



画像AI-CNNがよくわかる資料

作成者 岡田 隆之

目次

- 本日の概要
- 講師について
- 画像データとは
- ニューラルネットワーク(ディープラーニング)について
- CNN詳細

概要

目的

画像認識技術の概要を知って貰う

開催日時

2022年12月21日 18:15～19:15

対象者

- 主に理工系学部で基本的な数学の基礎計算ができる方
- 画像等のAIを学ぶ意欲がある方

講義概要

- 画像認識技術について
 1. 理論概要紹介
 2. Tensorflowによる画像認識プログラミング

講師

岡田隆之(1990年4月生まれ)

講師について

自己紹介

- 名前: 岡田 隆之(たかゆき)
- 性別: 男
- 年齢: 30台
- 学歴: 東京大学大学院(修了)
- 職務経験: AI・データサイエンティストとして3年半。フリーランスで勤める。
画像AIについても、工業製品の異常検知や文字認識、
内観写真の分析経験あり
- 趣味: 検定受験、e-Sports、1人旅行
- 住所: 中野

今日の目次

CNNの仕組みについて

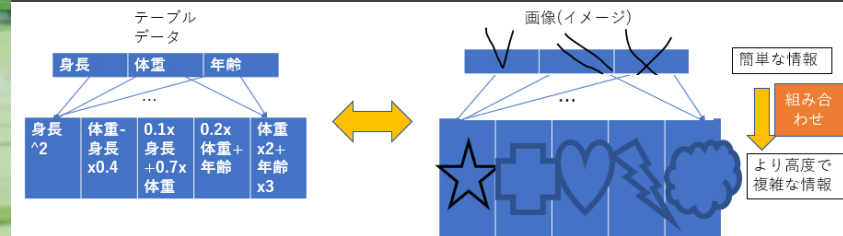
子記事：平均プーリング(Average Pooling)

12	20	33	130
8	12	20	120
7	42	155	202
45	168	142	183

カーネル内ごとに平均を出力
■カーネルサイズ = $[2 \times 2]$
■ストライド幅 = 2

20	75
65	170

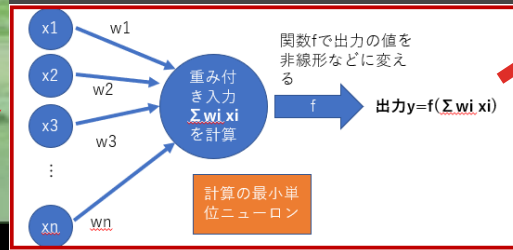
ニューラルネットワークの仕組みについて



画像データについて



ニューラルネットワーク & ディープラーニングとはなにか



画像データについて



画像データとは

- .pngや.jpgなどの拡張子のついたファイルのことである。例えば、'cat.png'は左下図。
- 各画像の最小単位=ピクセルについて、0,1,2,3,...255の整数値をもつ**数値データ**である。⇒欠損値処理やダミー変数化が全く必要ない
- 多くの場合numpyのarrayで、shape=(縦のピクセル数, 横のピクセル数, チャンネル数)
* チャンネル数=1(白黒のとき), 3(カラーのとき)となっている。

