

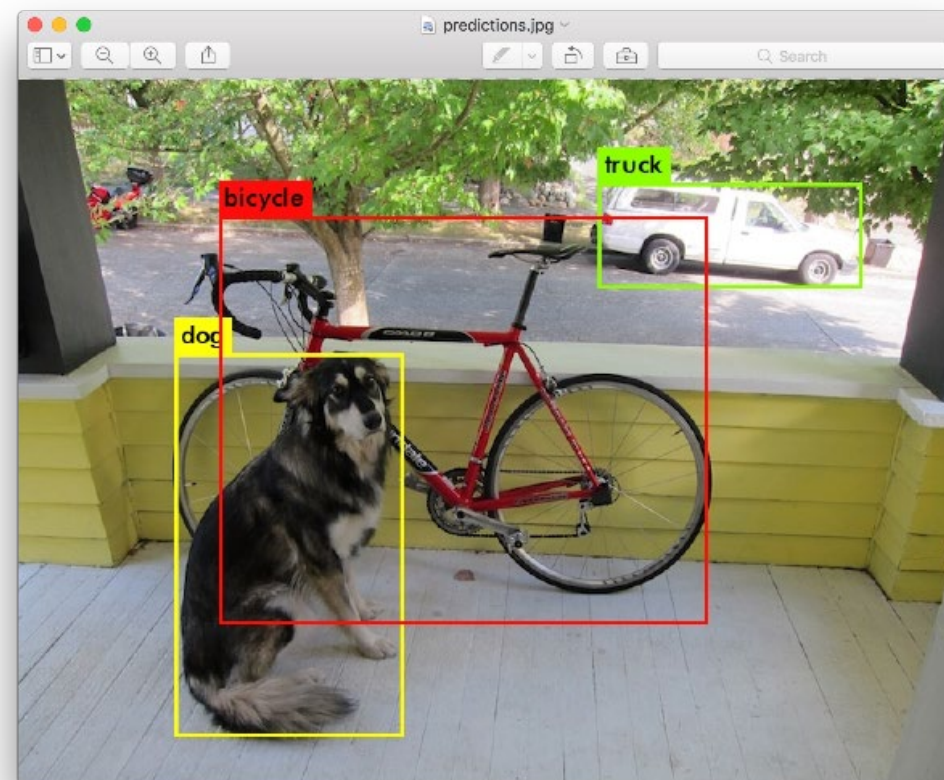
計算知能 (COMPUTATIONAL INTELLIGENCE)

第 1 1 回 畳み込みニューラルネットワーク
教員： 谷口彰

第11回 畳み込みニューラルネットワーク

■ 畳み込みニューラルネットワーク

- 畳み込み層
- プーリング層

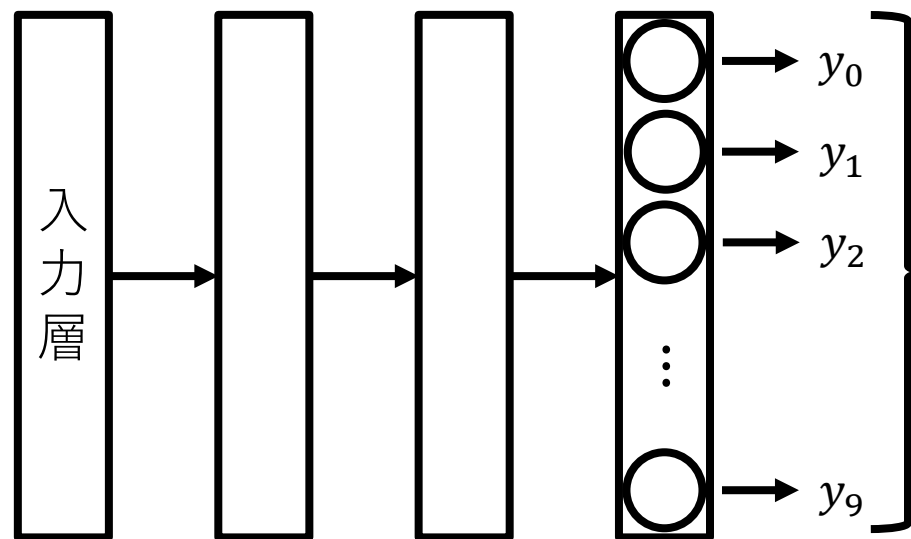


畳み込みニューラルネットワーク

- 畳み込みニューラルネット（Convolutional Neural Network, CNN）は、
畳み込み層と**プーリング層**という特別な2種類の層を含む
順伝播型（順伝搬型）ネットワーク（feedforward neural network）で、
画像認識・クラス分類に応用される
- 通常の順伝播型ネットワークと同様に、誤差逆伝播法を用いて学習する
- 局所受容野および重み共有と呼ばれる特別な層間結合を持つことが
これまでのネットワークとの違い

ニューラルネットワークによるクラス分類

- 入力層：これまでの順伝播型ネットワークと同様
- 出力層：各ユニットの値域は $[0,1]$ 、全ユニットの総出力は1
- 教師信号：入力データのクラスラベルを表す1hotな信号

