

# 深層学習による動画像からの表情認識手法の開発

7 月 17 日 (水)

小松 大起

## 1 CNN とは

- CNN とは多層 NN の一つで、畳込み NN と呼ばれる。ネットワーク上に畳込み層とプーリング層の 2 層を交互に持つネットワークである。  
対象とする画像を  $W \text{ pixel} \times W \text{ pixel}$  の RGB の階調値とし、 $y_{ij}^{(k)}$  と表す。  
ここで  $k$  は RGB の色調を表し  
 $k = 1, 2, 3$  ( $N = 3$ ) の値をとる。  
畳込み層  $(i, j)$  の素子の内部状態  $x_{ij}$  は以下で与えられる。

$$y_{ij}^{(a)} = \sum_{k=1}^N \left( \sum_{p=1}^w \sum_{q=1}^w x_{i+p, j+q}^{(k)} h_{ij}^{(k)(a)} + b^{(k)(a)} \right) \quad (1)$$

ここで  $a$  はフィルタの数を示し、 $h_{ijk}^{(k)(a)}$  は受容野における各素子の係数であり CNN は学習によってこの  $h_{ijk}^{(k)(a)}$  を変化させて

$$\tilde{y}_{ij}^{(a)} = f(y_{ij}^{(a)}) \quad (2)$$

$$f(x) = \max(x, 0) \quad (3)$$

であるが多クラス分類のために出力層には活性化関数  $g(x)$  を用いる。  
入力  $y_j$  に対して、最終的な出力  $z_j$  は以下で与えられる。

$$z_i = \frac{e^{\tilde{y}_i}}{\sum_{k=1}^n e^{\tilde{y}_k}} \quad (4)$$

## 2 先週までの作業

- 地方会の予稿の作成。

## 3 今週の作業

- 地方会の予稿の作成。

## 4 来週以降の作業

- 地方会での発表スライドの作成。
- CNN についての勉強。