

# Transformer モデルによる自律移動の視覚的説明 と拡張現実による提示

○尹文韜 板谷英典 真野航輔 平川翼 山下隆義 藤吉弘亘 (中部大学)

ロボットの自律移動を獲得する技術として深層強化学習がある。深層強化学習は高い自律移動性能を獲得している一方で、ロボット動作に対する判断根拠が不明瞭という問題がある。この問題に対し画像による視覚的な解析を可能とした手法が数多く提案されている。しかし、画像では実空間と物理的な結び付きがなく、ユーザがロボットの動作を理解するには不十分である。そこで本研究は Transformer による Encoder-Decoder モデルを導入することで深層強化学習の視覚的説明を実現し、拡張現実によりロボット動作の判断根拠を 3 次元空間に可視化することでユーザのロボット理解を促進させる。

## 1. はじめに

ロボット技術の進歩に伴い、人とロボットが共存する社会に向けた技術が求められている。ロボットの高い自律移動性能を獲得する技術として深層強化学習がある。深層強化学習とは、エージェントが環境とのインタラクションを通じて得る経験をもとに、報酬を最大化するための最適な振る舞いを学習する技術である。この技術は未知環境においてもエージェントが経験から最適な振る舞いを学習可能であることから、様々な状況に対応した柔軟なロボットの自律移動を実現できる。このような高性能な制御の実現により、配送や生活支援などでは自律移動型ロボットが活躍し始めている。一方で、深層強化学習はモデルのブラックボックス問題がある。エージェントモデルは膨大な数のパラメータにより構成され、モデル内部処理を直接解析することができない。そのため、ユーザがエージェントであるロボットの思考プロセスを理解することは非常に困難である。またロボットの思考プロセスが不透明な場合、ユーザのロボットに対する信頼も失いやすいことが知られている [1]。そのため、人とロボットの共存にはロボットの信頼性が重要であり、ロボットへの不信感人は人とロボットの共存関係に悪影響を与える。

本研究では、自律移動ロボットと人の共存促進を目的とし、自律移動ロボットの思考プロセスの可視化、および拡張現実 (AR) 技術によるユーザに対するロボット思考の提示を提案する。ロボットの高い自律移動性能を獲得できる深層強化学習に Transformer モデル [2] を導入し、ロボットの注視領域を可視化することでロボット動作に対する判断根拠を解析する。また、実風景に仮想的な視覚情報を重畳し表示する AR 技術を活用することで、物理的に制限されることなくユーザが理解しやすいロボット思考の提示を実現する。ユーザがロボット思考を理解できるか確認するため、複数の被験者にアンケート調査を実施することで本手法の有効性を確認した。

## 2. 関連研究

### 2.1 説明可能な強化学習

深層強化学習は様々なタスクで高い性能を獲得している反面、エージェントの意思決定に対する判断根拠が不明確である。そのため、この問題の解決を目的と

した説明可能な強化学習 (XRL) 手法が研究されている。XRL 手法は、エージェントモデルに対し学習後に何らかの処理を行い解析する手法と、解釈可能な構造としてエージェントモデルを予め設計する手法がある。

学習後に処理を行う手法として、Greydanus らは入力画像の摂動に対する出力の揺らぎからエージェントの注視領域を可視化している [3]。この方法は、入力画像に対し様々なパターンの摂動を付与することでエージェントの注視領域を特定する。そのため、計算コストが高くリアルタイム処理が困難である。

解釈可能な構造とする手法として、板谷らは深層強化学習手法 Rainbow [4] に Transformer モデルを導入した Action Q-Transformer (AQT) を提案している [5]。Transformer は入力トークンと他トークン間の類似度を算出でき、入力に対する重要な情報を Attention により強調できる。この Attention を用いることで、エージェントの注視領域を画像として表現した Attention map の可視化を実現している。AQT では Transformer Decoder へ入力する Query を行動情報とすることで、各行動に対する注視領域が算出可能であり、動作に直結した Attention map の獲得を実現している。また、Attention の計算コストも低くリアルタイム処理が可能である。

