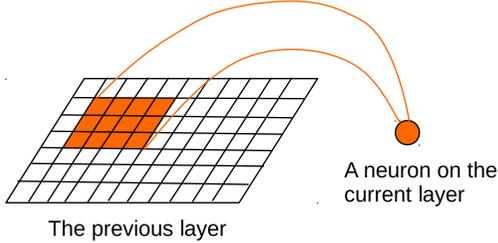


<p>正確な訓練データを使ったニューラルネットワークによる ゲーム局面の学習</p> <p><b>Learning Game Positions by Neural Networks with Accurate Training Data</b></p> <p>神保 秀司 (岡山大学, jimbo-s@okayama-u.ac.jp) Shuji Jimbo (Okayama University)</p> <p>2018年3月7日 大阪府立大学 中百舌鳥キャンパス 第13回 組合せゲーム・パズル研究集会</p>	<p>序論 (1)</p> <ul style="list-style-type: none"> <li>■ 畳み込みニューラルネットワーク (CNN) は、パターン認識と画像認識に広く使われている</li> </ul> <p>畳み込み層中のニューロンは、直前の層の限られた小さい領域中のニューロンとだけ接続している</p>  <p>The previous layer</p> <p>A neuron on the current layer</p> <ul style="list-style-type: none"> <li>■ CNN は、ゲーム局面の評価においても非常に有効に使われている</li> </ul> <p>例 AlphaGo, Google DeepMind 社により開発された現時点で最強の囲碁 AI</p>
<p>正確な訓練データを使ったニューラルネットワークによる</p> <p>2018年3月7日 - 1 / 20</p>	<p>正確な訓練データを使ったニューラルネットワークによる</p> <p>2018年3月7日 - 2 / 20</p>
<p>序論 (2)</p> <ul style="list-style-type: none"> <li>■ 研究の目的: ニューラルネットワーク, 特に CNN による盤ゲームの局面の形勢判断</li> </ul> <p>対象とする盤ゲーム: <b>ペンタゴ (Pentago)</b>, Tomas Flodén により発明された 2 人零和完全情報思考ゲーム</p> <p>6 × 6 の盤で行う五目並べを拡張した形のゲーム</p> <p>手の選択肢が多数存在することが特徴</p> <ul style="list-style-type: none"> <li>■ ペンタゴは, 2014 年に Geoffrey Irving により, NERSC 所有のスーパーコンピュータ Cray Edison を使って<b>完全解析</b>が達成されている</li> </ul> <p>例えば, ペンタゴは先手必勝である。</p> <p>後退解析により得られた完全解析結果の一部がパブリックドメインで公開されている</p> <p>そのデータのサイズは, 約 4TB である</p>	<p>序論 (3)</p> <ul style="list-style-type: none"> <li>■ ペンタゴは, MindTwister 社から商品として販売されている</li> </ul> <p>下図は, 商品としてのペンタゴの例</p>  
<p>正確な訓練データを使ったニューラルネットワークによる</p> <p>2018年3月7日 - 3 / 20</p>	<p>正確な訓練データを使ったニューラルネットワークによる</p> <p>2018年3月7日 - 4 / 20</p>

## 序論 (4)

- **CIFAR-10 チュートリアル CNN** CIFAR-10 画像分類問題のための CNN モデルが TensorFlow チュートリアルのデモとして公開されている  
CIFAR-10 分類は  $32 \times 32$  ピクセルの RGB 画像を 10 種類のカテゴリに分類するよく知られたベンチマーク問題である  
ペンタゴの局面の形勢判断は、 $6 \times 6$  ピクセルの 3 色画像を 3 つのカテゴリ (黒勝ち, 引き分け, 白勝ち) に分類する問題と見なすことができる
- 我々は, CIFAR-10 チュートリアル CNN を修正した CNN によるペンタゴの局面の形勢判断を試み, 次の結果を得た  
スライス 12 の局面に対する正答率は約 85 パーセントであり, スライス 11 の局面に対する正答率は約 75 パーセントだった  
ペンタゴにおける**スライス  $n$  の局面**とは, 盤上に両対局者の石が合計  $n$  個置かれた局面のことである