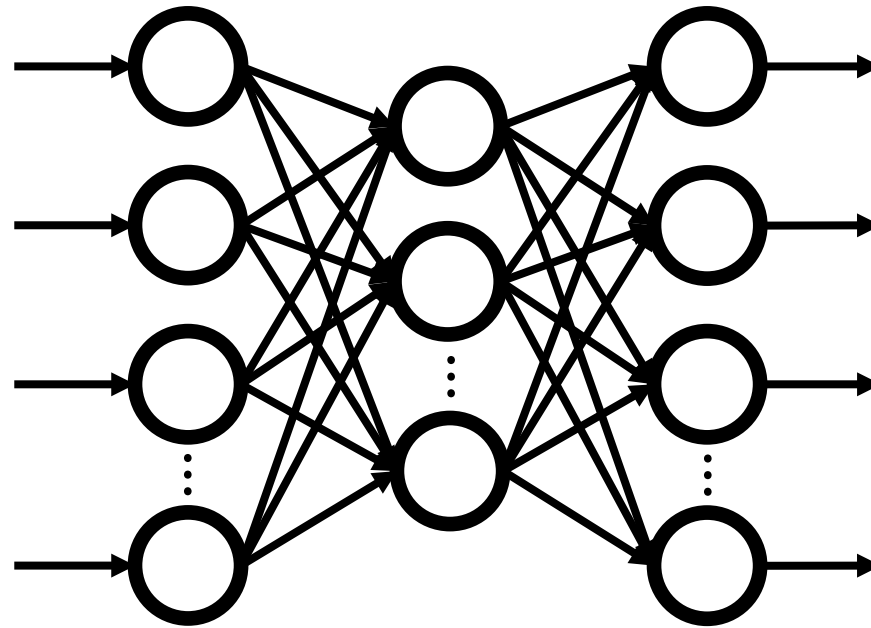


各ユニットが
クラスラベルに相当

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

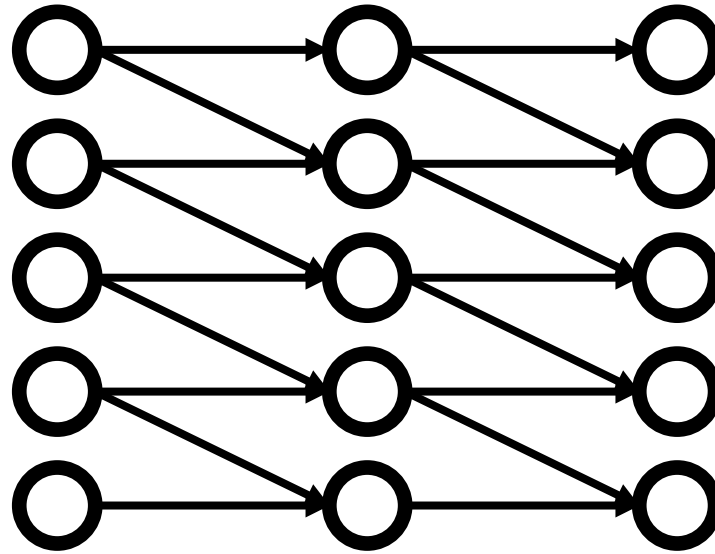
一般的な順伝搬型ネットワークとの違い(1/3)

- 一般的な順伝搬型ネットワークは各層、すべてのユニット間において結合が存在



一般的な順伝搬型ネットワークとの違い(2/3)

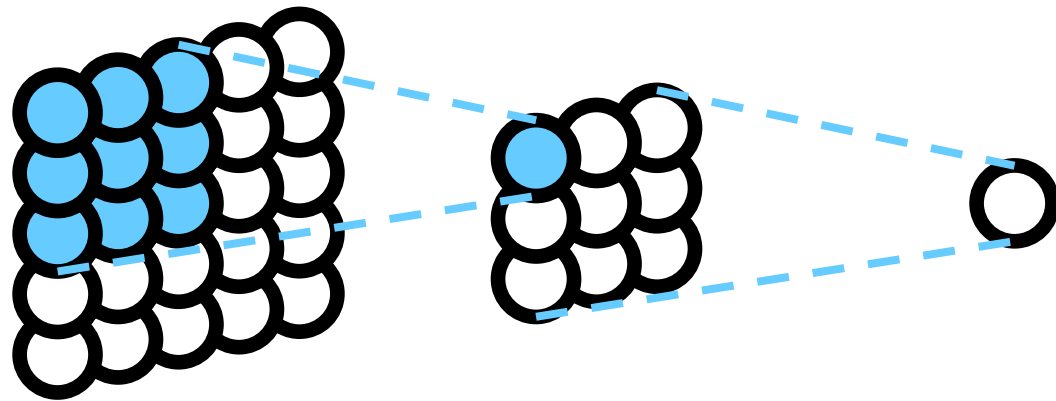
- 畳み込みニューラルネットワークは隣接層間の特定のユニットのみが結合を持つ特別な層を持つ



- これらの層では、畳み込みとプーリングという画像処理の基本的な演算を実施（生物の脳の視覚野に関する知見を応用）

一般的な順伝搬型ネットワークとの違い(3/3)

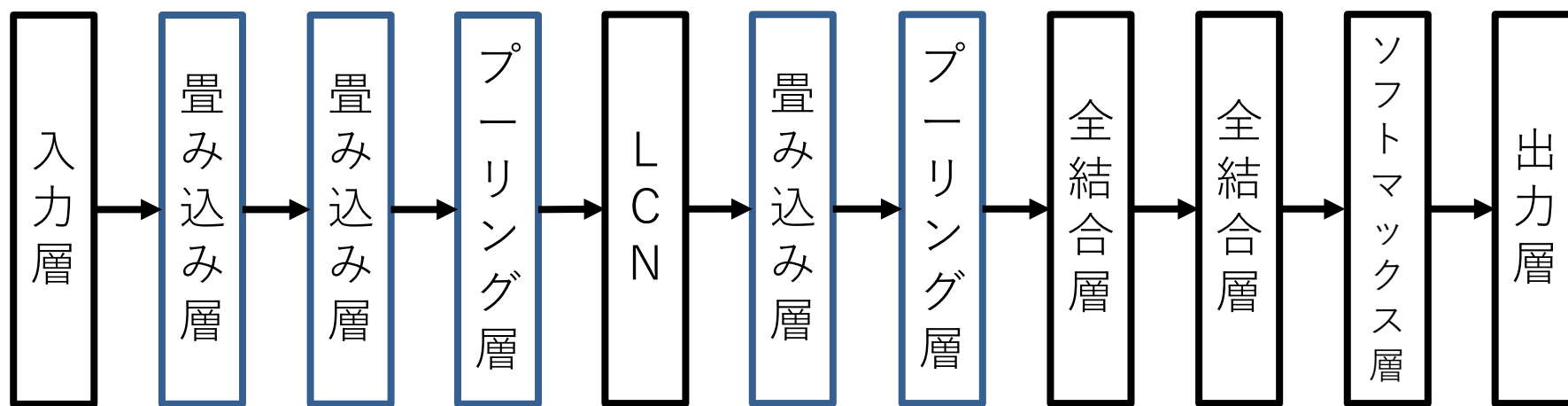
- 各層のユニットは2次元的に並んでおり、以下では中間層の各ユニットは、入力層の 3×3 のユニット群とのみ結合を持ち、そこに特定のパターンが入力されるとそれに反応して活性化する



一部のユニット間のみ、結合が存在

CNNの全体構造(1/2)

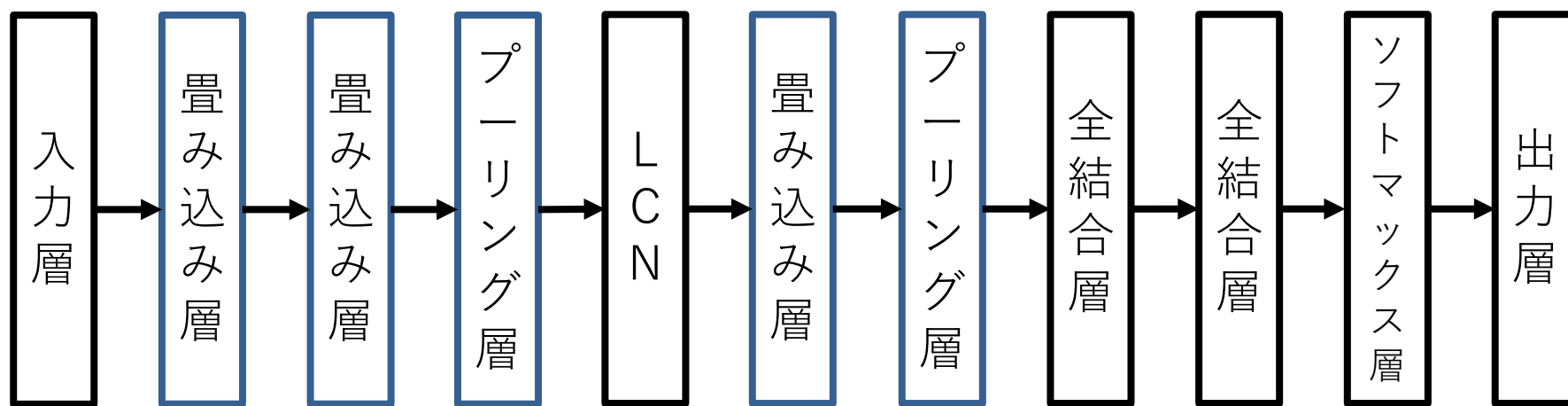
- 入力層から出力層に向けて、畳み込み層とプーリング層がペアで並び、複数回繰り返される
(畳み込み層だけが複数回繰り返された後、プーリング層が1層続く場合もある)



