部にわたる高精度な

データが得られる。このため、橋梁や建物の精密なモデル作成に広く利用されている。これらの計測手法を利用して、橋梁などのインフラストラクチャの詳細なモデリングや解析が可能となった。辻井らは、土木構造物の維持管理のために点群を利用して数値解析やモデリングを行うための情報を抽出するための方法を開発した<sup>1)</sup>。米山らは鉄筋結束マシンに取り付けたカメラの動画から点群を生成し、鉄筋の3次元モデルを生成した<sup>2)</sup>。関ら<sup>3)</sup>は橋梁の点検現場作業の支援するために、橋梁の3次元点群データの変化抽出による損傷箇所の検出技術を提案した。その他には、山崎ら<sup>4)</sup>の道路面検出による、床板の高さ測定の自動化などが報告されている。

他分野では軽量な点群データを扱う例も存在するが<sup>5)</sup>、土木分野で扱う点群データは、ファイルサイズが非常に大きく、解析が難しいという課題も存在する。LiDAR点群から物体を検出する手法としては、PointPillars<sup>6)</sup>が広く知られている。また、点群の各点を分類するセマンティックセグメンテーションの手法として PointNet<sup>7)</sup>、PointNet++<sup>8)</sup>、DGCNN<sup>9)</sup>などが存在する。しかしこれらの深層学習技術を用いて点群を解析することには、いくつかの課題が存在する。点群データは一般的にファイル

サイズが大きく膨大な計算資源が必要となる。さらにデータの順番に関わらず同じ対象を表現するなど、従来の2次元画像処理とは異なる特性を持つ。また点群は非構造化データであり、一定の間隔やパターンで配置されていない。各点が独立しており、隣接する点との関係が明確でないため、データ間の空間的関連性を理解するのが難しく、物体認識や分類を行う上で大きな障壁となる。加えて、3次元点群データの解析では、3次元的なアノテーションの作成が必要である。これは2次元画像と比較して、はるかに多くの労力と時間を要する。正確なアノテーションデータを生成するためには、高い精度と専門知識が求められ、これがさらに解析の難易度を高めている。しかし、地上型レーザースキャナでは、点群の取得時にスキャナに内蔵されたカメラにてRGB画像も同時に取得していることが多い。カメラとLiDARのキャリブレーションを実施することで、それらのセンサーフュージョンを行うことができ、画像の各画素と点群の対応関係を取得することができる。そのため点群中の物体を認識する際にカメラ画像の情報を利用できる可能性がある。これにより、点群のみの解析では難しいようなより詳細な情報を自動でできるだけでなく、画像にて物体認識の処理を行うことで、計算負荷を抑えることができる可能性がある。そのため、センサーフュージョンは点群処理の課題である計算資源および詳細な解析という両面を解決できる可能性がある。Liuら<sup>10</sup>はLiDARとカメラのセンサーフュージョンを行い、自動走行のための物体認識を行った。これにより、精度よくかつ、リアルタイム性の高い物体の周辺の物体認識が可能となった。また、Hosoiら<sup>11</sup>は地上型レーザースキャナより取得した点群に対して、475-840 nmの情報を取得できるマルチスペクトルカメラの画像を投影した。マルチスペクトルカメラの情報からクロロフィル量を相関関係により予測した。そしてその情報を点群に反映することで、クロロフィルの3次元的な分布を計算した。また、熱画像と3次元点群のセンサーフュージョンを行い、果樹園の植物の生理的な状態に関わる情報の3次元的な可視化を行った<sup>12</sup>。

画像の画素ごとの分類情報を求めるセマンティックセグメンテーションを行う手法としてはU-Net<sup>13</sup>やDeepLabv3<sup>14</sup>、DeepLabv3+<sup>15</sup>、Segnet<sup>16</sup>などが知られている。劉らは、DeepLabv3+を利用して、橋梁の腐食損傷部のセグメンテーションを行った<sup>17</sup>。さらに、複数のスケールの画像を利用し高解像度の画像から損傷を検出する手法も報告されている<sup>18,19</sup>。また、斎藤らは、DeepLabv3+を利用して農作物の病害の領域のセグメンテーションを行った<sup>20</sup>。これらの手法は高い精度を持ち、これらのネットワークをうまく利用することで点群から対象を効果的にセグメンテーションし、そこから情報を

抽出することが可能となる。また、Segment Anything Model<sup>21</sup>のような画像のセグメンテーションに特化したモデルも存在する。大規模なデータセットで学習し、その学習していない物体に対してもうまく輪郭などを抽出することができる。セグメンテーションのためのモデルやSegment Anything Modelのような学習済みのモデルを併用することで、より精緻な物体の輪郭を抽出できる可能性がある。これまで筆者らは、橋梁を中心として点群を対象にノイズ除去を行ってきた<sup>22</sup>。次の段階として、橋梁などの構造物の情報を取得する必要がある。最も基本的な構造物の情報としては、橋梁の寸法などが挙げられる。このような基本的な情報を高密度な3次元点群から自動的に抽出することができれば、より高度な情報も抽出できる可能性がある。

そこで、本研究では、画像から橋梁の寸法情報の計測につながる情報を自動的に取得し、さらに点群にその情報を投影することで、そのような寸法情報を自動的に取得する手法を提案する。具体的には、地上型レーザースキャナで取得した点群データと、キャリブレーションされたRGB画像を組み合わせて、深層学習技術を活用し、有効幅員や高さなどの基本的な構造情報を抽出することを目指す。

## 2. 実験データと方法

### (1) 点群の計測について

本研究で利用した実験データは2023年12月に福島県石川郡平田村にて取得された。橋梁を主な計測対象とし、LiDARによる計測においては橋梁の床板だけでなく、橋梁から離れた場所や下部からも計測した。そのため、対象とする点群には橋梁だけでなく、樹木や川の水面、斜面なども含まれる。本研究の実験に利用した橋梁の情報を表1に示す。計測には、Matterport Pro3 (Matterport Inc., USA)を使用した。以下Matterportとする。図-1は、Matterportを用いて橋梁をスキャンする際の様子を示している。対象全体をスキャン照射できるように複数の地点

