

1D-CNNによる全方位LiDARデータのセマンティックセグメンテーション

鈴木貴大† 村瀬卓也† 平川翼† 山下隆義† 藤吉弘亘†

†中部大学

E-mail: stkin26@mprg.cs.chubu.ac.jp

1 はじめに

自動車の自動運転技術では、車両周辺の環境を把握することが必要不可欠である。そのため、自動運転に関する研究は非常に注目されており、運転支援システムとして、車両周辺に存在する物体検出や経路予測などが代表的な機能である。特に、物体検出手法は、車載カメラ映像のRGB画像を用いる手法 [1-3]、Light Detection and Ranging (LiDAR) で取得した3次元点群を用いる手法 [4-7] が広く用いられている。LiDARは赤外線レーザを照射し、物体に反射して光が返ってくるまでの時間を計測し、3次元情報や反射強度を取得する技術である。LiDARの3次元情報を用いる手法は、画像ベースでは困難な夜間における物体検出やセマンティックセグメンテーションも可能である。LiDARの中でも、全方位LiDARは回転しながらレーザを周囲に照射し、LiDARを原点とした360度全方位分の3次元情報を取得することができるため、全方位LiDARを用いた手法は多数提案されている [4-10]。一方で、全方位LiDARを用いる場合、撮影に時間がかかり撮影から処理完了までに物体の位置が大きくずれることがある。この問題に対して、1Dimensional-Convolutional Neural Network(1D-CNN)を用いる手法 [11] が提案されている。この手法では、全方位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に入力することで歩行者検出を行う。これにより、LiDARでのデータ取得中に逐次処理が可能となり、位置ずれを最小限に抑えることを可能としている。しかし、この手法は検出対象が歩行者のみであり、自動運転を実現する手段としては不十分である。

本研究では、複数のクラスに対応するために1D-CNNをセマンティックセグメンテーションタスクに応用する。これにより、従来手法 [11] の利点を残したまま、歩行者だけでなく多クラスの識別を可能とする。また、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。反射強度は物体の材質により異なる性質があるため、これを入力として用いることで物体の質感を考慮した識別が可能となる。したがって、物体の反射強度を考慮することによる識別精度の向上を図る。

2 関連研究

全方位LiDARから取得した3次元情報を用いたセマンティックセグメンテーションとして、様々な手法が提案されている。3次元情報を直接畳み込み層に入力する手法 [4, 5] や鳥瞰図、正面図、RGB画像を用いて各入力に対して畳み込み処理をするMulti-View [6] などがある。また、精度を維持しつつ高速な処理を実現したSqueezeSeg [7] も提案されている。さらに、LiDARからデータを取得する過程から、逐次的に歩行者領域か否かを識別する1D-CNNによる歩行者検出 [11] がある。本章では、LiDARデータを用いたアプローチにおける代表的な手法を説明する。

2.1 3次元点群を直接入力する手法

3次元点群情報を畳み込み層に直接入力する手法として、PointNet [4] がある。PointNetでは、3クラス分類やセグメンテーションなどのタスクに対応することができるネットワークが提案されている。PointNetは、Spatial Transformer Network(STN)とClassification Network, Segmentation Networkの3つから構成される。まず、STNで入力の点群に対してノイズを軽減する。次に、Classification Networkによる畳み込み処理から各点群の特徴抽出を行う。その後、Max Poolingにて全体の特徴を抽出することで、クラス分類を行う。セグメンテーションの場合は、Classification Networkにて抽出した全体特徴と各点群ごとの局所特徴を結合し、Segmentation Networkに入力する。再度畳み込み処理を複数回行い、各点群に対してセグメンテーションを行う。

PointNetには、詳細な空間の情報が欠落してしまう可能性があり、結果として局所構造を捉えられないことがある。この問題に対して、PointNetを改良することで精度を向上したPointNet++ [5] がある。PointNet++では、PointNetを階層的に適用することで局所的な特徴を捉えることが可能なネットワークが提案されている。PointNet++では局所特徴の抽出にPointNetを用いる。また、入力する点群にクラスタリングを行なった近傍点を入力することで、擬似的な局所特徴を抽出することが可能となる。これにより、PointNetの問題点を解決し、クラス分類およびセグメンテーションの精

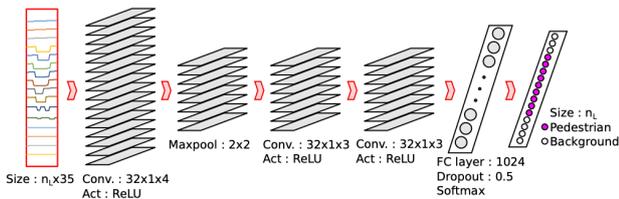


図 1: 1D-CNN の構造. 文献 [11] より引用

度向上を実現した.

