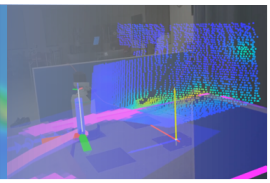
機械知覚&ロボティクスグループ
中部大学　工学研究科　ロボット理工学専攻

MPRG
MACHINE PERCEPTION AND ROBOTICS GROUP

2023年度　藤吉研究室　修士論文発表　アブストラクト

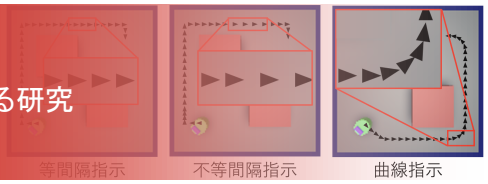
Deep Reinforcement Learning, Explainability, Robotics

深層強化学習によるロボット移動の視覚的説明と拡張現実による提示に関する研究
尹 文韜



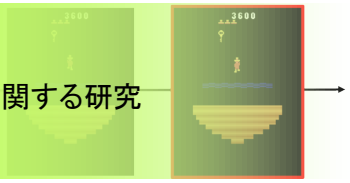
Deep Reinforcement Learning, Sketch Sequence, Robotics

スケッチ指示を用いた深層強化学習によるロボット動作の獲得に関する研究
本多 航也



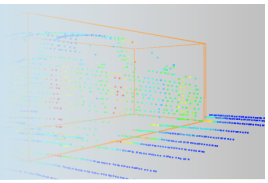
Deep Reinforcement Learning, Curiosity, Explore

深層強化学習における状態遷移を考慮した内発的動機付けによる学習効率化に関する研究
大鹿 海都



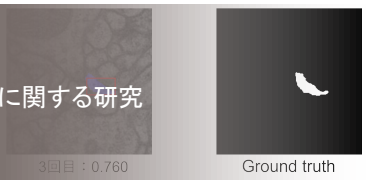
Object Detection, PointCloud, Visual Explanation

点群データを対象とした物体検出の視覚的説明に関する研究
三原 一真



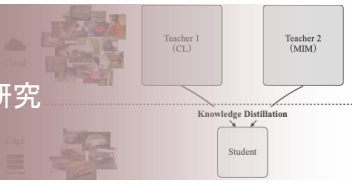
Segmentation, Prompt tuning, EM image

SAM のプロンプトチューニングと繰り返し推論による細胞画像セグメンテーションの高精度化に関する研究
舩井 祥吾



Knowledge Distillation, Self-Supervised Learning

複数の自己教師あり学習モデルを用いた下流タスクにおける知識蒸留に関する研究
鈴木 渉起



深層強化学習によるロボット移動の視覚的説明と拡張現実による提示に関する研究

TR21003 尹 文韜　　指導教授：藤吉 弘亘

1. はじめに

深層強化学習はロボットの高い自律移動性能を獲得できる．一方で，学習した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ており，モデルの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行動選択に対する判断根拠を理解することは非常に困難である．そこで本研究では，自律移動ロボットと人の共存促進を目的とし，ロボットの行動選択に対する判断根拠の可視化，および拡張現実（AR）技術によるユーザに対するロボット動作の判断根拠を提示する手法を提案する．ロボットの高い自律移動性能を獲得するために，深層強化学習に Transformer [1] を導入する．Transformer の Attention をもとに深層強化学習モデルの注視領域を獲得する．そして，AR 技術を活用することで注視領域をユーザに提示する．これにより，ユーザがロボット動作の判断根拠をより容易に理解できるようになる．

2. 深層強化学習の視覚的説明

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．板谷らは深層強化学習に Transformer を導入した Action Q-Transformer (AQT) を提案している [2]．Transformer は入力トークンと他トークン間の類似度を算出して，入力に対する重要な情報を Attention として獲得している．この Attention を可視化することで，注視領域を把握することができる．AQT では Transformer Decoder に入力する Query を行動情報とすることで，各行動に対する注視領域を獲得できる．

3. 提案手法

本研究は深層強化学習手法 Deep Q-Network (DQN) [3] に Transformer 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動作の判断根拠を効率的にユーザへ提示する方法を提案する．

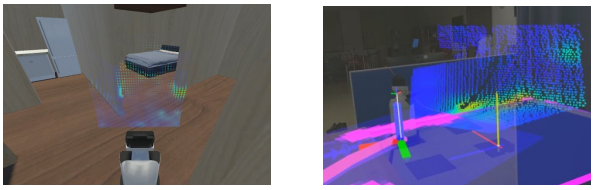
3.1 深層強化学習による自律移動能力の獲得

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．以下で提案手法による行動算出までの流れを述べる．

はじめに，ロボット視点の画像上の物体を学習済みセマンティックセグメンテーションモデルにより壁，床，家具，人の 4 クラスに分類し，それぞれ緑，青，赤，灰色で表示する．これは，シミュレーション環境と実環境とのテクスチャの相違が生じるドメインギャップを吸収するためである．そして，この画像を 84×84 ピクセルにリサイズ後，12×12 の個のパッチ画像に分割する．つまり，パッチ画像のサイズは 7×7 ピクセルである．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 Encoding により位置情報を付与し Encoder へ入力する．また，Decoder は Encoder の出力を Value と Key とし，Action query を Query とする．Action query は Query branch で算出した各動作とゴール情報を含む特徴ベクトルである．最後に，Decoder の各 Query に対する出力を Q Heads へ入力し，各動作の Q 値を算出する．ここで，エージェントであるロボットの動作は停止，前進，左右旋回の 4 種類である．

3.2 AR を用いたユーザへの提示

Transformer Decoder では行動情報を含む Action query 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，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．シミュレーション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．



(a) シミュレーション環境　　(b) 実環境

図 2: AR による Attention 可視化例

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，HSR ロボットによる屋内環境における自律移動タスクを対象とする．実環境でロボットを自律移動させて学習させるコストは高い．そのため，Unity 上で屋内環境を再現し，人や机等の障害物を設置した．また，ドメインギャップを対処するため，シミュレータ内のテクスチャをオブジェクトごとに色分けした．本環境では，エピ

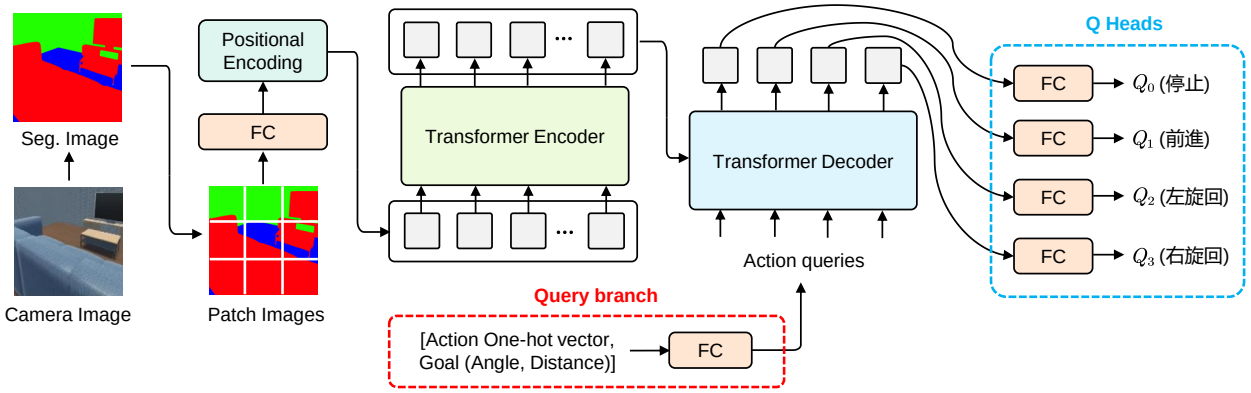


図 1: 提案手法のネットワーク構造

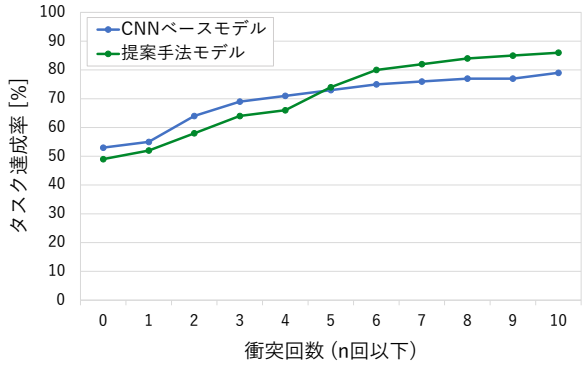
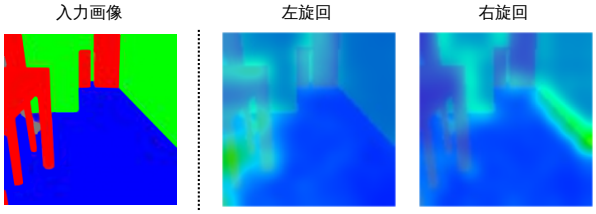
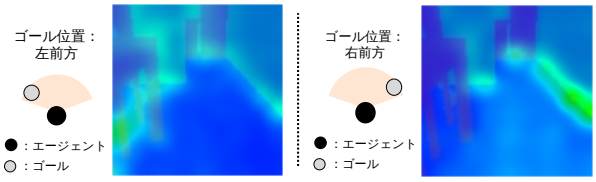


図 3: 100 エピソード間のタスク達成率



(a) 行動の変化による可視化: ゴール位置が左前方である時の “左旋回” と “右旋回” に対する Attention.



(b) ゴール位置の変化による可視化: ゴール位置のみ変更した時の “停止” に対する Attention.