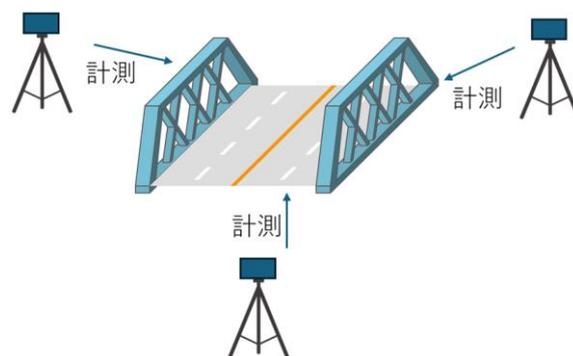


図-1 Matterport Pro3 を用いた橋梁の3次元計測の様子

で計測を行う。Matterport は、屋内に加え、屋外の日光下でも高精度な計測を行うことができる。また、1 度のスキャンで、広域を計測することができ、スキャン時間も 20 秒未満であるため、短時間で広域を計測できるという特徴も有している。さらに、対象を複数の地点から計測し、その点群を統合（レジストレーション）することで、対象のあらゆる角度の 3 次元構造を取得することもできる。この計測データのレジストレーションは付属のソフトウェアにて自動で行われる。レジストレーションされた点群は E57 形式で出力し、本研究の解析に利用された。使用するレーザーはクラス 1 であり、波長は 904 nm、視野は水平方向に、360°、垂直方向に、295°、精度は ±20 mm（対象までの距離 10 m）、深度分解能は毎秒 100,000 点で 1 スキャンあたり 1,500,000 点、最小計測範囲は 0.5m、最大計測範囲は 100m である。また、点群の計測時に画像も取得される。水平方向は前後および左右の 4 方向、上下方向では上下の 2 方向にて画像が取得される。これにより 1 回のスキャンにつき、6 枚の画像が取得される。この画像のサイズは、4096×4096 ピクセルである。

本節の冒頭で説明した橋梁のデータを本研究での訓練データとして利用した。点群ではなく、Matterport にて計測した時に同時に取得された画像を利用した。合計 82 枚を訓練データとして利用した。また、表-1 の H, I, J 橋のデータをテストデータとして利用した。テストにおいては、そこで取得された画像および点群を利用した。テストにおける処理については後述する。

表-1 本論文にて利用した橋梁の情報

| 本論文における<br>橋梁の名称 | 橋長(m) | 幅員(m) | 橋梁形式                                 |
|------------------|-------|-------|--------------------------------------|
| A                | 17    | 8.2   | RC 床版橋<br>逆 T 式橋台                    |
| B                | 14    | 6.6   | H 形鋼桁                                |
| C                | 23    | 6.2   | H 形鋼桁<br>逆 T 式橋台                     |
| D                | 5     | 4.6   | プレテン PC 桁橋                           |
| E                | 19    | 5.4   | H 形鋼桁<br>重力式橋台                       |
| F                | 19    | 3.8   | PCT 桁橋<br>逆 T 式橋台                    |
| G                | 68    | 9.5   | 3 径間単純 PCT<br>桁橋<br>逆 T 式橋台<br>張出式橋脚 |
| H                | 15    | 8.2   | T 桁橋                                 |
| I                | 11    | 9.4   | RC 床版橋                               |
| J                | 8     | 5     | RC 床版橋                               |

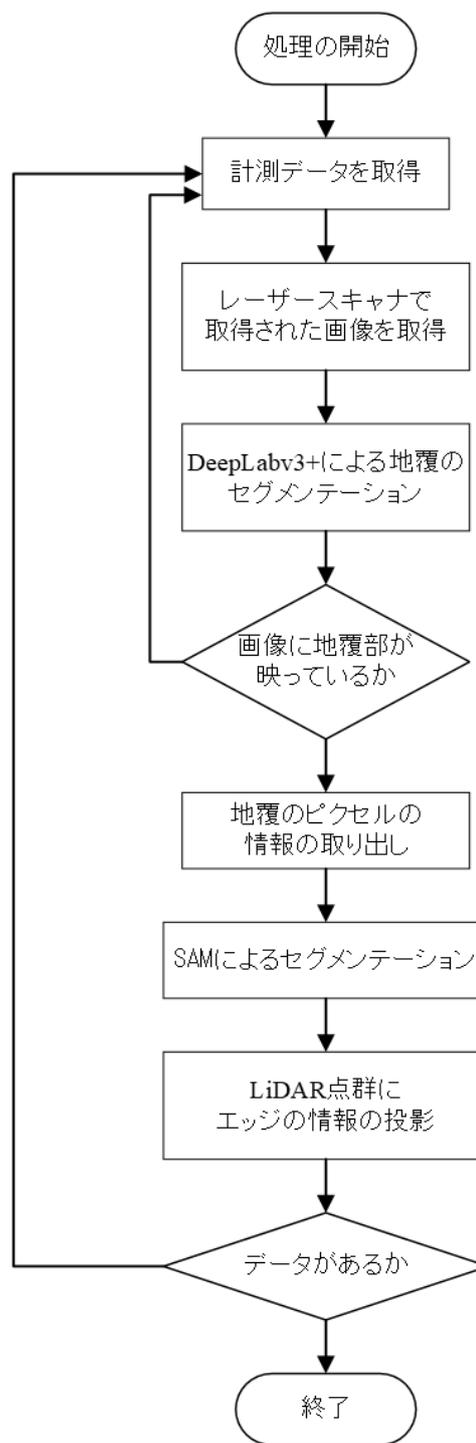


図-2 本研究の解析のフロー図

## (2) 解析の手順

以下に、本研究での解析の手順を述べる。図-2 に解析のフロー図を示す。それぞれの詳細については後述する。本研究では、前節の計測によって、複数の地点から計測した、橋梁の点群および、それに対応する 6 方向の画像を取得した。E57 形式で保存された点群ファイルから、点群の XYZ 座標の情報や、画像を抽出することができる。初めに、各計測点での点群データや画像、キャナ



図3 本研究にてセグメンテーションする対象である地覆および地覆と道路面の境界

の位置や向きなどの情報を取得する。図-2での「計測データを取得」に相当する。その次のステップとしてレーザースキャナに内蔵されたカメラで取得された画像を処理のために読み込む。このステップを「レーザースキャナで取得された画像を取得」として示している。ここで取得した画像に対して、以下の処理を実行した。1点だけでなく、各計測地点のデータに対して、それぞれ実行した。まず各計測地点で取得した、水平方向の画像に対して、DeepLabv3+を利用して地覆のセグメンテーションを行う。地覆は図-3において青色で示されている。また図-3において、赤線で示されている地覆と床板の境界を点群上で特定することを目的とする。この線分の中心どうしを結ぶことで、有効幅員を計算する。スキャナの上方方向の空を映した画像や真下方向の地面の画像には、地覆が映っていないため、利用しないようプログラムした。この処理によって、画像を地覆と、背景の領域に分割することができる。基本的に、画像中には一つの地覆しか存在しない。そのため、分割された地覆の領域は一つしか存在しないはずである。しかし、画像中に地覆が一つしか存在しない場合でも、地覆の領域が複数の領域に分割された形でセグメンテーションされる場合がある。また、セグメンテーションの結果誤った箇所が地覆として判定されることがある。そのようなセグメンテーションの誤差による影響を最小にするために、最も大きな領域のみを地覆の領域とした。また、最も大きな領域を処理するため、両端の地覆が一枚の写真に写っていた場合は、片方の地覆のみ処理される。そして、もう片方は別の計測点で取得された画像に映っている場合、同様の処理が実行される。

そして、地覆のセグメンテーションの精度を向上させるため、上記のセグメンテーションに加えて、SAM (Segment Anything Model) によるセグメンテーションを行う。SAMの入力として、先ほど地覆として判定した領域からランダムで50点選択した。そして、SAMによ