CNNの全体構造 (クラス分類)

CNNの全体構造(2/2)

- 畳み込み層とプーリング層の間に局所コントラスト正規化層 (Local Contrast Normalization, LCN)を挿入することもある
- 畳み込み層とプーリング層を繰り返した後は、隣接層間のユニットが全結合した層が配置される



CNNの全体構造 (クラス分類)

畳み込み層(1/7)

■ 畳み込みの演算を行う単層ネットワーク

画像

1	2	2	4	3	4	5	4	5	5
1	2	3	4	3	4	5	5	4	6
2	2	2	3	3	4	4	5	6	6
3	3	3	4	4	5	6	5	6	7
4	4	4	3	4	6	7	8	7	8
3	3	3	4	4	5	6	8	7	8
4	3	5	4	6	7	7	8	8	9
4	5	6	5	7	6	7	9	8	9

フィルタ($w=1$)

$1/9$	$1/9$	$1/9$
$1/9$	$1/9$	$1/9$
$1/9$	$1/9$	$1/9$

積和

結果画像

畳み込み層(4/7)

- 画像は i 層の出力、フィルタは $i \sim j(: i + 1)$ 層間の重みに対応する

画像

1	2	2	4	3	4	5	4	5	5
1	2	3	4	3	4	5	5	4	6
2	2	2	3	3	4	4	5	6	6
3	3	3	4	4	5	6	5	6	7
4	4	4	5	5	6	7	6	7	8
3	3	3	4	4	5	6	5	6	7
4	3	5	4	6	7	7	8	8	9
4	5	6	5	7	6	7	9	8	9

ニューラルネットワークの
 i 層での出力 o_i

フィルタ($w=1$)

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

ニューラルネットワークの
 $i \sim j(: i + 1)$ 層間の重み w_{ji}

畳み込み層(5/7)

- 畳み込み結果は $j(:i+1)$ 層の入力に対応する
- i 層の一部のユニットとのみ、結合を持つ

結果画像（CNNでは特徴マップと呼ぶ）

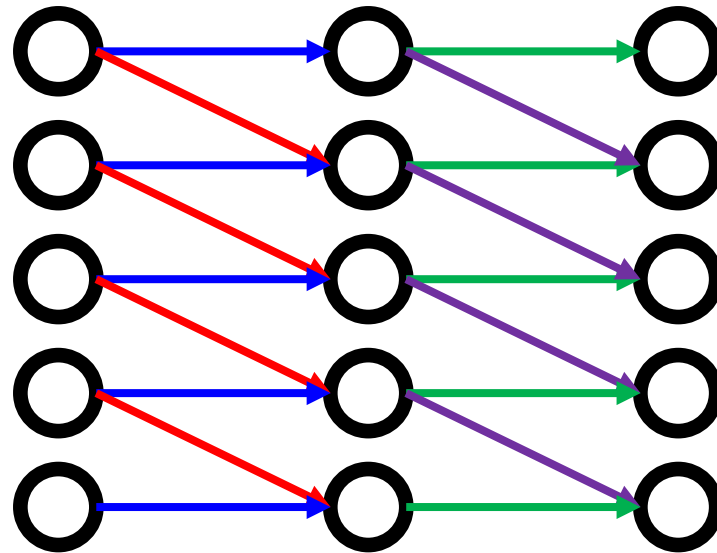
0	0	0	0	0	0	0	0	0	0
0	2	3	4	3	4	5	5	4	0
0	2	2	3	3	4	4	5	6	0
0	3	3	4	4	5	6	5	6	0
0	4	4	5	5	6	7	6	7	0
0	3	3	4	4	5	6	6	7	0
0	3	5	4	6	7	7	8	8	0
0	0	0	0	0	0	0	0	0	0

ニューラルネットワークの
 $j(i+1)$ 層での入力 u_j

畳み込み層(6/7)

