

1. はじめに

自動運転では、全方位 LiDAR で取得した 3 次元点群から、物体検出やセマンティックセグメンテーションを行う。國貞らは、全方位 LiDAR でのデータ取得中に逐次処理可能な、1Dimensional-Convolutional Neural Network (以下 1D-CNN) による歩行者検出 [1] を提案した。本手法は、検出対象が歩行者のみであるため、車両周辺の詳細な状況を把握するために、本研究では、1D-CNN をセマンティックセグメンテーションに応用する。

2. 1D-CNN

1D-CNN による歩行者検出では、LiDAR から取得した距離値を 1 次元波形データとみなし、1D-CNN に入力する。そして、各レーザ ID に対して、歩行者か否かを判定している。これにより、LiDAR の回転に合わせて逐次処理できるため、撮影から処理完了までの時間差による物体の位置ずれを最小限に抑えることを可能としている。

3. 提案手法

従来手法は、LiDAR から得られる距離値データを 1D-CNN に入力し、歩行者のみを検出対象としている。本研究では、複数のクラスに対応するために、全方位 LiDAR による撮影中に得られる反射強度も入力データとして用いる。そして、歩行者以外のクラスにも対応するために、1D-CNN をセマンティックセグメンテーションタスクに拡張する。

3.1. 反射強度の追加

反射強度は、金属など反射率の高い物質では光が拡散するため弱くなり、布など反射率の低い物体では強くなるといった性質がある。これにより、物体の質感を考慮した識別が可能となる。また、反射強度は距離の 2 乗に反比例して減衰する。そのため、距離値を用いて反射強度を式 (1) により正規化する。

$$I' = I * (2 * d)^2 \tag{1}$$

I' は正規化後の反射強度、 I は正規化前の反射強度、 d は距離値である。LiDAR の照射光は物体に反射して返ってくるため、距離値の 2 倍の平方を乗算する。

3.2. ネットワーク構造

ネットワークは、図 1 に示すように 3 層の畳み込み層および 1 層の全結合層で構成される。LiDAR から取得した一定範囲の 1 次元波形データを、ネットワークに入力する。入力データの縦方向のサイズおよび出力ユニット数は、LiDAR のレーザ数に合わせて 64 である。また、横方向のサイズは、歩行者に対応した 39 点分 (約 7 度) を基準とし、 n 倍した範囲とする。入力チャンネル数は、距離値と反射強度の 2 チャンネルである。そのため、入力サイズは $2 \times 64 \times 39n$ となる。

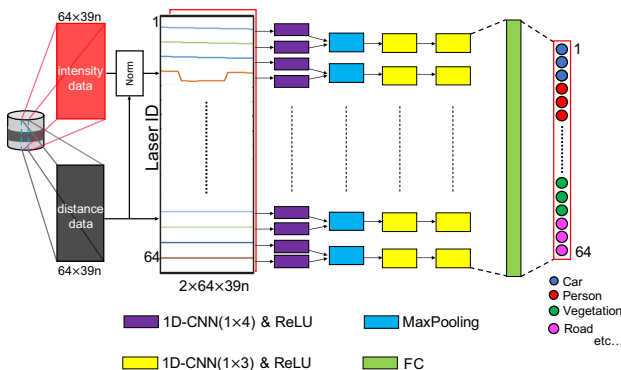


図 1: ネットワーク構造

4. 評価実験

セマンティックセグメンテーションに適した入力サイズの調査及び反射強度の有効性を検証するために、比較実験を行う。

4.1. 実験概要

学習時の更新回数を 500,000 回、バッチサイズを 256 とする。最適化手法には SGD を使用し、学習率は 0.01 とする。損失関数にはクロスエントロピー誤差を用いる。学習および評価は SemanticKITTI [2] データセットで行う。対象となる物体クラスは、歩行者、車、建物など、20 クラスである。SemanticKITTI は、10 シーンで構成され、このうち 9 シーン 19,310 フレームを学習用、1 シーン 4,071 フレームを評価用とする。評価指標には IoU (重なり率) を用いる。

4.2. 入力範囲の比較

入力範囲ごとの mIoU の比較結果を表 1 に示す。表 1 より、入力範囲が 28° のとき mIoU が最も高いことがわかる。

表 1: 入力範囲ごとの mIoU 比較結果

n	1	2	3	4	5	6	7	8	9	10	11	12	13
角度 [°]	7	14	21	28	35	42	49	56	63	70	77	84	91
mIoU [%]	28.6	27.8	27.9	28.9	28.8	28.7	28.4	28.8	28.4	28.0	28.4	26.7	26.3

4.3. 反射強度の有効性の評価

最も高精度であった範囲を入力とし、従来手法 [1] との mIoU および一部のクラスの IoU の比較結果を表 2 に示す。

表 2: 従来手法との IoU 比較結果 [%]

入力	入力範囲	mIoU	person	car	truck	bicycle	motorcycle
距離値	28°	27.1	5.2	55.5	3.1	0.02	0.6
距離値+反射強度	28°	28.9	6.0	57.9	9.6	0.2	0.7

表 2 より、反射強度を導入することで、mIoU が 1.8pt 向上している。また、人や車など動的物体に対する IoU が向上している。特に、反射率が高いと考えられるトラックの IoU が 6.6pt 向上しており、最も精度向上が見られる。これにより、反射強度はセマンティックセグメンテーションにおいて有効であるといえる。

結果の可視化例を図 2 に示す。図 2 は、クラスごとに色付けしており、各物体の分類ができていことがわかる。また、反射強度無しの場合、図 2(a) の赤枠に示すように、車に誤識別しているクラス等が多数あるが、反射強度を入力に追加した場合は誤識別が減り、トラックなど正しく識別できている。

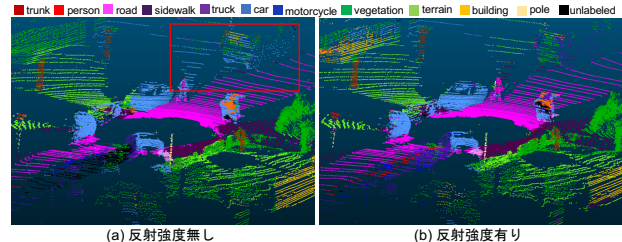


図 2: 入力範囲 28° における結果の可視化例

5. おわりに

本研究では、1D-CNN をセマンティックセグメンテーションタスクに拡張し、最適な入力データのサイズを調査した。また、反射強度を入力に追加することで、精度向上を実現した。今後は、他の手法との比較、最適なネットワーク構造の検討等を行う。

参考文献

[1] K.Yuki, et al., "Pedestrian-Detection Method based on 1D-CNN during LiDAR Rotation", ITSC, 2018.
 [2] B.Jens, et al., "SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences", ICCV, 2019.