囲碁AIの仕組み

第1回

囲碁方策ネットワーク(教師付き学習)

講師 : Sakurai

囲碁 AIの仕組み 目次

- 1 アルファ碁の衝撃
- 2. 人間の脳・記憶とAI
- 3. ディープラーニング
- 4. 例題:手書き数字の認識
- 5. 囲碁方策ネットワーク(教師付き学習)

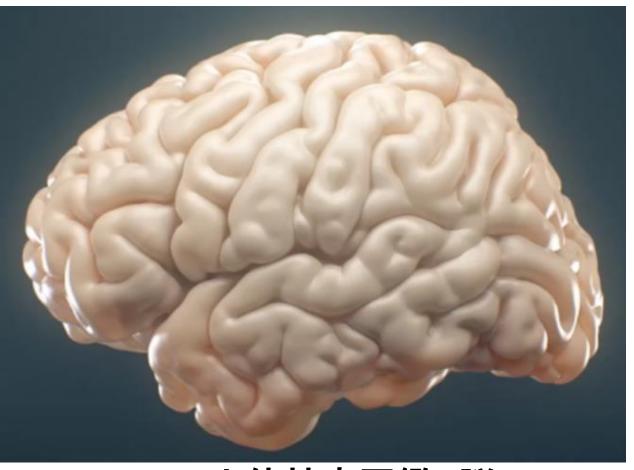
今後: 囲碁AIの強化学習(教師なし学習)

1. アルファ碁の衝撃

- 2015年, 16年 AlphaGo の論文発表(国際学会, 学術誌)
 Move Evaluation in Go Using Deep Convolutional Neural Networks (ICLR, 2015, 10月)
 Mastering the game of Go with deep neural networks and tree search (Nature, 2016, 1月)
- 2016年3月 AlphaGo がイ・セドル九段に4勝1負で勝ち越し
- 2017年1月 AlphaGo Master がネット碁でプロ棋士に60連勝
- 2017年5月 AlphaGo Master が世界トップの柯潔九段に3局全勝
- ・2017年10月 人間の棋譜を使わない AlphaGo Zero の論文発表 Mastering the game of Go without human knowledge (Nature, 2017, 10月)
- 2017年12月 将棋・チェスに対応した Alpha Zero の論文発表 Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm, arXiv (論文が公開されているウェブサイト): 1712.01815[cs.Al] (2017, 12月)

出典: 山岡忠夫「将棋 A I で学ぶディープラーニング」(ナイナビ出版, 2018年)

2. 人間の脳・記憶とAI



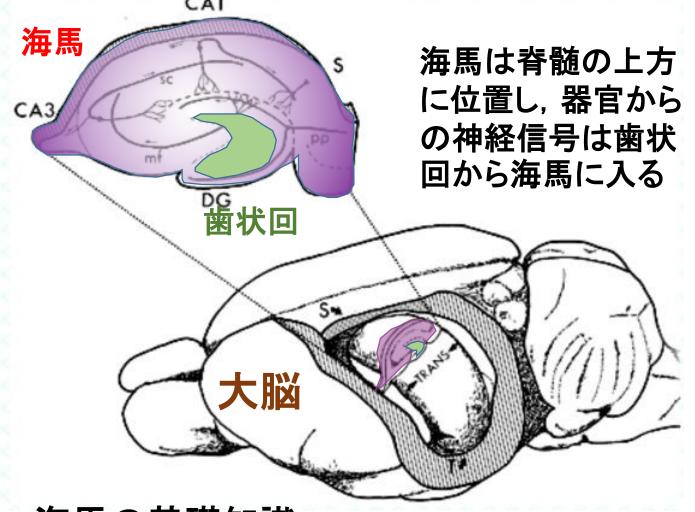
NHK 人体精密図鑑 脳



脳の中の神経線維(CG)

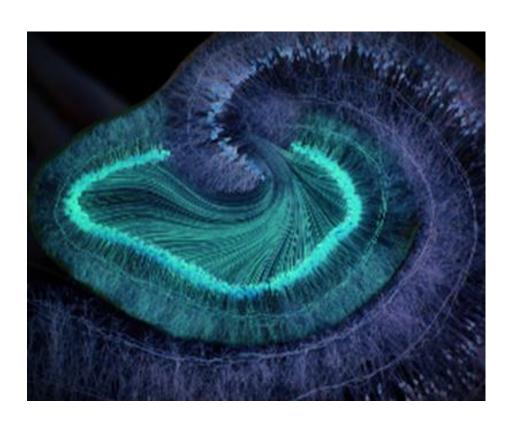
大脳皮質には神経細胞(ニューロン)とよばれる細胞がおよそ140億個も密集している。内側にある白質は、神経線維(神経軸索)が無数に走行している。 出典: 脳神経外科 馬場元毅 http://jbf.or.jp/iframe/71/

