

2021年度 山下研究室 修士論文発表 アブストラクト

Deep Learning Robotics

生活支援ロボットによる物体配置タスクにおける他物体との衝突予測に関する研究
河合 竹彦



Deep Learning Object Detection

Grid-wise-attentionを導入した物体検出と視覚的説明に関する研究
木村 秋斗

Deep Learning Human Pose Estimation

Transformerによる中間特徴を考慮した人体姿勢推定に関する研究
小松 悠斗

Deep Learning Semantic Segmentation Point Cloud

Self-Attentionを用いた点群データからのセマンティックセグメンテーションに関する研究
鈴木 貴大

Deep Learning Semantic Segmentation

セマンティックセグメンテーションにおけるマルチドメイン学習に関する研究
正木 翔大

Deep Learning Reinforcement Learning Multi-agent

マルチエージェント強化学習を用いたデッドロック問題の解決に関する研究
五藤 強志

1.はじめに

セマンティックセグメンテーションは、ピクセルレベルでのクラス識別を行うことで、オブジェクトの種類だけでなく形状も認識することができる。学習と異なるシーンなどドメインの変化が生じると認識精度は著しく低下するため、対象となるドメインごとにモデルを学習する必要がある。そのため、様々な地域で運用される自動運転システムに用いる場合、地域ごとにそれぞれのデータで学習したモデルが必要となり、メモリコストが高くなる。そこで本研究では、単一のモデルで複数のドメインデータを学習するために、マルチヘッド構造を導入したセマンティックセグメンテーションを提案する。各ヘッドは、データセット固有のクラスを出力する。これにより、オブジェクトクラスが異なるデータセットを同時に学習することが可能となる。提案手法は、ベースモデルと比べて 1.28 倍のパラメータ増加で複数のドメインに対応可能なモデルを実現できることを確認した。

2.セマンティックセグメンテーション

セマンティックセグメンテーションの代表的な手法として、DeepLab v3+[1] がある。DeepLab v3+は、Encoder-Decoder 構造と、Atrous Spatial Pyramid Pooling (ASPP) を採用している。ASPP は、異なる Dilation の畠込み処理を並列で行い、特徴マップを統合することで、異なる Receptive field を考慮する処理である。従来のセマンティックセグメンテーションは、未学習のドメインに対して認識精度が低下する問題がある。MSeg[2] は、データセットごとで付与されているラベルを共有ラベルに変換して学習することでマルチドメイン学習を実現している。しかし、共有ラベルの再定義にはコストがかかることや、一部のクラスは削除される問題がある。

3.提案手法

本研究では、データセットを再構築することなく複数のドメインを同時に学習するために、ドメイン固有の出力が可能なマルチヘッド構造を導入したセグメンテーション手法を提案する。提案手法のネットワーク構造を図 1 に示す。

3.1.共有ネットワーク

共有ネットワークは、单一のモデルで複数ドメイン情報を共有する部分である。この共有ネットワークには、ResNet101 を Backbone にした DeepLab v3+[1] を用いる。また、ResNet101 の Residual Block に Domain Attention (DA) module[3] を適用する。DA module は、ドメインに対する Attention により、共有したドメイン情報から入力に対して重要な特徴表現を獲得する機構である。DA module の構造を図 2 に示す。DA module は、SE Adapter と Domain Assignment によって構成される。SE Adapter は、複数の SE module で構成されている。それぞれの出力を連結することで全ドメインの表現空間を形成できる。Domain Assignment は、GAP と全結合層とソフトマックス層で構成されており、ドメインに適応した重みを獲得する。GAP の後の全結合層からの出力は、SE Adapter の SE module の数と同じになる。獲得した重みは SE Adapter の出力と乗算し、Sigmoid 関数で算出する。これにより、SE Adapter からドメインに適応した重みベクトルが獲得できる。

3.2.ドメインに対応したマルチヘッド構造

一般的なセグメンテーションネットワークは、あらかじめ定義したクラスに対する出力を行うシングルヘッド構造であるため、クラス数が異なるデータセットを同時に学習できない。そこで、異なるオブジェクトクラスを持つデータセットも同時に学習するためマルチヘッド構造を提案する。共有ネットワークで獲得した特徴マップをデータセット固有の出力ヘッドに入力する。そのためデータセット固有のパラメータは、出力ヘッドのみになる。これにより、特徴抽出器は共有となるため、データセット増加によるパ

