

CNNベースの強化学習エージェントの行動精度がよくない（＝学習がうまく進まない、性能が出ない）場合、主な原因は以下のようなものが考えられます。

1. 入力画像の前処理が不適切

- 画像サイズが小さすぎたり、大きすぎたり、ノイズが多い。
- グレースケール化や正規化（0～1スケーリング）がされていない。
- 画像のレイアウトや情報がCNNで抽出しにくい形になっている。

2. CNNの構造がタスクに合っていない

- 層が浅すぎる／深すぎる。
- チャンネル数やカーネルサイズが不適切。
- 入力に対して出力次元が合っていない。

3. ハイパーパラメータの問題

- 学習率（lr）が大きすぎる／小さすぎる。
- バッチサイズが小さすぎる／大きすぎる。
- エージェントの探索率（ ϵ -greedyの ϵ ）が低すぎてすぐに局所解に陥る。

4. 強化学習アルゴリズムの設計ミス

- 報酬設計が不適切（Sparse rewardやnegative rewardばかりなど）。
- エピソード長が短すぎて十分な経験が得られていない。
- 行動空間や状態空間の設計が複雑すぎる、または単純すぎる。

5. 学習データの多様性不足

- エピソード数やステップ数が少ない。
- 探索が十分でなく、特定のパターンばかり学習している。

6. 実装上のバグ

- 損失計算やbackwardのミス。
- 勾配クリッピングやoptimizerの設定ミス。
- replay bufferや経験蓄積のバグ。

7. 環境自体の問題

- 環境の状態遷移や報酬が正しく設計されていない。
 - renderやresetのタイミングが間違っている。
-

よくあるチェックポイント

- 入力画像を実際に表示して、人間が見て状態が分かるか確認する
- CNNの出力（特徴量）や中間層を可視化する

- 学習曲線（報酬、損失）をプロットして、学習が進んでいるか確認する
- 探索率（ ϵ ）や学習率（lr）を変えてみる
- 報酬設計を見直す
- 学習ステップ数やエピソード数を増やす

まとめ

CNNベースの強化学習エージェントの精度が上がらない場合、

「入力前処理」「CNN構造」「ハイパーパラメータ」「報酬設計」「探索」「実装バグ」

のいずれかに問題があることが多いです。

まずは入力画像の可視化・学習曲線の確認・ハイパーパラメータの見直しから始めるのがおすすめです。

Conv2d

```
torch.nn.Conv2d(in_channels, out_channels, kernel_size, stride=1, padding=0,
dilation=1, groups=1, bias=True, padding_mode='zeros', device=None, dtype=None)
```

In the simplest case, the output value of the layer with input size (N, C_{in}, H, W) と出力 $(N, C_{\text{out}}, H_{\text{out}}, W_{\text{out}})$ は簡単に以下のように示すことができる。

$$\text{out}(N_i, C_{\text{out}_j}) = \text{bias}(C_{\text{out}_j}) + \sum_{k=0}^{C_{\text{in}}-1} \text{weight}(C_{\text{out}_j}, k) \star \text{input}(N_i, k)$$

```
# With square kernels and equal stride
m = nn.Conv2d(16, 33, 3, stride=2)
# non-square kernels and unequal stride and with padding
m = nn.Conv2d(16, 33, (3, 5), stride=(2, 1), padding=(4, 2))
# non-square kernels and unequal stride and with padding and dilation
m = nn.Conv2d(16, 33, (3, 5), stride=(2, 1), padding=(4, 2), dilation=(3, 1))
input = torch.randn(20, 16, 50, 100)
output = m(input)
```

別取り組みの結果

CartPole-v1について確認すると。

<https://github.com/lmarza/CartPole-CNN>