2.2 各視点からの特徴を用いる手法

視点ごとの特徴を用いる手法として Multi-View [6] がある. Multi-View では, 鳥瞰図と正面図, RGB 画像の複数の視点の特徴を組み合わせることで 3 次元境界を検出するネットワークが提案されている. Multi-View は 3D Proposal Network と Region-based Fusion Network の 2 つから構成される. まず, 3D Proposal Network にて鳥瞰図, 正面図, RGB 画像の各入力に対して畳み込み処理を行い, 3 次元境界の候補を作成する. そして, Region-based Fusion Network にて, 3 次元境界の候補と各視点の特徴を組み合わせることで 3 次元境界を検出する.

2.3 パラメータの削減による処理の高速化を目的とした手法

パラメータの削減することで精度を維持しつつ処理の高速化を実現した手法として SqueezeSeg [7] がある. SqueezeSeg では, SqueezeNet [12] で特徴の抽出を行うことで, 元来の精度を維持しつつ処理の高速化が可能なネットワークとなっている. SqueezeNet はニューラルネットワークに Fire Module を導入することでパラメータ数を削減し, ダウンサンプリングを後方で行うことで精度を保持する. SqueezeSeg はこの SqueezeNet に FireDeconv を導入し, 特徴マップのアップサンプリングを行う. FireDeconv により, さらなる処理の高速化を実現している. また, Conditional Random Field as RNN にてラベルの修正を行うことでセグメンテーションの精度を向上させている.

2.4 LiDAR の回転とともに逐次処理を行う手法

全方位 LiDAR の回転とともに逐次的に処理を行う手法として 1D-CNN による歩行者検出 [11] がある. 1D-CNN による歩行者検出手法では, LiDAR から取得できる距離値をレーザ ID ごとに 1 次元波形データとみなし, 逐次的に 1D-CNN に入力することで LiDAR の回転とともに歩行者検出を可能としている. 1D-CNN の構造を図 1 に示す. ネットワークは, 3 層の畳み込み層と 1 層の全結合層から構成される. 入力データのサイズは $n_L \times 35$ である. n_L は LiDAR のレーザ ID 数を表す. また, 歩行者の平均幅が水平およそ 7 度であるため, 7 度分の点群数 35 が横軸となる. 実際の車両の

走行中に, LiDAR の回転とともに歩行者分の点群数が確保できたら逐次ネットワークに入力し, 歩行者か背景かの識別を行う. さらに, 歩行者として識別した点群に対してクラスタリングを行うことで, 高精度化を実現している. このように, LiDAR の回転とともに逐次処理を行うことで, 走行時でも歩行者を検出した位置と実際の位置のずれを最小限に抑えることを可能としている.

2.5 従来手法の問題点

従来手法 [4,6,7] の問題点として, 検出する際に全方位 LiDAR から 3 次元情報を抽出してから検出処理を行うため, 自動車など高速で移動するものに搭載して検出する場合, 検出した位置と実際の位置にずれが生じる可能性がある. 一方で, 1D-CNN による歩行者検出では, 従来手法 [4,6,7] の問題点である位置ずれに着目し逐次処理することで, 位置ずれを抑えることが可能な歩行者検出手法を実現している. しかし本手法では, 検出対象が歩行者のみとなっている. 実際の自動運転では歩行者以外の物体を検出することも重要である. また, 入力データを距離値のみとしているため, 物体の形状だけで判断してしまう. 同形で異なる物体が存在した場合, 誤識別をしてしまう可能性があるという問題点がある.

3 提案手法

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, 1D-CNN をセマンティックセグメンテーションに応用する. LiDAR から得られた距離値と反射強度を 1 次元波形データとみなし, これらを結合し 1D-CNN により各点群のクラスを判定する. これにより, 従来手法 [11] の利点を残したまま多クラス識別を実現できる.

3.1 1 次元波形データの作成

本研究で用いる全方位 LiDAR は, Velodyne 社製の HDL-64E である. HDL-64E は, 64 本のレーザを搭載する. 0.18 度間隔でレーザを全方位に照射するため, 1 フレームにつき縦 64 本, 横 2000 点の合計 128,000 点の 3 次元情報を取得することができる. 3 次元情報の距離値と反射強度をレーザ ID ごとに結合することで, レーザ ID ごとの 1 次元波形データを作成する.

