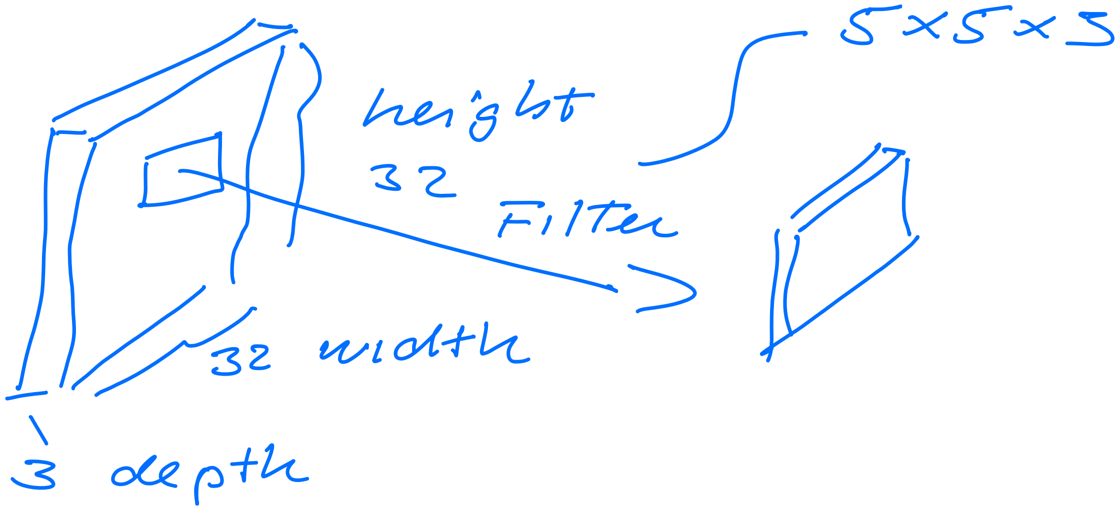
FYS-STK 4155, OCT 27, 2022

CNN

- architecture
  - conv size (filter), stride, padding, depth
  - ReLU & other non-linearities
  - Pooling methods
  - # layers
  - Flattening layer.
  - Fully connected layers
- cost function & optimization
  - type of cost function
  - Regularization
  - Gradient descent
  - SGD batches & epochs

CNN

32 x 32 x 3 image



conv-stage { parameter, Backpropagation

=> Relu stage

=> pooling stage

=

conv filter	Relu	pooling

→

Convolutional stage 1

→

conv	Relu	Pooling

Conv stage 2

$\rightarrow$ . . . $\boxed{\text{output}}$ $\rightarrow$ Flattening

$\rightarrow$ conventional NN

$\rightarrow$ Final output y

$\rightarrow C(y, t, \epsilon)$

Convolution stage

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{bmatrix}$$

input 3x4

stride = 1

$$\begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

Filter 2x2

$$\begin{bmatrix} aw + bx + ey + fz & bw + cx + fy + gz & cw + dx + gy + hz \end{bmatrix}$$

$ew+fx+iy+yz$	$fw+gx+fy+kz$	$gw+hx+ky+lz$

2×3

- accept a volume of

  $W_1 \times H_1 \times D_i$ $(3\times4)\times1$

  width height depth

- Four hyperparameters
  - K = number of filters
  - F = the spatial extent (2)
  - S = stride (=1)
  - The amount of zero padding P (=0)

- produces a volume of size

  $W_2 \times H_2 \times D_2$

$$W_2 = (W_1 - F + 2P)/S + 1$$

$$H_2 = (H_1 - F + 2P)/S + 1$$

$$D_2 = K$$

Example $W_1 \times H_1 \times D_1 =$

$3 \times 4 \times 1$

$S = 1 \quad P = 0$

$F = 2$

$K = 1$

in total $W_2 \times H_2 \times D_2 =$

$$\underline{2 \times 3 \times 1}$$

parameters:

$(F \cdot F \cdot D_1) \times K$ = weights

and K - biases.

in our example: $2 \times 2 + \underset{bias}{1}$

$= 5$

Typical hyperparameters:

$F = 3$, $S = 1$, $P = 1$

$F = 5$, $S = 1$, $P = 2$

$F = 5$, $S = 2$, $P = ?$

Example

input $32 \times 32 \times 3$

10 Filters ($K = 10$) 5×5 Filter with $S = 1$ and $P = 0$

output

$$(32 - 5)/1 + 1 = 28$$

$$\boxed{28 \times 28 \times 10}$$

include 3 color channels

we have $5 \times 5 \times 3 + 1 =$

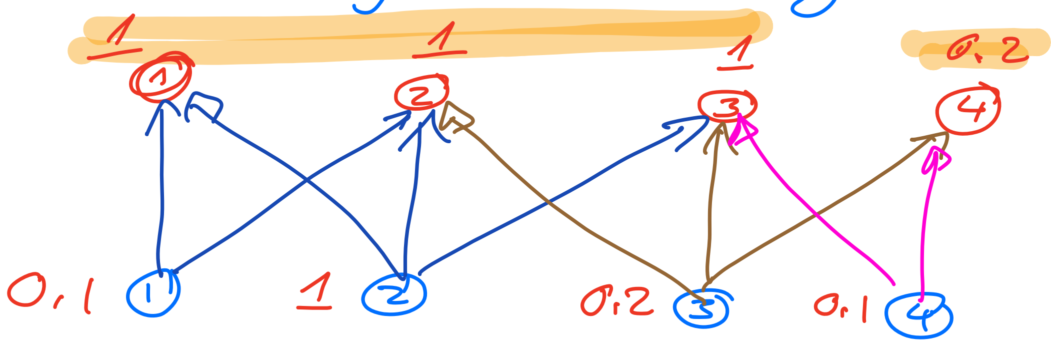
bias

76 parameter.

in total $76 \times 10 = 760$ parameters.

Pooling stage

Max pooling: reports the max output within a rectangular neighborhood



Detection stage (RELU)

pooling with downsampling shrink to less nodes when several neighboring nodes have same value.

$$\begin{bmatrix} i_{00} & i_{01} & i_{02} \\ i_{10} & i_{11} & i_{12} \\ i_{20} & i_{21} & i_{22} \end{bmatrix} * \begin{bmatrix} W_{00} & W_{01} \\ W_{10} & W_{11} \end{bmatrix}$$

$$
\begin{bmatrix}
i_{00} w_{00} + i_{01} w_{01} + i_{10} w_{10} + i_{11} w_{11} & i_{01} w_{00} + i_{02} w_{01} + i_{11} w_{10} + i_{12} w_{11} \\
i_{10} w_{00} + i_{11} w_{01} + i_{20} w_{10} + w_{11} i_{21} & i_{11} w_{00} + i_{12} w_{01} + i_{21} w_{10} + i_{22} w_{11}
\end{bmatrix}
$$

$$
I' \in \mathbb{R}^9
$$

$$
I' = [i_{00}\ i_{01}\ i_{02}\ \ldots\ i_{21}\ i_{22}]^T
$$

$$
W' = \begin{bmatrix}
w_{00} & w_{01} & 0 & w_{10} & w_{11} & 0 & 0 & 0 & 0 \\
0 & w_{00} & w_{01} & 0 & w_{10} & w_{11} & 0 & 0 & 0 \\
0 & 0 & 0 & w_{00} & w_{01} & 0 & 0 & w_{10} & w_{11} \\
0 & 0 & 0 & 0 & w_{00} & w_{01} & 0 & w_{10} & w_{11}
\end{bmatrix}
$$

$$
\text{output} = W' I'
$$

$$
W' \in \mathbb{R}^{4 \times 9}
$$