### 2.2 ロボティクスへの AR 応用

ロボットの高性能化に伴い、人とロボットの相互作用に関する分野である Human-Robot Interaction (HRI) に関心が集まっている。HRI において AR 活用は、人のロボットに対する直感的な理解をサポートすることができる。Makhataeva らはロボット周辺の作業空間における危険性を AR を用いて可視化している [6]。この手法により、AR 上でロボット周辺空間を危険性に基づいて色付けすることで、ユーザが危険領域を回避しながら作業することが可能となる。Hedayati らはユーザによる無人飛行ロボットの操作をサポートする AR システムを提案している [7]。このシステムは、無人飛行ロボットに搭載したカメラ映像や経路など必要な情報を AR デバイスを通じてユーザに表示している。これにより、ユーザは無人飛行ロボットの直感的な操作が可能となり、遠隔操作性の向上およびクラッシュの減少を実現している。

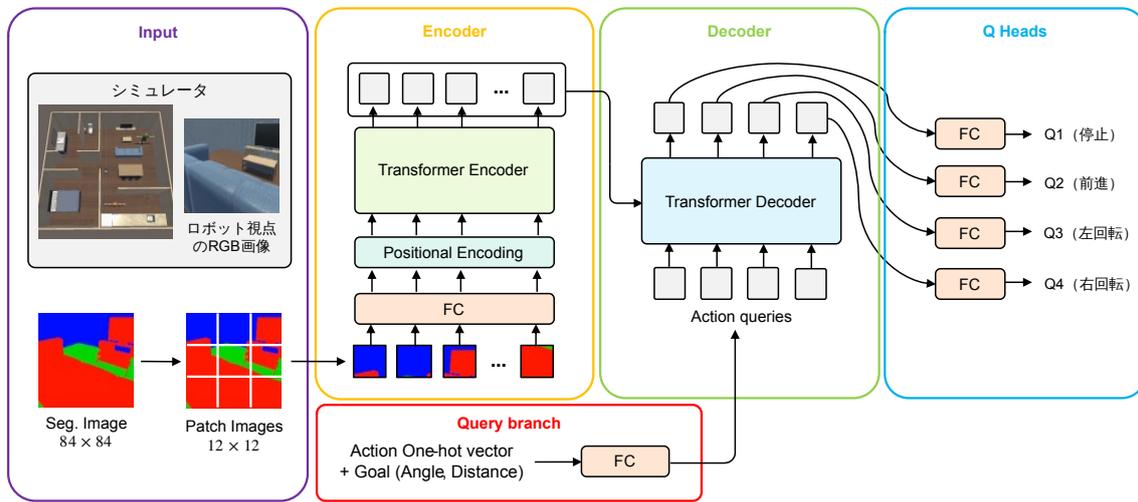


図 1: 提案手法のネットワーク構造

### 3. 提案手法

ユーザがロボットの思考プロセスを理解できないことは、ロボットの信頼性に影響する。そこで本研究は深層強化学習手法 Deep Q-Network (DQN) [8] に Transformer Encoder-Decoder モデルを導入することで、ロボットの自律移動と同時にロボットの注視領域を示す Attention を獲得する。加えて、この Attention を AR により実空間に投影することでロボットの思考を効率的にユーザへ提示する方法を提案する。

#### 3.1 深層強化学習による自律移動能力の獲得

本研究では深層強化学習によりロボットの高性能な自律移動動作を獲得する。

**ドメインギャップの吸収** 実環境下で多種多様な学習データを収集することは容易ではないため、シミュレーターによる CG 環境を用いる。しかし、CG 環境では実空間とのテクスチャの相違などドメインギャップが存在するため、CG 環境下で学習したモデルは実環境下で著しく精度が低下する。そこで本研究では、モデルへの入力をセグメンテーション画像とすることで、テクスチャ情報をラベル情報へと置き換えドメインギャップに対処する。セグメンテーション画像の生成には、代表的なセマンティックセグメンテーション手法である SegNet [9] を用いる。

**ネットワーク構造** 代表的な深層強化学習手法である DQN に Transformer モデルを導入することでロボットの判断根拠の可視化を実現する。DQN とは行動価値 (Q 値) をニューラルネットワークによる関数近似により最適化する手法である。提案手法のネットワーク構造を図 1 に示す。ロボット動作に対する解釈性の高いネットワーク構造とするため、AQT をベースとした Transformer Encoder-Decoder 構造を採用する。提案手法による行動算出までの流れを述べる。はじめに、ロボット視点の RGB 画像を SegNet により中間表現であるセグメンテーション画像に変換する。そして、この画像を  $84 \times 84$  ピクセルにリサイズ後、パッチサイズ  $12 \times 12$  でパッチ画像に分割する。各パッチ画像は全結合層により埋め込みベクトルへ変換後、Positional

