

### 1. はじめに

セマンティックセグメンテーションは、自動運転車の走行可能領域を把握するために必要な技術である。セマンティックセグメンテーションの学習には、画像の各ピクセルに対してアノテーションする必要があるため、人的コストが高い。そのため、教師ありデータと教師なしデータの両方を用いて学習を行う半教師あり学習が注目されている。一方で、教師ありデータの割合が少ないと、過学習を起こす問題がある。そこで本研究では、教師なしデータの一部に擬似ラベルを付与した半教師あり学習法を提案する。

### 2. 半教師あり学習

セマンティックセグメンテーションを対象とした半教師あり学習法として、CCT[1]が提案されている。本手法は通常の教師あり学習と並行する形で、ラベルなしデータを用いた一貫性学習を行う。一貫性学習は、ラベルなしデータに対するエンコーダ出力に摂動を加えたものを入力する補助デコーダを使用し、摂動を適用していないメインデコーダの出力と一貫性を持たせるように学習する。これにより、単純な構造のネットワークモデルを用いた場合でも高い精度を得ることが可能となる。

### 3. 提案手法

本研究では、擬似ラベルを用いた半教師あり学習法を提案する。擬似ラベルは教師なしデータに対して付与したラベルであり、一部を教師あり学習に組み込むことで、過学習の回避と精度向上を目的とする。

#### 3.1. 計測情報によるデータ選択と擬似ラベルの付与

擬似ラベルは、事前に少量の学習データで学習したモデルを用いて付与する。そのため、学習データと大きく異なるデータは最適ではない。そこで提案手法では、図1に示すように選択条件により教師ありデータと類似する教師なしデータを選択する。選択条件は、データ撮影時に獲得した計測情報をもとに、位置と速度から以下のように定義する。

**位置情報：**教師ありデータと位置情報の近いデータを選択する。これにより、教師ありデータのドメインに似たデータを用いた学習が期待できる。

**速度情報：**一定以上の速度で走行しているデータを選択する。教師ありデータには直線の道路を走行中のデータが多いため、速度を基準にすることによって似たデータを選択することが期待できる。

上記の条件により選択されないデータは教師なし学習に用いる。

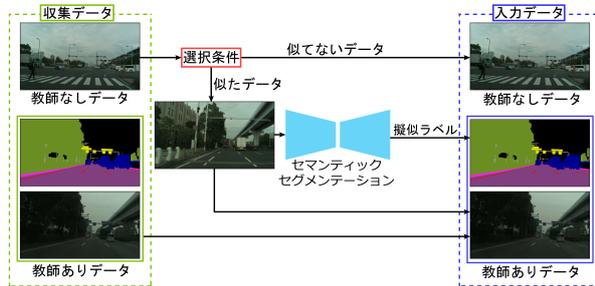


図1：データの選択と擬似ラベルの付与

#### 3.2. 損失関数

半教師あり学習手法として、CCTを用いる。CCTは、教師あり学習と教師なし学習を同時に行う。教師あり学習は、過学習を抑制するために確率によって損失を0にする annealed bootstrapped-CE (ab-CE) 損失  $L_s$  を用いて学習を行う。ab-CE 損失は式(1)により求める。

$$L_s = \frac{1}{|D_s|} \sum_{x_i^s, y_i \in D_s} \{f(x_i^s) < \eta\} H(y_i, f(x_i^s)) \quad (1)$$

このとき  $D_s$  は教師ありデータ集合、 $x^s$  は教師ありデータ、 $y$  はラベルデータ、 $\eta$  はしきい値、 $f$  はネットワーク

出力、 $H$  は Cross-Entropy loss である。教師なし学習は、入力データ  $x_i^u$  から獲得した特徴マップ  $z_i$  に対するセグメンテーション結果  $g(z_i)$  と摂動を加えたものに対する結果  $g_a^k(z_i)$  の平均二乗誤差を基に学習する。摂動は、特徴に基づく摂動、予測に基づく摂動、ランダムな摂動の3種類である。教師なし学習における損失  $L_u$  は式(2)となる。

$$L_u = \frac{1}{|D_u|} \frac{1}{K} \sum_{x_i^u \in D_u} \sum_{k=1}^K d(g(z_i), g_a^k(z_i)) \quad (2)$$

このとき  $D_u$  は教師なしデータ集合、 $K$  は摂動を適用した補助デコーダの数、 $d$  は平均二乗誤差である。

$L_s$  と  $L_u$  の総和を複合損失とする。このとき、 $L_u$  には重み  $\omega_u$  を与える。

### 4. 評価実験

評価実験では、データ選択の手法による精度を検証する。

#### 4.1. 実験概要

学習及び評価には、東京臨海部を走行して収集したデータを用いる。教師ありデータは280枚、教師なしデータは5747枚、検証用データは100枚である。評価指標には mIoU を用いる。ベースラインはデータの選択をしない場合とし、各選択手法との比較を行う。CCTの補助デコーダの数は30個である。

#### 4.2. 実験結果

定量評価を表1、定性的評価を図2に示す。表1より、速度情報でデータ選択した提案手法の精度が53.0%と最も高く、ベースラインと比較して5.1pt向上した。教師ありデータは、走行中のデータが多いため、位置情報による低速時の選択よりも、速度情報による定速走行時の選択によって似たデータを選択でき、高精度になったと考えられる。定性的評価より、図2の赤丸で示すように、速度情報を用いる方が信号機や電柱などの小さなオブジェクトの認識精度の向上が確認できる。

表1：実験結果

選択条件	教師ありデータ数		教師なしデータ数	mIoU[%]	
	実ラベル	擬似ラベル			
-	280	0	5,747	47.9	
全て	280	5747	0	47.8	
提案手法	位置情報	280	280	5,467	51.3
	速度	280	280	5,467	<b>53.0</b>

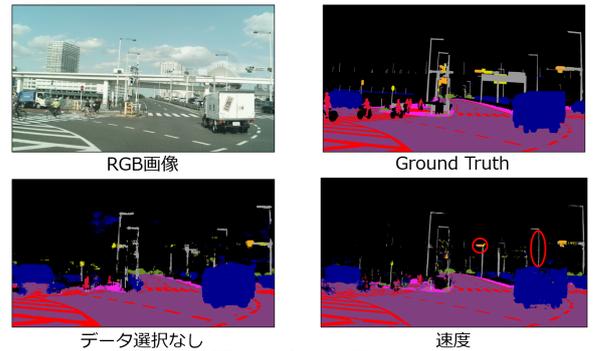


図2：定性的評価

### 5. おわりに

本研究では、擬似ラベルを用いた半教師あり学習において、計測情報を用いたデータ選択の有効性を示した。速度情報による選択によって、選択をしない場合と比べて5.1ptの精度向上を確認した。今後は、異なるデータ選択方法を考案し、より高い精度を得られる条件を検討する。

#### 参考文献

[1] Y. Ouali, et al., "Semi-Supervised Semantic Segmentation with Cross-Consistency Training", CVPR, 2020.