図 4: Decoder の Attention 可視化例

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．

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を比較する．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に到達，かつ衝突が一定回数以下とした．比較手法は，従来の CNN ベースモデルと提案手法の Transformer ベースモデルである．各モデルは同じ学習条件により 1.0×10^6 ステップの学習を行った．図 3 から，提案手法の Transformer ベースモデルは CNN ベースモデルと比較し，エピソードごとの衝突回数がやや増加しているが，最終ゴールへの到達率が向上していることが分かる．このことから提案手法は高い自律移動性能を獲得できていると言える．

Attention の可視化　獲得した Attention 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，Decoder の行動やゴール位置に対する注視領域の変化を確認した．図 4 (a) から，ロボットは，左旋回動作では左の家具を，右旋回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，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 は動作毎と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いることが分かる．

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を調査した．

表 1: 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 2: ロボット動作に関する理解度アンケートの集計結果

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

ユーザのロボット動作予測による調査　本調査では，図 2 (a) に示すシミュレーション環境を使用し，33 名の被験者を対象とした．被験者を 2 グループに分け，それぞれ Attention 提示なしとありの状態でロボットが自律移動するデモ動画を視聴し，練習問題に回答することで，ロボットの動作について学んでもらう．その後，すべての被験者に Attention なしの状態でロボットの動作予測問題を 10 問回答してもらった．動作予測問題に対する全被験者の平均正答率を表 1 に示す．表 1 から，AR を用いた提示ありのグループが提示なしグループよりも正答率が 7pt 向上している．この結果から，AR を用いた Attention 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 可視化が有効であるか確認するためにアンケート調査を行った．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，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った．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行った．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 2 に示す．表 2 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行動選択に対する判断根拠の可視化，およびユーザに対する AR を用いた効率的なロボット動作の理解を提案した．提案手法は Transformer 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，ユーザのロボット動作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．

参考文献

[1] A. Vaswani, *et al.*, “Attention is all you need”, Advances in neural information processing systems, 2017.

[2] 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.

[3] V. Mnih, *et al.*, “Human-level control through deep reinforcement learning”, Nature, Vol. 518, No.7540, pp. 529–533, 2015.

研究業績

[1] 尹文韜，板谷英典，真野航輔，平川翼，山下隆義，藤吉弘亘,“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”, 日本ロボット学会学術講演会, 2023.

スケッチ指示を用いた深層強化学習によるロボット動作の獲得に関する研究

TR22011 本多 航也　　指導教授：藤吉 弘亘

1. はじめに

少子高齢化の拡大に伴い，介護業界では労働者不足が深刻な課題となっている．そのため，高齢者や障がい者の生活支援を目的としたロボットの実用化が期待されている．生活支援ロボットの操作を容易にするには，直感的な指示を可能とすることが重要である．これまでに，自然言語指示とスケッチ指示を用いた研究が取り組まれている．Zitkovich らは，自然言語指示を用いてロボットの行動を指示する RT-2 を提案している [1]．ロボットに移動を指示する場合，自然言語指示はどのように移動させるか指示が曖昧になりやすい．一方で，スケッチによる指示は，曲がる位置や止まる位置などを直感的かつ適切に指示できる．

そこで本研究では，スケッチ指示からのロボット動作の獲得のために，Action Q-Transformer (AQT)[2] に，スケッチ指示を考慮するための機構である Instruction Phase を導入した手法を提案する．評価実験により，提案手法はスケッチ指示を考慮するロボット動作の獲得に有効であることを示す．

2. Action Q-Transformer

AQT[2] は Transformer 構造を導入した深層強化学習手法である．AQT のモデル構造を図 1 に示す．AQT は，環境から取得した画像を CNN を用いて特徴量に変換して，位置情報を付与して Transformer Encoder に入力する．そして，Transformer Encoder からの出力を Key，Value として Transformer Decoder に入力する．Query には Action query を入力し，入力画像と Action query の関係性を算出する．Action query は行動ごとに作成した One-hot Vector を全結合層を用いてトークンにしたものである．Action query を行動数分用意することで行動ごとの Q 値を算出できる．また，AQT は入力画像と Action query との Attention を計算しているため，Attention を可視化することで行動選択に寄与する画像内の領域を解析可能となる．

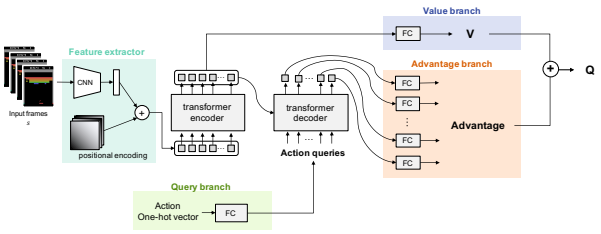


図 1: AQT のモデル構造（文献 [2] より引用）

3. 提案手法

本研究では，生活支援ロボットのナビゲーションタスクにおいてスケッチ指示を用いたロボット動作の獲得を目的とし，新たなモデル構造を提案する．

3.1 提案手法の構造

本手法は，AQT をベースにスケッチ指示を考慮できる機構を導入する．提案手法のモデル構造を図 2 に示す．提案手法は，Sketch Sequence から Instruction Feature を算出する Instruction Phase と，ロボット動作に対する Q 値を算出する Control Phase から構成されている．深層強化学習アルゴリズムとして Deep Q-Network (DQN) [3] を用いる．

Instruction Phase

Instruction Phase は事前に作成したスケッチ指示と Transformer Encoder を用いて，時系列情報を考慮しながら Instruction Feature を算出する段階である．スケッチ指示から時系列を考慮した Sketch Sequence を作成する．そして，全結合層を用いて特徴量に変換して，位置情報を付与して Transformer Encoder へ入力する．Transformer Encoder の Self-Attention を用いて Sketch Sequence 間の関係性を考慮した Instruction Feature を獲得する．

Control Phase

Control Phase はロボット動作を獲得するための段階である．提案手法は，Transformer Encoder-Decoder 構造を用いており，入力には各ステップの 128 × 128 ピクセルの画像，ロボットの位置と向き，Instruction Phase で算出した各ステップに該当する Instruction Feature を用いる．入力画像をパッチ分割したあと，位置情報を足し合わせて Transformer Encoder に入力する．Transformer Encoder の出力を Key，Value として Transformer Decoder へ入力する．Query は行動ごとに作成した One-hot Vector とロボットの位置と向き，Instruction Feature を連結して全結合層を用いて特徴量に変換した Action query を用いる．Transformer Decoder からの出力を Q 値を算出する Q-Head に入力して，最終的な Q 値を求める．

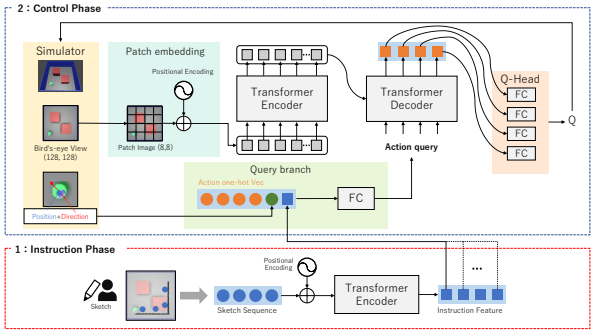


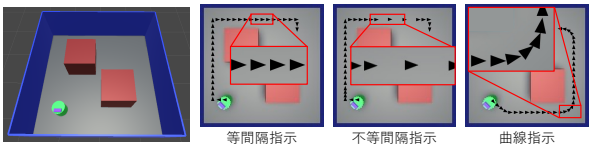
図 2: 提案手法のモデル構造

4. 評価実験

提案手法の有効性を確認するため，生活支援ロボットのナビゲーションタスクを用いて Displacement Error による定量的評価，ロボット動作と Attention の可視化による定性的評価を行う．

4.1 実験概要

本研究では，物理シミュレータである Unity3D を用いて作成した生活支援ロボットのナビゲーションタスクを対象とする．学習で使用する実験環境とスケッチ指示の例を図 3 に示す．移動可能なロボットをスケッチ指示の通りに移動させることを目標とする．エージェントの行動はその場で停止する Noop，前方へ進む Forward，右回転の Right，左回転の Left の 4 種類である．報酬は，各ステップのスケッチ指示の位置を waypoint として，waypoint とロボットの位置のユークリッド距離の差を負の報酬とする．



(a) 実験環境　　(b) スケッチ指示の例

図 3: 実験環境とスケッチ指示の例

本研究の学習で用いるスケッチ指示は，等間隔のスケッチ指示を 4 種類，不等間隔のスケッチ指示を 2 種類，曲線のスケッチ指示を 4 種類，計 10 種類のスケッチ指示である．等間隔指示，不等間隔指示，曲線指示それぞれのスケッチ指示の例を図 3(b)に示す．

提案手法の比較手法として，スケッチ情報を用いた CNN ベースのモデルを用いる．スケッチ情報を用いた CNN ベースのモデルは，提案手法と同じく Instruction Phase と Control Phase で構成されている．Instruction Phase は，提案手法と同様の構造を用いる．Control Phase は画像から特徴量を算出する Image-Feature extractor，ロボットの

