

## 第4章 発展的なニューラルネットワーク

### 4.1 畳み込みニューラルネットワーク

須賀勇貴

茨城大学理工学研究科

April 17, 2023

# はじめに

近年の深層学習の二枚看板

## 畳み込みニューラルネットワーク (CNN)

画像認識に特化したニューラルネットワーク

ex. 画像認識, 物体検出

## 再帰的ニューラルネットワーク (RNN)

時系列データの学習に特化したニューラルネットワーク

ex. 音声認識, 言語翻訳, 気象予報

まず, 畳み込みニューラルネットワーク (CNN) について説明する.

- 説明すること

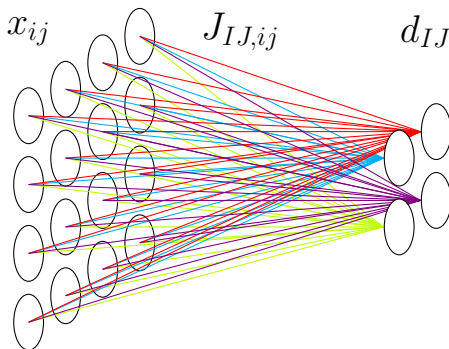
- ① 画像 (猫の画像) と特徴 (猫の写っている場所) が相関するモデルを考える
- ② 適切なハミルトニアンの構築
- ③ 畳み込み演算の導出

## 4.1.1 畳み込み

$i \approx I, j \approx J$  の辺りに猫が写っている確率 (4.1.1)

この場合の最も素朴なハミルトニアン

$$H_{J,x}(\{d_{I,J}\}) = - \sum_{IJ} \left( \sum_{ij} d_{IJ} J_{IJ,ij} x_{ij} + d_{IJ} J_{IJ} \right) \quad (4.1.2)$$



## 4.1.1 畳み込み

(4.1.1) を議論したい場合、全ての  $ij$  との相互作用を考えても無駄

$$J_{IJ,ij} = \begin{cases} \text{非ゼロ} & \begin{pmatrix} i = s_1 I + \alpha, \alpha \in [-W_1/2, W_1/2] \\ j = s_2 J + \beta, \beta \in [W_2/2, W_2/2] \end{pmatrix} \\ \text{ゼロ} & \text{その他} \end{cases} \quad (4.1.3)$$

$s_{1,2}$  : ストライド (stride) … 「何ピクセル飛ばし」で特徴を捉えるかのパラメータ

$W_{1,2}$  : フィルターサイズ … どのくらいの大きさの領域で特徴を捉えるかのパラメータ

これを実現できる結合定数

$$J_{IJ,ij} = \sum_{\alpha\beta} J_{IJ,\alpha\beta} \delta_{i,s_1 I + \alpha} \delta_{j,s_2 J + \beta} \quad (4.1.4)$$

## 4.1.1 畳み込み

$s_{1,2} = 3, W_{1,2} = 2$  のときを考える (右図)

$$J_{IJ,ij} = \sum_{\alpha, \beta=-1}^1 J_{IJ, \alpha\beta} \delta_{i, 3I+\alpha} \delta_{j, 3J+\beta}$$

$$H_{J,x}(\{d_{IJ}\})$$

$$= - \sum_{I,J=0}^1 \left( \sum_{i,j=0}^3 \sum_{\alpha, \beta=-1}^1 d_{IJ} J_{IJ, \alpha\beta} \delta_{i, 3I+\alpha} \delta_{j, 3J+\beta} x_{ij} + d_{IJ} \right)$$

$$= - \sum_{I,J=0}^1 \sum_{\alpha, \beta=-1}^1 d_{IJ} J_{IJ, \alpha\beta} x_{3I+\alpha, 3J+\beta} - \sum_{I,J=0}^1 d_{IJ} J_{IJ}$$

$$\begin{aligned} 1 \text{ 項目} &= d_{00} J_{00,00} x_{00} + d_{00} J_{00,01} x_{01} + d_{00} J_{00,10} x_{10} + d_{00} J_{00,11} x_{11} \\ &= d_{01} J_{01,02} x_{02} + d_{01} J_{01,03} x_{03} + d_{01} J_{01,12} x_{12} + d_{01} J_{01,13} x_{13} \\ &= d_{10} J_{10,20} x_{20} + d_{10} J_{10,21} x_{21} + d_{10} J_{10,30} x_{30} + d_{10} J_{10,31} x_{31} \\ &= d_{11} J_{11,22} x_{22} + d_{11} J_{11,23} x_{23} + d_{11} J_{11,32} x_{32} + d_{11} J_{11,33} x_{33} \end{aligned}$$

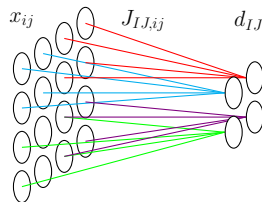


図: 結合定数の模式図

$$i, j = 0, 1, 2, 3$$

$$I, J = 0, 1$$

## 4.1.1 畳み込み

猫の写っている場所は真ん中だったり、右側だったりする

↓

捉えたい画像の特徴が「猫っぽさ」である場合、  
その特徴は  $I, J$  に依存しないはず

↓

結合定数に  $IJ$  依存性は入れないことにする

$$J_{IJ,ij} = \sum_{\alpha\beta} J_{\alpha\beta} \delta_{i,S_1 I + \alpha} \delta_{j,S_2 J + \beta} \quad (4.1.5)$$

ハミルトニアンは

$$\begin{aligned} H_{J,x}^{\text{conv}}(\{d_{IJ}\}) \\ = - \sum_{\{d_{IJ}\}} \exp \left[ \sum_{IJ} d_{IJ} \left( \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} + J \right) \right] \end{aligned} \quad (4.1.6)$$

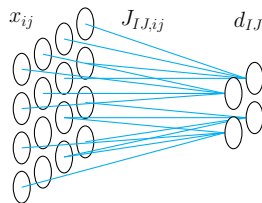


図: 結合定数の模式図

$$i, j = 0, 1, 2, 3$$

$$I, J = 0, 1$$

### 4.1.1 畳み込み

$$H_{J,x}^{\text{conv}}(\{d_{IJ}\}) = - \sum_{\{d_{IJ}\}} \exp \left[ \sum_{IJ} d_{IJ} \left( \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} + J \right) \right] \quad (4.1.6)$$

この状態のボルツマン重みを計算

$$\begin{aligned} Q_J(\{d_{IJ} = 1|x\}) &= \frac{\exp \left[ -H_{J,x}^{\text{conv}}(\{d_{IJ} = 1\}) \right]}{\sum_{\{d_{IJ}\}} \exp \left[ -H_{J,x}^{\text{conv}}(\{d_{IJ}\}) \right]} \\ &= \frac{\exp \left[ \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} + J \right]}{\sum_{\{d_{IJ}\}} \exp \left[ \sum_{IJ} d_{IJ} \left( \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} + J \right) \right]} \\ &= \sigma \left( \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} + J \right) \end{aligned} \quad (4.1.7)$$

ここで,

$$\{d_{IJ}\} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

## 4.1.1 畳み込み

$$Q_J(\{d_{IJ} = 1|x\}) = \sigma \left( \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} + J \right) \quad (4.1.7)$$

このときの演算

$$x_{ij} \rightarrow \sum_{\alpha\beta} J_{\alpha\beta} x_{s_1 I + \alpha, s_2 J + \beta} \quad (4.1.8)$$

を畳み込み演算という

畳み込み演算を含む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

現在、画像認識において畳み込みニューラルネットワークは常識となっている



## 4.1.2 転置畳み込み