3.2 ネットワーク構造

ネットワーク構造を図 2 に示す. ネットワークは 3 層の畳み込み層と 1 層の全結合層から構成する. 全ての畳み込み層において, 横軸方向にのみ畳み込み処理を行う. 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 のレーザ数に依存する. 本稿で用いる LiDAR は 64 本のレーザを照射するため 64 である. ま

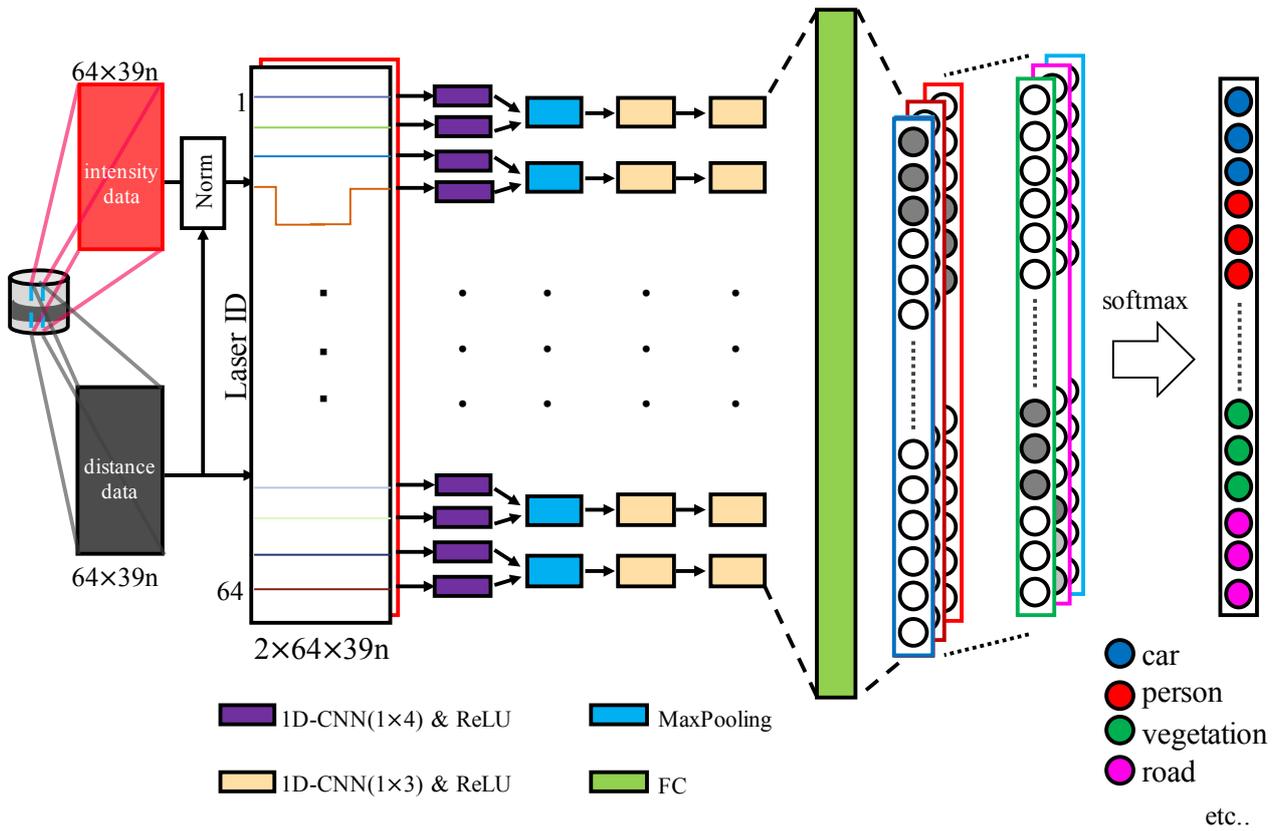


図 2: ネットワーク構造

た、横方向のサイズは歩行者に対応した7度分の点群数39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。入力データをレーザIDごとに1D-CNNにより畳み込み処理をする。その後、全結合層にて特徴を結合し、入力した1次元波形データの各レーザIDの中心に対して、softmax関数にてクラスごとの確率を算出し、識別結果を出力する。これを逐次的に行うことで、全方位に対するセマンティックセグメンテーションが可能となる。