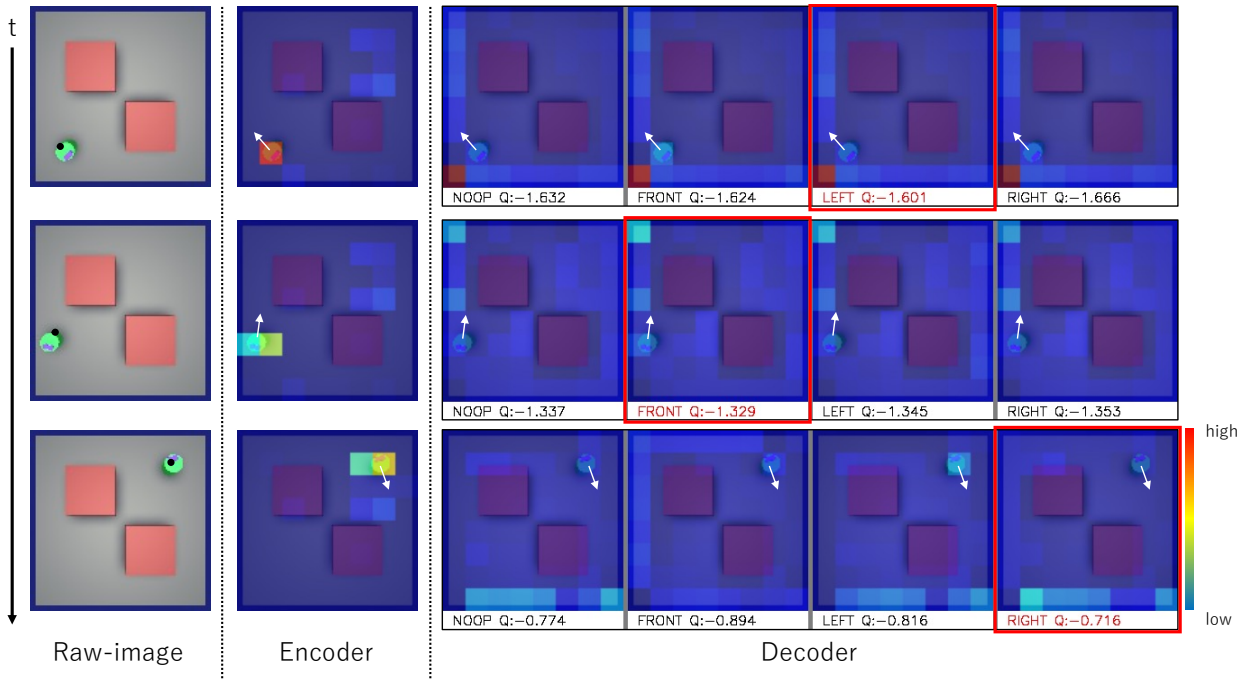


図 4: ロボット動作と Attention の可視化

位置と向きから特徴量を算出する Position-Head，Image-Feature extractor の出力と Position-Head の出力，各ステップに該当する Instruction Feature から Q 値を算出する Q-Head で構成されている．

学習条件は学習率を 0.0001，割引率を 0.99，replaybuffer のサイズを 250,000 とし，学習の終了条件は 1,000,000 エピソードに達した場合，エピソードの終了条件は 100 ステップが経過した場合のみとする．

4.2 定量的評価

本研究では評価指標に，Average Displacement Error (ADE) 及び， Final Displacement Error (FDE) の 2 つの評価指標を用いて評価を行う．CNN ベースモデルと提案手法を用いた評価結果を表 1 に示す．表 1 はスケッチ指示ごとの ADE および FDE の平均と全スケッチ指示の平均である．表 1 より，すべてのスケッチ指示において提案手法が CNN ベースモデルより ADE および FDE が小さい．このことから，Instruction Feature を用いた Action query を Decoder の Query に入力することで，直感的に作成したスケッチ指示にも対応することができると考えられる．また，提案手法のスケッチ指示ごとの ADE および FDE のばらつきが小さいことが確認できる．このことから，時系列を考慮した Sketch Sequence を Instruction Phase で解析し，Decoder に入力することで汎化性能が向上して，どのスケッチ指示であっても安定した精度を獲得することが可能である．

表 1: Displacement Error を用いた評価結果

比較手法	CNN モデル		提案手法	
評価指標	ADE	FDE	ADE	FDE
等間隔指示	0.41	0.25	**0.27**	**0.19**
不等間隔指示	0.40	0.24	**0.29**	**0.22**
曲線指示	0.92	3.10	**0.25**	**0.25**
平均	0.61	1.39	**0.26**	**0.22**

4.3 定性的評価

提案手法で獲得した図 3(b)の等間隔指示におけるロボット動作と Attention を可視化して，定性的評価を行う．提案手法で獲得したロボット動作と Attention の可視化例を図 4 に示す．Raw-image の黒い点はスケッチ指示から生成した waypoint，Decoder の各図下の文字は行動と Q 値を表しており，赤文字の行動が選択された行動，画像中の白矢印がロボットの向きである．

ロボット動作の可視化

図 4 の Raw-image から，スケッチ指示に追従するようにロボット動作を獲得していることが確認できる．このことから，AQT に Instruction Phase を導入することでスケッチ指示を考慮したロボット動作の獲得が可能であると言える．

Attention の可視化

図 4 の Encoder の Attention から，ロボットの領域に注目していることがわかる．このことから，Encoder は入力画像から適切なロボット動作の獲得に重要となる領域に注目していると考えられる．図 4 の Decoder の Attention から，選択した行動が強い Attention を獲得していることが確認できる．また，選択された Action query が「FRONT」の場合（2 行目の赤枠）はロボットの前方が，「RIGHT」の場合（3 行目の赤枠）はロボットの右側の領域に注目していることがわかる．このことから，Decoder の Query に Action query を用いることで行動選択に寄与する領域を注目できると言える．

5. おわりに

本研究では，スケッチの情報を考慮したロボット動作の獲得のために，Transformer を導入した深層強化学習モデルにスケッチ指示を考慮するための機構である Instruction Phase を導入した手法を提案した．定量的評価より，人が直感的に作成したスケッチ指示を用いる場合でも安定した精度を獲得できることを確認した．また，定性的評価からスケッチ指示を考慮するロボット動作の獲得に有効であることを示した．今後は，さらなる精度向上と大規模な環境を用いての実験に取り組む予定である．

参考文献

[1] B. Zitkovich, *et al.*, “RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control”, PMLR, 2023.

[2] H. Itaya, *et al.*, “Action Q-Transformer: Visual Explanation in Deep Reinforcement Learning with Encoder-Decoder Model using Action Query” , arXiv:2306.13879, 2023.

[3] V. Mnih, *et al.*“Human-level control through deep reinforcement learning”, Nature, 2015.

研究業績

[1] 本多航也 等, “Mask-attention 機構を導入した PPO による物体把持動作の視覚的説明”, 日本ロボット学会学術講演会, 2022.

深層強化学習における状態遷移を考慮した内発的動機付けによる学習効率化に関する研究

TR22003 大鹿 海都　　指導教授：藤吉弘亘

1. はじめに

深層強化学習はエージェントと環境との相互作用により経験を収集し，報酬をもとにエージェントの行動を学習する．しかし，Montezuma's Revenge のような報酬が疎な環境では，報酬に繋がる経験を獲得することが困難となり，その結果学習が停滞する．この問題を解決する手法として，エージェントの内発的動機付けによる探索手法が提案されている [1]．本手法は本来のタスクに関する外部報酬に加え，未知の状態への訪問を誘発する内部報酬を用いる．内部報酬により，報酬が疎な環境においてもエージェントによる環境の探索が促進され，それに伴い外部報酬に繋がる経験の獲得が可能となる．内発的動機付けによる深層強化学習手法はエージェントの現状態に着目して内部報酬を生成するため，到達までの経路が異なる場合でも同じ内部報酬が生成される．そこで，本研究では新たな内部報酬として状態遷移に着目する．これにより，内発的動機付けは経験の過程も考慮するため幅広い探索表現が可能となり，エージェントによる環境探索を効率的に行えると考える．評価実験では Atari2600 を用いた評価実験により，エージェント性能を解析することで状態遷移を考慮する有効性を示す．

2. Random Network Distillation

Random Network Distillation (RND) [1] は内発的動機を用いた強化学習手法の 1 つである．RND は状態を 2 つの畳み込みニューラルネットワークに入力し，それらが出力する値の差を内部報酬とする．Random feature network は教師の役割を担っており，学習開始時に初期化したパラメータで固定する．Predictor network は学習可能なネットワークであり，Random feature network の出力を近似するように学習する．Predictor network の学習によって同じ状態を観測し続けるほど 2 つのネットワーク間の出力誤差が小さくなり，内部報酬値が低下する．一方，未知の状態を観測した場合は未学習の状態が入力であるため出力誤差が大きくなり，大きな内部報酬値が出力される．このように RND は各ネットワークの出力誤差をもとに状態の真新しさを評価する．

3. 提案手法

深層強化学習に内発的動機付けを用いた環境の探索効率化において，状態遷移を考慮した新たな内発的動機付けを提案する．

3.1 提案手法の構造

図 1 に提案手法による内部報酬生成機構を示す．本手法は RND の入力次元を任意の時刻 $n+1$ まで拡張し，現状態から n 時刻前までの状態列を入力とし，各ネットワークの出力誤差を内部報酬として出力する．状態列に対する出力誤差を状態遷移の真新しさとして評価することで，同じ現状態に対しても過去に経験した状態の遷移によって異なる評価を行う．

3.2 状態列の作成

状態列は現在を含めた過去 n 時刻前までの queue 型の配列である．状態列の作成手順を以下に説明する．

1. エージェントの行動 a により遷移後の状態を観測する
2. 状態列の最も過去の状態 s_{t-n} を忘却する
3. 状態列の s_{t-i} の状態は s_{t-i-1} の状態として格納する
4. s_t に観測した状態を格納する

このように状態を観測するたびに状態列の 1 番新しい状態として格納する．また，エピソード終了後は状態列を全てリセットし，0 埋めした状態として扱う．

3.3 エージェントの学習に用いる報酬

式 (1) に内部報酬の算出式を示す．このとき，$\boldsymbol{c}$ は状態列 $[s_t, s_{t-1}...s_{t-n}]$，$f_1(\boldsymbol{c})$ は Predictor network の出力，

