

Attention Branch Network による 一貫学習・強化学習におけるアテンションの獲得

○森啓介 福井宏 丸山祐矢 平川翼 山下隆義 藤吉弘亘 (中部大学)

1. はじめに

一貫学習による自動運転制御 [1] では、画像を入力として制御値であるステアリングとスロットルの値を直接出力することで車両制御を実現している。また、強化学習を用いたロボットの自律動作の獲得 [2] では、ロボットの前方を撮影した画像を入力として、壁などの障害物に衝突することなく目的地への移動を実現している。これらは、Convolutional Neural Network (CNN) を学習することで、様々な問題に対して高い性能を実現している。しかし、ネットワークからの出力は何を根拠に決定されたのかが分からないという問題がある。特に、一貫学習や強化学習において CNN へ画像を入力した際の判断根拠を解釈することは、誤認識の理由を明確にするために重要である。

画像認識の分野では CNN の挙動を解釈する様々な手法が提案されている [3, 4, 5, 6]。その手法の一つに、画像中の注視領域である Attention を生成しつつ、認識結果を出力する Attention Branch Network (ABN) が提案されている [6]。ABN では、畳み込み層のみで構成された CNN が強く反応した領域を上位層まで伝播する Global Average Pooling (GAP) [7] の特性を利用し、CNN が注目している領域 (Attention map) を可視化しつつ、入力画像に対する注視領域を認識過程に利用して認識性能を向上させることができる。しかし、ABN は画像認識のような離散的な教師ラベルを持つ問題に対して適用する手法である。一方、一貫学習による自動運転制御はステアリングやスロットルのような連続値を推定する回帰推定であり、Deep Q-Network (DQN) [8] のような深層強化学習に基づく自律制御には教師ラベルが存在しない。そのため、上記の制御タスクに対しては ABN を直接適用することが困難である。

そこで、本研究では一貫学習および深層強化学習において、入力画像に対する動作の判断根拠を解釈することを目的とし、ABN を回帰推定および教師ラベルの存在しない深層強化学習に応用する手法を提案する。回帰推定に対しては、GAP の出力を畳み込むことで連続値に変換する。また、深層強化学習に対しては、あらかじめ学習したネットワークに Attention Branch を追加し、Attention Branch のみを学習する。評価実験では、一貫学習による自動運転制御および深層強化学習によるロボットの動作獲得を用いて、ABN が出力する Attention がネットワークの出力に対する理由付けや判断根拠を解釈に有効かを検証する。

2. 提案手法

Attention Branch Network (ABN) [6] は、入力画像から特徴を抽出する Feature extractor, Attention map を出力する Attention Branch, 出力された Attention

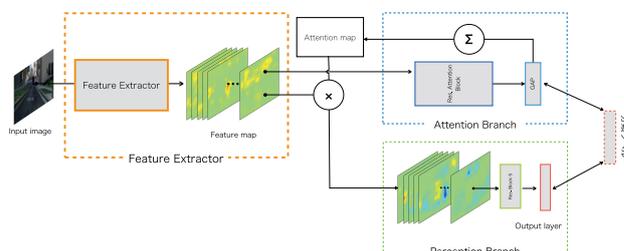


図1 Attention Branch および Perception Branch の構造

map を考慮して最終的な認識結果を出力する Perception Branch から構成されている。Attention Branch から出力される Attention map は、入力画像における注視領域を可視化できる一方で、Attention 機構への応用も可能である。しかし、ABN は画像分類にのみ用いられており、出力が連続値である回帰推定や教師なし学習の強化学習へは応用が困難である。本章では ABN の構造を述べた後、提案する ABN の回帰推定と強化学習への適用について述べる。

2.1 Attention Branch Network の構造

ABN は、ベースのネットワークを Feature extractor と Perception Branch に分割し、Feature extractor の後層に Attention Branch を配置することで構築する。Attention Branch および Perception Branch の構造を図 1 に示す。Attention Branch は全ての層を畳み込み層で構築する必要があるため、Perception Branch の全結合層を Fully Convolutional Network [9] のように 3×3 のカーネルサイズの畳み込み層に変換する。認識結果を出力する際には、出力層に GAP を配置することで各クラスの認識結果を出力する。ここで、GAP は 1 枚の特徴マップの平均値を特定クラスの応答値とする手法であり、GAP の応答値に対して softmax 関数を施すことで最終的な認識結果を得ることができる。Attention Branch から Attention map を出力した後は、Feature extractor の特徴マップに Attention map を施す。これにより、Perception Branch において特定の領域の特徴に対して重点的に学習できる。