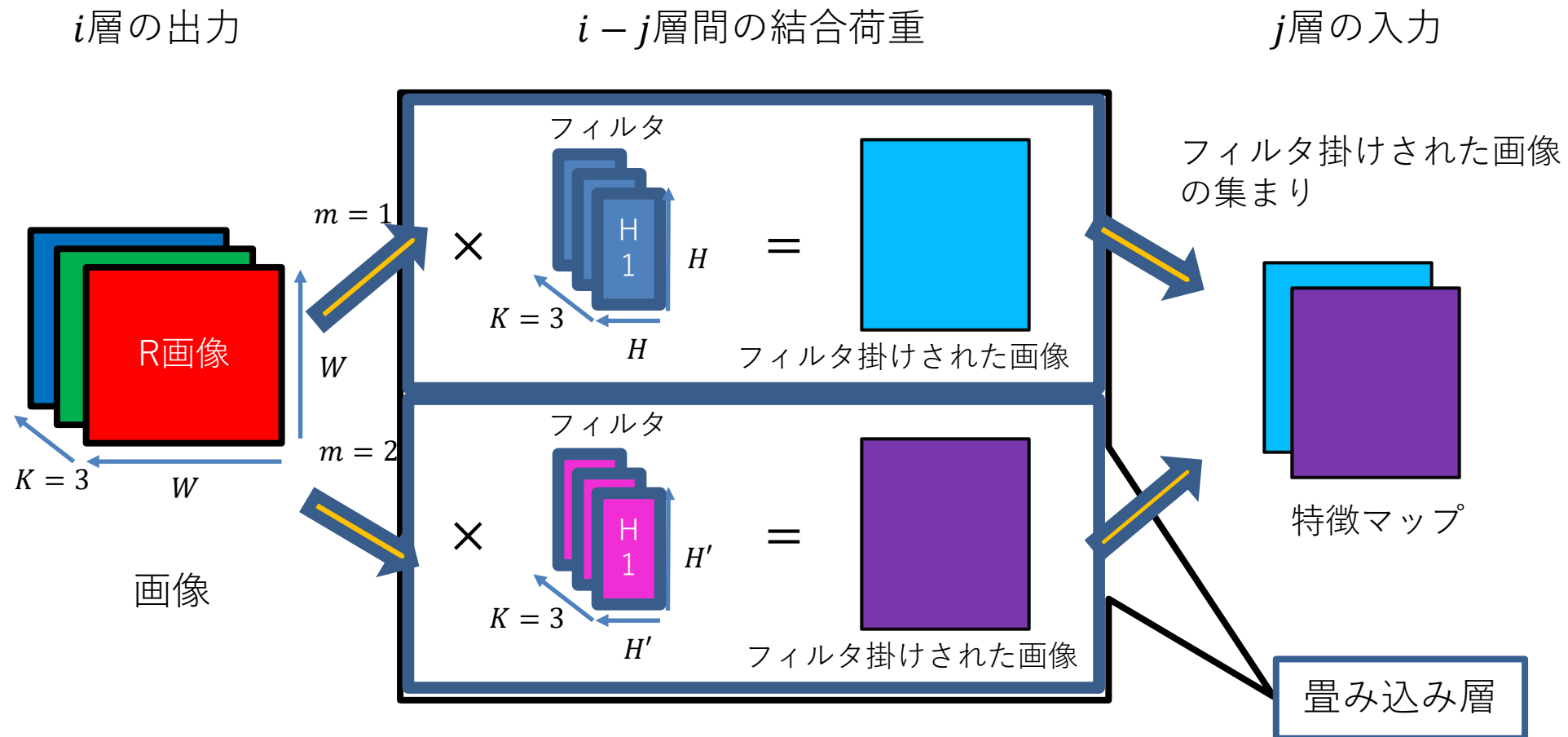
■ 重み共有

- 複数のユニット間接続において、同じ重み（フィルタ）が共有される



同じ色の接続は同じ重みを表現

畳み込み層(7/7)



入力として受け取った画像から特徴マップを出力する層

プーリング層(1/4)

- 局所における最大値を抽出 (MAXプーリング)

特徴マップ

	2	3	4	3	4	5	5	4	
	2	2	3	3	4	4	5	6	
	3	3	4	4	5	6	5	6	
	4	4	3	4	6	7	8	7	
	3	3	4	4	5	6	8	7	
	3	5	4	6	7	7	8	8	

プーリング結果

3	4					

プーリング層(2/4)

- 最適化する重みは存在しない

特徴マップ

	2	3	4	3	4	5	5	4	
	2	2	3	3	4	4	5	6	
	3	3	4	4	5	6	5	6	
	4	4	3	4	6	7	8	7	
	3	3	4	4	5	6	8	7	
	3	5	4	6	7	7	8	8	

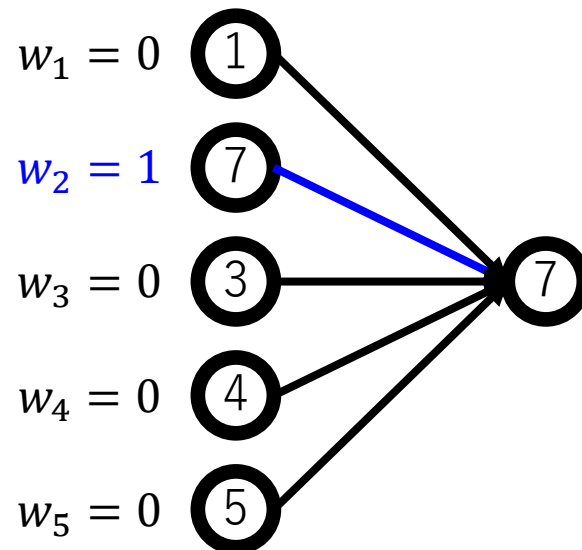
プーリング結果

3	4	4	4	5	5	6
3	4	4	5	6	6	6
4	4	4	6	7	8	8
4	4	4	6	7	8	8
5	5	6	7	7	8	8

プーリング層(3/4)

■ MAXプーリング

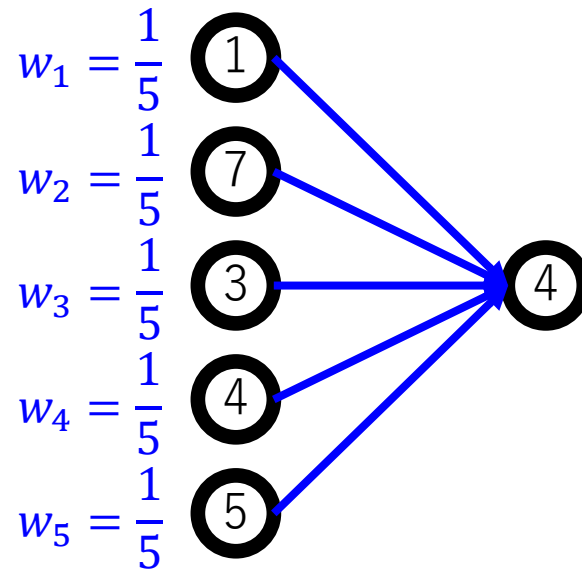
複数のユニット間結合のうち、最大値を持つユニットの結合のみ1、残りはすべて0



プーリング層(4/4)

■ Averageプーリング

n 個のユニット間結合はすべて $1/n$



畳み込み層の意味

- 画像の畳み込み演算は画像とフィルタの相関を表す

$$g(i, j) = f * h = \sum_{n=-w}^w \sum_{m=-w}^w \underbrace{f(i + m, j + n)}_{\text{red}} \cdot \underbrace{h(m, n)}_{\text{green}}$$

両方の値が大きいときだけ
大きな値を返す

- 相関が強くなる（絶対値の大きい値を返す）ように
重み（フィルタ）の学習が進む

プーリング層の意味(1/2)

- 画像中の値の位置が多少変化しても、演算結果は同じ
→ 画像中で、物体の存在する位置が変わっても
同じものと判断することが可能

5	3
2	1

MAXプーリング



5

1	5
1	2

MAXプーリング



5

プーリング層の意味(2/2)

- 画像中の値の位置が多少変化しても、演算結果は同じ
→ 画像中で、物体の存在する位置が変わっても
同じものと判断することが可能



画像認識



猫



画像認識



猫

局所コントラスト正規化(1/2)

- 一般的な画像認識問題では画像ごとの全体的な明るさやコントラストの違いをうまく吸収する必要がある



色は違っても撮影対象は同じ（同じクラスに分類したい）

- 局所コントラスト正規化は減算正規化と除算正規化に分けられる

局所コントラスト正規化(2/2)