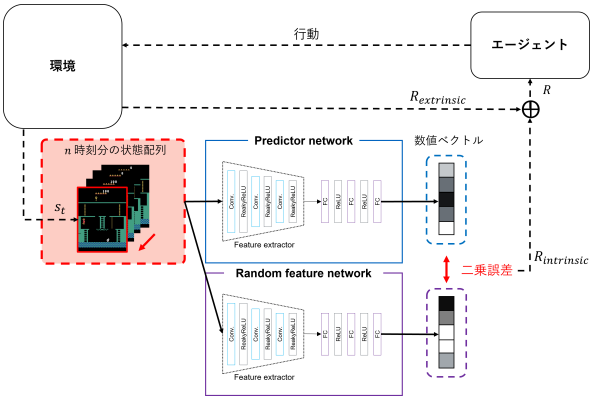


図 1: 提案手法による内部報酬生成機構

$f_2(\boldsymbol{c})$ は Random feature network の出力を示す．

$$R_{intrinsic} = ||f_1(\boldsymbol{c}) - f_2(\boldsymbol{c})||^2 \tag{1}$$

エージェントが学習に用いる報酬値は式 (2) で求める．

$$R = R_{extrinsic} + R_{intrinsic} \tag{2}$$

環境から得られた外部報酬と式 (1) によって求めた内部報酬の総和を報酬値とすることでエージェントは状態遷移を考慮した探索的な行動選択を実現する．

4. 評価実験

Atari2600 のビデオゲームタスクを用いて，状態遷移を考慮した内発的動機付けの有効性を確認する．

4.1 実験概要

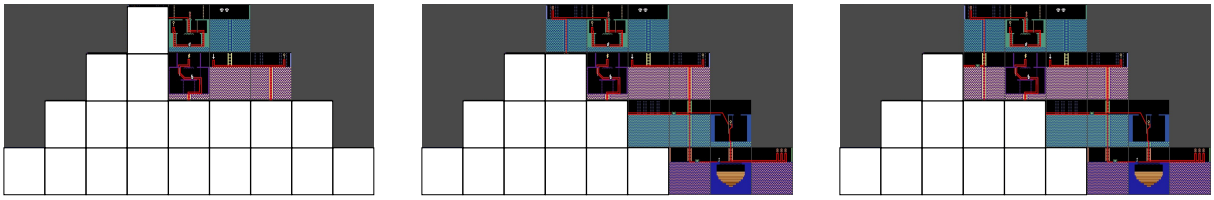
本実験ではまず，内発的動機付けなしのモデル (PPO)，RND による内発的動機付けを用いたモデル (PPO+RND)，提案手法を用いたモデル (PPO+Ours) を用いて学習済みモデルによる獲得スコア平均及び学習時の外部報酬推移を比較する．評価環境は Atari2600 の 報酬が疎な環境である Montezuma's Revenge(MR)， Venture(VT)， 報酬が密な環境である Breakout(BO) の 3 つである．比較の際は，Montezuma's Revenge は 1.5×10^3 エピソード，Venture は 1.0×10^3 エピソード，Breakout は 4.5×10^3 エピソード学習したモデルを使用する．次に Montezuma's Revenge 環境を用いて学習中に訪れた部屋の可視化を行うことで各手法に対するエージェントの探索範囲を確認する．最後に，学習した提案手法による内部報酬生成モデルを用いて観測状態に対する内部報酬を確認することで出力される内部報酬の特徴を明確に解析する．

4.2 ゲームスコアによる比較

図 2 に学習時の外部報酬推移，表 1 に各学習モデルを 100 エピソード実行したときの平均スコアを示す．このとき，表 1 の random はエージェントがランダムで行動選択を行った際の平均スコアである．Montezuma's Revenge では PPO+Ours が最も高いスコアを獲得している．さらに図 2(a) においても Ours が最も高い報酬を獲得していることが分かる．この結果から状態遷移を考慮した網羅的な探索が高いスコアの獲得に貢献できていると考えられる．Venture でも PPO+ours が最も高いスコアを獲得している．しかし，図 2(b) より報酬推移に大きな変化は見られなかった．Venture の環境は部屋数が Montezuma's Revenge と比較してかなり少なく，部屋の切り替わりによる経験の種類が限られている．提案手法では部屋の切り替わり等による大きな状態変化に着目しているため，部屋数が少ない Venture では RND と類似した推移が見られたと考えられ

(a) Montezuma's Revenge (b) Venture (c) Breakout

図 2: 学習時の報酬推移



(a) Montezuma's Revenge (b) Venture (c) PPO+Ours

図 3: 各手法毎の探索範囲

表 1: 各環境の平均スコア

env	methods			
	Random	PPO	PPO+RND	PPO+Ours
MR	0	2500	4600	**6600**
VT	0	0	1660	**1810**
BO	2	**422**	417	402

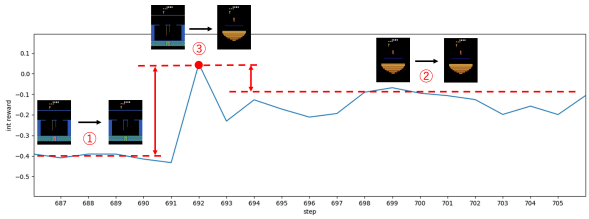


図 4: 各ステップの内部報酬の推移

る．Breakout では 3 つの手法間でスコアと図 2(c) による報酬推移ともにに大きな変化が生まれなかった．これは Breakout が報酬が密な環境であり，表 1 の Random のスコアからも，1 以上獲得できていることが分かる．そのため，外部報酬のみで十分学習可能であり，内部報酬導入による効果は低いことが考えられる．

4.3 Montezuma's Revenge の探索範囲の可視化

図 3 に各手法における学習時の探索範囲を示す．学習を通して PPO は 5 種類，PPO+RND は 12 種類，PPO+Ours は 13 種類の部屋を学習全体を通して訪れた．この結果から，提案手法である Ours が最も多くの状態を訪れていることが分かる．また，提案手法のみが訪れている部屋に着目すると，学習によって進んでいる右下の部屋ではなく左方向の部屋であることから探索的な行動の促進が行われていると考えられる．

4.4 内部報酬の可視化による比較

図 4 に Montezuma's Revenge にてエージェントが部屋を移動する瞬間における内部報酬推移を示す．図 4 の①と②に着目すると，②で高い内部報酬が出力されていることが分かる．このことから②の部屋が①の部屋より珍しいことが分かる．さらに③に着目すると，②より高い内部報酬が出力されていることが分かる．これは③は部屋が切り替わる際の内部報酬であり，部屋が切り替わる際の経験は切り替わった後の経験より少ないからだと考えられる．この結果から部屋の変化に対して変化した際とその後に対して異なる内部報酬が出力されていることが分かる．これにより長期間の探索により状態への探索促進が低下しても，状態遷移を考慮した探索促進により探索の継続が期待できる．

5. おわりに

本研究では，深層強化学習において過去の状態にも着目した幅広い探索表現を実現するため，状態遷移を考慮した内発的動機付けを提案した．提案手法では，現状態を始めとした複数時刻分の状態を RND による内部報酬生成機構に入力することで状態遷移を考慮した内部報酬の出力を実現した．評価実験では 5 時刻分の時系列情報を用いて状態遷移を表現したモデルを Atari2600 の環境を用いて評価した．その結果，状態変化に対して変化の際と変化後で異なる評価ができており，変化の際に対する探索促進の低下が抑制されていることを確認した．また，探索部屋の比較にて提案手法が一番多くの部屋の探索が出来ていることを確認するとともに，提案手法による網羅的な探索が外部報酬の獲得に繋がることを確認した．今後は，状態遷移を用いた内部報酬生成機構に Transformer encoder 構造を導入することで状態遷移の表現力向上を図る予定である．

参考文献

[1] Y. Burda, *et al.*, "Exploration by Random Network Distillation", 2019.

研究業績

[1] 大鹿海都 等, "深層強化学習における状態遷移を考慮した内発的動機付けによる探索の効率化", 電子情報通信学会技術研究報告, パターン認識・メディア理解研究会, 2024

点群データを対象とした物体検出手法の視覚的説明に関する研究

TR22012 三原一真　　指導教授：藤吉弘亘

1. はじめに

自動運転システムを実現するために，周囲の障害物を検知する物体検出は重要なタスクである．自動運転システムのセンサには，RGB カメラ，LiDAR センサ，Radar センサが用いられる．LiDAR センサはレーザ光を照射し，障害物などの周囲の物体から跳ね返ってきた反射光を受光し，ToF 原理により 3 次元座標を取得する．取得した 3 次元座標は点群データとして，セマンティックセグメンテーションや物体検出などのタスクに利用される．点群データを入力とする物体検出手法として，PointPillars[1] がある．PointPillars は，点群データを Pillar と呼ばれる xy 平面に垂直な格子状の空間に分割し，局所的な特徴抽出をすることで，疑似画像に変換している．疑似画像に変換することで，畳み込み処理の適用が容易となり，高速な推論が可能となる．一方で，点群ベースの手法の問題点として，入力した点群データに対する予測処理過程はブラックボックスとして扱われており，人間による判断根拠の判別が困難であるという点がある．そこで本研究では，点群データを入力とする物体検出モデルの視覚的説明の実現を目指す．

2. 関連研究

本節では，点群データを入力とする物体検出手法の PointPillars と，画像データを入力とする判断根拠の視覚的説明手法の ODAM について述べる．

2.1. PointPillars

PointPillars[1] は，Pillar feature Net，Backbone，Detection head から構成される．PointPillars のネットワーク構造を図 1 に示す．Pillar Feature Net は，局所的な特徴を抽出するモジュールである．まず，3 次元座標と反射強度で構成する点群データ (x, y, z, r) を擬似画像へ変換する処理を行う．入力された点群データは，Pillar と呼ばれる xy 平面に垂直に定義される領域に分割され，(D, P, N) の 3 次元のテンソルに変換される．ここで，D は点の次元数，P は Pillar の数，N は Pillar 内の点の数である．Pillar 内の点群データは，Pillar 中心からのオフセットと Pillar 内の点群データの平均値から点の次元数を $D = 9$ に変換し，PointNet へ入力することで (C, P, N) の 3 次元テンソルとなる．ここで，C はチャンネル数である．次に，N 方向の最大値を取ることにより (C, P) の 2 次元テンソルへ変換し，Pillar を元の位置に戻すことで，(C, h, w) の Birds-Eye View の擬似画像となる．ここで，h, w は画像の縦と横の画素数である．Backbone では，Pillar Feature Net で特徴抽出した疑似画像に対して，大局的な特徴抽出を行う．Backbone は，3 層のストライドが 2 の畳み込み層から構成されており，特徴マップのスケールを変更して特徴抽出を行う．その後，それぞれの特徴マップに対して Deconvolution によりアップサンプリングした後，連結して出力する．Detection head では，Backbone の出力から，3D Bounding Box のパラメータとクラススコアの予測結果を出力する．予測結果に対して，Non-Muximum Suppression(NMS) を行い，冗長な検出結果を削除する．