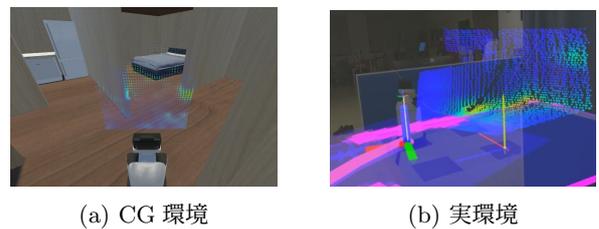


図 2: AR による Attention 可視化例: (a) は AR デバイスを使用した Attention の可視化をシミュレータ上で再現した環境である。

Encoding により位置情報を付与し Encoder へ入力する。また Decoder への入力は Encoder の出力と Action queries である。Action queries はロボットの各行動を表現した One-hot vector と、ロボットからゴールまでの距離と角度のゴール情報を組み合わせた Query である。Decoder の各 Query に対する出力を Q Heads へ入力し各行動の Q 値を算出する。

#### 3.2 AR を用いたユーザへの提示

提案手法では Transformer Encoder-Decoder 構造を導入しているため、Encoder と Decoder のそれぞれにおいて Attention を算出している。Decoder では行動情報を表す Action queries を用いているため、ロボットの動作に関連した Attention を獲得している。そこで、本手法では出力した Q 値が最も高い動作に対応する Decoder の Attention を AR により可視化する。ここでの可視化方法は、Depth センサーから獲得したロボット視点での深度情報を点群に変換し、その点群を Attention の値に応じて色付けすることで可視化する。ここで Attention はヒートマップで表現し、値が大きいほど赤く、小さいほど青く色付けする。CG 環境と実環境における AR による Attention の可視化例を図 2 より示す。実環境では、AR ヘッドセットを用いてユーザの視野に重畳させるように Attention を可視化する。

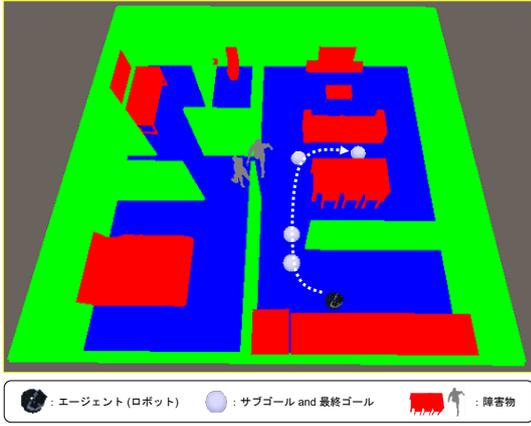


図 3: Unity による再現環境

## 4. 評価実験

提案手法の有効性を確認するため、深層強化学習により獲得した自律移動性能の評価、および AR を用いたユーザへの提示に関する有効性調査を行った。

### 4.1 シミュレーション環境

本研究では、生活支援ロボット Human Support Robot (HSR) による屋内環境における自律移動タスクを対象とする。そのため Unity を使用し屋内環境を再現し、人や机等の障害物を設置した。3.1 節で述べたように、提案手法ではドメインギャップの吸収にセグメンテーション画像を用いる。そこでシミュレータ内のテクスチャを図 3 に示すようにオブジェクトごとに色分けした。

自律移動タスクにおける目標は、ロボットが障害物を避けながら最終ゴールに効率的に辿り着くことである。そこで本環境では、エピソード毎でランダムにスタート位置と最終ゴール位置を設定し、スタートと最終ゴール間にはサブゴールを生成する。その他の設定として、エージェントであるロボットの動作は停止、前進、左右回転の 4 種類とする。また本タスクを解くために到達、接近、衝突に関する報酬を設計する。各報酬を式 (1) に示す。

$$r = \begin{cases} 10 (\text{reach the goal}) \\ -0.01 (\text{not moving forward}) \\ -3 (\text{crash}) \end{cases} \quad (1)$$

