

色差が DQN に及ぼす影響の分析

Analysis of the influence of color difference on DQN

張重陽

指導教員 亀田弘之・渡邊紀文・相田紗織

東京工科大学大学院 バイオ・情報メディア研究科 コンピュータサイエンス専攻

Abstract

本研究では、ヒトの目の認識機構を参考に「色差実験環境」を構築するとともに、知能エージェントを「色差実験環境」において深部強化学習手法により学習させた。その学習結果を分析し、深部強化学習構築の知能エージェントにおいては、ヒトの目と同じ認識処理方法で画像を処理したにも関わらず、人間の目のような色収差に敏感ではなかった。

キーワード: 人工知能、強化学習、色差、CNN

【背景・目的】

近年、機械学習の研究・応用が盛んに行われており、主として「画像認識」に関わる分野で大きな成功を収めている。一方、Q 学習の分野でも DQN (Deep Q-network) のように深層学習を適用する手法が提案されている。

本研究では、色テスト環境を例として色差が DQN の訓練に対して影響を与えるのかについて確認することを目指している。具体的には、二つの色テスト環境で行われた DQN の訓練で本研究を論証する。機械学習において、畳み込みニューラルネットワーク (Convolutional Neural Network) は、順伝播型人工ディープニューラルネットワークの一種であり、画像や動画認識に広く使われているモデルである。強化学習 (Reinforcement Learning) とは、ある環境内におけるエージェントが、現在の状態を観測し、取るべき行動を決定する問題を扱う機械学習の一種である。DQN (Deep Q-network) [1] とは、2015 年にディープマインド社が開発した Q 学習と畳み込みニューラルネットワークを利用した深層強化学習アルゴリズムである。本研究では色差を DQN で識別するエージェントを「知能 Agent」と呼ぶ。

次に本研究で DQN を利用して学習する色差分析について説明する。ヒトには青、赤、緑の 3 種類の「錐体細胞」が存在し、加法混合と減法混合の組み合わせで写真や自然などの色を認識する。色

差があるとき、例えば、白い紙に色差が大きい黒や赤などの正方形がある場合は、白い紙に注意を向けずに正方形に注意を向けることができる。

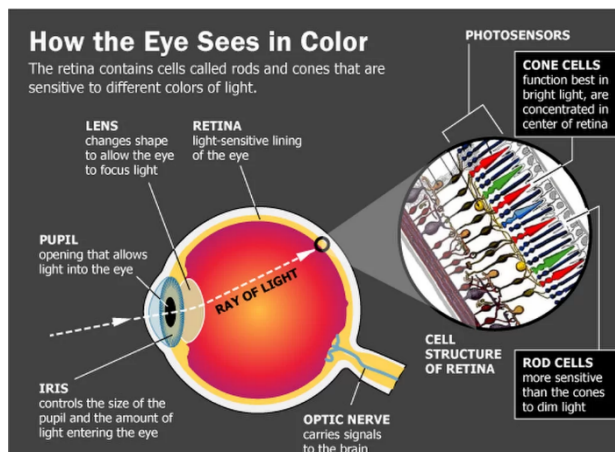


図 1. 目の色の認識

色見本	R	G	B	色の名前
■■■■	0	0	0	黒
■■■■	255	255	255	白
■■■■	224	224	224	薄いグレー
■■■■	128	128	128	グレー
■■■■	64	64	64	濃いグレー
■■■■	255	0	0	赤
■■■■	255	96	208	ピンク
■■■■	160	32	255	紫
■■■■	80	208	255	水色
■■■■	0	32	255	青
■■■■	96	255	128	黄緑
■■■■	0	192	0	緑
■■■■	255	224	32	黄色
■■■■	255	160	16	オレンジ
■■■■	160	128	96	茶色
■■■■	255	208	160	薄紅色

図 2. RGB カラー例

コンピューターは色の認識で人間と同じ 3 つの色 (青・赤・緑) を使っている。そこで、コンピューターに色を認識させるために色差が重要なものであるのかを探求した。特に深層強化学習の訓練

方法と結果は人間の色認識の仕組みに近いと思われるので、本研究ではDQNを採用した。

【方法】

研究環境：

Ubuntu16.04 Anaconda5.2 (Python3.6)