(1) 短期的記憶の場所: 海馬



海馬の基礎知識

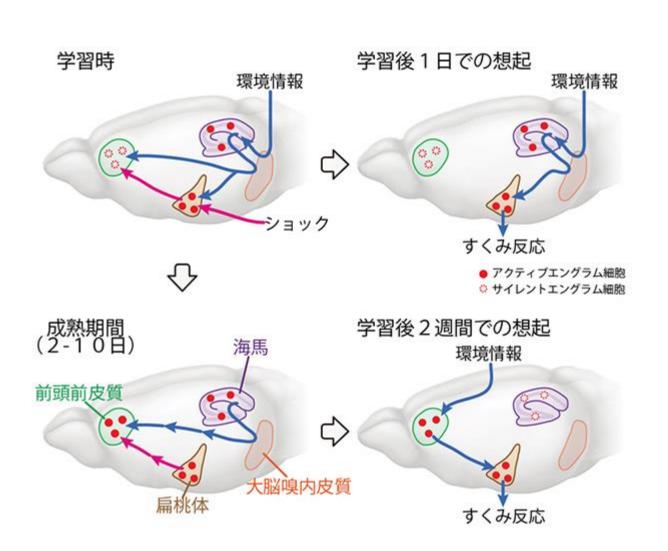
http://gaya.jp/research/hippocampus.htm



NHK海馬の中の「歯状回」(CG) 青く光っているのが歯状回の細胞、 紫色が海馬の細胞

神経細胞が生まれ続けている。

(2) 長期的記憶の場所: 大脳皮質



出典: 理化学研究所 -

MIT 神経回路遺伝学研究センター

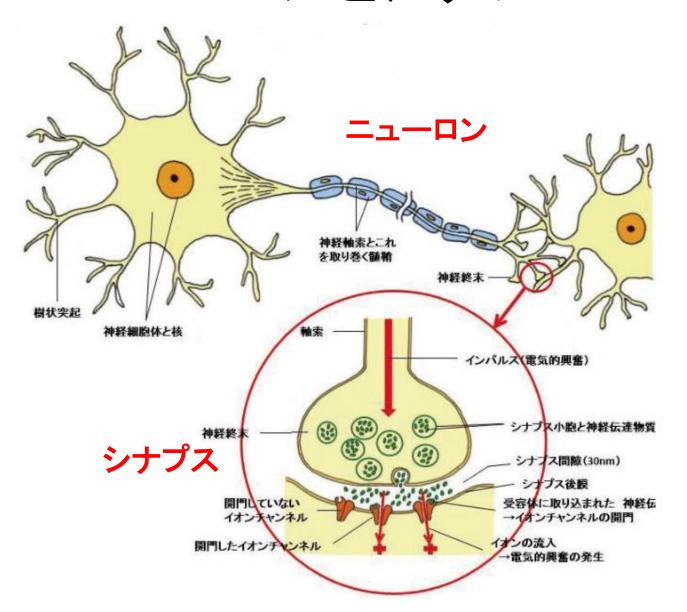
学習時に海馬においてエングラム 細胞は最初に形成される。さらに、 恐怖記憶に関わる扁桃体の細胞と ともに、前頭前皮質のエングラム細 胞を生成する。

学習後1日の記憶想起では、「海馬 →大脳嗅内皮質→扁桃体」の神経 回路が使われる。

成熟期間:長期的記憶の醸成

学習後2週間以降の記憶想起では、 「前頭前皮質(大脳皮質)→扁桃 体」の神経回路が使われる。

ニューロンとシナプス



出典: 脳神経外科 馬場 元毅

http://jbf.or.jp/iframe/71/

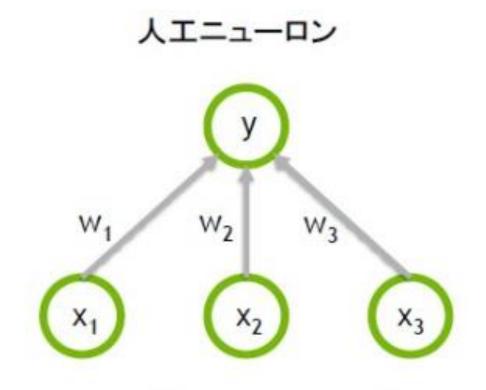
ニューロンは、樹状突起と軸索から成り立っている。樹状突起は複数出ており、これがさらに数百本に枝分かれしている。

シナプスでは、狭いすき間に 特殊な化学物質が放出され、 これが二次ニューロンの細胞 の表面にある受容体に働きか けて、そこに電気的興奮を引き 起こし、指令が伝えられる。

3. ディープラーニング

ニューラルネットワークとは、人間の脳内にある神経細胞(ニューロン)とそのつながりを人工ニューロンという数式的なモデルで表現したものである。それは、入力層、出力層、隠れ層から構成される。

人間の脳の中にあるニューロンは、 電気信号として情報伝達を行う。 その際、 シナプスの結合強度によって、情報の伝わりやすさが変わってくる。 人工ニューロンでは、結合強度を 重みw で表現する。



$$y=F(w_1x_1+w_2x_2+w_3x_3)$$

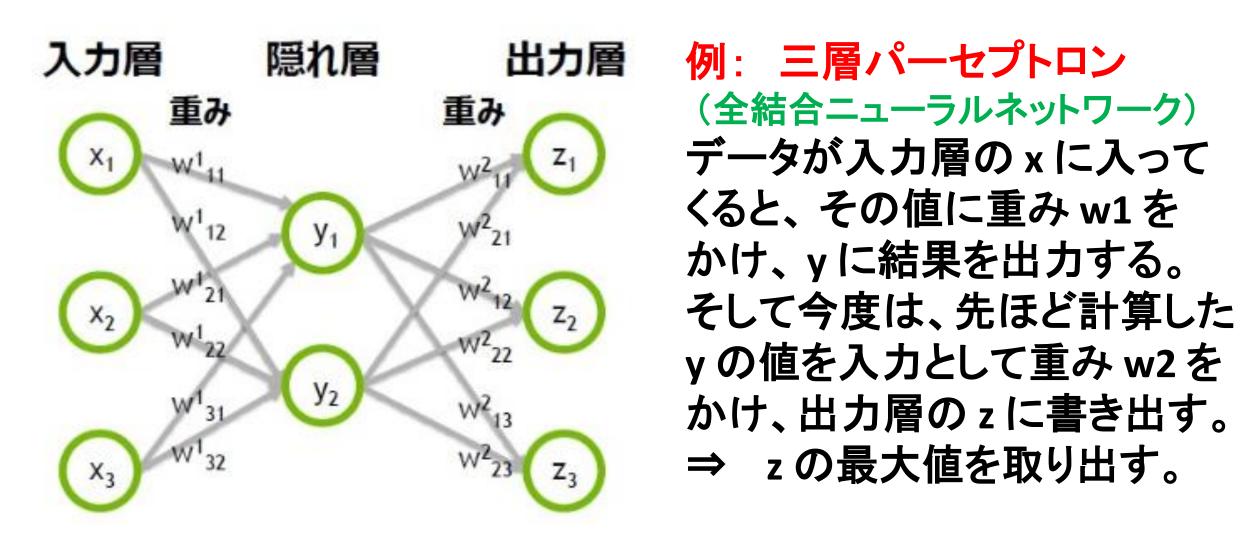
F(x)=max(0,x)