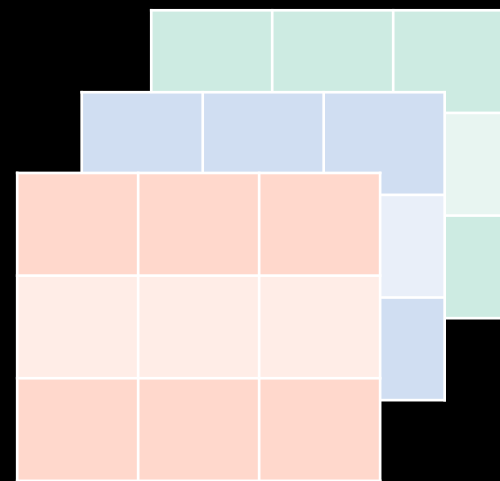
画像データイメージ



白黒画像

20	42	33
		12

カラー画像



画像の特有の前処理

- 画像データ特有の前処理がある。

1. 拡大・縮小・回転

画像の拡大・縮小・回転を行う。様々なパターンのデータを作成する際に行う。

2. ノイズ除去・先鋭化

ノイズ除去とは、カメラ不調などで現れる不要な黒点などを表し、それらを除去する。
先鋭化とは、画像をよりくっきり移すことである。

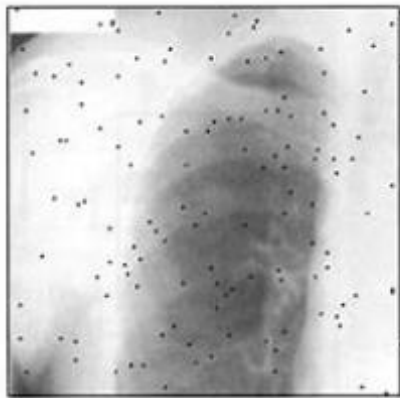
3. 画像認識精度向上に向けた加工(⇒コーディング編で説明)

