

## 2048 における盤面の対称性を考慮した CNN プレイヤの学習

1225119 近藤 直季 【高度プログラミング研究室】

## Training CNN Players Considering Board Symmetry in Game 2048

1225119 Naoki Kondo 【High-Level Programming Lab.】

## 1 はじめに

「2048」は G.Cirulli によって 2014 年に公開された一人用のパズルゲームである [1]。2048 におけるコンピュータプレイヤは N-tuple ベースのものやニューラルネットワーク (NN) ベースのものがある。現在において N-tuple ベースのプレイヤの研究はいくつも報告されており、人間を凌駕する域に到達している。一方で、NN ベースのプレイヤの研究はほとんど報告されておらず、N-tuple ベースに劣るものである。

2048 における CNN プレイヤに関する研究論文は Guei らによるもの [2] が最初である。結果は優れなかったものの、2048 のゲーム盤面を CNN の入力としてどのように変換して与えるかのアイデアは本研究において欠かせないものである。

本研究では、CNN の畳込み層数に着目した実験と 2048 における盤面の対称性を考慮した CNN の実験を行う。CNN の学習には N-tuple ベースプレイヤのプレイログ (6 億盤面) を教師データとして教師あり学習を行う。

## 2 本研究における CNN の構成

ここでは、本研究における CNN の構成について述べる。入力層の入力には  $4 \times 4 \times 16$  の形で与える。これは  $4 \times 4$  のゲーム盤面に対して「空白, 2, 4, 8, ..., 32768」のタイル (計 16 種類) の有無による 0 か 1 かの 2 値を持たせている。中間層は  $k_1$  層の畳込み層からなる。畳込みには conv2d を使用し、ストライド 1, フィルタサイズは  $2 \times 2$  で各層統一で  $k_2$  枚のフィルタを持つ。畳込む際には、右側と下側にゼロ埋めを行い、 $4 \times 4$  のサイズを維持させる。このため、畳込み層における入力と出力は 1 層目を除き、 $4 \times 4 \times k_2$  となる。また、各畳込み層の出力に対して ReLU 関数を適用する。出力層は 1 層の全結合層からなる。入力は  $4 \times 4 \times k_2$  で出力は 4 (上, 右, 下, 左) である。出力に対して softmax 関数を適用し、各方向の確率 (入力盤面に対してどの方向を選ぶべきか) に変換する。また、各層における出力に対してバイアス項を活性化関数の適用前に加えている。最適化アルゴリズムは Adam を使用する。本研究では  $k_1 =$  畳込み層数と  $k_2 =$  各畳込み層のフィルタ数 (ch 数) を変化させることで性能向上を目指した。

表 1 5 層におけるプレイ方法による結果

プレイ方法	平均得点	最高得点	クリア率
1 盤面プレイ	24,105	182,296	35.8%
8 盤面プレイ	86,203	386,972	86.5%

表 2 畳込み層の層数差による結果一覧

畳込み層	ch 数	平均得点	最高得点	クリア率
2 層	436	25,669	175,628	45.6%
3 層	312	69,840	332,868	79.4%
4 層	256	80,284	343,496	83.3%
5 層	222	86,203	386,972	86.5%
6 層	200	83,791	387,376	83.5%
7 層	182	79,812	401,912	83.1%
8 層	168	74,787	363,916	81.1%
9 層	158	68,129	358,736	75.9%

3 章は上記の構成であるが、4 章では、入力に対称な 8 盤面を 1 盤面ずつ計 8 回与え、全結合層の入力で  $4 \times 4 \times k_2 \times 8$  とすることで 8 回の入力に対して 1 つの出力を得られる構成に改変したものとなる。

## 3 畳込み層の層数に着目した実験

総パラメータ数が約 82 万となる条件のもとで、畳込み層が 2 層から 9 層からなる 7 種類の CNN を構成し、それぞれ学習を行った。学習後に 1,000 ゲームをプレイさせ、平均得点、最高得点、クリア率 (2048 到達率) を比較する [3]。

表 1 は同じ学習済み CNN に対して異なるプレイ方法による性能差を示している。学習した CNN プレイヤに 2048 をプレイさせる際に、現盤面を与えるだけ (以降、1 盤面プレイとする) よりも、現盤面を回転・反転によって得られる対称な 8 盤面を与える (以降、8 盤面プレイとする) 方が結果が優れた。このため表 2 はすべて 8 盤面プレイによる結果である。

平均得点が最も優れたのは 5 層であった。最高得点は 7 層が最も高く、クリア率は 5 層が最も高かった。畳込み層が 2 層は他に比べ大きく劣り、3 層から 5 層は層数が増加するごとに性能が向上したが、それ以降は性能が