ディープラーニングとは、隠れ層が多数存在する

多層構造のニューラルネットワークのことである。

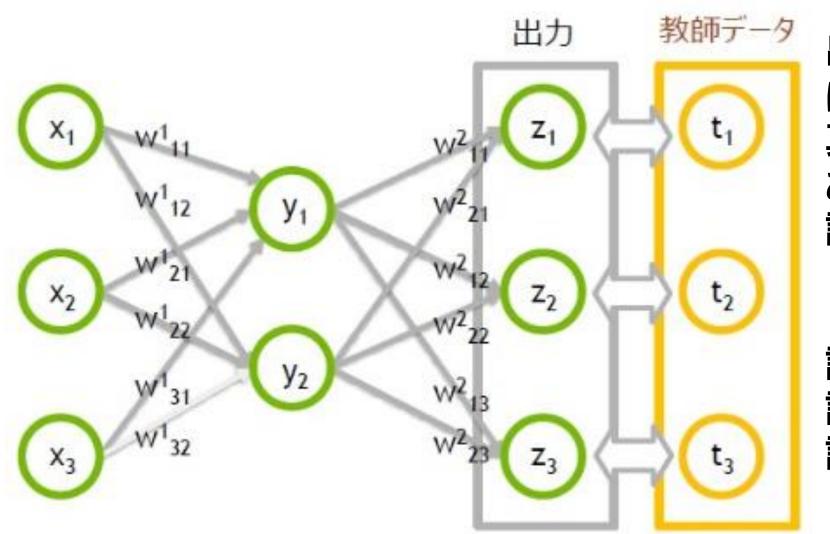
出典: ニューラルネットワークの基礎解説 https://www.sbbit.jp/article/cont1/33345

人エニューロンの推論方法(判断)



出典: ニューラルネットワークの基礎解説 https://www.sbbit.jp/article/cont1/33345

学習方法: 教師付き学習(Supervised Learning)

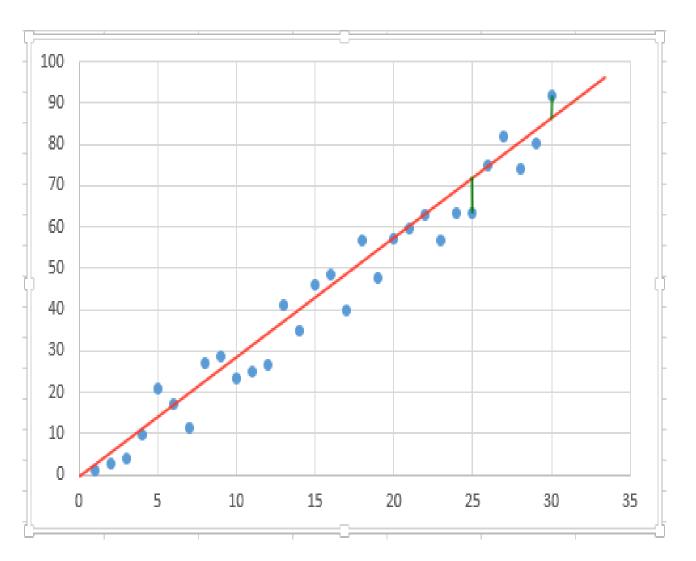


出力 z1 が 0.8 という出力になったとする。 しかし,真値(教師データ)は 0.5 であるとき,重み w1 と w2 を調整していく。 (誤差を最小化する。)

調整の方法: 誤差の評価(2乗和誤差) 誤差最小化(勾配降下法)

出典: ニューラルネットワークの基礎解説 https://www.sbbit.jp/article/cont1/33345

誤差の評価(2乗和誤差)



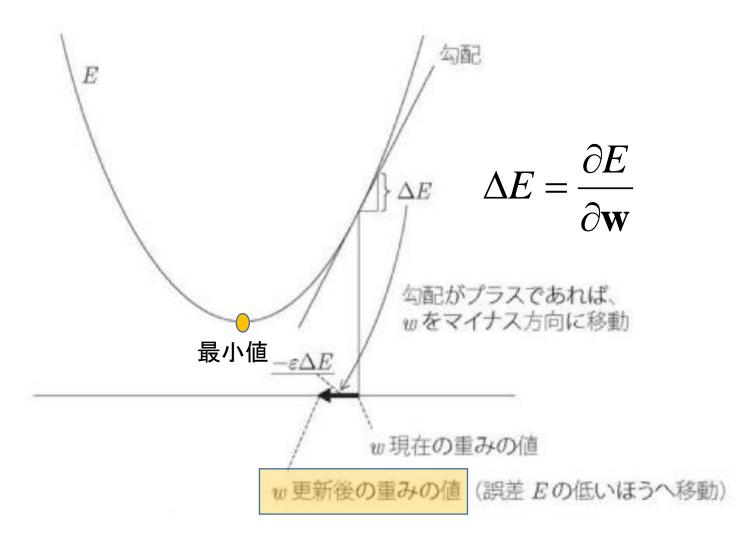
どれだけ学習できているかということを判断するための指標を 「誤差関数」(損失関数)という。

出力 Z_k が左図の青い点で, 真値 t_k が赤い直線であるとしたとき, 誤差には正負があるので2乗して, それら全部を足して得られる次の 2乗和が使われる。

$$E = \frac{1}{2} \sum_{k=1}^{n} (z_k - t_k)^2$$

出典: 白猫のメモ帳 http://shironeko.hateblo.jp/entry/2016/10/29/173634

誤差最小化(勾配降下法)