- 減算正規化

画素 (i, j) を中心とする $H \times H$ の正方領域における平均値を $\overline{x_{ij}}$ とし、各画素値 x_{ij} から差し引く

$$z_{ij} = x_{ij} - \overline{x_{ij}}$$

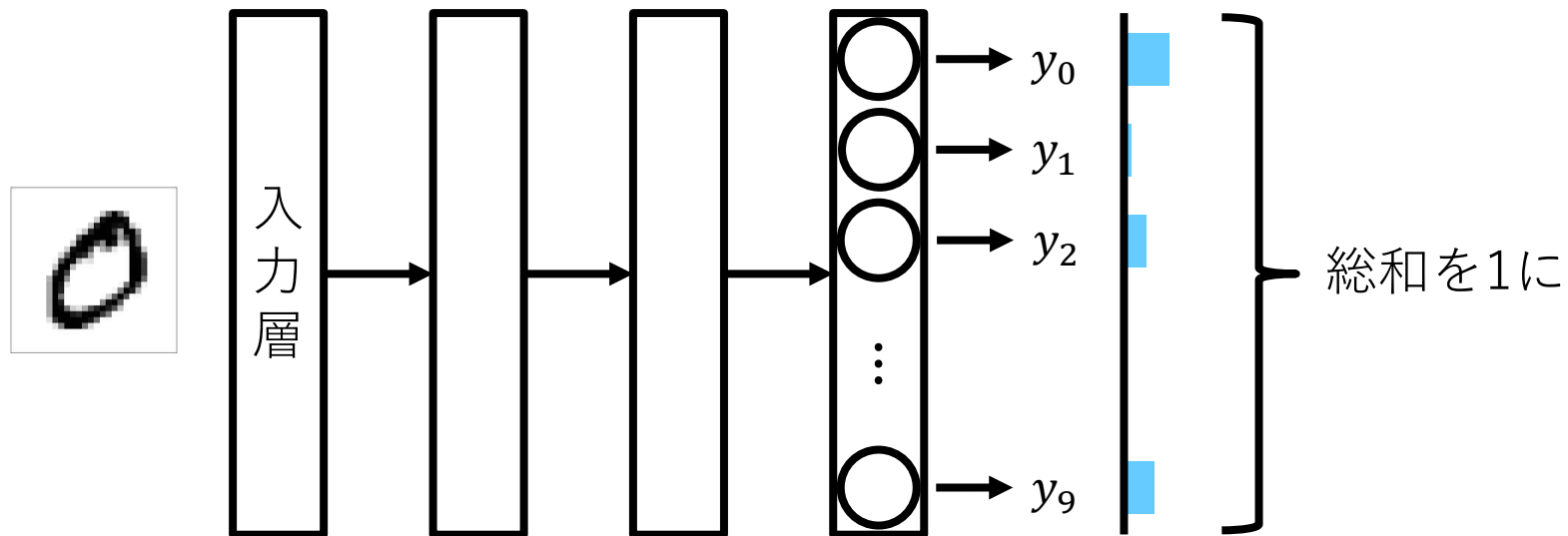
- 除算正規化

各領域内での画素の分散を σ_{ij}^2 とし、減算正規化ののちに各領域内での分散を揃える

$$z_{ij} = \frac{x_{ij} - \overline{x_{ij}}}{\sigma_{ij}}$$

ソフトマックス関数(1/2)

- ニューラルネットワークによる多クラス分類に使用される
- 出力層の k 番目のユニットにおける総入力 u_k をもとに出力層の全ユニットの合計出力が1となるよう値を調整



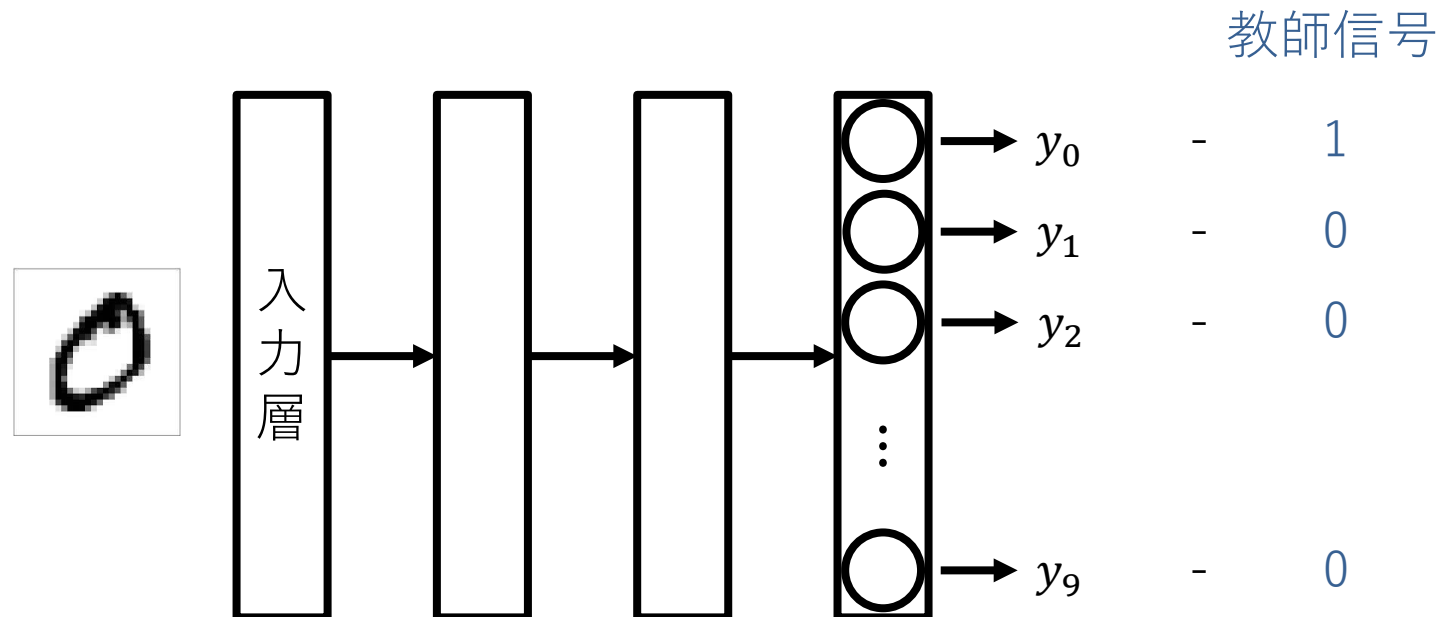
ソフトマックス関数(2/2)

- ニューラルネットワークによる多クラス分類に使用される
- 出力層の k 番目のユニットにおける総入力 u_k をもとに出力層の全ユニットの合計出力が1となるよう値を調整

$$y_k \equiv \frac{\exp(u_k)}{\sum_{j=1}^K \exp(u_j)}$$

CNNの学習

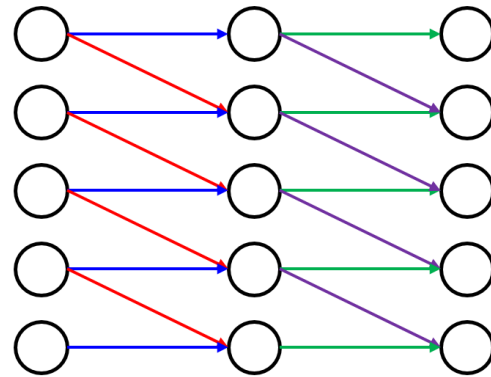
- 出力層の出力値と、教師信号の差を最小化するように学習
- 誤差の逆伝搬により学習するが、畳み込み層とプーリング層では注意が必要



畳み込み層の学習

- 同じ重みが複数のユニット結合に存在する
- バッチ学習により複数結合をまとめて更新

$$E(\mathbf{w}) = \sum_{n=1}^N E_{n(\mathbf{w})} \quad \text{複数ユニットにおける誤差の総和を最小化}$$