これら 3 つの報酬の総和がエージェントが環境から受け取る報酬値である。

### 4.2 自律移動性能の評価

**タスク達成率** 提案手法の性能を評価するため、100 エピソード間のタスク達成率により比較する。ここでのタスク達成条件は、ロボットが 100 ステップ以内に最終ゴール到達、かつ衝突回数が 6 回以下とした。比較手法としては、代表的な深層強化学習手法である DQN と提案手法である。各手法の学習ステップ数は  $1.0 \times 10^6$  ステップとした。表 1 から、提案手法は DQN と比較し 6% 向上し、8 割以上と高いタスク達成率であること

表 1: 100 エピソード間のタスク達成率

モデル	DQN	提案手法
タスク達成率 [%]	77	83

が分かる。このことから提案手法は高性能な自律移動動作を獲得できていると言える。

**Attention map の可視化** 獲得した Attention map が正しくロボットの注視領域を示しているか確認するため、行動やゴール位置に対する注視領域の変化を確認した。Decoder は行動情報である Action one-hot vector とゴール情報を Query として用いている。そこで、各行動とサブゴール位置の変化に対する Decoder の Attention map を確認する。

図 4 (a) から、ロボットは左回転動作では左の家具を、右回転動作では右の壁を注視していることが分かる。また図 4 (b) から、サブゴール位置を左前方から右前方へ変化することで、ロボットの注視対象も左の家具から右の壁に変化している。これらのことから、Decoder の Attention map は動作毎とサブゴール位置に対するロボットの注視領域を正しく示しており、Attention の可視化によりユーザがロボット動作に対する判断根拠の理解が容易であると言える。

### 4.3 AR を用いた人への提示に対する有効性調査

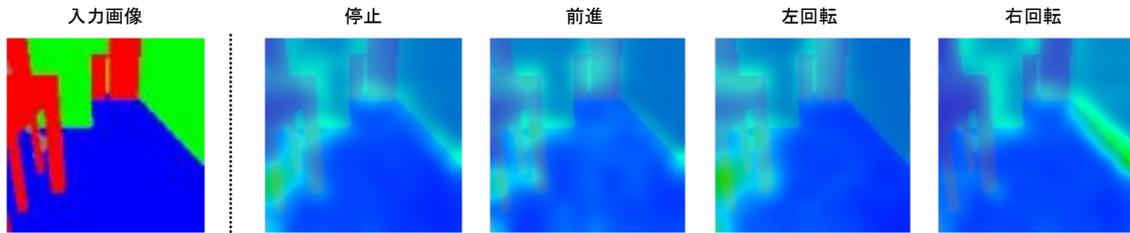
本実験では 2 つの方法により、AR を用いた Attention 可視化がロボット動作に対するユーザの理解に有効か調査した。

**ユーザのロボット動作予測による調査** AR を用いた Attention 可視化による提示の有無で被験者を 2 グループに分け、被験者がロボットの動作を予測できるか調査した。本調査では、図 2 (a) に示す CG 環境を使用し、33 名の被験者を対象とした。調査手順を以下で述べる。

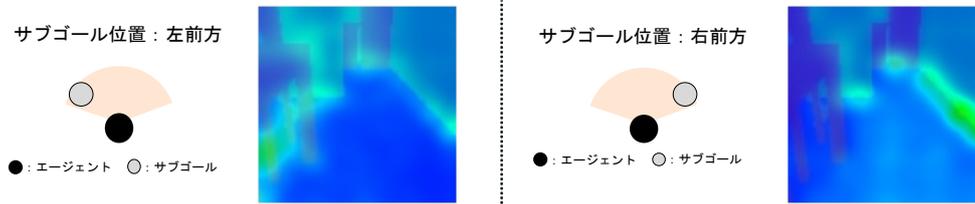
1. 被験者はロボットの自律移動に関するデモ動画を視聴する。視聴する動画として、提示ありグループは AR により Attention を可視化したデモ動画、提示なしグループは通常のデモ動画である。
2. 被験者は、特定シーンの静止画からロボットの次動作を予測する問題に 10 問取り組み、ロボット動作を学ぶ。静止画は、1 で視聴したデモ動画と同様の方法で可視化したものである。つまり、提示ありグループは Attention を含めロボット動作を学ぶ。
3. 被験者は、通常の静止画 (Attention なし) からロボットの次動作を予測する問題に 10 問取り組む。これら問題の正答率を集計し、AR による提示が有効か調査する。