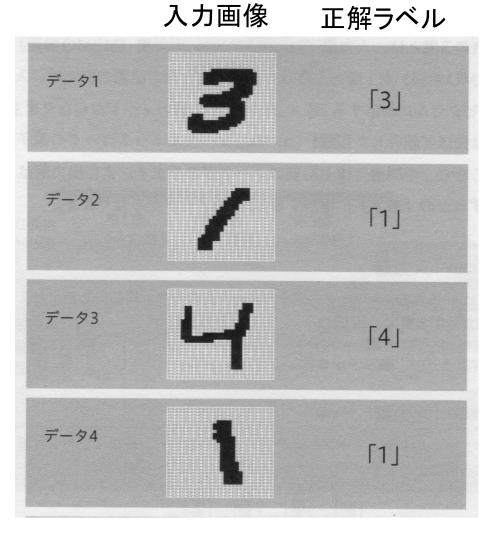
教師付き学習の図から判るよ うに. 出力 z は重み w の関数 であるから、前項の誤差関数 Eはwの関数である。 誤差関数 E(w) はw を1変数 としたとき、wの変化に対して 左図のような放物線になる。 その最小値を求める方法と して勾配降下法がある。

$$\mathbf{w} = \mathbf{w} - \varepsilon \, \frac{\partial E}{\partial \mathbf{w}}$$

 ε は更新の量を表す。

出典: http://yaju3d.hatenablog.jp/entry/2017/08/27/233459

4. 例題:手書き数字の認識



MNIST データセット は、手書きの数字「0~9」に 正解ラベルが与えられているデータセットである。

入力画像は, 手書き文字の縦28×横28ピクセルで, 各ピクセルの値は0~255であるが, 左図では128 以上を黒で, 未満を白で示した。

MNIST には、このような画像と正解ラベルを組にしたデータが7万枚あり、学習画像が6万枚、テスト画像が1万枚用意されている。

一般的には、学習画像を使って学習を行い、学習したモデルでテスト画像に対して正解率を計測する。

出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」 (翔泳社, 2017年)

手書き数字認識に三層パーセプトロンを適用

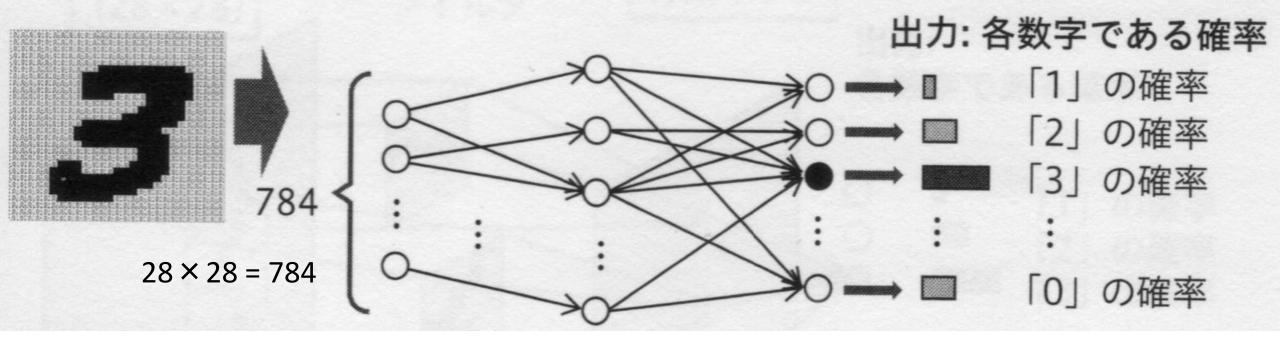
入力画像 (28×28 ピクセル) 第1層 (入力層) 784ノード

第2層 (中間層) 100ノード

第3層 (出力層) 10ノード

重みパラメータ数 784×100 = 約8万個

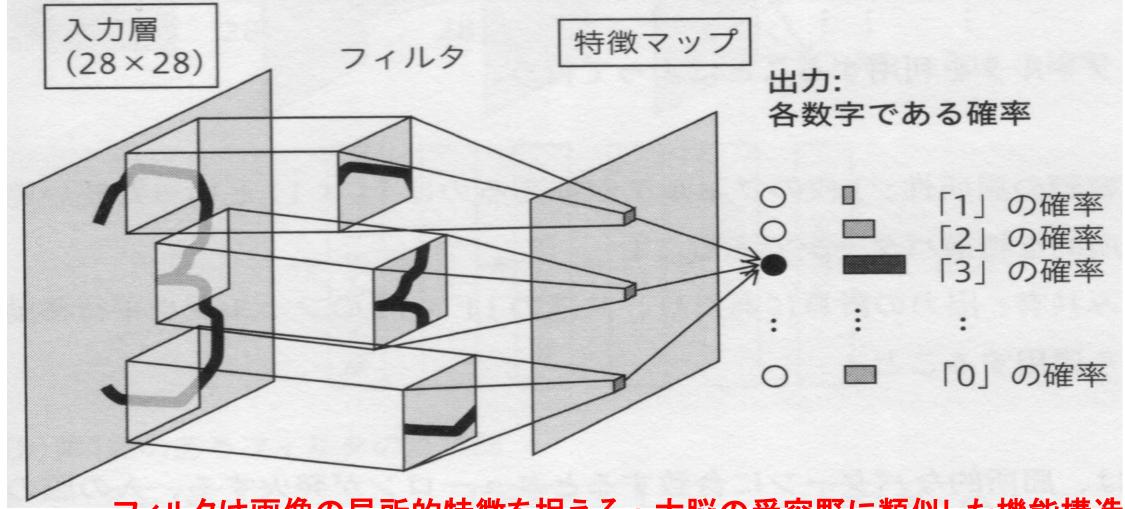
全結合ニューラル ネットワークでは、 階層を増やすと、 重みパラメータ数 が急激に増える。