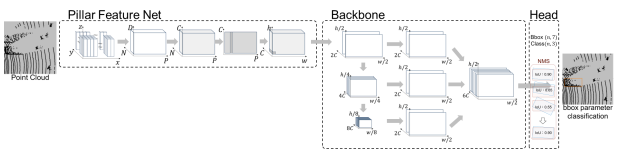


図 1: PointPillars のネットワーク構造

2.2. ODAM

ODAM[2] は，物体検出モデルの中間特徴の勾配情報から Attention Map を作成する．ODAM は，中間特徴に対する Attention Map を Bounding Box の 4 つのパラメータとクラススコアを逆伝播して求める．そして，中間特徴の要素ごとに Attention Map の最大値を取り，1 つの Attention Map にする．ODAM は，検出した Bounding Box に対して個々の Attention Map を求めることができる．

3. 提案手法

本研究では，点群データを入力とする物体検出手法の視覚的説明の実現を目的とする．提案手法はである Grad ODAM は，点群データを物体検出手法である PointPillars の Pillar 特徴による疑似画像に視覚的説明手法である ODAM を導入し，さらに Guided BackPropagation による Attention を統合することで，点群データに対する視覚的説明の獲得を目指す．提案手法を図 2 に示す．

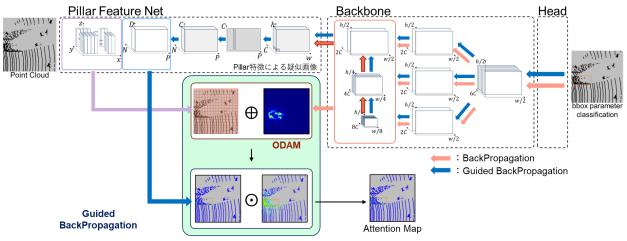


図 2: 提案手法のネットワーク構造

3.1. PointPillars への ODAM の導入

本研究では，PointPillars が，点群データを Birds-Eye View の疑似画像へ変換して特徴抽出する処理に着目し，疑似画像に対する Attention Map を求める．点群データを入力する際に Pillar Feature Net の出力である Pillar 特徴に対して，ODAM を用いて Attention Map を獲得する．次に，Pillar Feature Net のインデックス情報を利用して，Attention Map の値を対応する Pillar 内部の点群データに Attention として付与する．この場合，Attention を付与された点群データは，Pillar 内部で同値の Attetnion となる．

3.2. Guided ODAM

ODAM によって求めた Attention Map は，図 2 の Pillar Feature Net の最初の処理で分割した Pillar に対する Attention であるため，点ごとの視覚的説明ではない．そこで，逆伝播の際に，負の勾配に ReLU 関数を使用して 0 へ変換する Guided BackPropagation の Attention を統合する．これにより，点ごとに異なる Attention を付与することが可能となる．提案する Guided ODAM の式を式 (1) に示す．

$$A_{GuidedODAM} = (A_{ODAM} + g_{input})^2 \qquad (1)$$

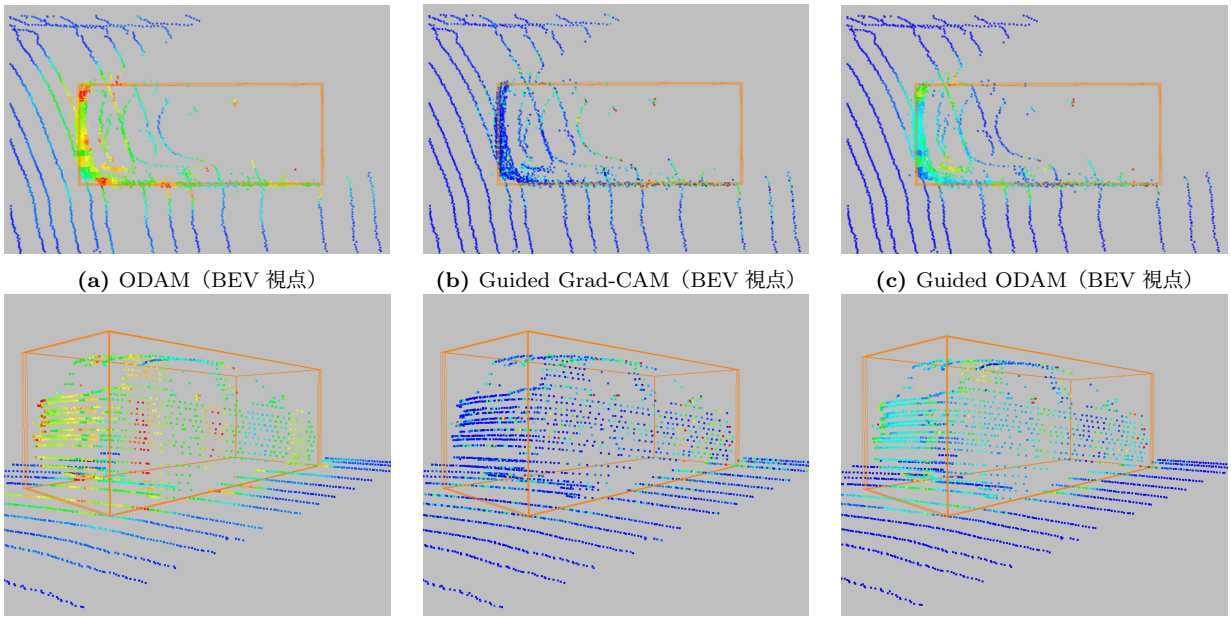
ここで A_{ODAM} は，ODAM の Attention を付与した点群データ，g_{input} は Guided BackPropagation のによるモデルの入力データの勾配情報である．Attention を付与した点群データに，Guided BackPropagation による勾配情報を加算し，2 乗することで Bounding Box 外の低い値の Attention を抑制する．これにより，Bounding Box 内部に高い値の Attention が発生し，点ごとの視覚的説明が可能な Attention Map を求めることができる．

4. 評価実験

本実験では，提案手法に導入する視覚的説明手法を，ODAM，Guided Grad-CAM，Guided ODAM に変更し，それぞれの視覚的説明手法を比較する．

4.1. KITTI dataset

本実験では，KITTI dataset[5] を使用する．KITTI dataset は，自動車に取り付けられた LiDAR センサや RGB カメラ等のセンサから作成された実環境データセットである．KITTI 3D object detection benchmark は，Car，Cyclist，Pedestrian の 3 クラスに対して Easy，Moderate，Hard



(a) ODAM（BEV 視点） **(b)** Guided Grad-CAM（BEV 視点） **(c)** Guided ODAM（BEV 視点）

(d) ODAM **(e)** Guided Grad-CAM **(f)** Guided ODAM

図 3: 視覚的説明手法ごとの定性的評価の比較

の 3 段階の難易度での識別精度の比較を行うベンチマークである．KITTI dataset の学習用データを分割し，3,712 枚，評価に 3,769 枚のデータを使用する．

4.2. 実験概要

本実験ではエポック数を 80，ミニバッチサイズを 8，初期学習率を 0.003，最適化手法を Adam に設定して学習する．

4.3. 評価指標

物体検出の視覚的説明において，Bounding Box 内部の物体に対して局所的に Attention を付与することで，物体検出モデルに対する説明性を可視化できる．本研究では，各予測結果に対する Attention Map 全体の Attention の総和と Bounding Box 内部の Attention の総和の割合の平均をクラスごとに計算することで，Attention Map の Bounding Box に対する局所性を求める．評価指標の計算式を式 (2) に示す．

$$mAPB = \frac{1}{M}\sum_{M}\frac{\sum^{n} Attention_{bbox}}{\sum^{N} Attention_{Map}} \tag{2}$$

M は Attention Map の枚数，n は Bounding Box 内の点群数，$Attention_{bbox}$ は Bounding Box 内の点群データの Attention，$Attention_{Map}$ は Attention Map のすべての点群データの Attention である．

4.4. 視覚的説明の評価結果

視覚的説明の定量的的評価を表 1 に示す．表 1 より，提案手法である Guided ODAM がすべてのクラスに対して最も高い精度であることが確認できる．これは，予測結果の Bounding Box 内部に高い値の Attention が発生し，Bounding Box 外に発生する Attention が抑制されている結果だと考えられる．

表 1: 視覚的説明における定量的評価の比較

手法	Car	Cyclist	Pedestrian
ODAM	71.39	42.65	53.27
Guided Grad-CAM	64.81	57.68	60.32
Guided ODAM	**80.56**	**68.00**	**73.69**

次に，定性的評価を図 3 に示す．図 3(a)，図 3(d) より，ODAM は，Bounding Box 内部が全体的に高い値の Attention が発生し，高さ方向に Attention を確認したときに点が同じ色となっていることが確認できる．これは，ODAM が Pillar 内部の特徴を点群データに付与した結果だと考えられる．そして，Bounding Box の外部にも高い値の Attention を確認できる．図 3(b)，図 3(e) より，Guided Grad-CAM は，Bounding Box 内にまばらに高い値の Attention を確認できる．また，Bounding Box の外部にも高い値の Attention を確認できる．図 3(c)，図 3(f) を見ると Guided ODAM は，Bounding Box 内部が全体的に高い値の Attention が発生し，高さ方向で Attention が変化していることを確認できる．これは，Guided BackPropagation による点ごとの視覚的説明によるものであると考えられる．また，ODAM と Guided Grad-CAM と比較すると，Bounding Box の外部に発生する Attention が低い値であることが確認できる．

4.5. おわりに

本研究では，点群データを入力とする物体検出手法に対する視覚的説明を行う手法を提案した．評価実験では，Guided ODAM は，定量的評価において最も評価指標の値が高く，定量的評価において点ごとの視覚的説明，Bounding Box 外の Attention の抑制による説明性の向上を確認した．今後は，ベースとなる物体検出手法や視覚的説明手法の変更を検討し，さらなる視覚的説明の向上を図る．

参考文献

[1] Alex H. Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, Oscar Beijbom. PointPillars: Fast Encoders for Object Detection from Point Clouds. CVPR,2019.