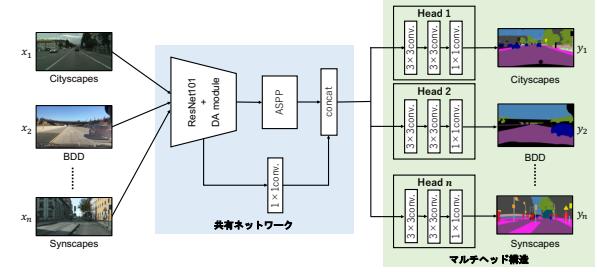


図 1: 提案手法のネットワーク構造

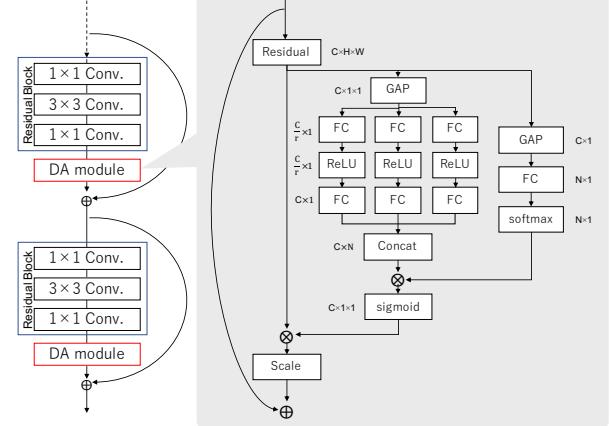


図 2: Domain Attention module の構造

パラメータ数の増加を抑えることが可能とする。

3.3.損失関数

一般的なマルチドメイン学習 [3] では、異なるドメインを持つデータを順番に入力して、各ヘッドで損失を算出し、逐次逆伝播を行う。このとき、データセット毎にパラメータが更新されるため、逆伝播させる順番によって特定のデータセットにバイアスがかかる可能性がある。そこで、全てのドメインデータを入力して、各ヘッドで出した損失を合計してから逆伝播する Mix Loss を使用する。 x を各ドメインの入力サンプル、 L_{CE} をクロスエントロピー損失関数として、 N 種類のドメインを学習する際の、損失関数 L は式 (1) のように求めることができる。

$$L = \sum_{n=1}^N L_{CE}(x_n) \quad (1)$$

全てのドメインの損失から逆伝播することで、ドメイン固有の偏りを防ぐことができる。

4.評価実験

本章では、複数のデータセットを用いて提案手法の有効性を示すために、複数のデータセットを用いて評価実験を行う。

4.1.実験概要

実験には、同一のオブジェクトクラスで構成されている Cityscapes, BDD, Synscapes の組み合わせと、異なるオブジェクトクラスで構成されている Cityscapes, A2D2, IDD の組み合わせを用いる。学習回数は 100 epoch、評価指標には mIoU を用いる。各データセットに含まれるデータ数は異なるため、1epochあたりに用いるデータ数を学習枚数が最も多いデータセットに合わせることでデータ数を調整する。最適化には、モーメンタムを 0.9、重み減衰を 0.0001 に設定した SGD を用いる。学習時のミニバッチは、同じデータセットのデータのみで構成する。

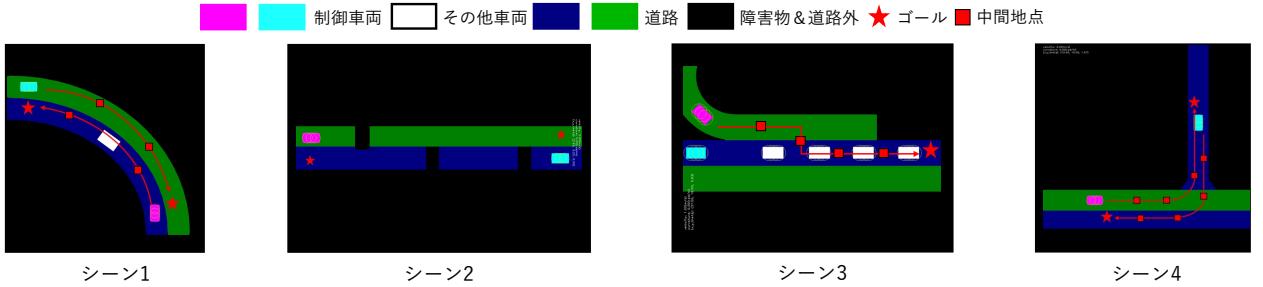


図 2: 各シーン毎の学習環境：片側一車線で、見通しの悪い道路に路駐車両が停車しているシーン（シーン 1）、走行車線に障害物があり、対向車を避けて走行するシーン（シーン 2）、幹線道路に合流するシーン（シーン 3）、狭い出入り口でのシーン（シーン 4）

表 3: 各シーンにおける獲得した累積報酬による比較 (100 エピソードの平均)

	シーン 1		シーン 2		シーン 3		シーン 4	
	エージェント 1	エージェント 2						
提案手法	961.5	580.2	997.0	994.4	785.0	585.3	773.4	589.6
A 3 C	324.7	275.3	76.1	84.8	195.0	140.8	134.8	137.1
MADDPG	170.0	165.1	109.1	120.0	109.5	126.7	154.4	134.0