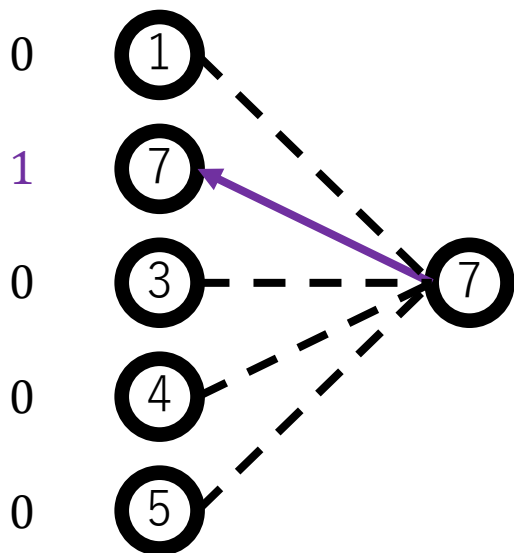
- 結合のないユニット間は誤差の逆伝搬を0とする

プーリング層の学習

- プーリング層には学習対象となる重みはないが、次の層へ誤差の逆伝搬計算は必要

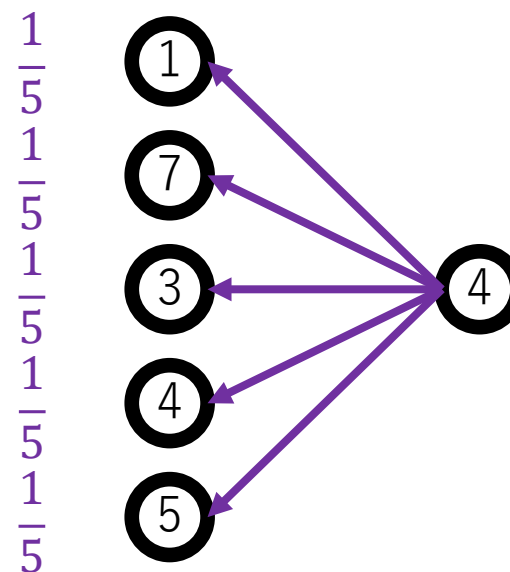
MAXプーリング

最大値を与えたユニットのみ1、
それ以外はすべて0倍の誤差を伝搬



Averageプーリング

すべてのユニットに $1/n$ 倍の誤差を
伝搬

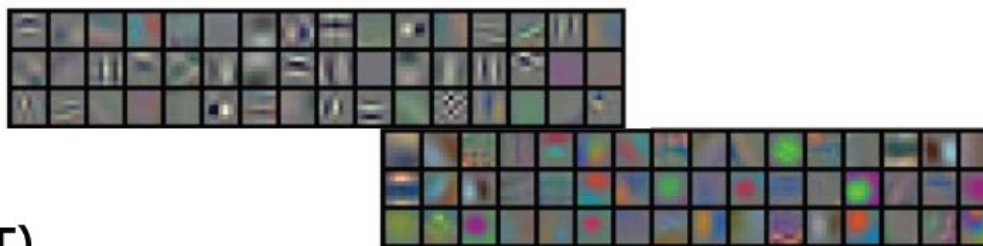


学習後のフィルタ

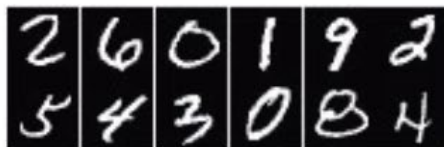
Object category (ImageNet)



第1層で学習された特徴 (フィルタ)



Handwritten digit (MNIST)



Material (FMD)