出典: 大槻知史「最強囲碁 A I アルファ碁解体新書」(翔泳社, 2017年)

全結合の利用 95%の認識率

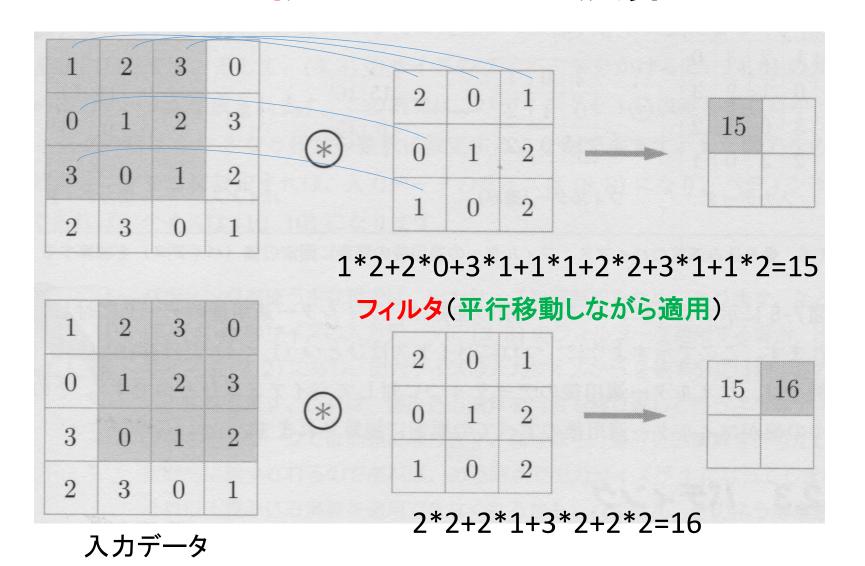
CNN(畳込みニューラルネットワーク)を適用する方法



フィルタは画像の局所的特徴を捉える:大脳の受容野に類似した機能構造初期段階では狭く単純な構造の受容野がみられ、高次になると広く複雑な構造の受容野がみられる。

出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」 (翔泳社, 2017年)

CNNの技法: 畳込み演算



全結合ニューラルネットワークの入力では、 画像の形状を無視して 1次元データにした。

CNNでは、形状を維持する。従って、画像の特徴を正しく評価できる(可能性がある)。フィルタ・畳込み演算の出力を特徴マップという。

フィルター: 重みに相当

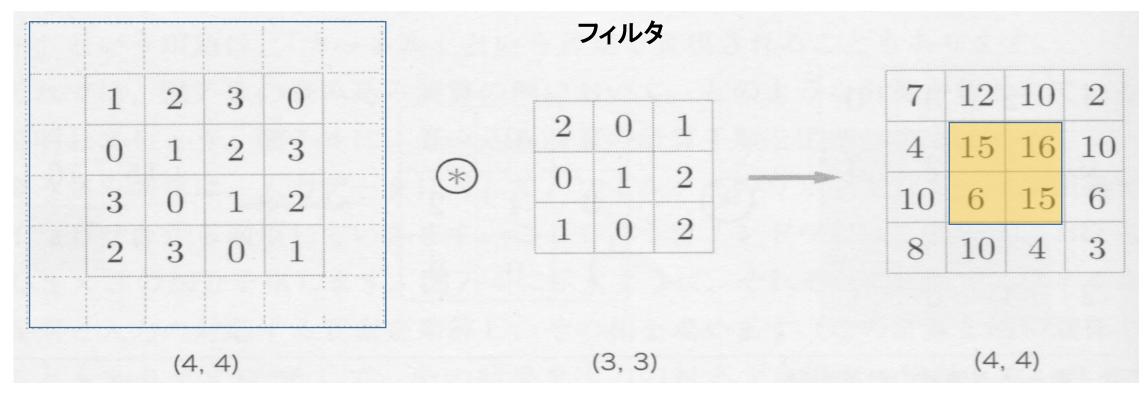
(重み共有)

出典: 斎藤康毅 「ゼロから作る Deep Learning」(オライリー・ジャパン,2016年)