3.3 反射強度の追加

セマンティックセグメンテーションの識別精度を向上させるため、入力データに物体の反射強度を追加する。反射強度には、金属など反射率の高い物体では光が拡散するため弱くなり、布など反射強度の低い物体では光が正確に返ってくるため強くなる、といった性質がある。したがって、反射強度をネットワークへの入力に追加することで、物体の質感を考慮した識別が期待できる。また、反射強度は物体への距離が遠くなるほど減衰し、返ってくる値が小さくなる。このとき、値が小さい理由が距離の関係からか物体の質感によるものかが分からない。そこで、LiDARからの距離値を用いて正規化することで値を補正し、距離による減衰の影響を抑制する。反射強度の正規化には、光の減衰の法

則を用いる。光の減衰の法則は、光の強さは光源から物体までの距離の2乗に反比例するというものである。距離値による反射強度の正規化の式を式(1)に示す。

$$I' = I * (2 * d)^2 \quad (1)$$

ここで、 I' は正規化後の反射強度、 I は正規化前の反射強度、 d はLiDARから得られた距離値である。式(1)を適用することで、距離により減衰した値を復元することが可能となり、反射強度を入力に用いることができる。

3.4 データの補間

全方位LiDARの3次元情報には、空や鏡面体など反射光が取得困難なシーンも含まれる。反射光が取得できなかった場合、その場所では値が取得できず外れ値となる。ネットワークへの入力データは $2 \times 64 \times 39n$ で固定しているため、外れ値の部分に値を補間する必要がある。従来手法[11]では、外れ値の影響を最小限抑えるため、レーザIDごとに値を補間する。照射角度が 0° のときは空と仮定し、レーザの最大照射距離、照射角度が 0° 未満のときはレーザが地面に当たったと仮定して値を補間する。また、本稿では反射強度の値も同様にして補間する。補間する値は、照射角度が 0° 以上のとき0.0、照射角度が 0° 未満のときは道路や歩道などの地面クラスの反射強度の平均値0.29とする。

表 2: 入力範囲ごとの IoU 比較結果 [%]

n	角度 [°]	mIoU	unlabeled	car	bicycle	motorcycle	truck	other-vehicle	person	bicyclist	motorcyclist	road	parking	sidewalk	other-ground	building	fence	vegetation	trunk	terrain	pole	traffic sign
1	7	28.6	56.1	56.8	0.4	0.2	8.4	4.8	7.6	8.3	0.0	76.2	9.8	56.9	0.7	65.7	28.5	71.1	26.2	59.8	24.8	9.4
2	14	27.8	55.1	56.8	0.09	0.5	5.8	4.3	4.7	9.2	0.0	76.2	10.4	55.6	1.1	65.6	27.2	69.2	26.9	56.9	22.4	7.7
3	21	27.9	56.7	57.4	0.1	0.4	8.1	3.4	5.7	6.3	0.0	76.5	12.8	56.8	1.4	65.8	25.8	69.4	26.4	56.1	23.6	5.4
4	28	28.9	56.4	57.9	0.2	0.7	9.6	3.6	6.0	10.4	0.0	77.0	14.8	56.9	1.4	65.9	28.5	70.1	25.6	57.6	22.6	9.9
5	35	28.8	57.6	59.0	0.2	0.7	6.8	3.3	2.5	5.4	0.0	78.2	13.8	58.7	1.2	65.9	31.5	70.8	26.5	59.5	25.6	9.1
6	42	28.7	57.6	58.6	0.1	0.2	10.2	3.7	4.0	4.4	0.0	77.8	12.1	58.8	1.7	66.4	31.9	69.8	25.4	57.6	23.4	11.1
7	49	28.4	57.8	59.1	0.06	0.2	9.0	4.6	3.1	2.4	0.0	78.0	14.1	59.1	2.3	66.9	32.9	69.2	25.1	55.6	21.9	7.1
8	56	28.8	57.3	59.0	0.4	0.3	15.2	3.8	4.0	3.5	0.0	78.0	15.4	59.0	2.1	64.7	30.8	69.8	23.4	57.2	24.6	7.5
9	63	28.4	57.9	58.9	0.1	0.2	9.8	2.4	4.0	2.2	0.0	77.7	13.5	58.5	2.4	66.1	32.5	68.9	25.1	55.0	23.1	8.9
10	70	28.0	57.4	58.5	0.03	0.2	10.4	1.9	2.5	3.7	0.0	77.5	15.3	58.1	2.1	65.6	31.7	69.1	23.6	56.3	19.9	5.7
11	77	28.4	58.5	59.3	0.07	0.07	9.7	2.9	2.9	1.4	0.0	77.7	14.2	58.8	2.5	66.2	34.2	69.6	21.2	56.6	22.4	9.5
12	84	26.7	56.4	56.9	0.0	0.09	8.3	1.4	1.2	0.8	0.0	76.2	13.2	55.8	2.3	61.9	24.3	67.7	22.0	56.2	22.4	6.3
13	91	26.3	57.0	56.6	0.0	0.1	6.7	4.1	1.5	0.6	0.0	76.5	13.4	56.3	2.4	61.4	24.6	66.4	20.1	53.9	18.2	6.7