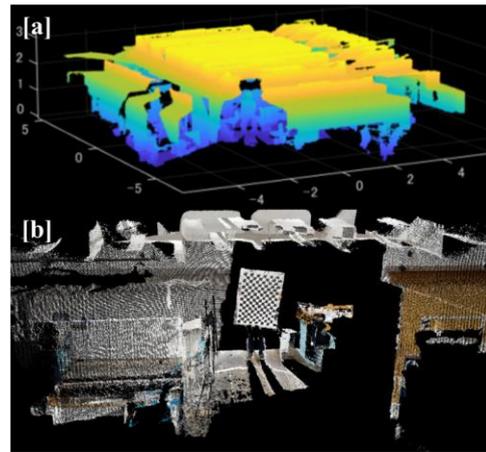


図4 [a] Matterport により取得した点群および、[b] LiDAR-Camera キャリブレーションの結果を利用して点群にカラー画像の情報をマッピングした時の結果。[a]は3次元点群を高さごとに色分けし表示した時の様子を示す。[b]はカメラの色情報を[a]の点群に投影し、色情報を表示した時の結果である。

て地覆であると判定された領域から、有効幅員を抽出するため、地覆と舗装面との接続線を抽出した。ここでの地覆と舗装面との接続線とは、地覆と床板が接する位置であり、図-3において赤色で表示された領域のことである。地覆と舗装面との接続線の抽出は、領域の最下点を取得することで行った。これは、橋を画像で見た際に、地覆と舗装面との接続線は、地覆の下側に存在するためである。こうして取得した画像上での地覆と舗装面との接続線と対応する点群を、LiDAR-Camera キャリブレーションにより取得する。取得した点群に対して、クラスタリングを行い、含まれる点数が多い、2つのクラスタを地覆と舗装面との接続線とみなした。また、各計測点にて同様の処理を繰り返し、全ての計測点での処理が完了した場合、本アルゴリズムの終了となる。図-2では、「データがあるか」に相当する部分で残りの計測点のデータが存在するかを確認するステップを記述している。

本研究での解析には、MATLAB 2024a (MathWorks, USA) を利用した。また利用したPCのCPUは第11世代Intel(R) Core(TM) i7-11800H @ 2.30GHzであり、RAMは64GBである。GPUはNVIDIA GeForce RTX 3070 Laptop GPUを利用した。

### (3) 解析の手順

チェッカーボードを利用し、LiDARとカメラの位置関係を求めた。本実験で使用した手法は、Jiaoら(2023)<sup>29</sup>や板倉ら<sup>29</sup>で用いられている手法を用いた。詳細については板倉ら<sup>29</sup>に記述されている。この手法を利用することで、以下のように点群に画像の情報をマッピングすることができる。以下の図-4の[a]はMatterportにより取得

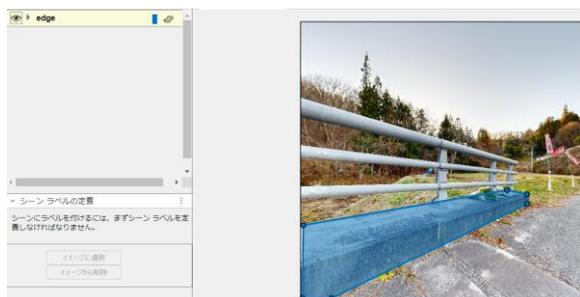


図-5 セマンティックセグメンテーションのためのラベリングを行っている様子。地覆のラベリングを行い、その領域を青で示している。

した点群を示す。ここでは色情報がない点群を表示しているが、形状を認識できるように高さごとに色を変更し、ヒートマップにて示している。しかし、この図においてはヒートマップの色自体は解析には利用されない。また [b]は、LiDAR-Camera キャリブレーションの結果を利用して、点群にカラー画像の情報をマッピングした時の結果である。[a]の画像では色情報がない点群を表示していたが、ここでは画像の色情報が付加されているため点群が色表示されていることがわかる。またカメラの情報を点群に付加できるため、画像のセグメンテーションの結果も同様に点群に付与することができる。本研究では、画像でセマンティックセグメンテーションをして、その情報を点群に投影する。セグメンテーションの方法については、次節以降で説明する。

#### (4) セマンティックセグメンテーションの方法

セマンティックセグメンテーションを行うために、DeepLabv3+を利用した。学習データの作成には、MATLAB の「イメージラベラー」機能を利用した。上記のアプリ起動し、セマンティックセグメンテーションのラベル付けに利用される「ピクセル」ラベルを選択した。そして、図-5 の青枠で囲まれたように、地覆に対して手でラベルを作成した。ラベリングは、82 枚の画像に対して行った。最後に「エクスポート」機能を利用し、ラベリング結果を保存した。このラベリング結果を学習に利用した。

上記に加えて、10 枚にラベリングを行い、テストデータとした。ラベリングした学習データを利用して、DeepLabv3+の学習を行った。そして、テストデータを利用してこのセグメンテーションの精度評価を行った。学習時の計算コストを低減するために、ネットワークの入力サイズは 512×512 pixel とした。また、深層学習では、過学習を防ぐために、データ拡張を行い、データのバリエーションを増やすことがよく行われる<sup>24)</sup>。本研究では、表-2 の範囲内でランダムなデータ拡張を行った。例えば、画像の彩度を-10%から 10%の範囲でランダムに変化させ

た。

学習データに含まれる、地覆と、背景のピクセルの数が等しいことが望ましい。しかし、本研究で使用した学習データには、地覆のピクセルの数が背景に比べ、非常に少なかった。このような不均衡データに対しては、クラスごとのピクセル数の逆数を、損失関数にかけることで、誤分類に対するペナルティの程度を調整することが多い<sup>25)</sup>。本研究でも、同様の手法を採用した。

#### (5) SAM を利用した地覆の情報の微調整について

SAM は、画像のセグメンテーションのためのモデルである。このモデルは、様々な物体を画像から効率的にセグメンテーションすることができ、訓練されていない新しい物体に対しても実行することができる。図-6 は SAM を利用し、画像中の物体をセグメンテーションしたときの結果を示している。SAM の実行は、MATLAB2024a および、MATLAB の Image Processing Toolbox Model for Segment Anything Model を利用した。

SAM の実行においては、背景や対象のピクセルの位置を指定することができる。ここでは、2.4 にて説明したセマンティックセグメンテーションの結果を利用して、対象のピクセルの位置を指定した。

#### (6) 後処理について

2(5)にてSAMの実行をしたのちに、点群に対して、その情報をマッピングする。この処理を複数の計測地点に対して行うと、以下のように地覆に相当する領域以外に、ノイズとして領域が発生する場合がある。その時の様子の例を図-7に示す。図-7は、地覆と舗装面との接続線の候補を示しており、対象ではないノイズのクラスタも含む。ここではそれぞれのクラスタを異なる色で示している。このような地覆と舗装面との接続線以外の領域を除