[2] Chenyang, ZHAO and Chan, Antoni B. ODAM: Gradient-based Instance-Specific Visual Explanations for Object Detection. ICLR,2023.

[3] Selvaraju, Ramprasaath R and Cogswell, Michael and Das, Abhishek and Vedantam, Ramakrishna and Parikh, Devi and Batra, Dhruv. Grad-cam: Visual explanations from deep networks via gradient-based localization. ICCV,2017.

[4] Petsiuk, Vitali and Das, Abir and Saenko, Kate. Rise: Randomized input sampling for explanation of black-box models. arXiv,2018.

[5] Geiger, Andreas, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. CVPR,2022.

研究業績

[1] 三原一真　等，”Self-Attention を用いた PointPillars による 3 次元物体検出の高精度化”　画像センシングシンポジウム，2022.

SAM のプロンプトチューニングと繰り返し推論による細胞画像セグメンテーションの高精度化に関する研究

TR22010 船井 祥吾　　指導教授：藤吉 弘亘

1. はじめに

医用画像や細胞画像のセグメンテーションは，生物学的に関連する形態学的情報を画素レベルで識別することができるため，正確な診断や治療計画のサポートに用いられる．一方で，細胞画像には多くの細胞が含まれているため，個々の細胞を正確に分離してセグメンテーションをすることは難しい．そこで，指定した領域ごとにセグメンテーションできる Segment Anything Model (SAM) [1] が注目されている．SAM は，モデルに与えた指示 (プロンプト) に基づいて前景と背景にセグメンテーションを行う基盤モデルである．例えば，プロンプトにバウンディングボックス (bbox) を使用すると，bbox で囲んだ領域をセグメンテーションすることができる．SAM は，様々な形状に対応することもできるため，細胞画像のセグメンテーションへの応用も可能である．一方，プロンプトを人が定義したり，物体検出結果を用いて与えると，サイズずれが生じることがある．SAM はプロンプトに基づいてセグメンテーションを行うため，プロンプトにサイズずれが生じた場合，セグメンテーションの精度が低下する．

そこで本研究では，SAM にサイズずれを許容する新たなトークンを追加する．さらに，繰り返し推論による最適化を導入し，SAM によるセグメンテーション精度の向上を図る．

2. 関連研究

セグメンテーションの基盤モデルである SAM は，Image encoder，Prompt encoder，Mask decoder から構成される．SAM には，プロンプトチューニング用の学習パラメータ (Learned embedding) が存在しており，Prompt encoder の Positional encoding を行った埋め込みベクトルに加算する．SAM は学習の際に，入力プロンプトに対して，最大 20pixel のランダムなサイズずれを加えているため，0〜20pixel 程度のサイズずれであれば許容できる．

Mazurowski らは，SAM のプロンプトとセグメンテーション能力に関する傾向を明らかにした [4]．主な傾向は以下の通りである．1 つ目は，SAM のセグメンテーションの精度はデータセットごとに差がある．2 つ目は，プロンプトが正確である場合は適切なセグメンテーションが得られる．3 つ目は，bbox をプロンプトとした場合，point をプロンプトにした場合よりも高精度なセグメンテーション結果が得られる．4 つ目は，複数の point をプロンプトに利用した場合，セグメンテーションの精度がわずかに向上する．これらの傾向から，SAM は与えるプロンプトが重要であることが分かる．

3. 提案手法

本研究では，サイズずれが生じた bbox に対して適切なセグメンテーションを出力するために，Promot encoder に新たなトークンを導入したプロンプトチューニングと，繰り返し推論を提案する．

3.1 プロンプトチューニング

プロンプトチューニングは，入力データに学習可能なパラメータを追加し，学習する手法である [5]．プロンプトチューニングでは，プロンプトチューニング用のトークンをモデルに追加するのが一般的である．そこで，提案手法ではサイズずれを考慮するための学習パラメータを新たにトークンとして SAM の Prompt encoder に追加する．追加したトークンを用いたプロンプトチューニングを図 1 に示す．また，このトークンを全結合層により 128 次元から 256 次元に変換することで，トークンはより複雑な表現を可能にする．これにより，サイズずれに対する頑健性に獲得が可能になる．学習時は，60%の確率で 0%，20%の確率で± 10%，± 20%のサイズずれした bbox をプロンプトとして入力する．学習時にサイズずれした bbox を学習することで，推論時にサイズずれした bbox への頑健背が向上する．

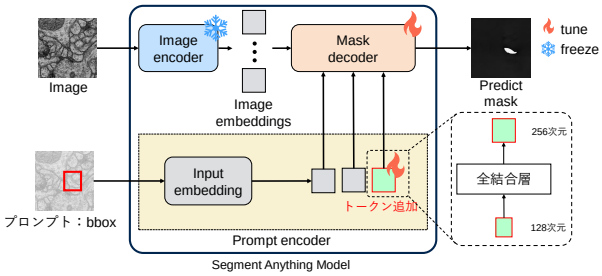


図 1： トークンの追加とプロンプトチューニング

3.2 繰り返し推論

推論時にサイズずれした不正確な bbox のプロンプトを用いた場合，1 度の推論で良いセグメンテーション結果を出力することは難しい．そこで推論時に，繰り返し推論を行う．繰り返し推論を行うことで，サイズずれした bbox を修正することができる．また，Prompt encoder は繰り返し推論の回数分の入力をする必要があるが，パラメータ数の多い Image encoder への入力は一度の入力でよいため，繰り返し推論の時間が大きく増えない．繰り返し推論の流れを図 2 に示す．繰り返し推論は以下のステップで行う．

step1 入力画像とプロンプトを用いて推論を行い，セグメンテーションマスク M_{t-1} を求める．プロンプトにはランダムなサイズずれを含む bbox を用いる．

step2 推論した M_{t-1} から外接矩形を bbox として設定する．

step3 入力画像と設定した bbox を用いて再度推論を行う．ここで，推論した Predict mask を M_t とする．

step4 推論した M_t と 1 回目に推論した M_{t-1} の変化率を求める．ここで，H_bW_b は bbox の面積，HW は画像サイズ，τ はしきい値である．

$$\text{変化率} = \frac{1}{H_bW_b}\sum^{HW}|M_t - M_{t-1}| \qquad (1)$$

step5 変化率に対して，しきい値より上場合，step2 に戻る．しきい値以下の場合，Predict mask を最終的なセグメンテーション結果として出力する．

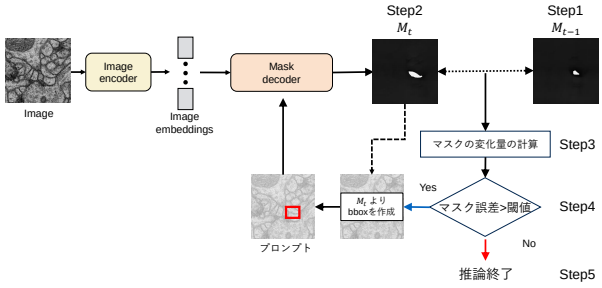


図 2： 繰り返し推論

4. 評価実験

本実験では，電子顕微鏡細胞画像を用いた細胞のセグメンテーションのデータセットである電子顕微鏡データセット，ISBI [2]，Electron Microscopy Dataset [3] で評価実験を行う．これらのデータセットは，bbox が用意されていないため，各細胞の Ground truth のセグメンテーションをもとに作成した外接矩形をプロンプトとして使用する．ベースラインは，SAM をファインチューニングしたモデルとする．SAM のファインチューニングでは，Mask decoder のみを学習する．プロンプトチューニングは，Prompt encoder の

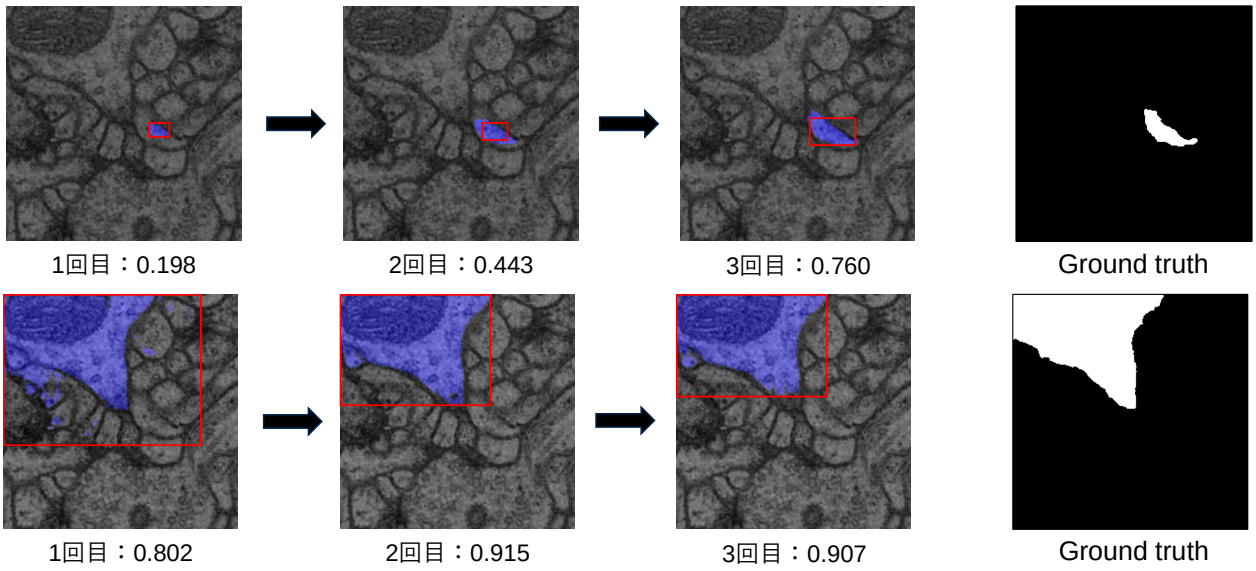


図 3：ISBI データセットの繰り返し推論による Predict mask と mIoU

表 1：電子顕微鏡データセットの実験結果

手法	P	K	bbox のサイズずれ				
			0%	10%	20%	30%	40%
SAM			0.827	0.822	**0.788**	0.710	0.595
	✓		0.827	0.822	0.787	0.709	0.594
提案手法	✓		**0.828**	**0.823**	**0.788**	0.714	0.600
	✓	✓	**0.828**	0.823	0.787	**0.734**	**0.653**
繰り返し推論回数の平均		1	1	1.02	2.42	3.33	