ノイズ除去・先鋭化

- 画像認識精度向上のため、関係のない電磁波などによって生じた黒いシミを取り除いたり、色の差を強調して、境界線を鋭くすることなどができる。

ノイズ除去

ある領域で値の平均値をとる



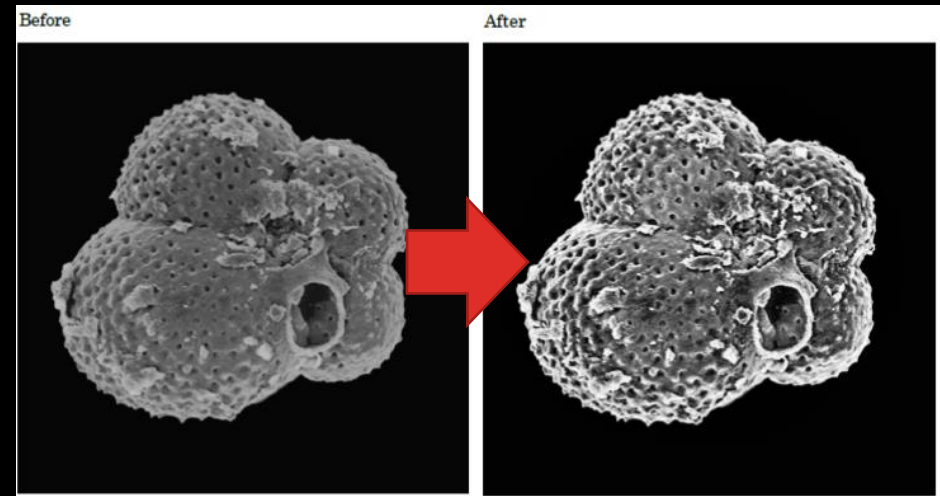
a: ノイズを含む画像



b: 処理後の画像

精鋭化

各値の差が大きくなる(N倍)ようにする

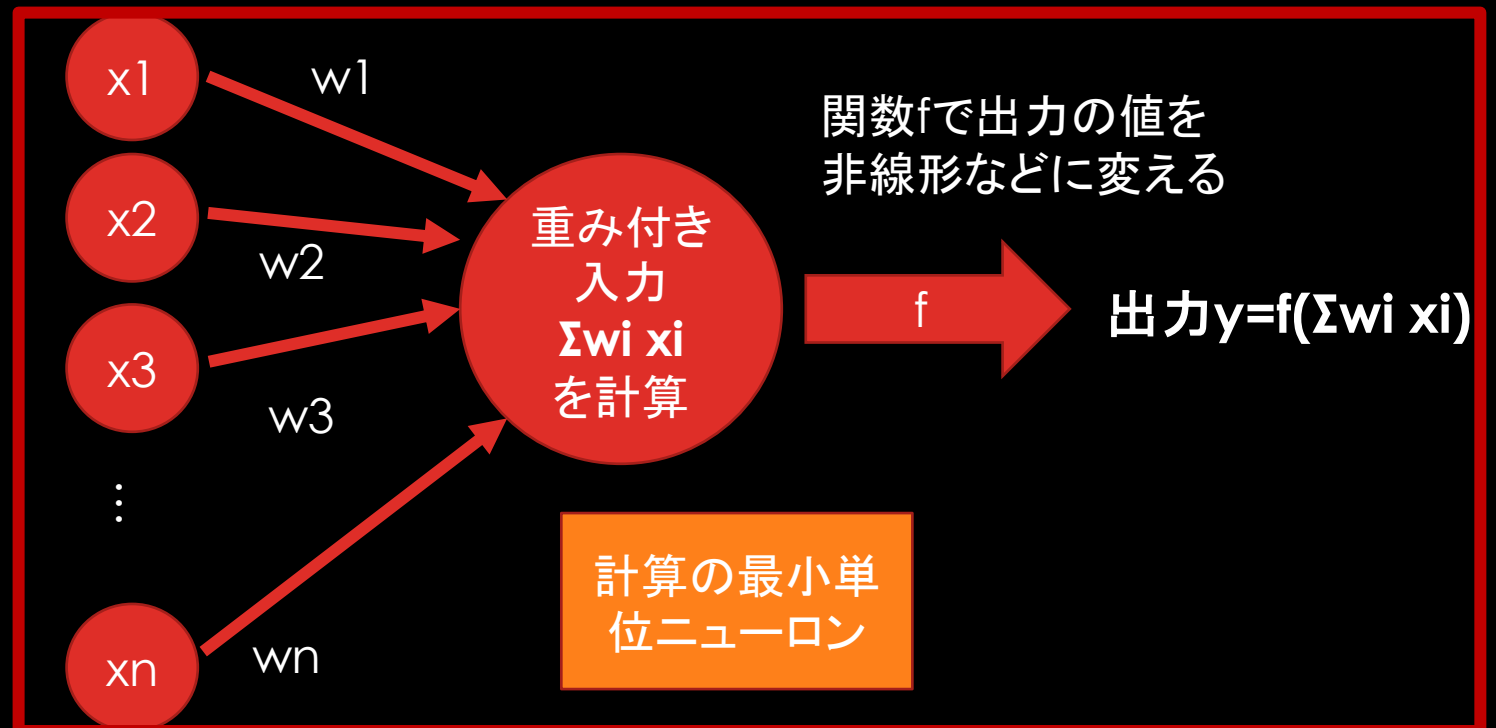
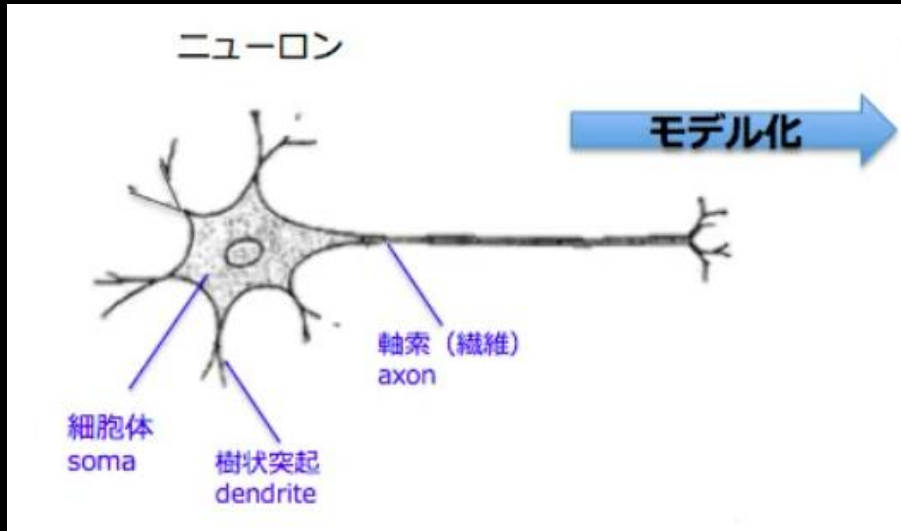