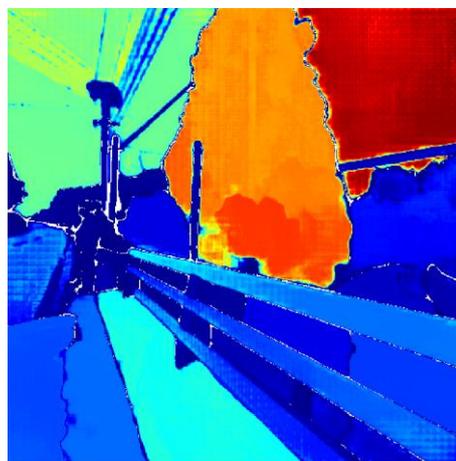


図-6 Segment Anything Model (SAM)を利用して画像のセグメンテーションを行った時の例。各物体がセグメンテーションされており、輪郭も適切に抽出されている。

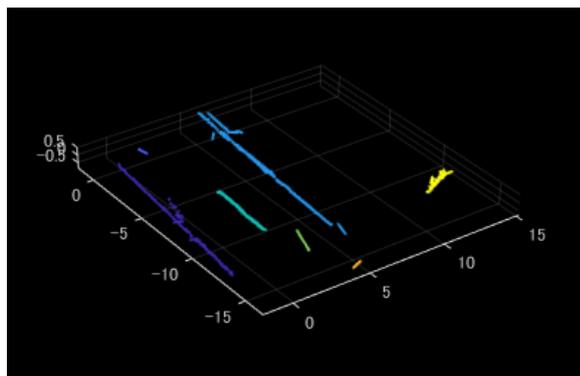


図-7 地覆と道路面の境界の情報を点群にマッピングしたときの様子。距離を利用してクラスタリングを行っており、それぞれのクラスタを異なる色で示している。地覆と道路面の境界以外のノイズのクラスタも含まれていることがわかる。後処理によりこれらのノイズのクラスタを除去する。

去するために以下の処理を行った。それぞれのクラスタに分割し、その点群に対して、主成分分析を利用して、直線性を計算した。主成分分析は、点群の特徴量として最もよく使われる手法の一つである<sup>29)</sup>。以下のような計算により、直線性の低いものをノイズのクラスタとして除去した。

点群データに対して主成分分析を行う際、まずはそのクラスタの点群の XYZ 値それぞれの平均値を各座標から差し引き、データの中心化を行う。次に、中心化されたデータの共分散行列を計算し、その固有値と固有ベクトルを求める。3次元の XYZ 座標系における点群の共分散行列  $C$  は以下のように表される (式(1))。ここで、 $\sigma_{ij}$  は座標  $i$  と  $j$  の共分散を示す。

表-2 DeepLabv3+によるセマンティックセグメンテーションの学習のためのデータ拡張におけるパラメータ設定

| データ拡張の方法 | 範囲        |
|----------|-----------|
| 彩度       | ±10%      |
| 明度       | ±30%      |
| スケール     | 0.5~1     |
| 回転       | -45° ~45° |
| 色相       | ±20%      |
| コントラスト   | ±40%      |

表-3 DeepLabv3+の学習のために利用されたパラメータ

| パラメータの種類  | 値   |
|-----------|---|
| 学習率       | 0.0001                                      |
| 最適化アルゴリズム | SGDM (Stochastic Gradient Descent Momentum) |
| ミニバッチサイズ  | 4   |
| エポック数     | 200   |

$$C = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \quad (1)$$

ここで得られる固有値を大きい順に  $\lambda_1, \lambda_2, \lambda_3$  とする。式(2)のように直線性 (Linearity) を定義する。

$$Linearity = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \quad (2)$$

この値が 0.95 以上であれば、そのクラスタの点群が高い直線性を持つとみなし、その領域は地覆と床板の境界とする。この閾値 (0.95) を大きくすると、より直線性の高いクラスタのみが抽出される一方で、見逃しのリスクも高まる。

### 3. 結果と考察

#### (1) セグメンテーションの結果について

図-8 に DeepLabv3+にて Matterport により取得した画像のセマンティックセグメンテーションを行った時の様子を示す。[a]が入力画像を、[b]がセグメンテーション結果を示している。[b]の紫色の領域が地覆を示しており、水色の領域はそれ以外の背景の部分を示している。[b]では、主に地覆をセグメンテーションしていることがわかる一方で、その周辺の地覆ではない場所もセグメンテーションしていることがわかる。本手法では、画像中の地覆の最下点を地覆と床板の境界から有効幅員を求める。そのため、境界の情報は正確である必要がある。そのため、ここでのセグメンテーション結果を点群に投影するのではなく、SAM などの方法により境界部の調整を行う。

また、予測ラベルと手動で作成した正解ラベルを比較することで精度検証を行った。表-4において Positive が地覆、Negative がそれ以外を表している。どちらの IoU も 0.9 を超えており、精度よくセグメンテーションを行うことができた。図-9 は、表-4 の精度検証の結果を混同行列で示したものである。Positive と Negative の割合を見ると、画像中には、地覆よりも、それ以外の領域の方が多

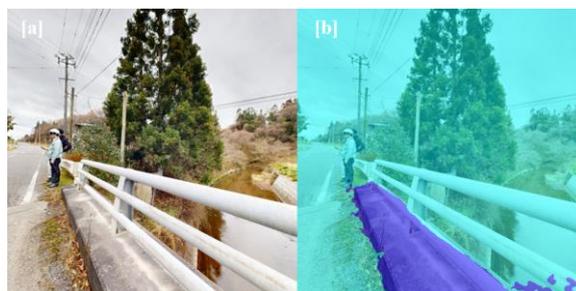


図-8 DeepLabv3+にて Matterport により取得した画像のセマンティックセグメンテーションを行った時の様子

|       |          |          |          |
|-------|----------|----------|----------|
| 真のクラス | Negative | 2390528  | 60416    |
|       | Positive | 24576    | 145920   |
|       |          | Negative | Positive |
|       |          | 予測されたクラス |          |

図-9 セマンティックセグメンテーションの結果をもとに作成した混同行列. Positive が地覆であり, Negative が背景を示す.

く映っているなど, データに偏りがあったことがわかる. しかし Positive も Negative も精度よく推定することができた.

表-5 はデータ拡張を実施せずに学習を行い, セグメンテーションした結果である. 地覆が精度よく抽出することができなかった (図-10). 訓練データに過適合し, 新規のデータに対してうまくセグメンテーションできなかったことが考えられる. このことから, データ拡張を行わなかった場合は, 過学習が生じており, 本手法で採用したデータ拡張は正しい処理であったと示唆される.

## (2) 地覆の抽出について

図-11 の[a]が 3.1 で示したセマンティックセグメンテーションの結果で[b]が SAM で地覆と舗装面との接続線の



図-10 データ拡張を行わずに学習をした Deeplabv3+モデルで推論をした場合の結果の例



図-11 [a] Deeplabv3+を利用したセマンティックセグメンテーションの結果. [b] Segment Anything Model (SAM)で領域の境界を調整した時の様子.