表 1: サンプルデータ数 (フレーム)

Training Data	19,310
Evaluating Data	4,071

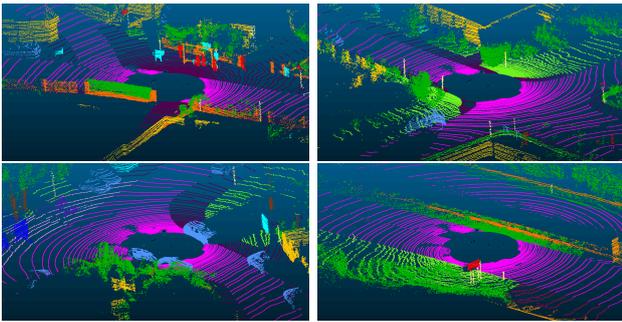


図 3: SemanticKITTI の可視化例

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範囲の調査および反射強度の有効性を検証するために、評価実験を行う。評価実験では、はじめにネットワークへのデータの入力範囲を変化させた場合の精度を比較する。そして、最も高精度な入力範囲を用いて反射強度の有無による精度比較を行う。また、従来手法との比較として、PointNet との比較を行う。

4.1 データセット

データセットには、SemanticKITTI [13] を用いる。SemanticKITTI は、自動走行を目的とした KITTI データセット [14] に基づいて作成された実環境のデータセットである。

3次元点群データは、HDL-64E で撮影している。SemanticKITTI はこの KITTI データセット全ての点群データに対してアノテーションが施されている。SemanticKITTI には人、車、建物など計 28 クラスが定義されている。人や車は動きの有無でもクラス分けがさ

れている。本研究にてセマンティックセグメンテーションを行う際は、動的物体を静的物体と同様に扱うため、識別対象は 20 クラスとなる。全 10 シーンで構成され、このうち 9 シーンを学習、1 シーンを評価に用いる。使用するデータ数を表 1、SemanticKITTI の可視化例を図 3 に示す。

4.2 ネットワークの学習

学習時の設定は、更新回数を 500,000 回、バッチサイズは 20 とする。また、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。

4.3 評価方法

評価指標には、重なり率 (IoU) を用いる。各点群に対してラベリングした結果と各点群の正解ラベルが合致しているかで評価を行う。IoU は、True Positive (TP), False Positive (FP), False Negative (FN) を用いて算出する。IoU の算出式を式 (2) に示す。

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (2)$$

また、全クラスに対する評価指標として mIoU を用いる。mIoU は、式 (3) に示すように全クラスの IoU の平均である。C はクラス数を示す。

$$\text{mIoU} = \frac{1}{C} \sum_{C=1}^C \frac{TP_C}{TP_C + FP_C + FN_C} \quad (3)$$

4.4 入力範囲ごとの精度比較結果

入力範囲ごとの精度比較結果を表 2 に示す。表 2 より、クラスごとの IoU を比較すると、例えば truck の IoU は 56 度のとき 15.2% と最も高く、bicyclist の IoU は 28 度のとき 10.4% と最も高く、person の IoU は 7 度のとき 7.6% と最も高くなっている。このように、クラスごとに最適な入力範囲が異なることがわかる。しかし、どの入力範囲においても motorcyclist の IoU は 0.0% となっ

表 3: 反射強度の有無による IoU 比較結果 [%]

入力	入力範囲	mIoU	unlabeled	car	bicycle	motorcycle	truck	othe-vehicle	person	bicyclist	motorcyclist	road	parking	sidewalk	other-ground	building	fence	vegetation	trunk	terrain	pole	traffic sign
距離値	28°	27.1	56.9	55.5	0.02	0.6	3.1	3.8	5.2	5.9	0.0	74.0	8.0	53.2	1.1	65.5	31.5	68.1	26.1	55.8	23.7	4.6
距離値+反射強度	28°	28.9	56.4	57.9	0.2	0.7	9.6	3.6	6.0	10.4	0.0	77.0	14.8	56.9	1.4	65.9	28.5	70.1	25.6	57.6	22.6	9.9