2.2 回帰推定への適用

ABN を回帰推定に応用する際に最も問題となるのが、Attention Branch である。全結合層で回帰推定をする場合、入力の特徴マップの各ユニットが重み付きで全結合しているため、入力の特徴マップ全体を考慮して回帰値を出力できる。しかし、畳み込み層は局所的な領域を畳み込むカーネルを複数枚用いることで応答値を出力するため、連続値を出力する回帰推定等の難易度の高いタスクには不向きである。

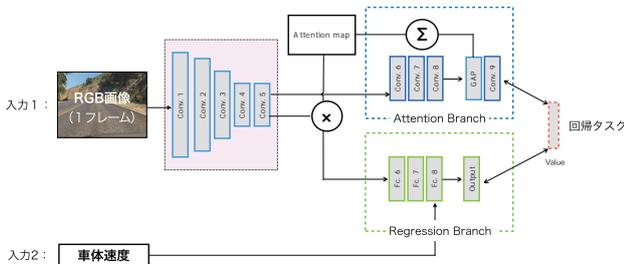


図2 ABNを導入した一貫学習による自動運転制御のネットワーク構造

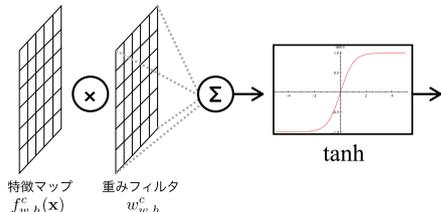


図3 Weighted Global Pooling の構造

そこで、Attention Branch を回帰推定に応用するために、GAP に重み付けを導入した Weighted Global Pooling (WGP) を提案する。ABN を導入した一貫学習による自動運転制御のネットワーク構造を図2に示す。また、WGP の詳細な構造を図3に示す。WGP は、GAP により特徴マップの平均を求める際に、式(1)のように特徴マップの解像度と同一の畳み込み層のカーネルを用いることでプーリング処理を行う。ここで、入力の特徴マップを $f(\mathbf{x})$ 、特徴マップの縦幅と横幅をそれぞれ w, h とする。図3に示すように、特徴マップ全体にカーネルを畳み込んだ後に \tanh 関数を施すことで、最終的な回帰値を出力する。Attention map は WGP の前層特徴マップを使用し、Feature extractor の特徴マップへ反映して Perception Branch へ入力される。Attention Branch 内のパラメータは $score_c$ と教師ラベルから L2 ノルム 2 乗誤差を算出し学習を行う。

$$score_c = \sum_h \sum_w w_{w,h}^c \cdot f_{w,h}^c(\mathbf{x}) \quad (1)$$

2.3 強化学習における可視化への適用

ABN を強化学習に応用する際に問題となるのが教師ラベルである。前述の分類や回帰の場合、教師ラベルが存在していたが、強化学習では環境から与えられる報酬を元にネットワークの重みを更新するため、教師ラベルが存在しない。そのため、教師ラベルとの誤差を元に Attention Branch の重みを更新する ABN を強化学習の枠組みに直接適用することは困難である。

強化学習への適用では、ロボットの動作獲得における注視領域の獲得のみ行う。ABN を導入した深層強化学習のネットワーク構造を図4に示す。まず、Attention Branch を導入しない従来のネットワークで動作獲得のための強化学習を行い、学習したネットワークの重みを固定 (Freezing) する。その後、Attention Branch を追加し、Attention Branch 内の重みのみを学習する。このとき、重みを固定したネットワークの出力結果を教