表-4 セマンティックセグメンテーションの結果

|          | Accuracy | IoU    | MeanBFScore |
|----------|----------|--------|-------------|
| Positive | 91.95 %  | 0.6634 | 0.5508      |
| Negative | 98.06 %  | 0.9767 | 0.8437      |

表-5 データ拡張を行わなかった場合のセマンティックセグメンテーションの結果

|            | Accuracy | IoU    | MeanBFScore |
|------------|----------|--------|-------------|
| Bridge     | 46.09 %  | 0.3634 | 0.3383      |
| Background | 98.65 %  | 0.9605 | 0.8388      |

認識結果を調整した結果である. [a]では, 地覆の周辺もセグメンテーションされていた一方で, 地覆と舗装面との接続線付近がより正確にセグメンテーションされていることがわかる. SAM を利用することで, 地覆と床板の境界をより正確に抽出することができたと考えられる.

また, DeepLabv3+を利用したセマンティックセグメンテーションや SAM で領域の境界を調整した結果を利用し, 地覆と床板の境界を抽出した時の例を図-12 に示す. 地覆の下部の部分を取得することでうまく境界部分を取得することができた. この結果を点群に投影することで, 点群でこの境界部分を特定することができる.

本研究では, 画像から地覆をセグメンテーションし, さらにその領域を SAM により微調整した. 微調整した領域で最も下の箇所を床板と地覆の境界とした. SAM は, 10 億個以上のマスクを含む非常に大規模なデータセットで訓練されている. この巨大なデータセットは, 多様なシーンと物体の種類をカバーしており, それにより地覆もうまく背景と分離できたと考えられる. 大量のトレーニングデータによって, モデルはより正確な地覆と舗装面との接続線の認識とセグメンテーションを実現する. このように, 本研究の手法では複数の手法を組み合わせている. しかし, そのような境界を直接的にセグメンテーションする方法も考えられる. 図-13 は, そのような境界をラベリングし学習させた場合の結果である. 直接的に境界部をセグメンテーションする方法も試したが, うまくセグメンテーションすることができず, 本研究のような手法を採用した. 地覆のほうが境界部よりも面積が大きく, かつ特徴も画像から比較的わかりやすく, セグメンテーションのための学習が可能であったためだと考えられる. また, 直線を検出するための深層学習ネットワークも存在する<sup>25)</sup>. 本研究では, 地覆をセグメンテーションすることで, 床板と地覆の境界を検出したが, このように直接的に直線を検出することも考えられる. 今後の課題として, このような他の手法とも比較し, さらにより多くのデータで訓練・検証することで, より信頼度の高いセグメンテーション手法を確立する必要がある



図-12 DeepLabv3+を利用したセマンティックセグメンテーションや Segment Anything Model (SAM)で領域の境界を調整した結果を利用し、地覆と道路面の境界を抽出した時の例。

る。

点群処理のために、画像情報を利用する例として、点群の各点を画像に投影する方法も知られている<sup>28,29)</sup>。ここでは、画像の情報ではなく、点群を2次元平面に投影し、その情報を利用して、対象の物体であるかどうかを判別している。点群の情報を投影する処理は、画像情報が利用できない状況でも行うことができ、かつ2次元の画像処理の手法を利用することができる。その他に、航空機 LiDAR より取得したデータから樹木を検出するために、点群を鳥瞰図のように変換する方法や<sup>30)</sup>、植物の葉の勾配を利用して歯の面積を計算するといった研



図-13 地覆と道路面の境界を正解データとしてラベリングし直接的にセマンティックセグメンテーションをした時の結果の例。適切に対象の領域をセグメンテーションすることができなかった。

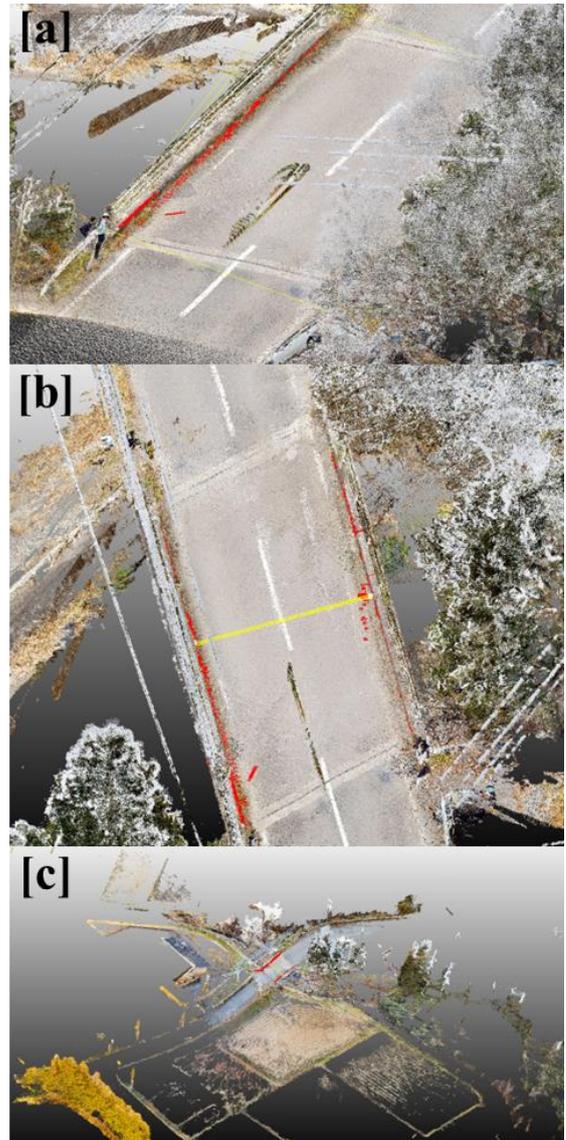


図-14 地覆と道路面の境界をセグメンテーションした時の様子。[a]および[b]は斜め上から対象の橋梁の床板を見た様子。[c]はより広域のビューを示す。

究も報告されている<sup>30)</sup>。一方で、本手法のように点群計測時に撮影された画像が利用できれば、線などの、より詳細で細かい情報まで取得できるというメリットがある。

本研究では、前後左右の4方向の画像をもとにLiDARとカメラのセンサーフュージョンを行った。全方位カメラから物体を検出し、写真測量と組み合わせることで3次元点群での物体検出を行った研究も知られている<sup>32)</sup>。全方位画像などの異なる種類の画像に対しても本手法を適用することが可能である。全方位画像では、1つの計測地点において解析する画像は1枚でよいいため、より計算が効率化する可能性がある。