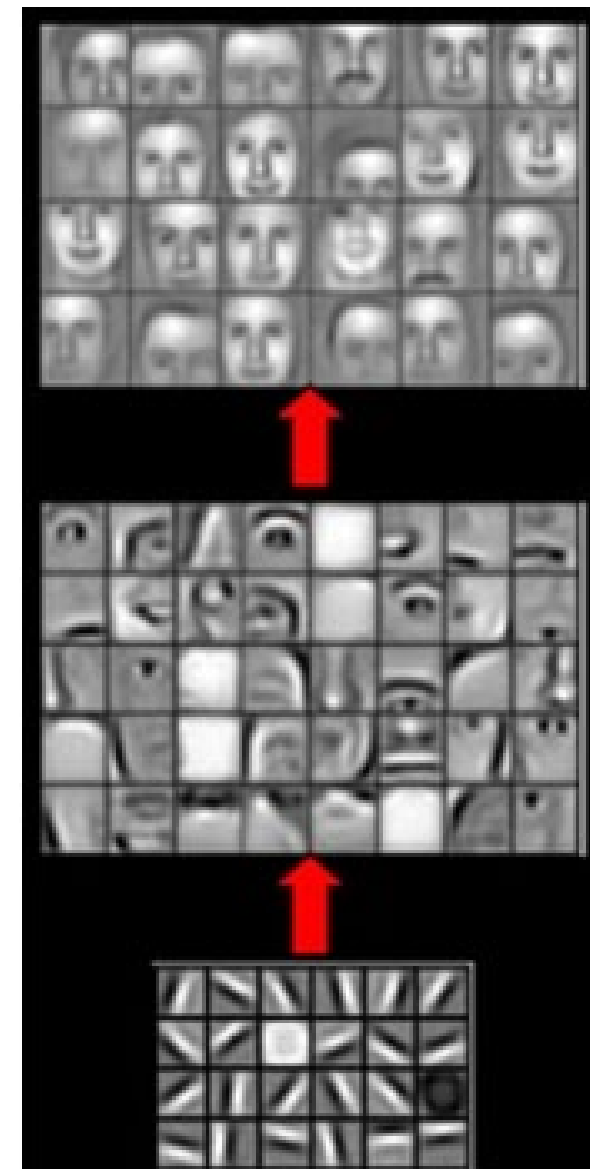


Glossiness

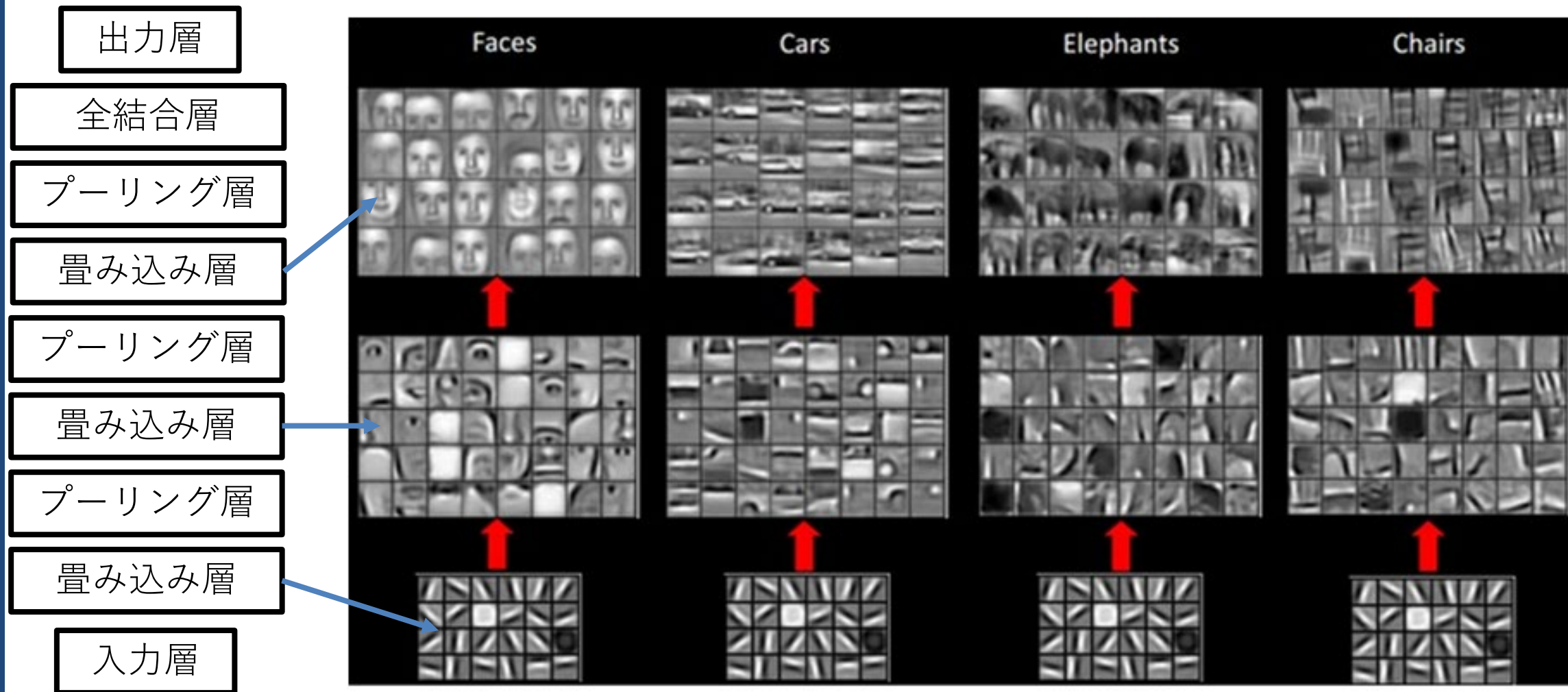


各畳み込み層の出力(1/2)

- 浅い層では単純な線、ついで細かい部品、最後は全体と、処理が進んでいる
- 多層にすることで、より抽象度の高い特徴を抽出できている

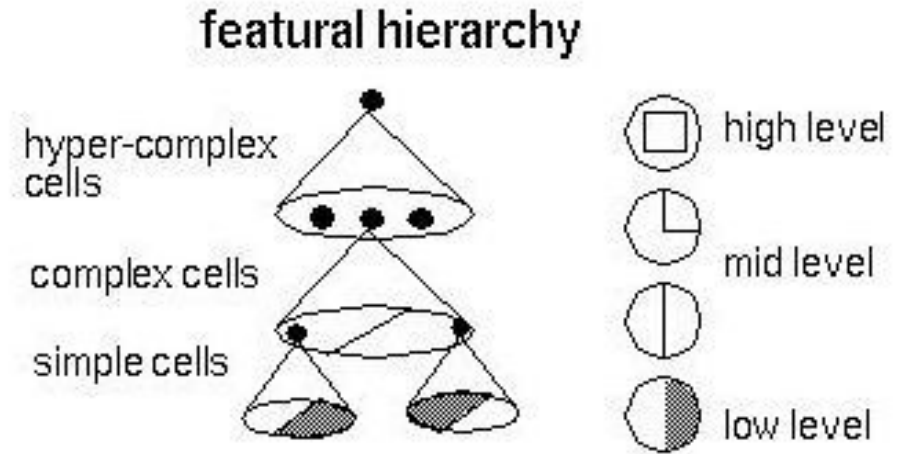
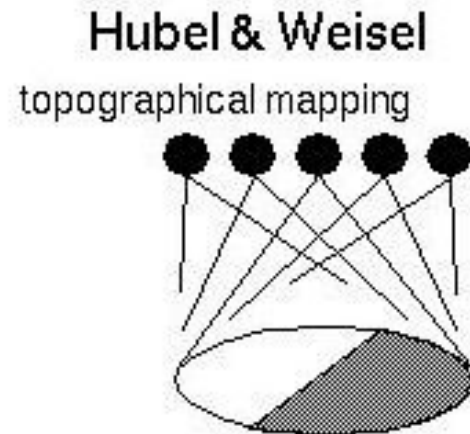
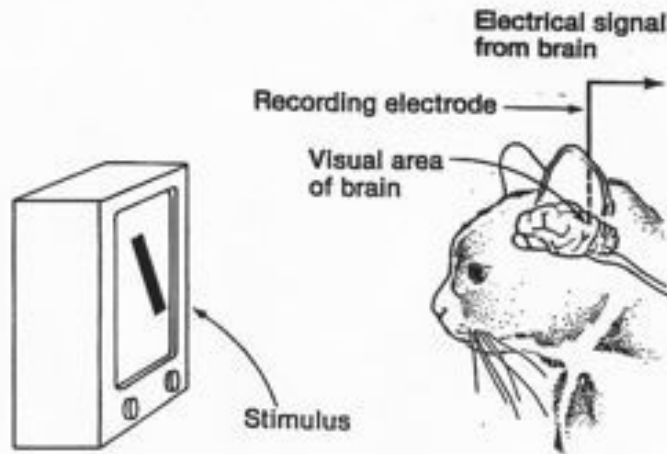


各畳み込み層の出力(2/2)