正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 5 / 20

## 序論 (5)

- ニューラルネットワークを関数近似の道具と見なしたとき, 与えられたペンタゴの局面から両対局者が最善を尽したときの勝敗結果の判定は, **中程度に困難な問題**と考えられる
- ニューラルネットワークにとって**暗号解読**は最も近似困難な問題の一つであり, 画像内の特定の形状の領域を認識する問題 (後述) は最も容易な問題の一つである
- **本研究の目的**  
(a) ニューラルネットワークによるゲーム局面の形勢判断能力を究明し, より性能の高いニューラルネットワークを開発する  
(b) ニューラルネットワークによる形勢判断が困難な局面が人間にとっても難解な局面であると予想し, **娯楽性の高い AI ペンタゴプレイヤー**を開発する

正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 6 / 20

## 戦略ボードゲーム, ペンタゴ (1)

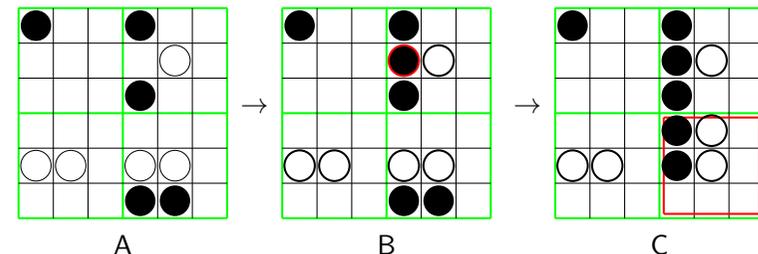
- **盤** ペンタゴの盤は  $6 \times 6$  のマス目を持ち, **90 度単位で自由に回転できる** 4 つの  $3 \times 3$  の小盤に分割されている
- **対局方法** 白と黒の色が割り当てられた 2 人の対局者が, 空の盤から始めて, 「自分の色の石を盤上に置き, その後任意の小盤 1 つを 90 度左右任意の向きに回転させる」という操作 (手) を交互に繰り返す  
どのような状況においてもパスは禁止
- **勝敗判定** 同色の石が 5 個縦, 横, あるいは対角線の直線上に続けて並んだら (5 連が出来たら) その色の勝ち  
ただし, 同時に両方の色の 5 連が出来たら引き分け  
全く 5 連が出来ずに次の手が打てなくなれば引き分け  
5 連が出来たことの判定は, **石を置いた直後と小盤を回転した直後の両方で行う**

正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 7 / 20

## 戦略ボードゲーム, ペンタゴ (2)

- 例: 盤 A は黒の手番である  
黒石を置いて盤 B にし,  
次に小盤を回転させて盤 C にすれば  
黒の勝ちである



正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 8 / 20

## 戦略ボードゲーム, ペンタゴ (3)

- **命題** ペンタゴで両対局者が最善を尽したら, 後手, すなわち白は勝つことができない  
**(証明の概略)** どちらの対局者にとっても現在の盤面に自分の色の石を置けばより有利になるので, 黒は白が取ろうとした戦略を自分の色の石が1つ多い状態で取ることができる
- ペンタゴの総局面数は,  $6 \times 6$  の五目並べと同程度  
 ペンタゴの特定の局面から生じ得る総局面数は,  $6 \times 6$  の五目並べの場合の高々128倍以下

## 完全解析結果のデータ

- データ中のスライス  $n$  の局面数 (重複を含む)  
 盤上の石数を**スライス**と呼ぶ

0	256	7	144645120	14	2180723700736
1	768	8	832507904	15	5720741273600
2	9216	9	3550828544	16	14930328007168
3	73216	10	15994707968	17	31389244375296
4	720896	11	60293679104	18	65007135675648
5	4037632	12	231269590016		
6	27040768	13	708482302976		

- $C$  で黒又は白の対局者を表し,  $\bar{C}$  で  $C$  の相手の対局者を表す  
 対局者  $C$  の手番のときの  $C$  の勝率は, 同程度のスライスの  $\bar{C}$  の手番のときの勝率よりも大きい  
 黒の手番のときの黒の勝率は極めて高い

	Slice 11	Slice 12
Draw	15116993859	20753197343
Black win	14013289093	199909936103
White win	31163396152	10606456570