表 4: 従来手法との IoU 比較 [%]

method	mIoU	unlabeled	car	bicycle	motorcycle	truck	othe-vehicle	person	bicyclist	motorcyclist	road	parking	sidewalk	other-ground	building	fence	vegetation	trunk	terrain	pole	traffic sign
PointNet	13.8	44.2	37.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	51.0	0.05	21.1	0.0	40.7	6.2	47.0	4.0	34.4	0.8	2.0
Proposed	28.9	56.4	57.9	0.2	0.7	9.6	3.6	6.0	10.4	0.0	77.0	14.8	56.9	1.4	65.9	28.5	70.1	25.6	57.6	22.6	9.9

ている。これは、学習データに含まれる motorcyclist が少ないためであると考えられる。全クラスによる mIoU では、表 2 に示すように 28 度のとき 28.9% と最も高くなった。

また、入力範囲ごとの結果の可視化例を図 4 に示す。図 4 から、入力範囲ごとに識別結果が異なっていることがわかる。例として、truck の IoU が最も低い 14 度のときと 56 度のときを比較する。各々を正解例と比較すると、56 度のときの方が truck の点群を正しく識別できている。さらに、84 度など入力範囲が広範囲になると誤識別の範囲が広がっている。一方で、21 度のとき識別できず unlabeled となっている点群に対して、91 度のときの方が正確に識別できている場合もある。このように、可視化例からもクラスごとに最適な入力範囲が異なることが確認できる。

以上の結果より、どの入力範囲にも識別に向き不向きな物体が存在し、入力範囲が広いほど誤識別の範囲も広がると考えられる。そのため本研究では、セマンティックセグメンテーションに最適な入力範囲は mIoU が最も高い 28 度であるとする。

4.5 反射強度の有効性の評価

反射強度の有無による精度の比較結果を表 3 に示す。入力範囲は 28 度とする。表 3 から、反射強度を導入することで mIoU が 1.8pt 向上していることがわかる。また、反射強度を考慮した場合、20 クラス中 14 クラスの IoU が向上している。しかし、fence や trunk など反射強度を考慮しても IoU が低下しているクラスもある。このことから、物質によっては反射強度を考慮しても精度が向上しない場合もあると考えられる。一方で、自動運転において重要であると考えられる car, person, motorcycle 等の動的物体に対する IoU は全て向上している。特に、反射強度の影響が大きいと考えられる truck の IoU が 6.5pt 向上し最も精度向上が確認できた。これにより、反射強度の導入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。

反射強度の有無の結果可視化例を図 5 に示す。図 5(b) と図 5(c) の左の可視化例の比較から、反射強度無しの場合に car に誤識別しているクラスが多数あるが、反射強度を考慮することで、truck と正しく識別することができる。その他にも、car など反射強度の影響が大きいと考えられる物体に誤識別している点群に対して、反射強度を考慮することで誤識別が削減されている。

4.6 従来手法との比較

提案手法による IoU と PointNet による IoU を表 4 に示す。表 4 から、PointNet は入力が x,y,z 座標なのに対し、提案手法は距離値と反射強度の入力のみで全てのクラスにて IoU が PointNet を上回っていることがわかる。特に、PointNet では truck の IoU が 0.0% となっており、認識できていないのに対し、提案手法では 9.6% となっており 9.6pt 向上している。提案手法では、物体の形状に加え反射強度も考慮するため、truck などを認識することが可能である。また、pole や fence などの IoU が向上していることから、28 度毎に逐次入力すると小物体に対する識別精度が向上する。

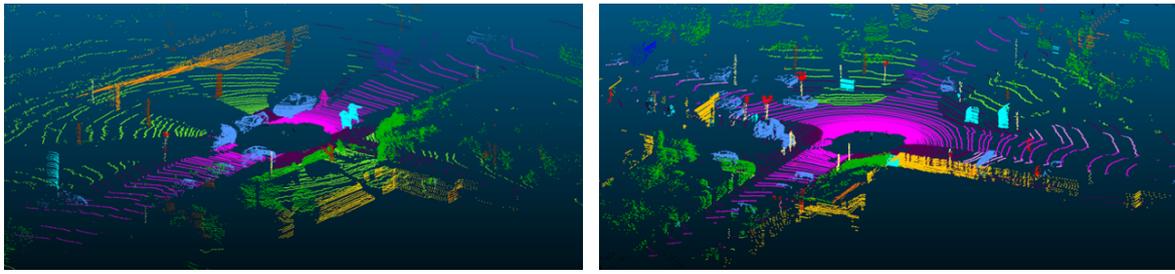
提案手法と PointNet の結果可視化例を図 6 に示す。PointNet は入力が 1 フレームなのに対し、提案手法は 28 度毎に逐次処理していくため、trunk などが正確に認識できている。