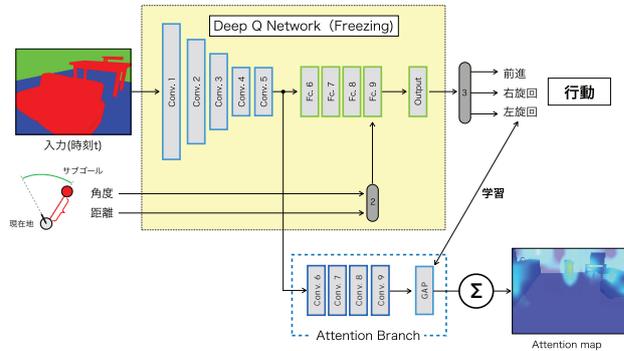


図4 ABNを導入した深層強化学習に基づくロボットの動作獲得のネットワーク構造

師ラベルとして用いる。これにより、強化学習における行動時の Attention map を獲得することができる。

3. 評価実験

前章で述べた ABN による一貫学習と強化学習に対する評価実験を、一貫学習として自動運転制御、強化学習としてロボットの自律動作獲得を用いて行う。以降では、それぞれの実験概要と自動運転制御における精度評価、両モデルから得られた Attention map についての定性的な評価について述べる。なお、強化学習では Attention map の可視化のみを行うため、精度評価は行わないことに注意されたい。

3.1 実験概要

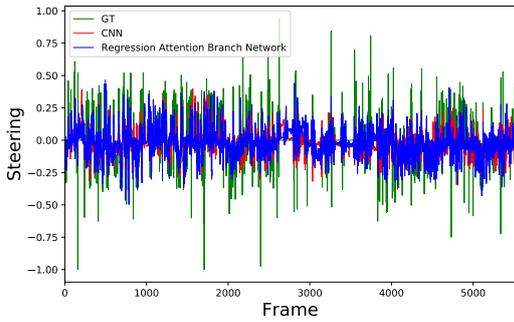
本実験に用いる自動運転モデルおよびロボット制御モデルのネットワーク構造とパラメータ、実験環境について述べる。

3.1.1 自動運転モデル

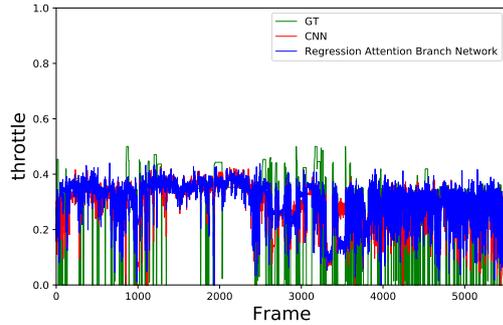
一貫学習による自動運転制御は文献 [10] に基づいて行う。ABN を用いた一貫学習による自動運転制御のネットワーク構造を図2に示す。3層の全結合層のうち、車両速度を追加しない2層にはドロップアウトを適用し過学習を抑制する。学習の最適化手法には RMSProp+Graves を使用し、学習率は 0.001、Exponential decay rate を 0.99 とした。また、学習回数を 150 エポック、バッチサイズを 4 とした。データセットの作成には、コンピュータゲームの Grand Theft Auto V (GTA V) をシミュレータとして使用する。ゲーム上の車両を人間が運転し、フロントカメラ映像を毎秒 10 フレーム分取得する。また、カメラ映像の各フレームに対応するステアリング、スロットル、車両速度も同時に取得する。本実験では作成したデータのうち、学習に 30,000 フレーム、評価に 16,556 フレーム使用する。

3.1.2 深層強化学習によるロボットの自律動作獲得

深層強化学習によるロボットの自律動作獲得には、文献 [11] と同様の環境を用いる。ロボットが動作する環境は人間が生活する空間を模した実験室とし、シミュレータ上で作成した。その際、シミュレータ上の家具、壁、床はそれぞれ赤、緑、青色のオブジェクトとして生成している。また、スタート地点とゴール地点をランダムに設定し、画像に加えて現在地から目的地まで



(a) ステアリングの出力比較



(b) スロットルの出力比較

図5 ステアリングとスロットルの正解値と出力の比較

	ステアリング	スロットル
CNN	0.014810	0.019124
ABN	0.015480	0.019331

の距離および角度を入力して動作を出力する。これらの数値は一貫学習の場合と同じく、出力層前の全結合層の出力に結合する。

ABNを適用する深層強化学習手法として、DQN [8]を使用する。実験に使用したネットワークを図4に示す。報酬 R は各ステップ毎に算出した r の累計とし、次式のように定義する。

$$r = \begin{cases} 30 & (\text{ゴールへ到達}) \\ D_c & (\text{前進時}) \\ -5 & (\text{衝突時}) \end{cases} \quad (2)$$

ここで、 D_c は各ステップでのゴールまでの相対距離の変化量である。また、エピソードの終了条件は、ゴールへ到達した場合および70ステップ以上かかった場合、累計報酬が -30 以下となった場合とする。ロボットの取り得る動作は、前進、左旋回、右旋回の三種類とする。