正確な訓練データを使ったニューラルネットワークによる

2018年3月7日 - 9 / 20

正確な訓練データを使ったニューラルネットワークによる

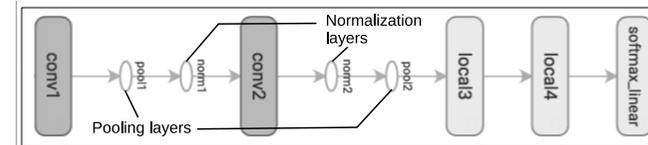
2018年3月7日 - 10 / 20

## 訓練と試験に使ったデータセット

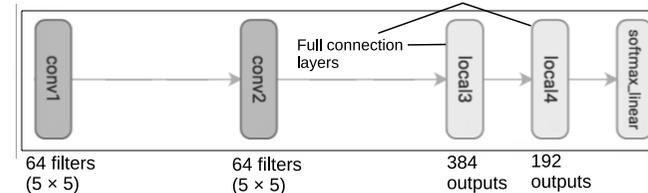
- 使用した計算機環境による制限のために**スライス 11 及びスライス 12 の局面だけ**を対象にした
- 局面は一様分布に従ってランダムに選んだ
- データ中の黒勝ち, 引き分け, 白勝ちの比率がほぼ等しくなるように一様分布に従ってサンプルを削除した
- ミニバッチのサイズは 128  
 訓練サンプル数は, スライス 11 と 12 のどちらについても 850,000 バッチ (108,800,000 サンプル) とした  
 試験サンプル数は, 1,000,000 とした

## ニューラルネットワークの構造 (1)

- CIFAR-10 に対する TensorFlow チュートリアル CNN の基本構造



- 今回報告の実験で使った CNN の基本構造



- 今回報告の実験で使った CNN では, 盤上の石の絶対位置に敏感になるようにプーリング層と正規化層を除去している

正確な訓練データを使ったニューラルネットワークによる

2018年3月7日 - 11 / 20

正確な訓練データを使ったニューラルネットワークによる

2018年3月7日 - 12 / 20

## ニューラルネットワークの構造 (2)

### ■ 構造の変更

- (a) 基本の CNN
- (b) 基本の CNN からフィルタ数を 64 から 128 に増加
- (c) 基本の CNN から畳み込み層の数を 2 層から 4 層に増加
- (d) (b) + (c)
- (e) 基本の CNN からフィルタ数を 64 から 32 に減少
- (f) 基本の CNN から畳み込み層の数を 2 層から 1 層に減少
- (g) 畳み込み層をすべて除去

- CNN のニューロン同士の結合の重みの初期値は、一様分布に従ってランダムに小さい値を選んだ

正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 13 / 20

## CNN のペンタゴの局面の形勢判断をする能力

- $X_i$  でスライス  $i$  のときの正答率を  
 $Y_i$  でスライス  $i$  のときの最後の 1000 バッチの平均交差エントロピーを表す  
( $i \in \{11, 12\}$ )

スライス	$X_{11}$	$Y_{11}$	$X_{12}$	$Y_{12}$
(a)	0.810	0.477	0.842	0.389
(b)	0.814	0.459	0.846	0.398
(c)	0.809	0.442	0.838	0.393
(d)	0.808	0.463	0.842	0.390
(e)	0.779	0.592	0.818	0.499
(f)	0.779	0.577	0.813	0.518
(g)	0.579	0.932	0.665	0.828

- 基本の CNN のフィルタと畳み込み層の数を増やしても、殆ど正答率と交差エントロピー誤差に変化がない
- 規模を縮小した場合は、それに応じて正答率が減少し、交差エントロピー誤差が増加する

正確な訓練データを使ったニューラルネットワークによる

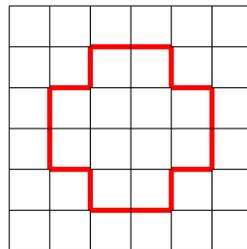
2018 年 3 月 7 日 - 14 / 20

## CNN の特定の形状を認識する能力

- ペンタゴの形勢判断の学習で使った訓練データと試験データのサンプルのラベルだけを特定の領域内の石数差を反映するように変更したデータを使って CNN の学習能力について実験した

正答率は 100 パーセントに達し、最後の 1000 バッチの交差エントロピーは 0.0 だった

- 変更後のラベルは、図の赤枠内の黒と白の石数に従って次のように決定した:  
黒石が白石よりも多ければ 1  
黒石が白石よりも少なければ 2  
黒石と白石が同数存在すれば 0



正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 15 / 20

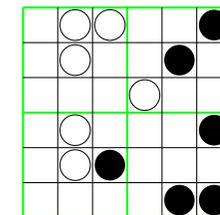
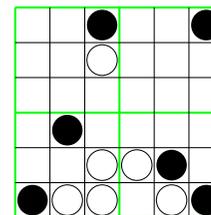
## CNN による形勢判断が困難な局面 (1)