5 おわりに

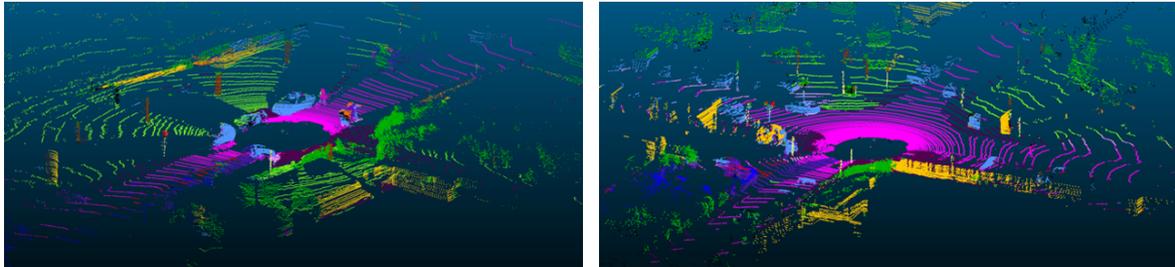
本稿では、1D-CNN による全方位 LiDAR のセマンティックセグメンテーション手法を提案した。提案手法は、LiDAR の回転とともに逐次処理が可能であるという 1D-CNN の利点を残しつつ多クラス識別を可能とした。また、反射強度を考慮することで精度の向上を実現した。SemanticKITTI を用いた評価実験により、ネットワークへの最適なデータの入力範囲の検討および反射強度の有無による精度比較を行なった。実験の結果、反射強度を考慮した場合に mIoU が 1.8pt 向上した。特に、車やトラックなど反射強度の影響が大きいと考え



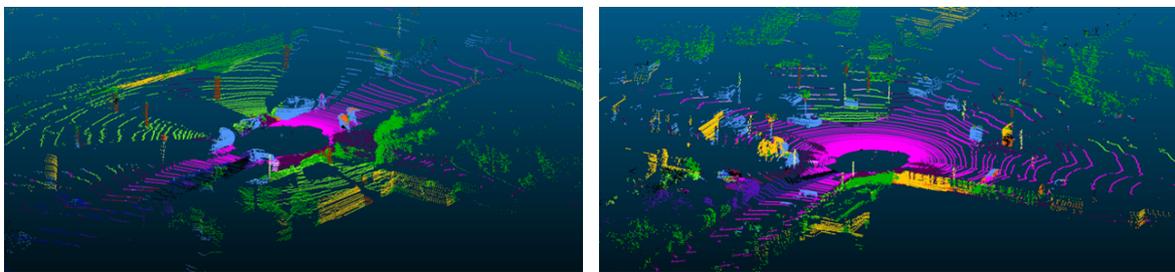
図 4: 入力範囲ごとの結果可視化例



(a) 正解例



(b) 反射強度有り



(c) 反射強度無し



図 5: 反射強度の有無による結果可視化例

られるクラスに対する識別精度の向上を示した。今後の課題として、ネットワーク構造の検討による識別精度の更なる向上などが挙げられる。

参考文献

- [1] N. Dalal, and B. Triggs, “Histograms of oriented gradients for human detection”, IEEE Conference on Computer Vision and Pattern Recognition, pp.886–893, 2005.
- [2] C. Papageorgiou, and T. Poggio, “A trainable system for object detection”, International journal of computer vision, pp.15–33, 2000.
- [3] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation applied to handwritten zip code recognition”, Neural computation, pp.541–551, 1989.
- [4] C. R. Qi, H. Su, M. Kaichun, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation”, IEEE Conference on Computer Vision and Pattern Recognition, pp.77–85, 2017.
- [5] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “Pointnet++: Deep hierarchical feature learning on point sets in a metric space”, Advances in Neural Information Processing Systems, pp.5099–5108, 2017.
- [6] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, “Multi-view 3d object detection network for autonomous driving”, IEEE Conference on Computer Vision and Pattern Recognition, p.3, 2017.
- [7] B. Wu, A. Wan, X. Yue, and K. Keutzer, “Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud”, ICRA, 2018.
- [8] L. Spinello, M. Luber, and K. O. Arras, “Tracking people in 3d using a bottom-up top-down detector”, IEEE Robotics and Automation Society,