CNNの技法: パディング処理

入力データ

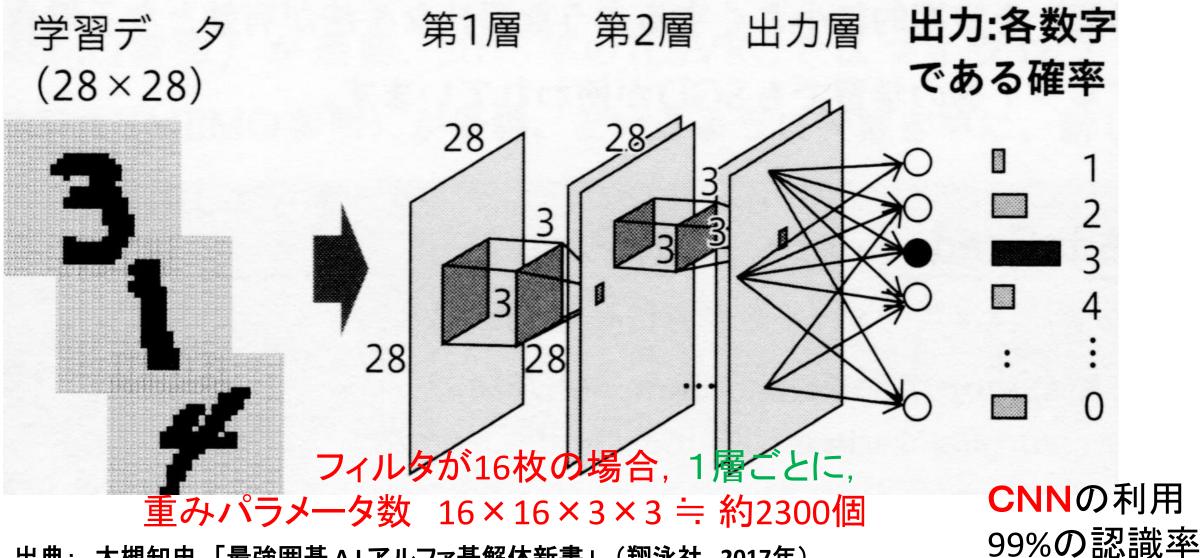
黄色の枠内は前図と同じ



入力データの周囲に0を埋める(図では0の記載を省略)。パディングを使うと、出力サイズが入力サイズと同じになる。

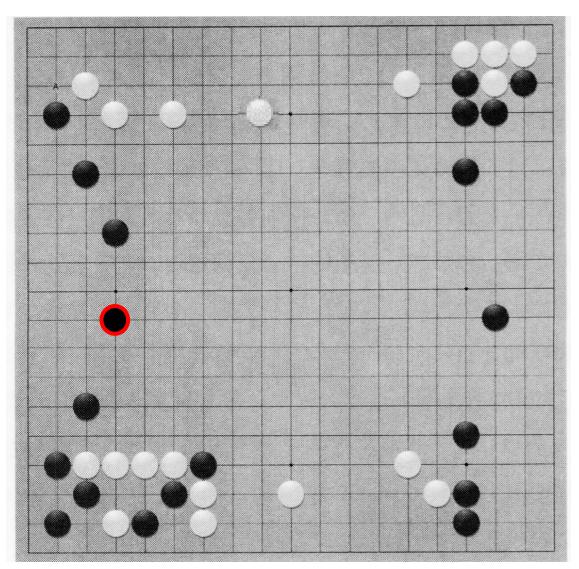
出典: 斎藤康毅「ゼロから作る Deep Learning」(オライリー・ジャパン, 2016年)

手書き数字認識への CNN 適用 重みパラメータ数が少ない



出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」 (翔泳社, 2017年) 出典: 斎藤康毅 「ゼロから作る Deep Learning」 (オライリー・ジャパン, 2016年)

5. 囲碁方策ネットワーク(教師付き学習)



AlphaGoが用いる2つのニューラルネットワーク

- 方策ネットワーク(policy networks) 与えられた局面に対して、各位置に打つ 確率の予測値を出力する。
- バリューネットワーク (value networks)与えられた局面に対して、その局面の勝率予測値を出力する。

今回は、方策ネットワークについて説明することにして、左図に例を示す。 現局面として、38手目までが与えられて、 39手目の黒番であるとき、方策ネットワークの予測値(あとで説明する)は、 ● の位置が最大になる。

出典: 大槻知史「最強囲碁 A I アルファ碁解体新書」(翔泳社, 2017年)

方策ネットワークの局面入力について

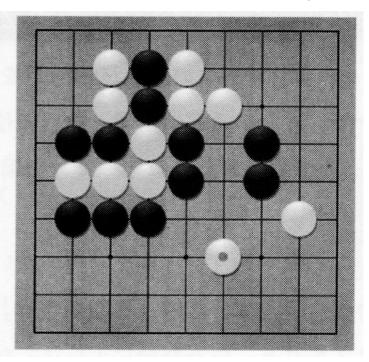
アルファ碁では、1つの局面に対して48チャネルの入力情報を作成する。

入力チャネルの種類	チャネル数
黒石の位置	1
白石の位置	1
空白の位置	1
k 手前に打たれた位置 (k = 1 ~ 8)	8
その石の当該連の呼吸点の数 (k = 1 ~ 8)	8
そこに打った時, 取れる石の数 (k=1~8)	8
そこに打った時, 当該連を取られる石の数 (k=1~8)	8
そこに打った時、隣接する相手の連をシチョウで取れるか?	1
合法手か?	1
その他	11
合計	48

出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」(翔泳社, 2017年)

局面と画像との類似性

(a)現局面 (黒番)



現局面 (a) から, 特徴量を取り出すと (b)~(f) が得られる。特徴量の図は, 手書き文字の図に類似している。 カラー画像で RGB 特性をチャネルと呼ぶのに準じて, 入力チャネルという。

入力チャネルの例

(b)黒石の位置

0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	1	1	0	1	0	1	0	0
0	0	0	0	1	0	1	0	0
0	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

(c)白石の位置

0	0	0	0	0	0	0	0
0	1	0	1	0	0	0	0
0	1	0	1	1	0	0	0
0	0	1	0	0	0	0	0
1	1	1	0	0	0	0	0
0	0	0	0	0	0	1	0
0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
	0 0 0 0 0 0	0 1 0 1 0 0 1 1 0 0 0 0 0 0	0 1 0 0 1 0 0 0 1 1 1 1 0 0 0 0 0 0 0	0 1 0 1 0 1 0 1 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0	0 1 0 1 0 0 1 0 1 1 0 0 1 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 1 0 1 0 0 0 1 0 1 1 0 0 0 0 1 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 1 0 1 0 0 0 0 1 0 1 1 0 0 0 0 0 1 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0

(e) n 手前の位置

0	0	0	0	0	0	_	_	_
0	0				0	0	0	0
	0	0	0	5	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	4	0	0
0	7	0	0	0	0	2	0	0
0	0	8	6	0	0	0	3	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