3.2 精度

評価実験における従来モデルとABNを適用したモデルの各フレームの出力値と正解値を図5に示す。(a)がステアリング、(b)がスロットルの結果を示している。このとき、停車時のスロットルの値は0.25と設定されていることに注意されたい。ABNおよび従来法のCNNによる出力結果を比較すると、Attentionの有無による大きな精度の差はないと言える。

定量的な評価のために、両モデルでの評価実験の結果から精度を求める。精度の指標として、モデルの予測値と正解値から平均二乗誤差を用いて比較を行う。結果を表1に示す。ステアリング、スロットルともにその差は0.001未満であり、ABNの導入による性能の劣化がないことを確認した。

3.3 Attention mapの可視化

提案手法により獲得したAttention mapを示し、入力画像に対する注視領域と出力の関係について定性的

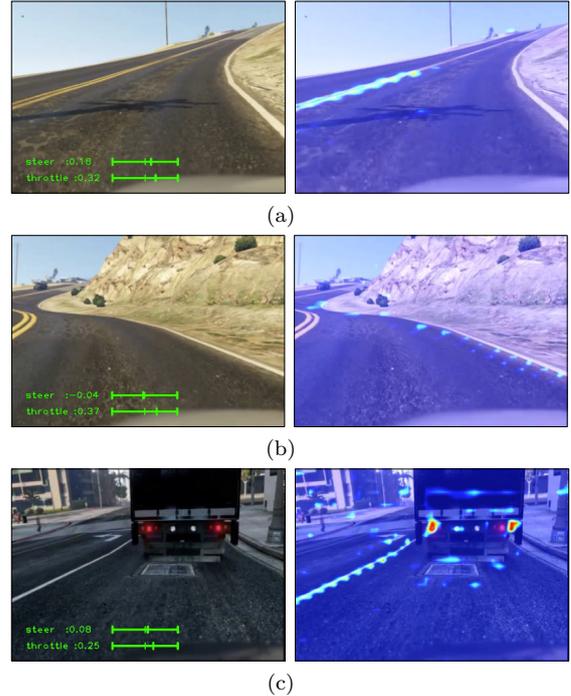


図6 一貫学習による自動運転制御で獲得した Attention map の例

な評価を行う。

3.3.1 一貫学習における獲得した Attention map

自動運転モデルから得たAttention mapの可視化画像を図6に示す。図6(a)では右方向へのカーブが必要となる場面を示している。この時、ネットワークの注視領域は走行車線左側の白線上に存在しており、ステアリングの出力値は右方向である正の値となっている。一方で、左方向へのカーブが必要な図6(b)の様な場面では、右側の車線上に注視領域が存在し、ステアリングは左方向である負の値が出力されている。したがって、ステアリングの決定には走行すべき領域を示す白線が影響しており、白線部分の特徴マップを考慮して出力を決定していることがわかる。図6(c)では車両が停止している場面を示している。この時の注視領域は前方車両の点灯したブレーキランプに存在しており、スロットルの出力は0.25となっている。そのため、ス

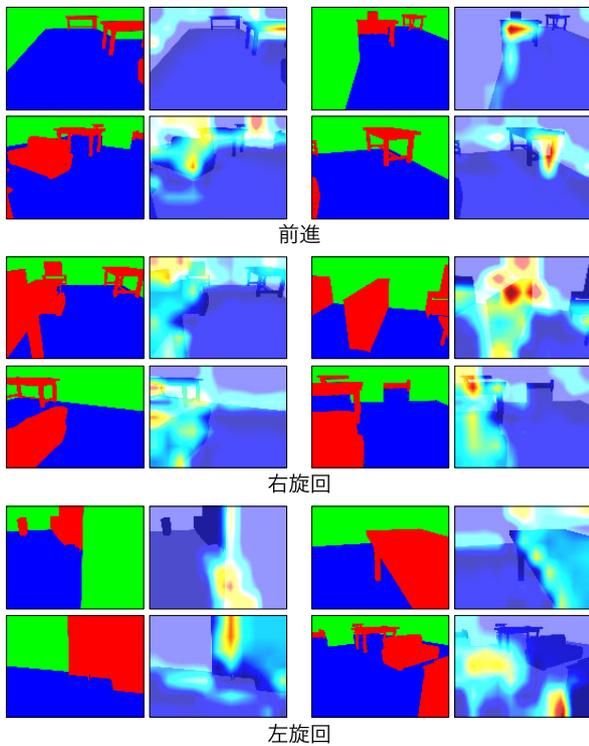


図7 各行動に対する Attention map の例

ロットの決定には前方の車両の状態が影響しており、ブレーキランプの状態を考慮して出力を決定していることがわかる。

3.3.2 深層強化学習における獲得した Attention map

図7に各行動に対する Attention map を示す。いずれの行動が選択される場合でも、赤色で示された家具に対して高い Attention map が生成されている。一方、青色で示された床の領域に対しては高い Attention map が生成されておらず、行動選択に影響していないと解釈することができる。