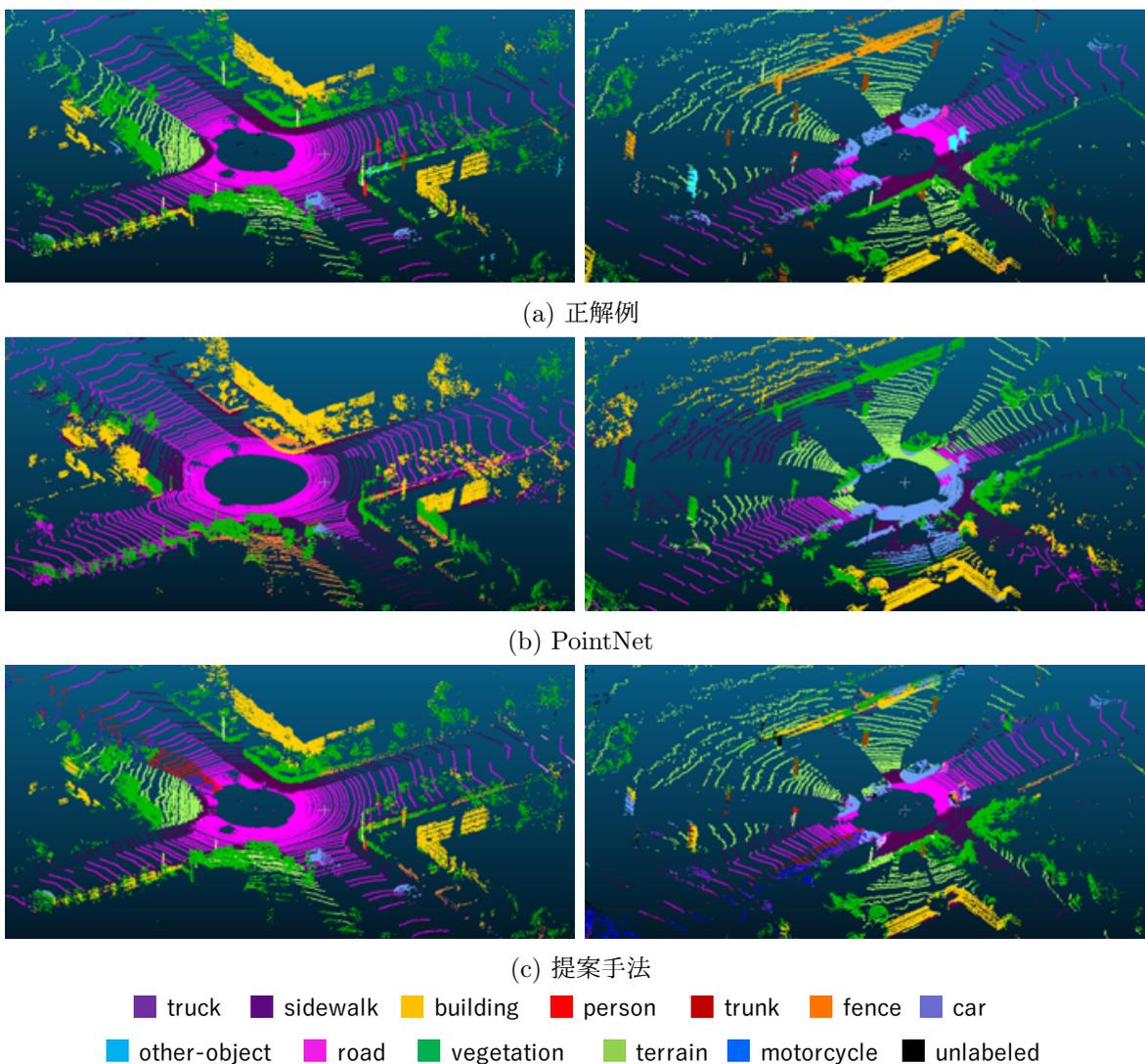


図 6: 従来手法との比較結果可視化例

pp.1304–1310, 2011.

- [9] C. R. Qi, W. Liu, C. Wu, H. Su, and L. J. Guibas, “Frustum pointnets for 3d object detection from rgb-d data”, IEEE Conference on Computer Vision and Pattern Recognition, pp.918–927, 2018.
- [10] M. Simon, S. Milz, K. Amende, and H. M. Gross, “Complex-yolo: Real-time 3d object detection on point clouds”, arXiv preprint arXiv:1803.06199, 2018.
- [11] Y. Kunisada, T. Yamashita, and H. Fujiyoshi, “Pedestrian-Detection Method Based on 1D-CNN During LiDAR Rotation”, The 21st IEEE International Conference on Intelligent Transportation Systems (ITSC), 2018.
- [12] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size”, arXiv:1602.07360, 2016.
- [13] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall, “SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences”, Proc. of the IEEE/CVF International Conf. on Computer Vision (ICCV), 2019.
- [14] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite”, Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp.3354–3361, 2012.