- 3 つのカテゴリに分類する CNN の直接の出力は logit と呼ばれる 3 次元ベクトルであり、確率分布になっている  
最も大きい確率値のカテゴリが CNN の解答となる

- logit により、CNN が自信をもって解答したか否かが推定できる

左図は、logit の最大成分  $v_1$  に対する最大成分と第 2 成分  $v_2$  の差の比  $\alpha = (v_1 - v_2)/v_1$  が 1000 万サンプル中最小 (0.000001) の盤面 (白勝ち)

右図は、logit の 3 成分の標準偏差が 1000 万サンプル中最小 (0.000236) の盤面 (白勝ち)



正確な訓練データを使ったニューラルネットワークによる

2018 年 3 月 7 日 - 16 / 20

CNN による形勢判断が困難な局面 (2)	CNN による形勢判断が困難な局面 (3)
<ul style="list-style-type: none"> <li>■ 実験 1           <ul style="list-style-type: none"> <li>(a) 黒勝ち, 引き分け, 白勝ちの局面を 8000 万ずつ含むスライス 12 の局面のデータをランダムに生成し, 最後の 1,536,000 サンプルを試験データ A とし, 残りを訓練データ A とした 訓練データ A を使って基本の CNN により学習した結果 CNN A を得た</li> <li>(b) 訓練データ A に対する CNN A の logit で最大成分 <math>v_1</math> に対する最大成分と第 2 成分 <math>v_2</math> の差の比 <math>(v_1 - v_2)/v_1</math> が 0.6 以下のサンプルが 10,220,000 個有り, それらを訓練データ B とし, 試験データ A に対する CNN A の logit で同じ条件を満たすものが 313,741 個有り, それらを試験データ B とした</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>■ 実験 2           <ul style="list-style-type: none"> <li>(a) 訓練データ B を使って基本の CNN により学習した結果 CNN B を得た</li> <li>(b1) 試験データ A に対する CNN A の正答率は 0.845</li> <li>(b2) 試験データ B に対する CNN A の正答率は 0.575</li> <li>(b3) 試験データ A に対する CNN B の正答率は 0.846</li> <li>(b4) 試験データ B に対する CNN B の正答率は 0.590</li> </ul> </li> <li>■ 結果 同じ構造の CNN を使い, 元の訓練データのうち難解なサンプルを取り出して作った訓練データを使って学習させた CNN は劣化せず, 難解な試験データに対する正答率が改善した</li> </ul>
<p>正確な訓練データを使ったニューラルネットワークによる <span style="float: right;">2018 年 3 月 7 日 - 17 / 20</span></p>	<p>正確な訓練データを使ったニューラルネットワークによる <span style="float: right;">2018 年 3 月 7 日 - 18 / 20</span></p>
<p>予想と今後の課題</p>	
<ul style="list-style-type: none"> <li>■ 予想 <math>i \in \{11, 12\}</math> について, 単純なニューラルネットワークでは形勢判断が殆ど不可能なスライス <math>i</math> の局面が全体の 1 割以上存在する</li> <li>■ 参考 画像認識の分野で最近注目を集め, AlphaGo Zero でも採用されている深い ResNet (CNN の拡張版) を使ってスライス 12 の局面についてペンタゴの局面の形勢判断の学習をしても基本の CNN と同様の結果だったという報告がある (学生による実験, 要確認)</li> <li>■ 今後の課題           <ul style="list-style-type: none"> <li>(a1) スライス 12 のペンタゴの局面の形勢判断の正答率が 90 パーセント以上のニューラルネットワークの開発を試みる</li> <li>(a2) ニューラルネットワークによるスライス 18 のペンタゴの局面の形勢判断の学習について実験する</li> <li>(b) 両対局者が最善を尽したとき負けの局面でも人間にとって手強い手を打つ AI ペンタゴ対局者の開発を試みる</li> </ul> </li> </ul>	<p style="font-size: 2em; color: blue; margin: 0;">御静聴どうも有難うございます</p> <p style="font-size: 0.8em; margin-top: 20px;">本研究は JSPS 科研費 JP15K00018 の助成を受けたものです 本研究は主に九州大学情報基盤研究開発センターの研究用計算機システムを利用しました</p>
<p>正確な訓練データを使ったニューラルネットワークによる <span style="float: right;">2018 年 3 月 7 日 - 19 / 20</span></p>	<p>正確な訓練データを使ったニューラルネットワークによる <span style="float: right;">2018 年 3 月 7 日 - 20 / 20</span></p>