ニューラルネットワーク(ディープラーニング) について



ニューラルネットワークとは

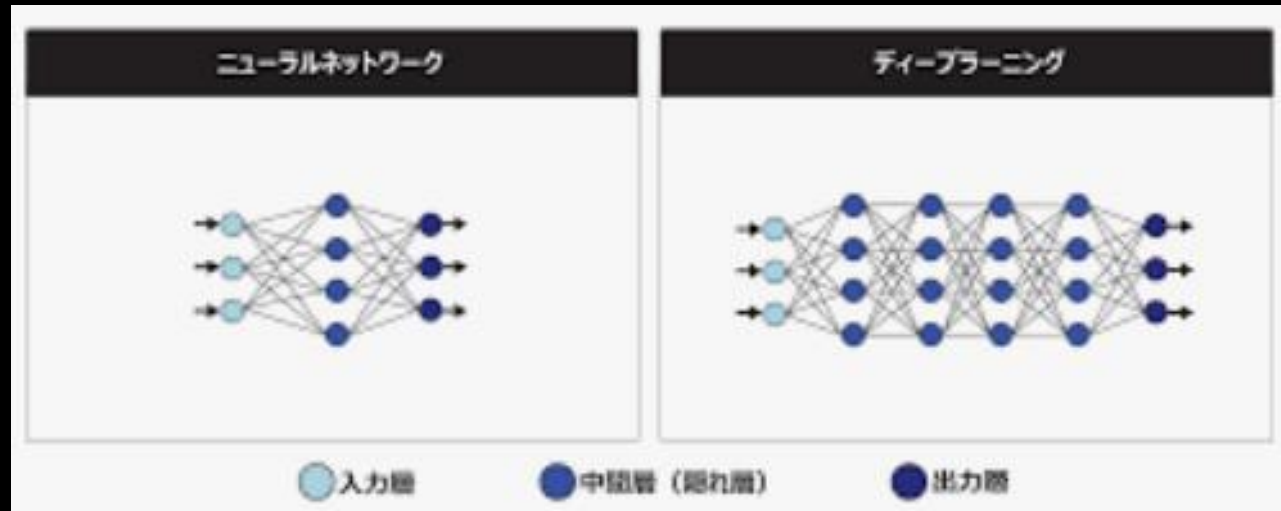
- ニューラルネットワークとは、人間の脳の神経伝達系を模倣してできた機械学習モデルの1つ。
- 神経伝達の単位はニューロンといわれる左下のような有機体。
- 樹状突起でほかのニューロンからの伝達を取り込み、刺激を組み合わせる軸索からほかのニューロンへ送り込む情報を出力する。
- ニューラルネットワークはこれを真似したもの(右下図)。



ディープラーニングとは

- ディープラーニングとは、**単に層の深い**ニューラルネットワークを使った"学習"のことをいう。
- このときどれくらい深ければディープラーニングというのかについては決まりはなく、計算リソースのなかった昔はたった4,5層でもディープと呼んでいた。
- 浅いニューラルネットワーク学習とは言わず、層が浅いものについても含めてディープラーニングといえよ(業界の感覚)。

左と右では、層が深くなるだけで使っている道具、仕組みは同じである点を理解する



ニューラルネットワークの種類

- ニューラルネットワークには大きく、次の3つの種類がある。

- ①p.8のニューロンをつなげ合わせただけの基礎のニューラルネットワーク(以降、**基礎ニューラルネットワーク**という)

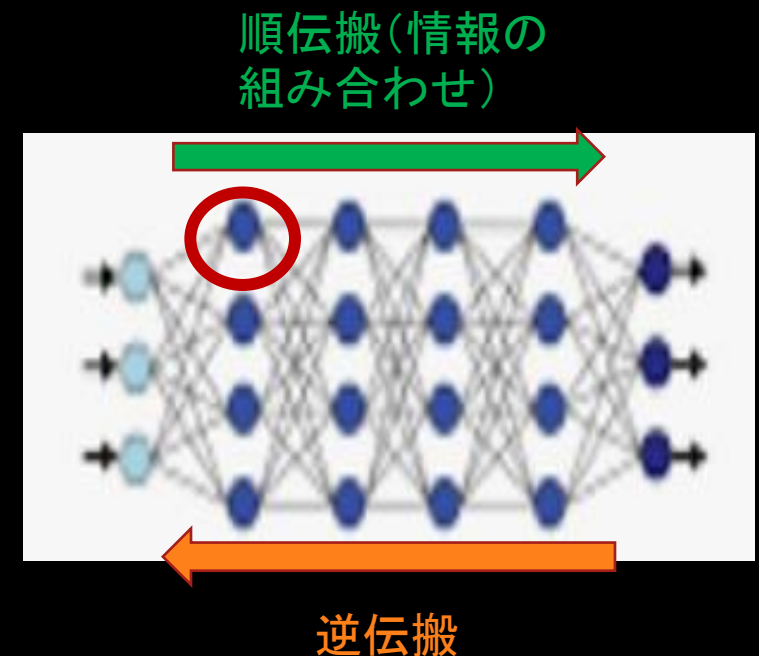
- ②画像用の特殊なニューロンを用意した**CNN**(畳み込みニューラルネットワーク)⇒後述する

