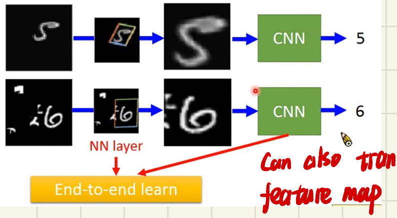
Spatial transformer 空间变换

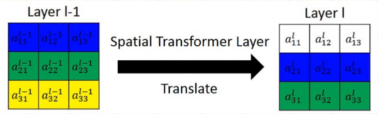
Spatial transformer layer

- CNN is not invariant to scaling and rotation. 缩放和旋转



Can also transform feature map

- How to transform an image / feature map



Layer l-1

a_{11}^{l-1}	a_{12}^{l-1}	a_{13}^{l-1}
a_{21}^{l-1}	a_{22}^{l-1}	a_{23}^{l-1}
a_{31}^{l-1}	a_{32}^{l-1}	a_{33}^{l-1}

Spatial Transformer Layer

Translate

Layer l

a_{11}^{l}	a_{12}^{l}	a_{13}^{l}
a_{21}^{l}	a_{22}^{l}	a_{23}^{l}
a_{31}^{l}	a_{32}^{l}	a_{33}^{l}

General layer : $a_{nm}^{l} = \sum_{i=1}^{3}\sum_{j=1}^{3} \underline{w_{nm,ij}^{l}}\, a_{ij}^{l-1}$

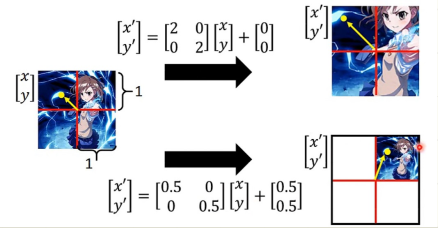
所以对图片进行伸缩变换就要改变 w 的值

eg: 如图所示平移的变换: 即 $a_{nm}^{l} = a_{(n-1)m}^{l-1}$

$$\begin{cases} w_{nm,ij}^{l} = 1, \text{时 } i = n-1,\ j = m \\ w_{nm,ij}^{l} = 0, \text{ otherwise} \end{cases}$$

- Image Transformation

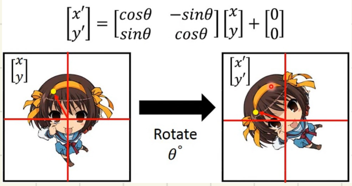
① Expansion Compression Translation
压缩



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

② Rotation (旋转)

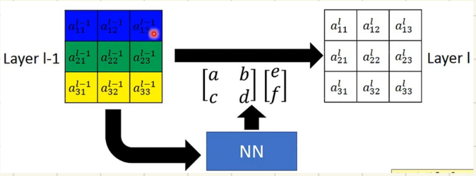


$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

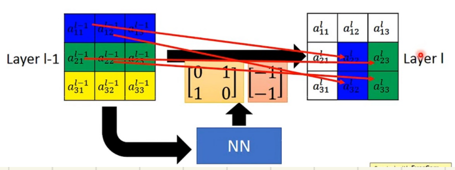
总结:

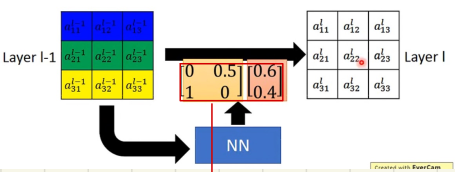
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \underline{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}\begin{bmatrix} x \\ y \end{bmatrix} + \underline{\begin{bmatrix} e \\ f \end{bmatrix}}$$

index of layer l-1　6个未知参数.　index of layer l



eg:

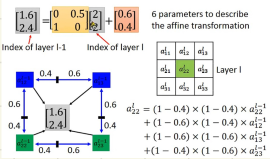




不是整数. 当对NN的参数做小小的改变时, x,y' 没有太大变化, 因为 gradient = 0

解决 gradient = 0 的问题

Interpolation (插值)


$$\begin{bmatrix} 1.6 \\ 2.4 \end{bmatrix} = \begin{bmatrix} 0 & 0.5 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$$

6 parameters to describe the affine transformation

Index of layer l-1

Index of layer l

$$\begin{aligned} a^{l}_{22} &= (1-0.4) \times (1-0.4) \times a^{l-1}_{22} \\ &+ (1-0.6) \times (1-0.4) \times a^{l-1}_{12} \\ &+ (1-0.6) \times (1-0.6) \times a^{l-1}_{13} \\ &+ (1-0.4) \times (1-0.6) \times a^{l-1}_{23} \end{aligned}$$

所以当NN参数变化时, Output 有明显变化, 可以用梯度下降.