(f)連の呼吸点の数

0	0	0	0	0	0	0	0
0	3	1	4	0	0	0	0
0	3	1	4	4	0	0	0
2.		2	3	0	6	0	0
	2	2	3	0	6	0	0
5	5	5	0	0	0	4	0
0	0	0	0	4	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
	0 0 5 0	0 3 2 2 2 2 5 5 5 0 0 0 0	0 3 1 0 3 1 2 2 2 2 5 5 5 0 0 0 0 0 0	0 3 1 4 0 3 1 4 2 3 3 2 2 3 5 5 5 0 0 0 0 0 0 0 0 0	0 3 1 4 0 0 3 1 4 4 2 2 3 0 2 2 3 0 5 5 5 0 0 0 0 0 0 4 0 0 0 0 0	0 3 1 4 0 0 0 3 1 4 4 0 2 2 3 0 6 2 2 3 0 6 5 5 5 0 0 0 0 0 0 0 4 0 0 0 0 0 0 0	0 3 1 4 0 0 0 0 3 1 4 4 0 0 2 2 2 3 0 6 0 2 2 3 0 6 0 5 5 5 0 0 0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0

出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」 (翔泳社, 2017年)

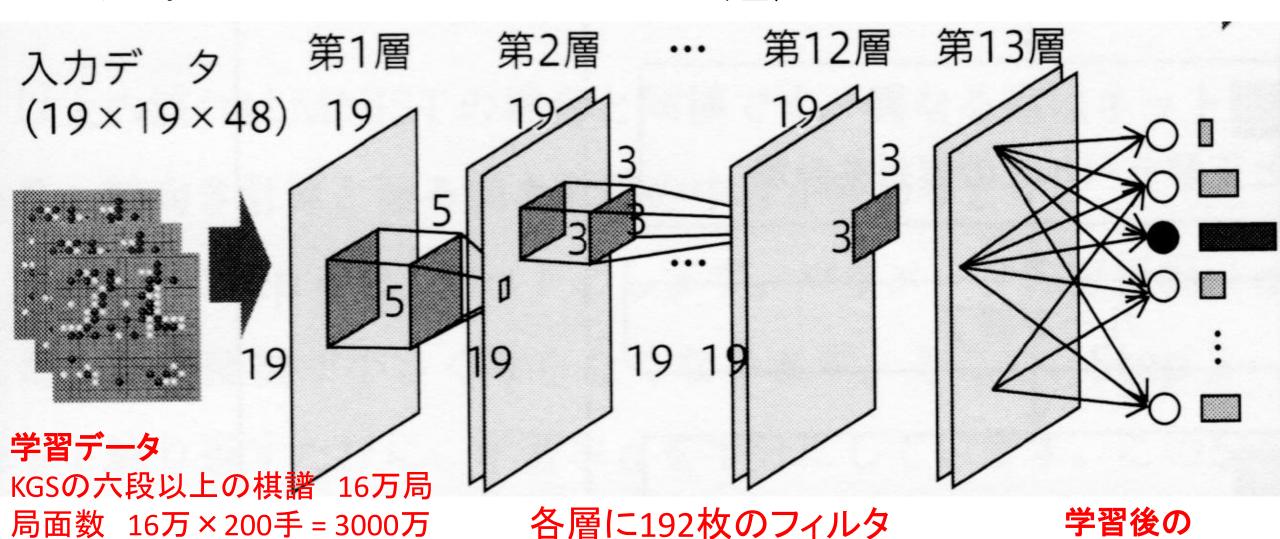
方策ネットワークと手書き数字との類似性

アルファ碁の「次の一手」タスクは、手書き数字認識とよく似ている。

	手書き数字	囲碁の手の選択
対象	手書き数字	囲碁の局面情報
入力層	28×28ピクセルの画像 グレースケールの1チャネル 値は0~255のいずれか	19×19 の盤面 48チャネル 値は 0 ~ 1 のいずれか
出力層	ラベルの 0~9 のいずれか	19×19 の盤面位置のいずれか
手法	4層のディープラーニング	13層のディープラーニング
学習データ	手書きした数字	6段以上の棋譜局面
教師データ	正解ラベル	局面の次の手

出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」(翔泳社, 2017年)

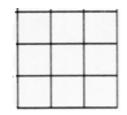
方策ネットワークへの CNN 適用



出典: 大槻知史「最強囲碁 A I アルファ碁解体新書」(翔泳社, 2017年) 出典: 斎藤康毅「ゼロから作る Deep Learning」(オライリー・ジャパン, 2016年) 一致率 57%

フィルタの機能: CNNの各層に192 枚のフィルタ

フィルタサイズ 3×3 のパターン



CNNがすごいのはこのフィルタを開発者が手動で設計するのではなく、 学習によって自動獲得できるという 点にある。 重みパラメータの

求め方:

調整の方法:

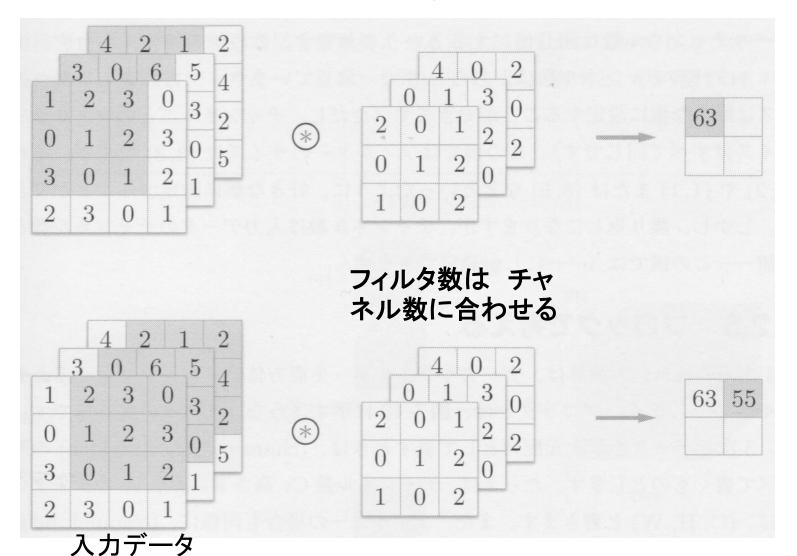
誤差の評価(2乗和誤差)

誤差最小化(勾配降下法)