Pygame1.9 Pytorch0.4

色テスト環境は4色と距離の異なる色が含まれる8色の二種類である(図3)。色正方形をターゲットとし、知能Agentは真ん中の黒い正方形からスタートして、各色の報酬は1に設定している。

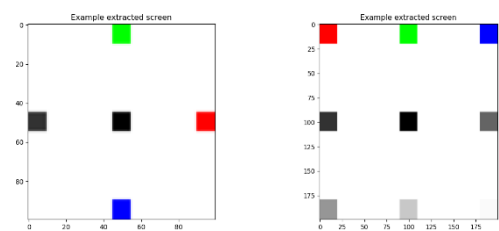


図3. 色テスト環境

畳み込みニューラルネットワークは1層、3層、6層の各層で作成し、ネットワークアーキテクチャはmobile Netを使用した(図4)。各テスト環境で800万~2000万回の訓練を行った。

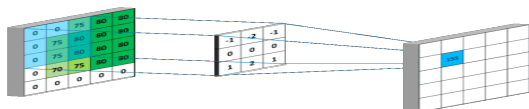


図4. 1層畳み込みニューラルネットワーク

【結果】

横軸は訓練回数であり、縦軸は10000回の訓練あたりの対応色と衝突の回数を示す。

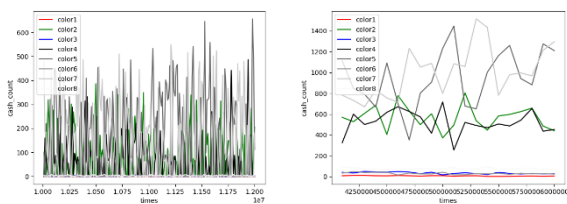


図5. 8色の環境 ランダム動作あり

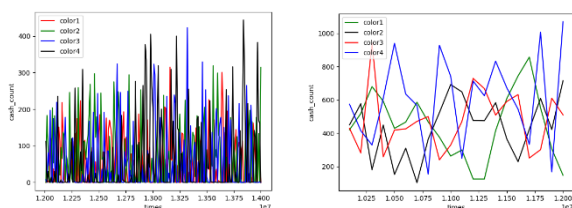


図6. 4色の環境 ランダム動作あり

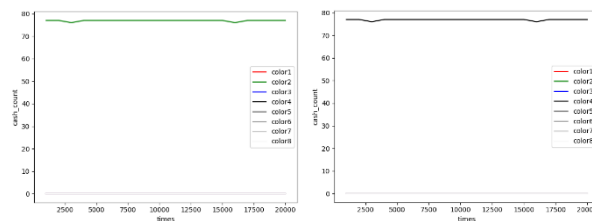


図7. 8色の環境 ランダム動作なし

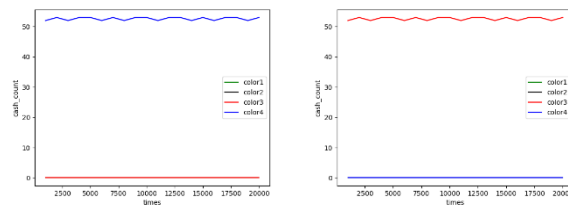


図8. 4色の環境 ランダム動作なし

各色の報酬が1になる場合。

8色環境のテスト：ランダム動作がある場合は、知能エージェントは色差を注意していない、距離が近い4色を選択することが多い。ランダム動作がない場合は、知能エージェントは距離が近い4色のみを選択し、訓練回数により、選択した色が異なるということが判明した。

4色環境のテスト：ランダム動作がある場合は、4色をランダムで選択する。ランダム動作がない場合は、訓練回数により、選択した色が異なった。

これらの結果より、エージェントとターゲットの距離が異なる時は、ランダムで近いターゲットを選択し、距離が同じである時は、色に関係なくランダムで選択することが確認された。

【結論】

DQNはターゲットの色に影響を受けず、ターゲットの距離に影響を受けている。つまり、色差はDQNの訓練に影響を及ぼさないことが判明した。

本研究は他のDQNなど深層強化学習研究の参考となることを願っている。すべてのモデルとImageはGitHub[2]にアップロードした。

【参考文献とGitHub】

- [1] David Silver, 他
Playing Atari with Deep Reinforcement Learning <https://arxiv.org/abs/1312.5602>.
- [2] <https://github.com/zcyln/The-Consortium-of-Universities-in-Hachioji>.