表 2：ISBI データセットの実験結果

手法	P	K	bbox のサイズずれ				
			0%	10%	20%	30%	40%
SAM			0.822	0.810	0.774	0.699	0.604
	✓		0.825	0.810	0.777	0.702	0.606
提案手法	✓		**0.830**	**0.820**	**0.785**	0.714	0.622
	✓	✓	**0.830**	0.817	0.784	**0.727**	**0.650**
繰り返し推論回数の平均		1	1.01	1.05	2.13	3.16	

Learned embedding を学習に追加したモデルである．提案手法のプロンプトチューニングは，Learned embedding のパラメータを固定し，代わりに追加したトークンと Mask decoder を学習する．また，通常のプロンプトチューニングは，既存のモデルのパラメータを凍結させる．しかし，SAM は Mask decoder の学習を行うことで精度が向上することが分かっており，Mask decoder は軽量であるため同時に学習を行う．サイズずれに対する評価では，0%から 40%のサイズずれした bbox をプロンプトとして入力した場合の Predict mask と Ground truth 間の mIoU を比較する．繰り返し推論ではしきい値を 0.05 とする．

4.1 実験結果

電子顕微鏡データセットの実験結果を表 1，ISBI データセットでの実験結果を表 2，Electron Microscopy Dataset での実験結果を表 3 に示す．ここで，各表の P はプロンプトチューニング，K は繰り返し推論を示す．提案手法のプロンプトチューニングのみを導入した場合，全てのデータセットにおいてサイズずれを 30%から 40%加えた場合に SAM を超える認識性能を達成した．どのデータセットにおいてもサイズずれを 30%以上加えた場合で SAM を超える認識性能を達成していることから，プロンプトチューニング時に学習していないサイズずれ度合いに対して頑健であると言える．繰り返し推論では，10%から 20%で精度が若干低下した．しかし，学習を行っていない範囲で大きく精度が向上した．プロンプトチューニングに加えて繰り返し推論を導入することで，さらに認識性能が改善することから，繰り返し推論も有効であると言える．

4.2 繰り返し過程の可視化

繰り返し推論における 1 回目から 3 回目の Predict mask(青色領域)，Ground truth を図 3 に示す．推論を繰り返すことで Predict mask が Ground truth へ近づいていくことから，繰り返し推論が有効であると言える．細胞よりも小さい bbox の繰り返し推論では，bbox を適切な大きさに修正できたため精度向上した．オブジェクトよりも大きい bbox の繰り返し推論では，2 回目の繰り返し推論時に高いセグメンテーション結果が得られた．また，サイズずれが 30%の時の繰り返し推論回数の平均は 2.4 回，40%の場合は 3.2 回となり，サイズずれが大きいほど繰り返し推論回数が増加する結果となった．

表 3：Electron Microscopy Dataset の実験結果

手法	P	K	bbox のサイズずれ				
			0%	10%	20%	30%	40%
SAM			**0.885**	**0.875**	0.844	0.772	0.673
	✓		**0.885**	0.874	0.844	0.772	0.673
提案手法	✓		0.884	0.874	0.845	0.779	0.683
	✓	✓	0.882	0.867	**0.855**	**0.824**	**0.762**
繰り返し推論回数の平均		1.03	1.37	1.54	2.72	3.11	

5. おわりに

本研究では SAM におけるサイズずれのある不正確なプロンプトによる性能低下を改善するプロンプトチューニングと繰り返し推論を提案した．トークンのプロンプトチューニングはサイズずれした bbox に対して精度低下を低減した．また，繰り返し推論をすることにより，さらに精度低下を低減した．繰り返し推論を可視化した結果，bbox が最適化されていることを確認した．今後は，マスクから作成した bbox に対する損失設計によるプロンプトの最適化を行う．

参考文献

[1] A. Kirillov, *et al.*, “Segment anything”, ICCV, 2023.

[2] A. Cardona, *et al.*,“An integrated micro-and macroarar-chitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy”, PLOS Biology, vol.8, no.10, pp.1-17, 2010.

[3] A. Lucchi, *et al.*, “Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features”, IEEE Trans Med Imaging, vol.31, no.2, pp.474-486, 2012.

[4] M. Mazurowski, *et al.*, “Segment Anything Model for Medical Image Analysis: an Experimental Study” , arXiv preprint arXiv:2304.10517, 2023.

[5] B. Lester, *et al.*, “The Power of Scale for Parameter-Efficient Prompt Tuning”, ICLR, 2021.

研究業績

[1] 舩井祥吾, 平川翼, 山下隆義, 藤吉弘亘,“SAM のプロンプトチューニングと繰り返し推論による細胞画像セグメンテーションの高精度化”, 動的画像処理実利用化ワークショップ, 2024.

複数の自己教師あり学習モデルを用いた下流タスクにおける知識蒸留に関する研究

TR22007 鈴木 渉起　　指導教授：藤吉 弘亘

1. はじめに

自己教師あり学習 (SSL) は，大量のラベルなしデータから下流タスクに転移しやすい特徴表現を獲得する事前学習法である．大規模なデータセットを用いた SSL は，教師あり学習を凌駕する性能を達成できるため注目されている．基盤モデル等の SSL モデルは，通常大規模であり，計算コストが高くエッジデバイスでの推論時間やメモリコストが問題となる．そのため，計算コストの低い小規模なモデルが必要とされる．しかし，小規模なモデルは表現能力に欠けているため性能が低い．

本研究では，小規模なモデルの性能向上を目的として，知識蒸留 (KD) [1] による SSL モデルからの知識の転移法を提案する．提案手法は，予備実験により確認した SSL の学習法による特徴表現の差異を基に，複数の SSL モデルを用いて，下流タスクにおける KD を行う．これにより，表現能力の高い小規模なモデルの獲得を目指す．

2. 自己教師あり学習 (SSL)

SSL の学習法には，Contrastive Learning (CL) と Masked Image Modeling (MIM) がある．CL は同じ画像間の特徴量は近づけ，異なる画像間の特徴量は遠ざけるように学習する．一方，MIM は画像の一部をマスクし，マスクした箇所の画素値または特徴量を予測するように学習する．これらの学習法は，獲得する特徴表現が大きく異なることが知られている．そこで，予備実験として CL と MIM の特徴表現の差異を確認する．

予備実験　SSL モデルの Attention Weight（Self-Attention における query と key の関係性）から，Attention distance を算出した結果を図 1 に示す．Attention distance は，層ごとに Attention Weight とピクセル間距離を乗算したものである．これにより，深い層に着目すると，赤色で示す CL は大域的な領域，青色で示す MIM は局所的な領域を認識していることがわかる．この結果から，SSL の学習法によって特徴表現に差異があることがわかる．そこで，それぞれの Attention Weight を小規模なモデルに伝達することで，小規模なモデルにおける表現能力の改善を図る．

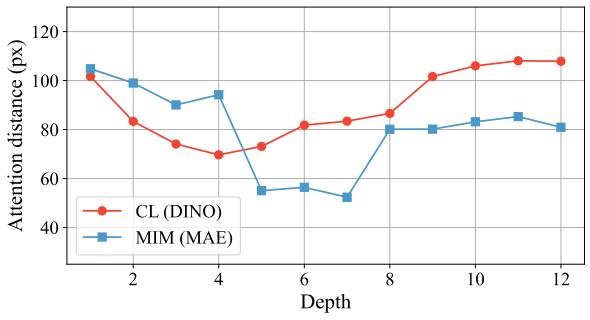


図 1: SSL モデルの Attention distance

3. 提案手法

予備実験を基に，複数の SSL モデルを用いて，下流タスクにおける KD を行うことで，より表現能力の高い小規模なモデルの獲得を目指す．KD は，学習済みモデル (Teacher) の知識を，未学習のモデル (Student) に伝達する手法である．図 2 に提案手法における KD を示す．提案手法では，Teacher として CL と MIM を用いて，2 つのモデルから KD をして Student に転移する．このとき，KD に用いる知識として，従来 KD で用いられるモデルの出力分布と，学習法による特徴表現の差異を確認した Attention Weight を考える．

3.1 Teacher の効率的な学習

学習済みモデルを下流タスクに適用する方法には，線形評価とファインチューニングがある．線形評価は，学習済みモデルにクラス分類のための全結合層 (Head) を結合し，Head 以外を固定して学習する．ファインチューニングは，学習済みモデルも含めて全てのパラメータを学習する．提案手法では，複数の大規模な学習済みモデルを Teacher として用いるため，メモリ消費量の少ない線形評価を用いる．しかし，MIM は線形評価で性能が低く，KD により Student に悪影響を与える．そこで，提案手法では Head に加えて，Transformer encoder の最終層も再学習する．これを Partial-1 と呼ぶ．これにより，MIM における出力分布を用いた KD の効果の改善を期待する．