- ③時系列データや自然言語処理など、入力の値に順番や方向性が存在する場合に適する**RNN**(再起型とニューラルネットワーク)

ニューラルネットワークについて 少しだけ深入りする

以下3点について、次のスライドから説明する。

- ニューラルネットワークはどのように自身の力で学習を行っているのか？⇒逆伝搬
- ニューラルネットワークが伝えている情報っていったい何だろう？⇒順伝搬(情報の組み合わせ)
- ニューラルネットワークの種類が違くと何が違うのか？
⇒**重みの並べ方と計算の仕方**
(丸の中身)



ニューラルネットワークに共通の値の更新(逆伝搬)

- CNNもニューラルネットワークの一種。機械が勝手に学習してくれるように誤差を出力側から伝えて値の更新を行っている(逆伝搬という)。

0.5	0.1	-0.2
-0.4	0.6	-0.4
0.7	0.1	0.3

逆伝搬

誤差関数の傾き(微分値)

例えば、誤差関数がMSE=2乗
平均誤差の場合、
 $-\partial W / \partial X = -\partial (y - a - bX)^2 / \partial X$
 $= -b(y - a - bX)$ が伝わるなど

ニューラルネットワークが伝えているもの(特徴量)

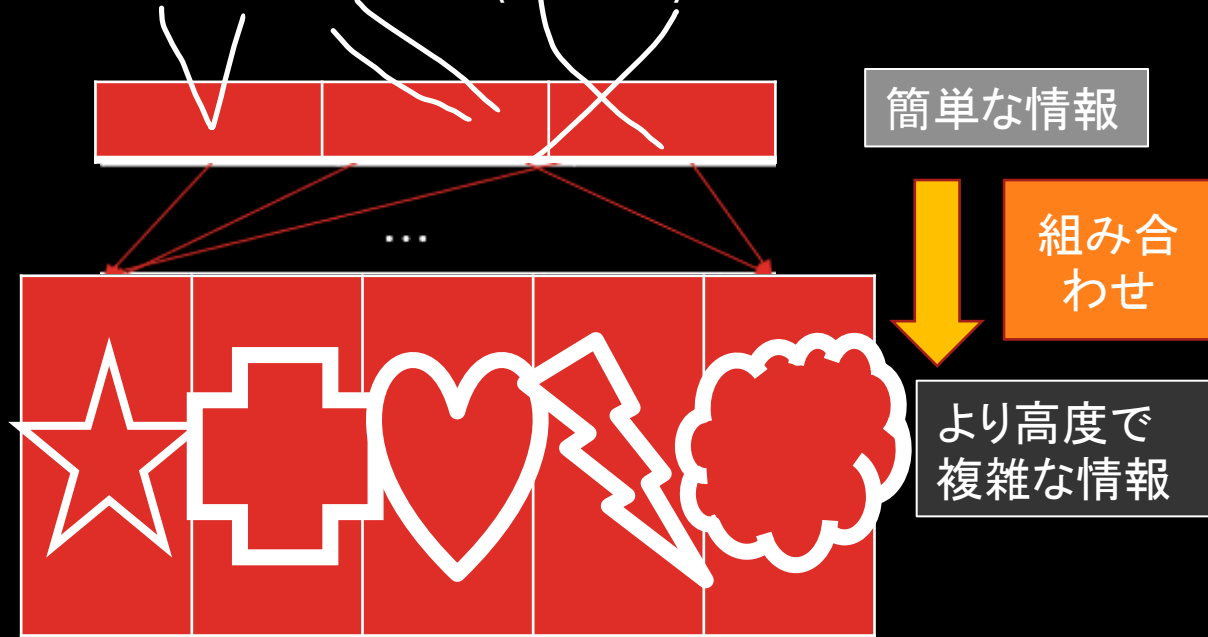
ニューラルネットワークは入力層の入力をどんどん組み合わせながら層を降りていくので、順々に複雑な組み合わせ・特徴量を考察する形になっている。

