

## 1. はじめに

セマンティックセグメンテーションは、画像内の全画素や点群の各点に対してラベルやカテゴリを関連付けるタスクである。学習には大量の実環境データが必要であり、不十分な場合は精度の低下を招く。そこで、CG により大量のデータを生成して学習に用いることで、実環境データを補填することが期待される。本研究では、点群データに対するセマンティックセグメンテーションにおいて、CG データによる事前学習が有効であるか検証する。

## 2. セマンティックセグメンテーション

点群データに対するセマンティックセグメンテーションの代表的な手法である SalsaNext[1] は、3 次元点群を疑似画像へ変換し処理を行う。SalsaNext は Context Module, Pixel-Shuffle Layer, Dilated Convolution を導入し、識別精度の向上とリアルタイム性を実現している。SalsaNext のネットワーク構造を図 1 に示す。

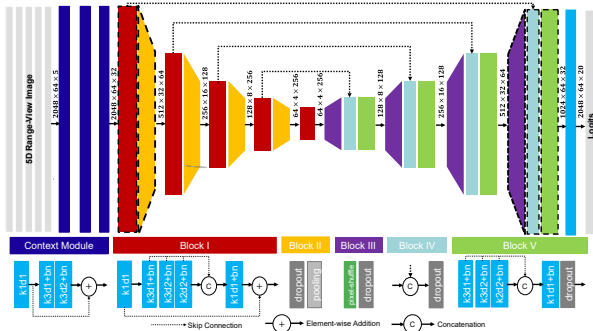


図 1: SalsaNext のネットワーク構造

## 3. 提案手法

本研究では、CG データセットを用いて事前学習モデルを作成し、実環境データセットでファインチューニングすることで事前学習の有効性を検証する。

### 3.1 CARLA を用いた CG データセット作成

セマンティックセグメンテーションの学習において、データセットが小規模であると、過学習などにより汎化能力が向上しない。そこで、実環境データセットに類似したデータを CG で作成する。CG データセットを用いて事前学習することで汎用的なモデルを獲得する。そして、実環境データセットでファインチューニングすることで実環境のドメインに適応させる。本研究では、CARLA を用いて工事現場を模したマップを作成する。作成したマップ例を図 2(a) に示す。マップ上では、ランダムに盛り土や建機、人を配置して点群を取得する。図 2(a) のマップで取得した点群データを図 2(b) に示す。生成したマップ数は 7、フレーム数は 3458 である。クラスは Background, Sandpile, Vehicle, Human である。

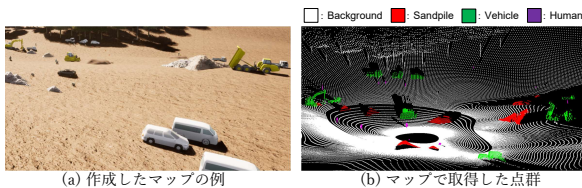


図 2: 取得した点群データの例

### 3.2 事前学習

事前学習モデルとは、大規模なデータセットを用いて学習したモデルである。本実験では、CG データセットを用いて学習を行い、事前学習モデルを作成する。

### 3.3 実環境データでの学習

CG データセットを用いて事前学習モデルを作成した後、実環境データセットを用いてファインチューニングを行い、実環境データに対する精度の向上を図る。実環境データは、工事現場で取得した点群データである。実環境データセット内のフレーム数は 383 である。

## 4. 評価実験

本実験では、事前学習モデルの有無による精度を比較する。

### 4.1 実験概要

事前学習モデルの学習には CG データセット内の 5 シーン (2568 フレーム) を用いる。学習回数は 100 エポックである。学習時のネットワークに SalsaNext、最適化手法に Adam、初期学習率は  $1.0 \times 10^{-4}$  とする。事前学習モデルの評価には 2 シーン (890 フレーム) を用いる。ファインチューニングには実環境データセット内の 350 フレームを用いる。学習回数は 300 エポックである。学習時のネットワークに SalsaNext、最適化手法に Adam、初期学習率は  $1.0 \times 10^{-4}$  とする。評価には残りの 33 フレームを用いる。評価指標には IoU を用いる。

### 4.2 実験結果

事前学習モデルを CG データで評価した結果を表 1 に示す。表 1 の Vehicle は、他クラスよりも精度が低い結果である。事前学習モデルの有無による評価結果を表 2 に示す。表 2 より事前学習モデルを使用することで Background は 0.4pt, Sandpile は 4.4pt, Vehicle は 1.6pt, Human は 24.6pt, Average は 7.8pt 精度が向上した。事前学習モデルの有無による検出結果を図 3 に示す。図 3 より、事前学習モデルを使用することで、Human の認識率が向上した。

表 1: 事前学習モデルの定量的評価

| Background | Sandpile | Vehicle | Human | Average |
|------------|----------|---------|-------|---------|
| 96.7       | 59.5     | 31.8    | 61.4  | 62.3    |

表 2: 事前学習モデルの有無による定量的評価の比較

| 事前学習 | Background  | Sandpile    | Vehicle     | Human       | Average     |
|------|-------------|-------------|-------------|-------------|-------------|
|      | 93.9        | 52.1        | 34.3        | 0.0         | 45.0        |
| ✓    | <b>94.3</b> | <b>56.5</b> | <b>35.9</b> | <b>24.6</b> | <b>52.8</b> |

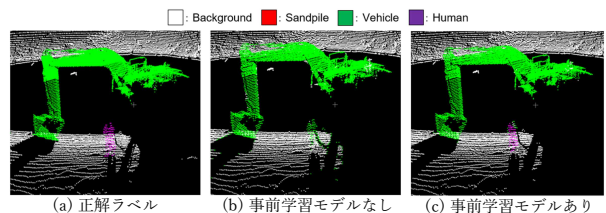


図 3: 事前学習モデルの有無による検出結果の比較

## 5. おわりに

本稿では、点群データに対するセマンティックセグメンテーションにおける CG データによる事前学習の有効性を検証した。評価実験により、CG データを用いた事前学習は有効であることを確認した。今後は、クラス間の点群数の不均衡問題について改善する方法を検討する。

## 参考文献

- [1] Tiago Cortinhal, et al, "SalsaNext: Fast, Uncertainty-aware Semantic Segmentation of LiDAR Point Clouds for Autonomous Driving", ISVC2020