### (3) 地覆の抽出について

以下の図-14 [a] では本手法により抽出した境界線が赤色で示されている。境界部がうまくセグメンテーション

されていることがわかる。様々な物体が存在する中でうまく対象の線分を取得することができた。また、[b]は上の視点から見たときの様子を示す。[c]はより広い視点で点群を表示させたときの様子を示す。多くの対象がある中からうまく線を抽出することができていることがわかる。本手法により自動的に予測した有効幅員は 8.13 m であり、手動で計測した値は 8.10 m であった。しかし、幅を計測する際の点により数 cm から数十 cm ほど変動することが考えられる。そのため今後の課題として、真値を計測するときの何らかの基準を設け、それにより計測した値と自動的に予測した値の比較を行う必要がある。また橋長に関しては、地覆と床板の境界の情報を利用して予測することができる。自動的に計算した結果は、11.71 m および 13.87 m であり、手動で計測した結果は、11.58 および 11.49 m であった。本手法では、地覆を過検出した場合があり、その結果、橋長が過大評価される場合があった。別のデータでの推論結果を以下の図-15 に示す。図-15 は J 橋のデータを利用している。

点群データは 1GB 以上の大容量になることが多い。ファイルサイズが大きくなるほど、一般的に点群の処理は難しくなる。本手法では計測点ごとに処理を行うため、そのような大容量データの処理も可能である。また、そのような点群データを閲覧用のソフトウェアにて開き、点の選択を手動で行う作業は非常に手間がかかる。一方で、本手法ではその作業を大幅に簡略化および自動化することができる。

画像から点群の中の物体を検出する方法として、ステレオカメラを利用する方法がある<sup>39)</sup>。この方法では、画像上で物体を検出し、その検出した情報を点群に反映させる。ステレオカメラは画像からステレオ画像法により 3 次元点群を得るため、センサーフュージョンの必要がなくシンプルである。しかし、計測できる範囲に限界があり、橋梁全体の解析を行う場合には難しい。小さな領域で対象が 1 つ程度の場合には有効な方法である。

RANSAC (Random Sample Consensus)<sup>39)</sup>を利用することで点群から平面を検出することができる。RANSAC は、外れ値を含むデータセットからモデルをフィットさせるためのアルゴリズムである。この手法は 3 次元点群にも応用することができる。3 次元点群から平面を検出するために、ランダムに点を 3 点以上選択し、その点を利用して平面を作成する。そして、点群の全ての点に対して、その点が定義した平面上または十分近い距離にあるかどうかを判断する。そのために、点と平面の距離を計算し、設定した閾値以下の点をインライアとみなす。本研究では、RANSAC を発展させた手法である、MLESC<sup>39)</sup>を使用して、平面抽出を行った。図-16 は、本研究で利用したデータから平面を検出した時の結果である。点群に対して、オレンジ色にて検出した平面を可視化している。

対象の点群が切り取られていたり、限定されていたりすると、うまく平面が検出でき、平面的な構造を有する物体が床板のみである場合は、そのような方法を利用して、床板を検出することができる。しかし今回のように点群全体に対して処理を行うとうまく床板のみを選択的に検出することができず、有効幅員の計測なども困難となる。

本研究では、画像上でセグメンテーションする部分を特定し、その情報を点群にマッピングすることで有効幅員の計測を行った。画像データはピクセル単位で非常に細かく対象の形状や輪郭などの情報が得られる一方で、点群データはスキャンの解像度やノイズの影響を受けやすい。特にスキャンの範囲や角度によりデータが不完全になることが多いため、物体の正確な検出が難しくなることも多い。また、画像処理には多くの既存のツールやアルゴリズムが利用可能で、そのような手法も利用することで、点群のセグメンテーションのための、画像処理を効率的に行うことができる。本研究では、地上型レーザースキャナの点群計測において画像も同時に取得されていることを利用して、それらのセンサーフュージョンを行った。それにより精度の良い点群のセグメンテーションを行い、橋梁の点群から有効幅員に相当する値を自動的に取得することができた。

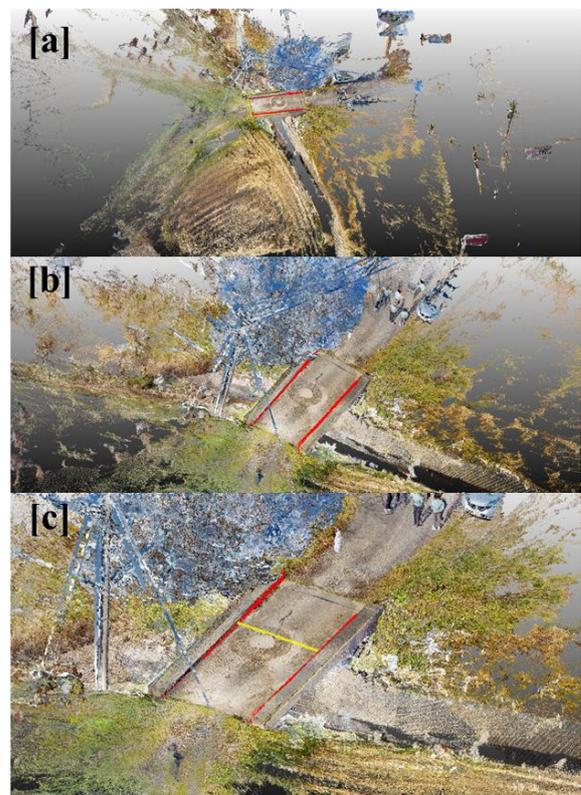


図-15 J 橋のデータに対して境界線をセグメンテーションした時の様子。[a]は対象の橋梁およびその周辺を閲覧した時の様子。[b]および[c]は対象の橋梁を中心に閲覧した時の様子を示す。[c]では、幅を計測するために選択された点を結んだ線分を黄色で示している。

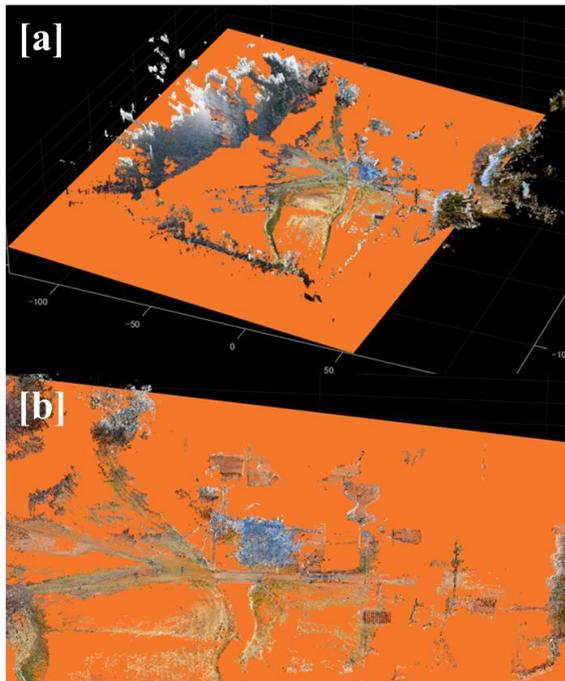


図-16 MLESAC 法を利用して対象の点群から平面を検出した時の様子。[a]は対象の点群に対して平面を検出した時の様子を示す。検出した平面全体がオレンジ色で示されている。[b]はより橋梁の床板に近いビューである。

