

1. はじめに

ピクセル毎にクラス識別を行うセマンティックセグメンテーションの学習には、大量の画像と教師データが必要である。そして、教師データのアンノテーションは多大な人的コストを要する。このコストを削減するアプローチとして、CG などの生成画像で代用することが考えられるが、実画像とドメイン間の差異があるため、認識性能向上が十分に見込めない。この差異を吸収する手法がドメイン適応であり、Domain Adversarial Neural Network (DANN)[1] が提案されている。DANN は、ドメイン間の違いを吸収するような学習を可能とした GRL を用いてドメイン適応を行う。しかし、DANN は、局所領域ごとのドメイン間の違いを考慮できない問題がある。そこで本研究では、領域ごとに GRL を用いることで、セマンティックセグメンテーションでのドメイン適応が可能な手法を提案する。

2. Domain Adversarial Neural Network

ドメイン適応は、異なるドメイン間の分布差を吸収し、認識精度を向上させる方法である。代表的な手法である DANN は、教師データを持つ Source データと教師データを持たない Target データを用いて学習する。DANN は特徴抽出器で特徴マップを出力し、ラベル予測器でクラス尤度と教師データの誤差を算出する。Gradient Reversal Layer (GRL) は、特徴抽出器とドメイン分類器の間に導入し、ドメイン適応を行う。GRL は、順伝播時に恒等関数として機能し、逆伝播時は勾配へ負の値を乗算して特徴抽出器に伝播する。

3. 提案手法

本研究では、領域ごとに GRL を用いたドメイン適応を提案し、セマンティックセグメンテーションに応用する。図 1 に提案手法のネットワークを示す。

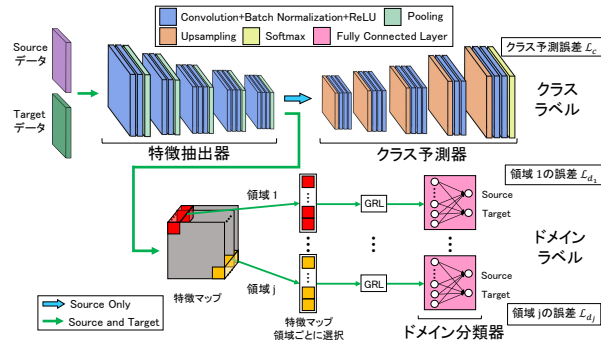


図 1: 本研究のネットワーク構造

ベースネットワークは SegNet[2] を用いる。提案手法における領域とは、特徴抽出器から出力される特徴マップにて最小領域をチャンネル方向に選択した範囲とする。ドメイン分類器は、領域ごとに用意し、それぞれ GRL と全結合の処理を行う。ここで、ある領域 j におけるドメイン分類誤差 $\mathcal{L}_{d_j}$ は式 (1) のように算出できる。ここで、領域内の特徴マップを F_j 、ドメイン分類器を G_{d_j} とする。また、 D_j はドメインラベルで、 $D_j \in \{0, 1\}$ である。

$$\mathcal{L}_{d_j}(G_{d_j}(F_j), D_j) = -(D_j \log[G_{d_j}(F_j)] + (1 - D_j) \log[1 - G_{d_j}(F_j)]) \quad (1)$$

逆伝播では、領域ごとに求めた式 (1) の誤差に対して、GRL にて負の値 (-1) を乗算する。

4. 評価実験

提案手法の有効性を確認するために、従来法とセマンティックセグメンテーションの精度比較を行う。

4.1. 実験概要

本実験では、Source データに CG 画像として図 2(a) の GTAV データセット、Target データに実画像として図

2(b) の Cityscapes データセットを用いる。従来法には、SegNet、DANN のドメイン分類器によるセマンティックセグメンテーション (DANN) を用いる。SegNet の学習は、Target データのみの学習を SegNet(T)、Source データのみの学習を SegNet(S) とする。DANN と提案手法は、教師データを持つ Source データと教師データを持たない Target データで学習をする。また、事前に Source データで学習した SegNet モデルを用いて提案手法を学習する実験 (提案手法+事前学習) も行う。ドメイン分類器より先にクラス予測器の学習を仕上げることで、全体の学習が上手くできることを調査する。学習データ数は、全て 2,964 枚である。評価は、全ての実験で Target データのテストデータ 500 枚を用いて精度比較を行う。



図 2: 各データセット画像

4.2. 実験結果

評価結果を表 1 に示す。提案手法の Mean IoU は、SegNet(S) より 3.6 ポイント、DANN より 0.4 ポイントの精度向上が確認できた。図 3 に各比較手法の Target データに対する推論結果を示す。これより、提案手法は Sidewalk (歩道) や Truck (トラック) などのクラスにおいて SegNet(S) や DANN より正しい結果を得ている。この結果から、特徴抽出器が領域レベルでドメイン間の違いを吸収でき、精度の向上に貢献していると考えられる。“提案手法+事前学習”は、Mean IoU が DANN や提案手法より精度が向上した。これより、クラス予測器の事前学習は、GRL によるドメイン適応に有効であることがわかる。

表 1: 各実験条件における Target の精度 [%]

Methods	Global Accuracy	Class Accuracy	Mean IoU
SegNet(T)	75.9	42.2	32.0
SegNet(S)	64.7	27.5	17.1
DANN	68.7	30.3	20.3
提案手法	68.9	30.7	20.7
提案手法+事前学習	68.7	31.7	21.2

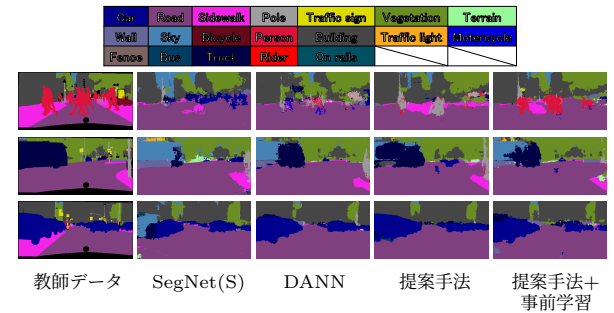


図 3: 各実験条件における Target の推論結果

5. おわりに

本研究では、GRL を用いたセマンティックセグメンテーションにて、領域ごとのドメイン適応を提案した。提案手法の Mean IoU は、Source データで学習した SegNet(S) や DANN より認識精度が向上した。今後は、提案手法に用いるセグメンテーションのベースネットワークを変更することで調査を行う。

参考文献

- [1] Y. Ganin, et al., “Domain-Adversarial Training of Neural Networks”, JMLR, vol. 17, no. 59, pp. 1–35, 2016.
- [2] V. Badrinarayanan, et al., “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation”, PAMI, vol. 39, no. 12, pp. 2481–2495, 2017.