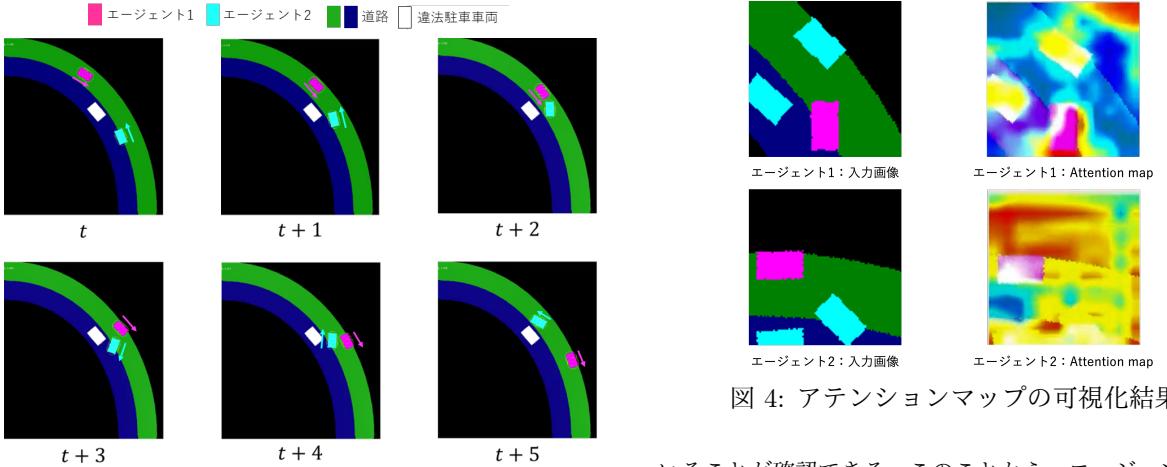


図 3: 行動の可視化結果：矢印は各車両の進行方向

は、 5.0×10^6 ステップ、シーン 4 では、 5.0×10^7 ステップ学習を行う。すべての環境において、エピソードの終了条件はエージェントが道路外、障害物及び他車両に衝突した場合と 1.0×10^5 ステップ経過で終了とする。

4.3. 累積報酬による比較

各シーン 100 エピソード評価を行った場合の平均累積報酬を表 3 に示す。表 3 から、全てのシーンにおいて提案手法を用いたエージェントが最も高い報酬を得ることができておらず、A3C 及び MADDPG と 300 点以上のスコア差がある。これは、提案手法のみデッドロックを回避することで、報酬を増やすことに成功しているため累積報酬が大きく向上している。逆に、A3C 及び MADDPG では、デッドロック状態を回避できず、局所解へ陥ってしまっているため（累積）報酬が低下している。これらのことから提案手法の有効性が確認できる。

4.4. 提案手法のモデルを用いた定性的評価

シーン 1 におけるデッドロック発生時のエージェントの行動とエージェントごとのアテンションマップを図 3、4 に示す。ここで赤車両がエージェント 1 であり、青車両がエージェント 2 である。図 3 から、エージェント 2 が $t+1$ から $t+3$ まで待機し、エージェント 1 が通り過ぎると動き出している。これより、デッドロックを回避するためにエージェント 2 はエージェント 1 が通り過ぎるのを待つという相手を考慮してデッドロックを回避する行動を獲得できていると考えられる。また、図 4 から、デッドロックが発生した場面において、エージェント 1 は周囲の道と先の道を注視し、エージェント 2 は、自身と他車両を注視して

いることが確認できる。このことから、エージェント 1 は相手を気にせず、先に道を進むエージェントであり、エージェント 2 は相手と自分を注視し、安全運転するエージェントであると考えられる。これらのことから、提案手法を用いることで、デッドロックが発生した場合、相手を考慮してデッドロックによる衝突を回避するエージェントを獲得することが可能となる。また、MARLにおいて、相手と異なる視点を持ち、異なる行動を行うことで、デッドロック問題を解決していることが確認できた。

5. おわりに

本研究では、他エージェントの学習を考慮するマルチエージェント強化学習の手法を用いることで、デッドロックが発生する環境における問題を解決した。自動運転においてデッドロックが発生する環境で提案手法の有効性を示した。また、Mask-Attention を用いてエージェント毎の方策を確認したところ、相手を考慮してデッドロックを回避する行動を獲得していることも確認できた。今後の予定としては、より多様な環境における最適な行動の獲得や、より相手を考慮し、協調行動を行うモデルの実現などが挙げられる。

参考文献

- [1] V. Mnih , et al., “Asynchronous methods for deep reinforcement learning”, ICML, 2016.
- [2] H. Itaya, et al., “Visual Explanation using Attention Mechanism in Actor-Critic-based Deep Reinforcement Learning”, IJCNN, 2021.

研究業績

- [1] 五藤強志 等, “インタラクションを考慮したマルチプランチャネットワークによる深層強化学習”, 人工知能学会全国大会, 2020.