#### 4. まとめ

本研究では、地上型レーザースキャナにより取得された点群と画像を組み合わせ、センサーフュージョンを行った。これにより、画像上でセグメンテーションし、特定した対象の情報を点群にマッピングすることができる。また点群上にマッピングした情報を利用して、有効幅員の計測を行った。

画像データはピクセル単位で非常に細かく対象の形状や輪郭などの情報が得られ、特定のオブジェクトの形状を詳細に理解する上で非常に有用である。一方、点群データはスキャンの解像度やノイズの影響を受けやすいため、一部の詳細が失われる可能性がある。しかし、これら二つのデータタイプを組み合わせることで、画像データの詳細性と、点群データの3次元的な情報を把握する能力を活用することができた。また、大きなファイルサイズを有する点群データの処理は一般的に難しいと考えられている。しかし、本研究では、計測点ごとに処理を行うことで、大容量のデータでも効率的に処理を行うことが可能となった。この手法の適用により、橋梁の点群から幅に相当する値を自動的に取得することが可能となった。さらに、本手法は橋梁以外の土地の境界線などその他の線分を点群から自動抽出し、シェープファイルなどの形式で出力することも可能である。これらの結果は、

地上型レーザースキャナの点群計測と画像取得の組み合わせ、そしてセンサーフュージョンにより、精度の良い点群のセグメンテーションや点群からの高度な情報抽出を実現する新たな可能性を示している。本研究では、1点に設置して点群を取得するタイプのレーザースキャナを利用した。移動しながらリアルタイムで点群を取得していく計測方法もある。この方法は自動運転などの用途が多い。ここでも画像を同時に取得し、画像と点群の情報を同期させ、センサーフュージョンを行うことで同様に物体検出が可能である。物体検出だけでなく、本研究のようなインフラ管理につながる解析もできる可能性がある。近年は、地上型レーザースキャナだけでなく、スマートフォンやタブレットなどの端末に搭載されたLiDARを使用した研究も報告されるようになった<sup>30)</sup>。将来の展望として、このような様々なLiDARのタイプにも適用するなど、本手法の適用範囲を広げ、様々なシーンで利用可能になるよう発展させていくことが望まれる。また、3次元点群からひび割れといった損傷を検出するために、SfM-MVSと画像認識を組み合わせる研究も報告されている。画像から損傷を検出し、さらにその画像撮影した位置や向きなどを利用し、その情報を点群に反映させる。このように、3次元点群からの情報抽出を行うために、様々なセンサーや計測条件を利用する研究も近年報告されている<sup>37,38)</sup>。これまで多く報告されてきた点群のみを入力として点群処理する方法とは異なり、センサーフュージョンやそれに類する手法が今後発展していく可能性が示唆される。

提案手法は、高解像度の点群データと画像データを組み合わせることで、橋梁下面の上部工の形状などの接近が困難な箇所の詳細な形状情報も取得できる。本研究では地上型レーザースキャナのデータを利用して検証を行ったが、本手法はUAV搭載型LiDARなどの別の種類のLiDARにも応用が可能である。このように、様々な種類のデータに対して適用することで、橋梁の全体的な健全性評価やメンテナンス計画にもつながる可能性がある。さらに、他の構造物や地形の解析にも応用でき、例えば3次元点群から構造物のエッジを自動抽出し、図化するといった課題にも利用できる可能性がある。また、土木分野でも仮想空間上で情報を加工して現実世界に反映するデジタルツインの概念が提唱されており研究開発が盛んに実施されている<sup>39)</sup>。このようなデジタルツインを作成するためにも、現実の3次元点群から情報を取得し、デジタルツインの構築に役立てることが考えられる。デジタルツインに必要な詳細な情報を本研究のような手法を利用し、抽出するといったアプリケーションも考えられる。

しかしながら、いくつかの課題も存在する。より複雑な対象を分析する際は、より高度な画像認識が求められ

る。また、より大規模で高密度な点群データと画像データの統合には、十分な計算リソースが必要である。さらに、より精度を高めるためには実際の橋梁環境におけるノイズや外乱の影響を最小限に抑えるための計測における工夫なども必要となる。このような課題を解決するためにも、より多くの対象にて本手法を検討し、より汎用性の高いアルゴリズムに発展させていくことが望まれる。

**謝辞**：本稿は、JSPS 科研費 JP21H01417 および 22H01561 の助成を受けた研究で行われた結果を含みます。また、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) 第 3 期「スマートインフラマネジメントシステムの構築」JPJ012187 (研究推進法人:土木研究所) で実施されている研究も含みます。これらに謝意を表します。