3.2 出力分布を用いた KD

1 つ目の知識の伝達手段として，Teacher と Student に画像を入力し，得られた出力分布を用いて転移する．損失関数には，一般的な KD と同様に KL ダイバージェンスを用いる．CL と MIM の出力分布を用いた KD を次のように定式化する．

$$\mathcal{L}_{\mathrm{KD}} = \lambda_1 \mathrm{KL}(p^{\mathrm{CL}} \| p^S) + (1-\lambda_1)\,\mathrm{KL}(p^{\mathrm{MIM}} \| p^S), \quad (1)$$

ここで，$p^S, p^{\mathrm{CL}}, p^{\mathrm{MIM}}$ は各モデルの出力分布，λ_1 は CL と MIM の損失の比率を調整する係数である．λ_1 は 0.5 とする．

3.3 Attention Weight を用いた KD

予備実験より，Attention Weight は学習法によって特徴表現に差異があることが判明した（図 1）．そこで，それぞれの特徴表現を小規模なモデルに伝達することで，小規模なモデルにおける表現能力の改善を考える．しかし，モデル構造によって Attention Weight の形状が異なるため，Teacher から Student に Attention Weight を転移することはできない．そのため，2 つ目の知識の伝達手段として，重要な Attention Weight を選択して転移する．

Attention Weight の選択方法　Attention Weight の選択方法として，Attention Confidence (AC) [3] を用いる．ここで，AC は Multi-head attention の重要度を測る指標である．AC による評価値を次のように定式化する．

$$C_h = \frac{1}{|\mathbb{Q}|} \sum_{q \in \mathbb{Q}} \max_{k \in \mathbb{K}} A_h(q,k), \quad (2)$$

ここで，h は Head のインデックス，$\mathbb{Q}$ は query の集合，$\mathbb{K}$ は key の集合，A は Attention Weight である．Teacher の Attention Weight を AC を用いて評価し，最も高く評価された Attention Weight が下流タスクにおいて重要だと考えて Student に転移する．

損失関数　Attention Weight を用いた KD では，KL ダイバージェンスを用いて Student の Attention Weight が Teacher の Attention Weight に近づくように学習する．Attention Weight を用いた KD を次のように定式化する．

$$\mathcal{L}_{\mathrm{Attn}} = \lambda_2 \mathrm{KL}(A^{\mathrm{CL}} \| A_1^S) + (1-\lambda_2)\,\mathrm{KL}(A^{\mathrm{MIM}} \| A_2^S), \quad (3)$$

ここで，A_1^S は Student の 1 番目の Attention Weight，A_2^S は Student の 2 番目の Attention Weight，A^{CL} と A^{MIM} は AC を基準に選択された CL と MIM それぞれの Attention Weight である．λ_2 は CL と MIM の損失の比率を調整する係数であり，λ_2 は 0 とする．

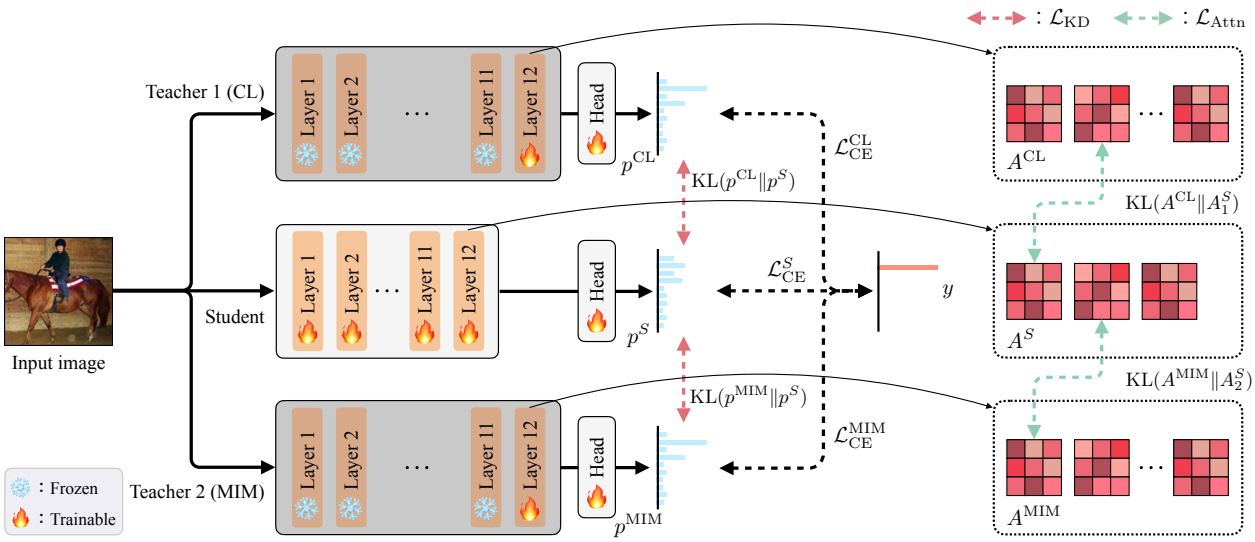


図 2: 提案手法の学習方法

3.4 Teacher の線形評価と KD の同時最適化

従来手法では，SSL モデルを線形評価やファインチューニングした後に KD を行うため，2 段階の学習になる．提案手法では，SSL モデルの線形評価と KD を同時に行うことで，1 段階の学習にする．全体の損失関数 $\mathcal{L}$ を次のように定式化する．

$$\mathcal{L} = \mathcal{L}_{\mathrm{CE}}^{S} + \mathcal{L}_{\mathrm{CE}}^{\mathrm{CL}} + \mathcal{L}_{\mathrm{CE}}^{\mathrm{MIM}} + \mathcal{L}_{\mathrm{KD}} + \mathcal{L}_{\mathrm{Attn}}, \qquad (4)$$

ここで，$\mathcal{L}_{\mathrm{CE}}^{S}, \mathcal{L}_{\mathrm{CE}}^{\mathrm{CL}}, \mathcal{L}_{\mathrm{CE}}^{\mathrm{MIM}}$ は正解ラベルと出力分布のクロスエントロピー損失である．

4. 評価実験

本節では，提案手法の有効性を検証するために様々な下流タスクにおける評価実験を行う．データセットとして，ImageNet-1k や CIFAR-10/100 などの多クラス分類用のデータセットを用いる．Student は ViT-Ti，Teacher は CL として DINO [4] と MIM として MAE [5] を用いる．

4.1 Teacher の学習方法の有効性

表 1 に SSL モデルの下流タスクへの適用方法を比較した結果を示す．その結果，提案手法に導入する Partial-1 は，最終層を再学習するのみでファインチューニングに匹敵する性能を達成している．

表 1: ImageNet-1k を用いた下流タスクへの適用 (%)

Method	Trainable Params (M)	CL	MIM
ファインチューニング	86.57	**81.92**	**81.70**
線形評価	0.77	76.31	54.02
Partial-1	7.86	80.26	79.30

4.2 従来手法との比較

ImageNet-1k を用いて評価した結果を表 2 に示す．Baseline は KD を行わずに Student が単体で学習したモデル，KD はファインチューニングした Teacher を用いて蒸留したモデル，Ours は複数の SSL モデルを用いて KD を行ったモデルである．Ours は Baseline から 2.2 pt，KD から 1.3 pt 向上した．

表 2: ImageNet-1k による評価結果 (%)

Method	Teacher	Top-1 Accuracy
Baseline (DeiT [2])	–	72.2
KD	DINO	73.1
	MAE	73.1
Ours	DINO, MAE	**74.4**

様々な下流タスクで評価した結果を図 3 に示す．Ours は，DTD と VOC2007 を除いたデータセットで KD よりも高い精度を達成した．

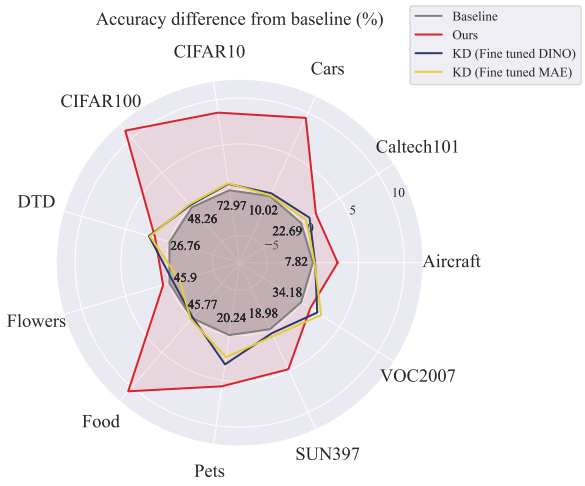


図 3: 様々な下流タスクにおける評価結果

5. おわりに

本研究では，小規模なモデルの精度向上を目的として，複数の自己教師あり学習モデルを用いた知識蒸留を提案した．評価実験により，様々な下流タスクにおいて提案手法が小規模なモデルの精度向上に有効であることを示した．今後は，さらなる精度向上を目指して，蒸留方法の多様化を行う．

参考文献

[1] G. Hinton, *et al.*, “Distilling the Knowledge in a Neural Network”, NIPS Workshop, 2015.

[2] H. Touvron, *et al.*, “Training data-efficient image transformers & distillation through attention”, ICML, 2021.

[3] E. Voita, *et al.*, “Analyzing Multi-Head Self-Attention: Specialized Heads Do the Heavy Lifting, the Rest Can Be Pruned”, ACL, 2019.

[4] M. Caron, *et al.*, “Emerging Properties in Self-Supervised Vision Transformer”, ICCV, 2021.

[5] K. He, *et al.*, “Masked Autoencoders Are Scalable Vision Learners”, CVPR, 2022.

研究業績

[1] 鈴木渉起 等, “Vision Transformer の相互学習による高精度化”, 画像の認識・理解シンポジウム, 2022.