手順 3 の問題に対する全被験者の平均正答率を表 2 に示す。表 2 から、AR を用いた提示ありが提示なしと比較し 7% 向上している。この結果から、AR を用いたロボット動作の提示はユーザがロボット動作を予測する上で有効と考えられる。

**ユーザのロボット動作に対する理解度** ユーザのロボット動作に対する理解において、AR を用いた Attention 可視化が有効か確認するためアンケート調査を



(a) 行動毎の可視化：サブゴール位置は左前方である。



(b) サブゴール位置の変化による可視化: サブゴール位置のみ変更し“停止”に対する Attention map を可視化した。また入力画像は (a) と同様のシーンである。

図 4: Decoder の Attention map 可視化例

表 2: ユーザによるロボット動作の予測

被験者グループ	提示なし	提示あり
平均正答率 [%]	31.8	38.8

表 3: ロボット動作に関する理解度アンケートの集計結果: 数値はそれぞれ回答した被験者の割合を示す。

評価段階	定点肉眼観測 [%]	AR を介した観測 [%]
理解できる	14.1	<b>30.6</b>
ある程度理解できる	41.3	<b>45.8</b>
あまり理解できない	<b>22.8</b>	18.1
理解できない	<b>21.7</b>	5.6

行った。定点カメラによるロボット動作シーンの動画 (定点肉眼観測) と、提案手法である AR による Attention 可視化を含むロボット動作シーンの動画 (AR を介した観測) を 23 名の被験者に視聴してもらう。その後、被験者に対しロボット動作に対する理解度アンケートをとる。本調査では理解度を 4 段階評価とした。また動画は図 2 (b) に示す実環境下で 4 パターン撮影し、AR デバイスには Microsoft 社が開発した HoloLens2 を用いた。

各動画に対する理解度アンケートの集計結果を表 3 に示す。表 3 から、AR を介した観測が定点肉眼観測と比較し、“理解できる”と答えた被験者の割合が増加している。この結果から AR を用いた Attention 可視化によるユーザへの提示は、ユーザがロボット動作を理解する上で有効と考えられる。

## 5. おわりに

本研究では、深層強化学習によるロボットの自律動作獲得と思考プロセスの可視化、およびユーザに対する AR を用いた効率的なロボットの思考理解を提案した。

提案手法は Transformer モデルの導入によりロボットの高い自律移動性能を獲得し、ロボット動作に対する視覚的説明を実現した。また、AR を用いてロボット動作の判断根拠を可視化することで、ユーザのロボット思考への理解促進を実現し、人とロボットの共存に貢献できる可能性を示した。今後は、モデルの性能向上と他タスクでの実験に取り組む予定である。

## 参考文献

- [1] T. Miller, “Explanation in artificial intelligence: Insights from the social sciences”, *Artificial Intelligence*, vol. 267, pp. 1–38, 2019.
- [2] A. Vaswani, *et al.*, “Attention is all you need”, *Advances in neural information processing systems*, 2017.
- [3] S. Greydanus, *et al.*, “Visualizing and understanding atari agents”, *Proceedings of international conference on machine learning*, 2018.
- [4] M. Hessel, *et al.*, “Rainbow: Combining Improvements in Deep Reinforcement Learning.”, In the *AAAI conference on artificial intelligence*, 2018.
- [5] H. Itaya, *et al.*, “Action Q-Transformer: Visual Explanation in Deep Reinforcement Learning with Encoder-Decoder Model using Action Query”, *arXiv preprint arXiv:2306.13879*, 2023.
- [6] Z. Makhataeva, *et al.*, “Safety Aura Visualization for Variable Impedance Actuated Robots”, *IEEE/SICE International Symposium on System Integration*, 2019.
- [7] H. Hedayati, *et al.*, “Improving Collocated Robot Teleoperation with Augmented Reality”, *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, 2018.
- [8] V. Mnih, *et al.*, “Human-level control through deep reinforcement learning”, *Nature*, Vol. 518, No.7540, pp. 529–533, 2015.
- [9] V. Badrinarayanan, *et al.*, “SegNet: A Deep Convolutional EncoderDecoder Architecture for Image Segmentation”, *IEEE transactions on pattern analysis and machine intelligence*, pp. 2481–2495, 2017.