## 参考文献

- 辻井純平, 合田哲朗, 中野雅章: 土木構造物の点群解析に向けた局所形状の畳み込みを伴う深層学習手法の適用, AI・データサイエンス論文集, 4 巻 3 号, pp. 442-450, 2023.
- 米山睦美, 高見澤拓哉, 眞部達也, 田尻大介: 鉄筋結束マシンに取り付けたカメラ動画から生成した点群による配筋検査への実用性検証, AI・データサイエンス論文集, 5 巻 1 号, pp. 239-244, 2024.
- 関和彦, 山口愛加, 窪田 諭: 3 次元点群データを用いた道路橋の損傷抽出とヒートマップ表示, 土木学会論文集, 79 巻 10 号, 2023.
- 山崎文敬, 前田幸祐, 柿市拓巳: 橋梁コンクリート床版平坦性自動検査システムの開発, AI・データサイエンス論文集, 5 巻 1 号, pp. 26-32, 2024.
- Kamiwaki, Y., and Fukuda, S.: A Machine Learning-Assisted Three-Dimensional Image Analysis for Weight Estimation of Radish, *Horticulturae*, Vol. 10, Issue 2, pp. 142, 2024.
- Lang, A. H., Vora, S., Caesar, H., Zhou, L., Yang, J., and Beijbom, O.: Pointpillars: Fast encoders for object detection from point clouds, *CVPR*, pp. 12697-12705, 2019.
- Qi, C. R., Su, H., Mo, K., and Guibas, L. J.: Pointnet: Deep learning on point sets for 3d classification and segmentation, *CVPR*, pp. 652-660, 2017.
- Qi, C. R., Yi, L., Su, H., and Guibas, L. J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space, *NeurIPS*, 30. (2017).
- Phan, A. V., Le Nguyen, M., Nguyen, Y. L. H., and Bui, L. T.: DGCNN: A convolutional neural network over large-scale labeled graphs, *Neural Networks*, Vol. 108, pp. 533-543, 2018.
- Liu, H., Wu, C., and Wang, H.: Real time object detection using LiDAR and camera fusion for autonomous driving, *Sci. Rep.*, Vol. 13, pp. 8056, 2023.
- Hosoi, F., Umeyama, S., and Kuo, K.: Estimating 3D chlorophyll content distribution of trees using an image fusion method between 2D camera and 3D portable scanning lidar, *Remote Sens.*, Vol. 11, Issue 18, pp. 2134, 2019.
- Narváez, F. J. Y., del Pedregal, J. S., Prieto, P. A., Torres-Torriti, M., and Cheein, F. A. A.: LiDAR and thermal images fusion for ground-based 3D characterisation of fruit trees, *Biosys. Eng.*, Vol. 151, pp. 479-494, 2016.
- Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, *MICCAI*, pp. 234-241, 2015.
- Chen, L. C., Papandreou, G., Schroff, F., and Adam, H.: Rethinking atrous convolution for semantic image segmentation, arXiv:1706.05587, 2017.
- Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation, *ECCV*, pp. 801-818. 2018.
- Badrinarayanan, V., Kendall, A., and Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation, *TPAMI*, Vol. 39, Issue 12, pp. 2481-2495, 2017.
- 劉佳明, 党紀, 全邦釘: DeepLabv3+を用いた橋梁腐食損傷とその精度の向上, AI・データサイエンス論文集, 3 巻 J2 号, pp. 802-810, 2022.
- Deng, L., Yuan, H., Long, L., Chun, P. J., Chen, W., Chu, H.: Cascade refinement extraction network with active boundary loss for segmentation of concrete cracks from high-resolution images, *Autom. Constr.*, Vol. 162, pp. 105410, 2024.
- Chu, H., and Chun, P. J.: Fine-grained crack segmentation for high - resolution images via a multiscale cascaded network. *Comput. Aided Civ. Infra. Eng.*, Vol. 39, No.4, pp. 575-594, 2024.
- 斎藤嘉人, 板倉健太, 山本一哉, 二宮和則, 近藤直: 可視・近赤外画像のセマンティックセグメンテーションによるパレイシヨ塊茎表面の病害検出, AI・データサイエンス論文集, 3 巻 J2 号, pp. 175-181, 2022.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W. Y., Dollár, P., Girshick, R.: Segment anything, *ICCV*, pp. 4015-4026, 2023.
- 板倉健太, 林拓哉, 上脇優人, 全邦釘: LiDAR とカメラのセンサーフュージョンによる点群からのノイズ除去, AI・データサイエンス論文集, 受理済み, 2024.
- Jiao, J., Chen, F., Wei, H., Wu, J., and Liu, M.: LCE-Calib: Automatic lidar-frame/event camera extrinsic calibration with a globally optimal solution, *IEEE/ASME Transactions on Mechatronics*, Vol. 28, No. 5, pp. 2988-2999, 2023.
- Wang, H., Xiao, C., Kossaifi, J., Yu, Z., Anandkumar, A., and Wang, Z.: Augmax: Adversarial composition of random augmentations for robust training, *NIPS*, Vol. 34, pp. 237-250, 2021.
- Ho, Y., and Wookey, S.: The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling, *IEEE access*, Vol. 8, pp. 4806-4813, 2019.
- 増田宏: 大規模点群のための処理技術, 計測と制御, 60 巻 10 号, pp. 716-720, 2021.

- 27) Pan, X., Shi, J., Luo, P., Wang, X., and Tang, X.: Spatial as deep: Spatial CNN for traffic scene understanding, *AAAI*, Vol. 32, No. 1, pp. 7276-7283, 2018.
- 28) Itakura, K., and Hosoi, F.: Three-dimensional tree monitoring in urban cities using automatic tree detection method with mobile LiDAR data, *AI Data Sci.*, Vol. 2, Issue 2, pp. 1-10, 2021.
- 29) Itakura, K., and Hosoi, F.: Automated tree detection from 3D lidar images using image processing and machine learning, *Appl. Opt.*, Vol. 58, Issue 14, pp. 3807-3811, 2019.
- 30) 板倉健太, 細井文樹: 画像処理や 3 次元深層学習を用いた航空機ライダー点群データからの樹木の検出, *AI・データサイエンス論文集*, 1 巻 J1 号, pp. 320-328, 2020.
- 31) Itakura, K., and Hosoi, F.: Voxel-based leaf area estimation from three-dimensional plant images. *J. Agri. Meteorol.*, Vol. 75, Issue 4, pp. 211-216, 2019.
- 32) Itakura, K., and Hosoi, F.: Automatic tree detection from three-dimensional images reconstructed from 360 spherical camera using YOLO v2, *Remote Sens.*, Vol. 12, Issue 6, pp. 988, 2020.
- 33) Itakura, K., and Hosoi, F.: Simple and effective tool for estimating tree trunk diameters and tree species classification, *Appl. Opt.*, Vol. 59, Issue 2, pp. 558-563, 2020.
- 34) Fischler, M. A., and Bolles, R. C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Com. ACM*, Vol. 24, Issue 6, pp. 381-395, 1981.
- 35) Torr, P. H., and Zisserman, A., MLESAC: A new robust estimator with application to estimating image geometry, *Comput. Vis. Image Und.*, Vol. 78, Issue 1, pp. 138-156, 2000.
- 36) 渡邊祥庸, 井上和真, 池田隆明, 志賀正崇, 小林雅人, 横山和佳奈: スマートフォンLiDARとネットワーク型RTK測位により作成した3次元地形モデルの小規模土工事への適用範囲の検討, *AI・データサイエンス論文集*, 5 巻 1 号, pp. 260-268, 2024.
- 37) Yamane, T., Chun, P. J., Dang, J., and Honda, R. (2023). Recording of bridge damage areas by 3D integration of multiple images and reduction of the variability in detected results. *Comput. Aided Civ. Infra. Eng.*, Vol. 38, Issue 17, pp. 2391-2407.
- 38) Yamane, T., Chun, P. J., and Honda, R.: Detecting and localizing damage based on image recognition and structure from motion, and reflecting it in a 3D bridge model. *Struct. Infrastruct. Eng.*, Vol. 20, Issue. 4, pp. 594-606, 2024.
- 39) 藤原圭哉, 佐藤誠, 山下千智, 黒田直樹, 亀田敏弘: 3Dデータと河床変動解析を活用した河川分野におけるデジタルツインの実現に向けての提案, *AI・データサイエンス論文集*, 5 巻 1 号, pp. 126-133, 2024.

## Semantic segmentation and sensor fusion in point cloud processing for estimation of bridge structural parameters

Kenta ITAKURA, Takuya HAYASHI, Yuto KAMIWAKI and Pang-jo CHUN

In this study, we implemented sensor fusion by combining point cloud data and image data obtained from a terrestrial laser scanner, and mapped the segmentation results from the images onto the point clouds. Semantic segmentation of the images was performed using DeepLabv3+ to classify into wheel guard and background. Also, the edge information was updated using the Segment Anything Model, then segmentation information was stored in the point clouds using the camera's external and internal parameters. Utilizing this information enabled the measurement of bridge widths. By leveraging the detailed information from the image data and the three-dimensional information from the point cloud data, we were able to achieve an analysis that extract detail and structural information, while also efficiently processing large files of point cloud data.