テーブルデータ

身長	体重	年齢
...		
身長 ^{^2}	体重-身長x0.4	0.1x身長+0.7x体重
	0.2x体重+年齢	体重x2+年齢x3

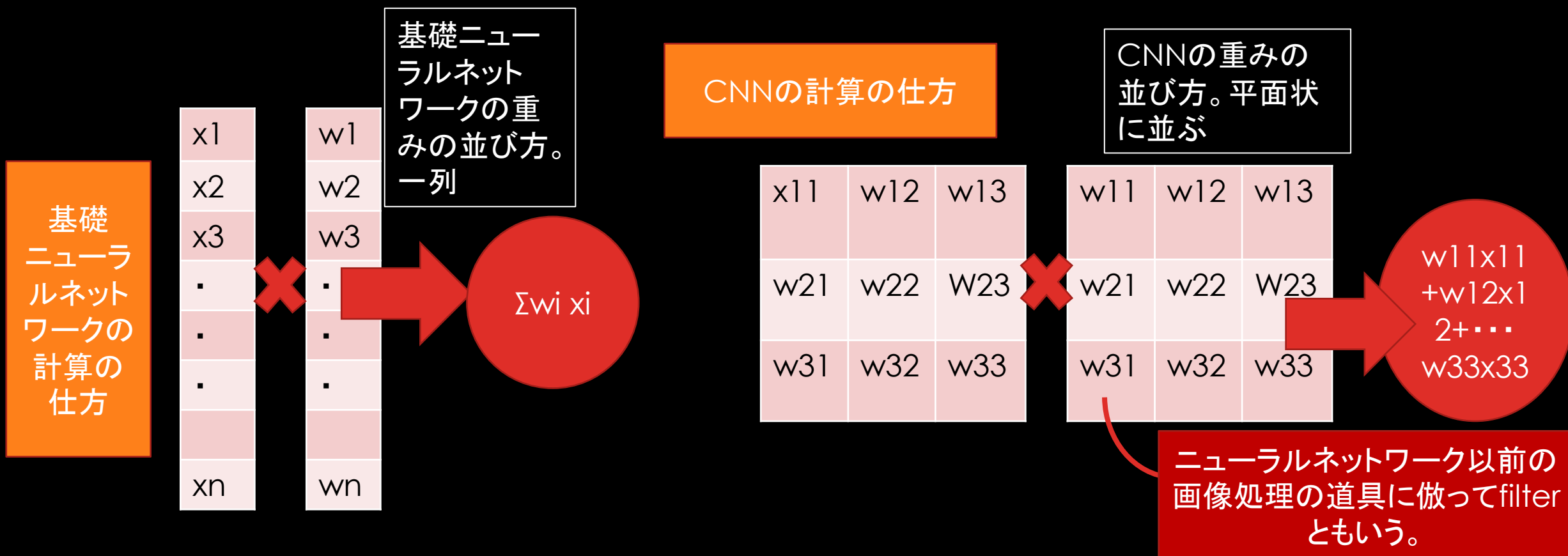


画像(イメージ)



基礎ニューラルネットワークとCNNとの違い

- 基礎ニューラルネットワークとCNN(、そしてRNN)とでは、伝達の最小単位のニューロンが違う！！
- 特に、ニューロンへの入力時の重みの並び方と計算の仕方が違う。



CNNについての詳細



CNNの順伝搬について詳しく見る

- CNNにも逆伝搬があることが分かった。飛ばしていた順伝搬(入力側から出力側への伝搬)の詳細について説明する。
- CNNの純伝搬に特有な、重要な役割を果たすのは次の2つ。