悪くなっていることがわかった。2層と3層の差に関しては、3層の方が明確な特徴抽出が行われていたという Matsuzaki ら [4] の報告がある。

同じ構成の CNN であるのに回転・反転によって得られる8盤面を用いた方が優れるということは、CNN プレイヤにとって重要な情報であると予測できる。また、8盤面プレイでは、CNN への入力を8回行うので8つの出力が得られるため「上3票、右3票、下0票、左2票」というような際には出力の合計が高い方を優先とした。

#### 4 学習時に対称性を考慮した CNN プレイヤ

前章の表1にて学習した CNN に対して1盤面プレイでは結果が優れないが8盤面プレイであれば結果が優れることが示された。そこで、学習後に8回評価する工程を CNN の構成を変更することで、学習時に組み込むことで CNN の内部に実現させ、1盤面プレイでも良い性能を出せるようにする。

図1は本章における CNN の構成である。合計  $k$  層の畳込み層は縦に8つ並んでいるが、8つの入力は1つのパラメータを共有している部分ネットワーク(破線枠)を通過する。よって、対称な8盤面を1回ずつ並行で同じ畳込み層を経由し、各出力を全結合層で繋ぎ合わせる形式となっている。

このため、前章では8回の入力に対して、各方向の入力すべき確率を8組を得たが、本章では8回の入力に対して各方向の入力すべき確率を1組を得る。つまり、本章における学習した CNN プレイヤに2048を1盤面プレイでプレイさせた場合、前章における8盤面プレイを行った結果と同等になることを期待する。

本稿では畳込み層5層の ch 数を調整した結果についてのみ示す。畳込み層5層に対して ch 数を100から500まで100ずつ増やして調査した。

表3, 4は学習後に1,000ゲームを1盤面プレイと8盤面プレイを行った結果である。ch 数が500の学習はうまく進まず、ランダムプレイと同等になっていること

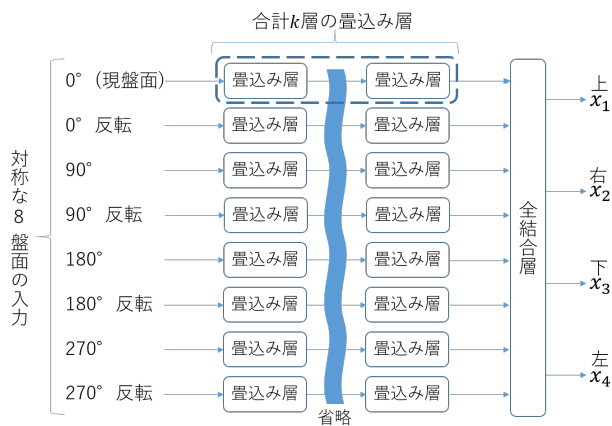


図1 対称性を考慮した CNN の構成概略図

表3 畳込み5層で各 ch 数における1盤面プレイ結果

畳込み層	ch 数	平均得点	最高得点	クリア率
5層	100	43,552	331,748	54.7%
5層	200	80,355	333,848	72.7%
5層	300	85,548	386,744	72.2%
5層	400	119,637	533,624	80.7%

表4 畳込み5層で各 ch 数における8盤面プレイ結果

畳込み層	ch 数	平均得点	最高得点	クリア率
5層	100	45,710	330,756	55.5%
5層	200	86,062	366,504	71.9%
5層	300	93,964	405,600	80.5%
5層	400	109,493	509,284	77.3%

を確認した。同じ畳込み層数においては上限があるが ch 数が多いほど平均得点が高くなることがわかった。また、1盤面プレイと8盤面プレイの平均得点の差はどれも1.1万点未満であった。ch 数400においては1盤面プレイの方が優れた結果となった。

#### 5 まとめ

クリア率は3章の CNN 構成の方が優れたが、平均得点と最高得点は4章の CNN 構成の方が優れた。また、4章の CNN 構成における1盤面プレイと8盤面プレイの結果は同等であった。同等の結果が出せるということは、CNN 内で等価な8盤面の情報をうまく取り込めたと考える。1盤面プレイであれば1試合にかかる時間が短いため、強化学習を行う際には効率良く行える。また、最高得点が50万点越えを達成することができ、その際には32,768のタイルに到達することができていた。

#### 参考文献

- [1] Cirulli, G.: 2048, <http://gabrielecirulli.github.io/2048/> (2014).
- [2] Guei, H., Wei, T., Huang, J.-B. and Wu, I.-C.: An Early Attempt at Applying Deep Reinforcement Learning to the Game 2048.
- [3] Kondo, N. and Matsuzaki, K.: Playing Game 2048 with Deep Convolutional Neural Networks Trained by Supervised Learning, *Journal of Information Processing*, Vol. 27, pp.340–347 (2019).
- [4] Matsuzaki, K. and Teramura, M.: Interpreting Neural-Network Players for Game 2048, *Proc. 2018 Conference on Technologies and Applications of Artificial Intelligence (TAAI 2018)*, pp.136–141 (2018).