出典: 人工知能に関する断創録 http://aidiary.hatenablog.com/

出典: 斎藤康毅「ゼロから作る Deep Learning」(オライリー・ジャパン, 2016年)

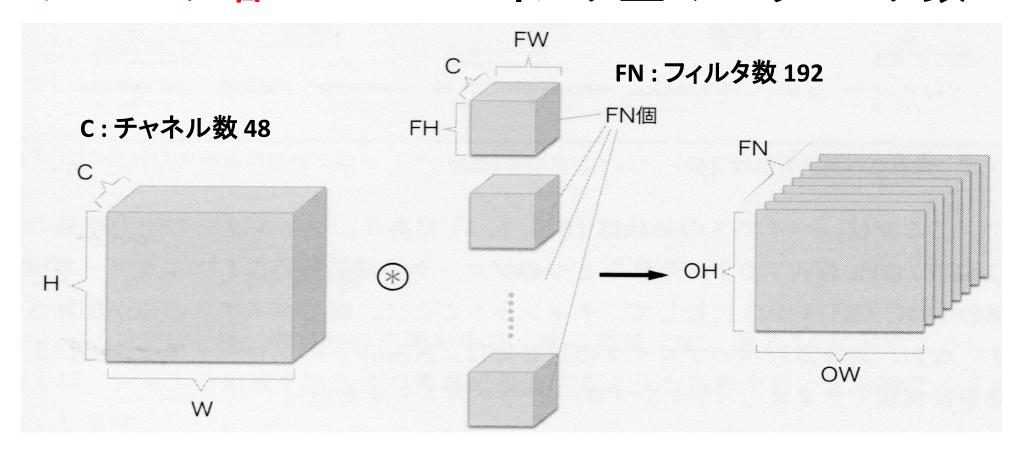
CNNの技法: 複数チャネルに対する畳込み演算



複数チャネルの計算方法 1チャンネル目の畳込み演算, 2チャンネル目の畳込み演算, 3チャンネル目の畳込み演算, それらの結果を加算して, 1つの値を求める。 以下,同様に計算する。

出典: 斎藤康毅「ゼロから作る Deep Learning」(オライリー・ジャパン, 2016年)

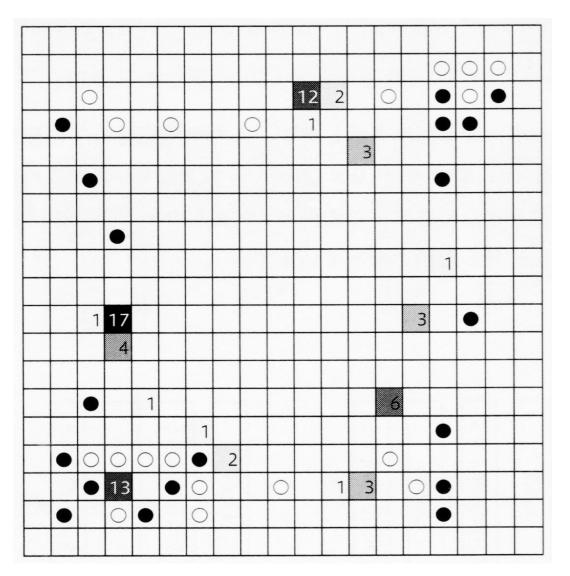
アルファ碁 CNN: フィルタ重みパラメータ数



フィルタ重みパラメータ数 3×3×192×192×12層 = 約400万

出典: 大槻知史「最強囲碁 A I アルファ碁解体新書」(翔泳社, 2017年) 出典: 斎藤康毅「ゼロから作る Deep Learning」(オライリー・ジャパン, 2016年)

方策ネットワーク(教師付き学習)の予測値



現局面として、38手目までが与えられて、 39手目の黒番であるとき、方策ネットワーク が出力した各位置に打つ確率の予測値を、 示した。空欄は1%未満である。

この例では、左真ん中の17%と表示された手が最大なので、着手として選ばれる。

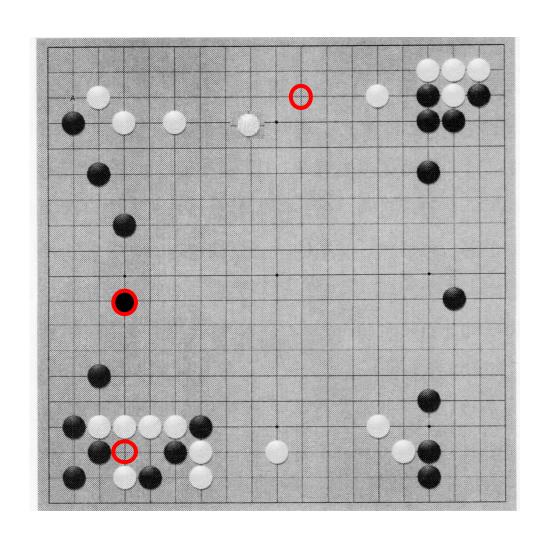
他の有力な着手としては,

- 左下隅の13%の位置
- ・ 上辺中央部の12%の位置

方策ネットワークだけを用いても、アマチュア 三段くらいの棋力となる。

出典: 大槻知史 「最強囲碁 A I アルファ碁解体新書」 (翔泳社, 2017年)

囲碁方策ネットワークのまとめ



方策ネットワークについて説明した。 左図の現局面が与えられて黒番であるとき、 予測値は ● の位置が示された。 他の有力な候補として ○ が示された。

方策ネットワークだけを用いたとき, 着手は アマチュア六段と較べて 57% であり, 三段 くらいの棋力とみられる。

AlphaGoが強くなったのは、これを踏み台にして、自己対戦を用いた強化学習の技術が成功したためである。今後、この強化学習(教師なし学習)について、説明したい。

出典: 大槻知史「最強囲碁 AIアルファ碁解体新書」(翔泳社, 2017年)