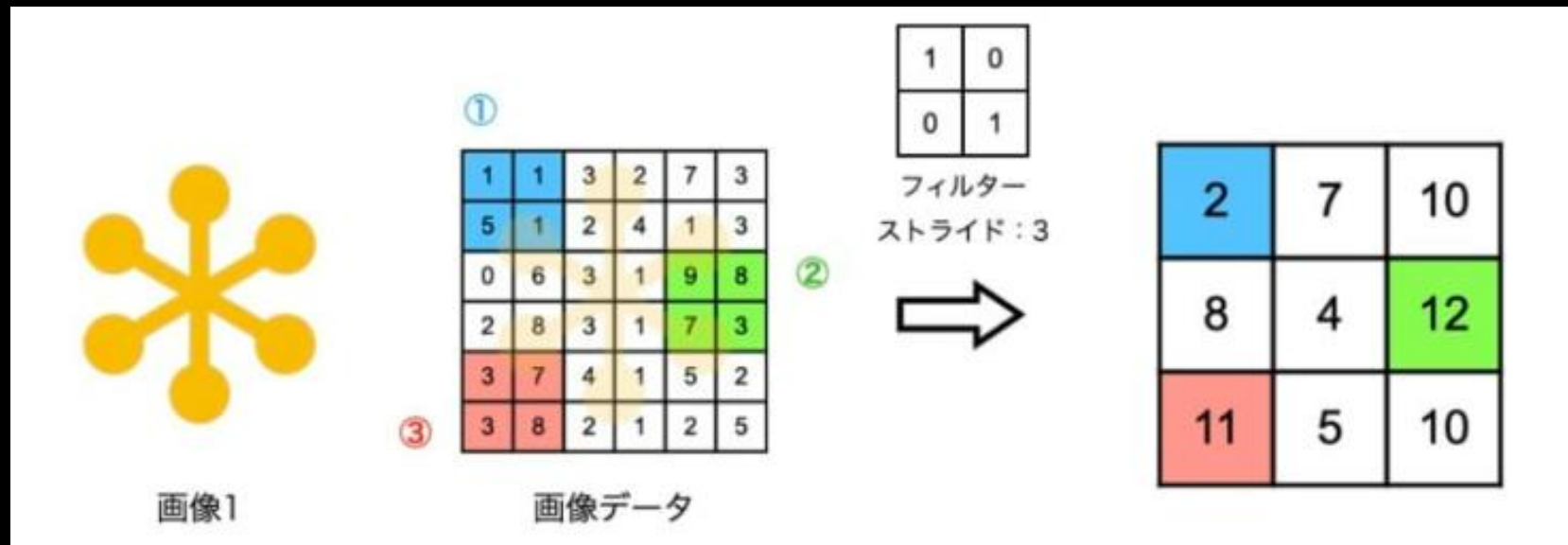
①畳み込み層(Convolution layer)

②プーリング層(Pooling layer)

順に紹介する

畳み込み層について

- 畳み込み層は情報を畳み込んで集約していくDLの層。
- 次のように計算する。



① $1 \times 1 + 1 \times 0 + 5 \times 0 + 1 \times 1 = 2$, ② $9 \times 1 + 8 \times 0 + 7 \times 0 + 3 \times 1 = 12$, ③ $3 \times 1 + 7 \times 0 + 3 \times 0 + 8 \times 1 = 11$

(練習問題)上の計算例は、filterを2つ飛ばしで移動させたときの計算である。
では、1つずつずらした場合はどのような結果になるか。計算せよ(⇒表を挿入する)。

参考資料: [CNN\(畳み込みネットワーク\)とは？図や事例を用いながら分かりやすく解説！](#)

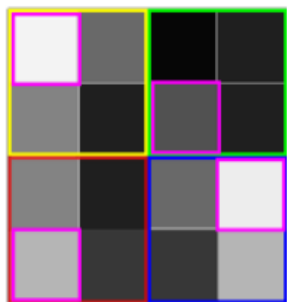
(練習問題)の答え

2	3	7	3	10
11	4	3	13	9
8	9	4	8	12
9	12	4	6	9
11	9	5	3	10

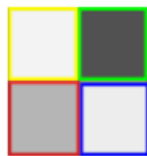
プーリング層について

- プーリング(Pooling)とは、ストレージ、ソフトウェアなどの物理資源を共有し管理するという意味
- CNNにおいては、主にMaxPooling, AveragePoolingの2つがあり、それぞれ各区画の最大値、平均値(だけ)を共有すること。
⇒一部の代表情報だけを共有する点に注意。
- 下のようにする。

子記事：最大値プーリング(Max Pooling)



カーネル内ごとに最大値を出力
■カーネルサイズ = $[2 \times 2]$
■ストライド幅 = 2



子記事：平均プーリング(Average Pooling)

12	20	33	130
8	12	20	120
7	42	155	202
45	168	142	183

カーネル内ごとに平均を出力
■カーネルサイズ = $[2 \times 2]$
■ストライド幅 = 2



20	75
65	170

まとめ

- 画像データは、主にnumpyのarrayですべて数値のだけのデータである。
- 画像の前処理には
 - ①拡大・縮小・回転
 - ②ノイズ除去・先鋭化
 - ③DL特有の変換といった変換がある。
- CNNはディープラーニングの1種であり、逆伝搬によって自己学習を行う。
- CNN特有の重要な要素として
 - ①畳み込み層(Convolutional layer)
 - ②プーリング層(Pooling layer)の2つがある。
- プーリング層には主に
 - 1. MaxPooling layer と 2. AveragePooling layer がある。