また、各行動が選択された時の入力画像および Attention map の平均画像を図8に示す。入力画像に対する平均画像では、左旋回の場合は右側に赤い領域が存在しており、右旋回では左側に赤い領域が存在していることがわかる。同様に、Attention の平均画像では左旋回の場合は右側に、右旋回の場合は左側に高い Attention が生成されている。入力画像および Attention の平均画像の両方から、このロボットの自律動作獲得では障害物の領域に Attention がかかり、それを避けるような動作を獲得していると解釈することができる。

4. おわりに

本研究では、ABN を回帰推定および深層強化学習へ応用する手法を提案した。回帰推定に対しては、GAP の出力を畳み込み、連続値へと変換することで、教師ラベルが連続値である問題での Attention map の獲得する。また、深層強化学習に対しては、事前学習したネットワークに Attention Branch を追加し、ネットワークの出力を教師ラベルとして Attention Branch のみを

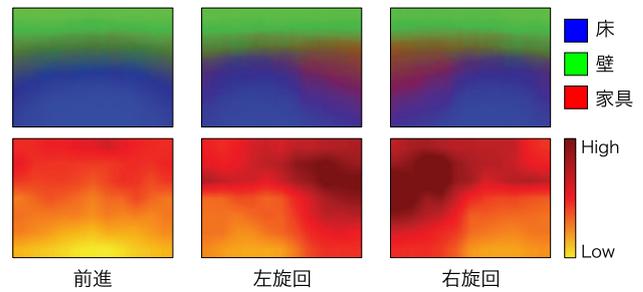


図8 各行動に対するシーンおよび Attention map の平均。上段: シーンの平均画像, 下段: Attention map の平均。

学習する。これにより、教師ラベルを用いることなく Attention map を獲得する。一貫学習による自動運転制御の評価実験では、従来の Attention map を獲得しない場合と比較し、同程度の精度を実現しつつ Attention を獲得することで、出力された制御値の判断根拠を示すことを可能とした。また、ロボットの自律動作の獲得では各行動が選択された時の Attention map を可視化することで、障害物を基準に行動を決定していることを視覚的に解釈することを可能とした。

今後の予定としては、深層強化学習において行動の出力および Attention Branch を同時に学習することなどが挙げられる。

参考文献

- [1] M. Bojarski *et al.*, “End to end learning for self-driving cars,” *arXiv preprint*, 2016.
- [2] Y. F. Chen *et al.*, “Socially aware motion planning with deep reinforcement learning,” *International Conference on Intelligent Robots and Systems*, 2017.
- [3] B. Zhou *et al.*, “Learning deep features for discriminative localization,” *Computer Vision and Pattern Recognition*, 2016.
- [4] D. Smilkov *et al.*, “Smoothgrad: removing noise by adding noise,” *arXiv preprint*, 2017.
- [5] R. R. Selvaraju *et al.*, “Grad-cam: Visual explanations from deep networks via gradient-based localization,” *International Conference on Computer Vision*, 2017.
- [6] 福井 他, “Global average pooling の特性を用いた attention branch network,” 画像の認識・理解シンポジウム, 2018.
- [7] L. Min, C. Qiang, Y. Shuicheng, “Network in network,” *International Conference on Learning Representations*, 2014.
- [8] V. Mnih *et al.*, “Playing atari with deep reinforcement learning,” 2013.
- [9] J. Long, E. Shelhamer, T. Darrell, “Fully convolutional networks for semantic segmentation,” 2015.
- [10] 村瀬 他, “自己状態を付与した CNN による自動運転制御の高精度化,” パターン認識・メディア理解研究会, 2017.
- [11] 丸山 他, “セマンティックセグメンテーションを用いた深層強化学習による自律移動の獲得,” 日本ロボット学会 学術講演会, 2017.