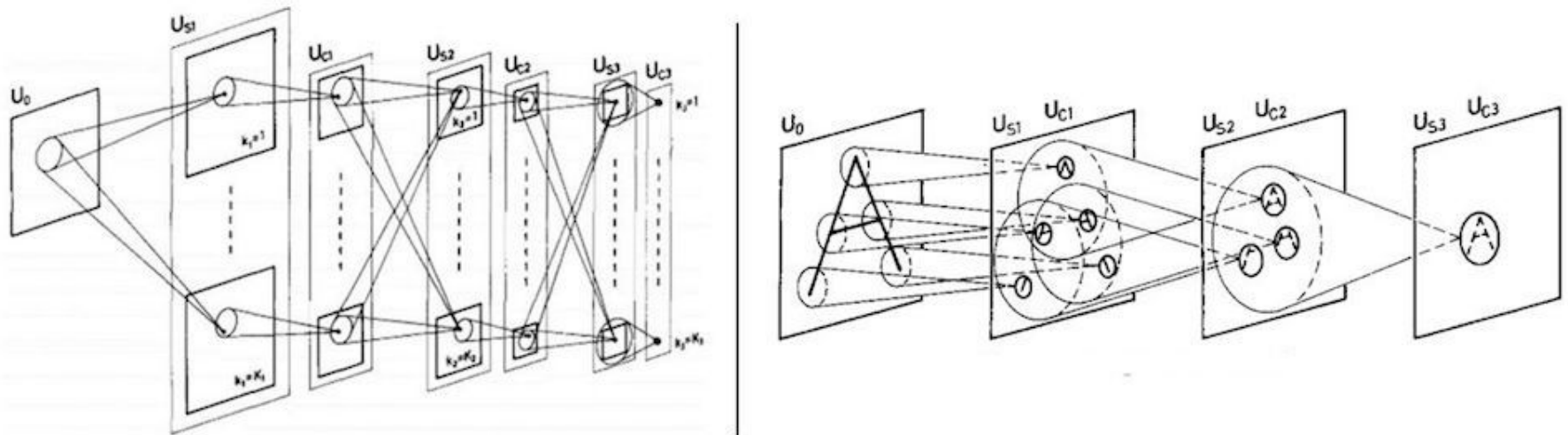
神経生物学における視覚野の実験

- 1960年代に動物の視覚野の特性を調べるため、神経生理学者のデイヴィッド・ヒューベルとトルステン・ウィーセルが一連の実験を行った。
- テレビスクリーンのシンプルなパターンで猫の脳を刺激してその電気反応を測定し、視覚野の初期段階におけるニューロンが次のように階層的に組織されていることを突き止めた。
- 猫の網膜につながっている最初のセルはエッジやバーといったシンプルなパターンの検出を担っており、後の層がそれら初期階層のニューロンの出力を組み合わせることでより複雑なパターンを検出する事がわかった。



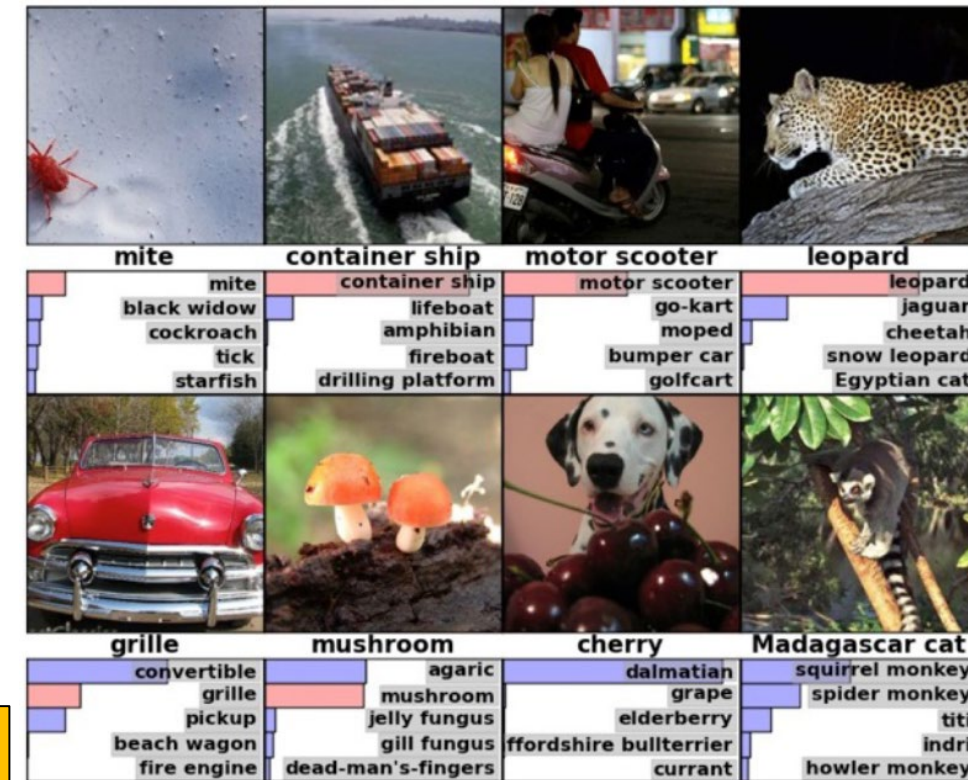
ネオコグニトロン

- ヒューベルとウィーセルの実験は、福島邦彦がニューラルネットワークで視覚野の階層的で構成的な視覚野の仕組みを模することを試みたネオコグニトロン（1982）を考案した際に、インスピレーションとして直接言及されている。
- ネオコグニトロンはスライディングウィンドウを用いることで、階層になったそれぞれの層が前の層から画像の中の位置に関わらずパターンを検出することのできる構造を持った初めてのネットワークだった。
- ネオコグニトロンはパターン認識においてある程度の成果を納め、畳み込みフィルタをニューラルネットワークに取り入れた。



Deep Learning(1/2)

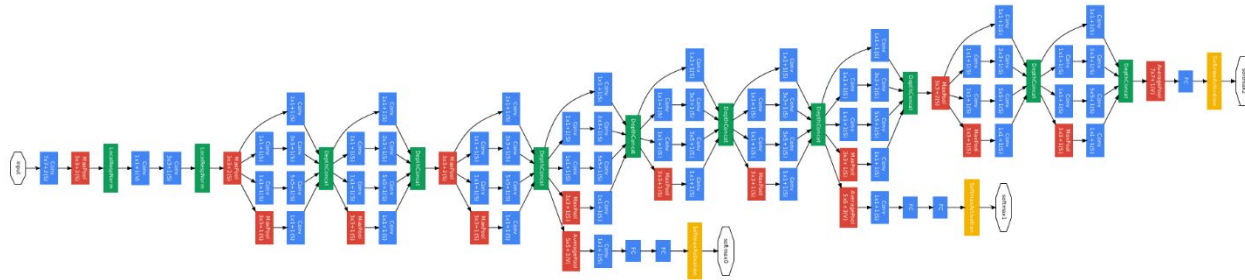
- 2012年 画像認識コンテストで、Deep Learningを用いたチームが優勝
 - 限界といわれた識別誤差を40%も減らした
 - それ以降、同コンテストの上位入賞者はほとんどDeep Learningを使用



ILSVRC2012優勝チームの結果

Deep Learning(2/2)

- 2012年以降、様々なコンペディションでDLを使用したチームが優勝
- 各企業がDL研究者の獲得競争
Google、FaceBook、Microsoft、Baidu
- 実サービスもDLベースに置き換えられている
Siri、Google等検索サービス



GoogLeNet



Google Brainによる
猫認識

まとめ

- 畳み込みニューラルネットワークの構造について学んだ。
- 畳み込み層とプーリング層について学んだ。
- 畳み込みニューラルネットワークに特有の学習アルゴリズムについて学んだ。
- 学習によって得られたフィルタ（結合荷重）の事例を紹介した。

復習問題

1. CNN特有の二種類の層を何というか？
2. フィルタ演算の結果として得られる出力を何というか？
3. 複数のユニットをまとめて学習する方法を何というか？
4. CNNは何型のニューラルネットワークか？

次回の講義

■ ニューラルネットまとめ

- ニューラルネットワークとパターン認識
- ニューラルネットワークの基礎
- 畳み込みニューラルネットワーク
- リカレントニューラルネットワーク