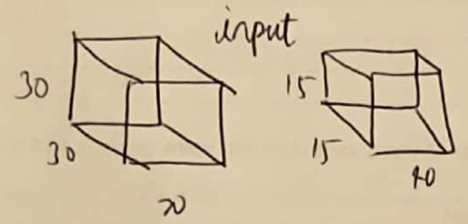
SML HW4

Goureesh Gunupati
00 2641 248.

(1)



from given input and output we could tell the variables that are involved

filter size = 3×3

Input size = $30 \times 30 \times 20$

Input channels = 20

Output size = $15 \times 15 \times 40$

output channels = 40

Comparing convolution and depthwise seperable convolution.

Convolutional Network -

In a convolutional network number of weights are given by product of filter size, input and output channels

No of weights = $3 \times 3 \times 20 \times 40 = 7200$.

number of computations is given by product of filter size, output height and width, out and input channels.

no of computations = $3 \times 3 \times 15 \times 15 \times 20 \times 40 = 1{,}620{,}000$.

Depthwise seperable Network -

This convolutional happens in two stage and the output is resultant of both operations.

(i) Depth wise convolution -

in this method we apply a . seperate filter to input and it on to the next layer. Number of weights is a product of filtersize and input channel

Number of weight $3 \times 3 \times 20 = 180$.

No. of computations is a product of output height, output width, filter size and input channel

$= 15 \times 15 \times 3 \times 3 \times 20 = 40,500$.

(ii) Point wise Convolution -

In this method, we apply do ~~seperate~~ output ~~filter to~~ input ~~and~~ layer.

here number of weights is product of input channels and output channels

No. of weights $= 20 \times 40 = 800$.

number of computations is product of output height, width, input channel and output channel

No. of computations $= 15 \times 15 \times 20 \times 40 = 180,000$.

Total weights $= 180 + 800 = 980$

Total computations $= 40,500 + 1,80,000 = 220,500$.

on observation. we can say that depth seperable convultion takes less weights (980) and less number of computations (220500) when compared to conventional convolutional network which weights are 7200 and computations are (1,620,000) and hence depthe seperable convultion will perform better than 3x3 convolutional network.

(2) Input volume : $W_1 \times H_1 \times D_1 = 20 \times 20 \times 10$

Output Volume $= W_2 \times H_2 \times D_2 = 20 \times 20 \times 10$

$K \rightarrow$ number of filters

$F \times F \rightarrow$ filter size

$S \rightarrow$ Stride

$P \rightarrow$ Amount of zero padding

$$W_2 = \frac{W_1 - F + 2P}{S} + 1 \quad -(1)$$

$$H_2 = \frac{H_1 - F + 2P}{S} + 1 \quad -(2)$$

$$D_2 = K \quad -(3)$$

(a) filter 1×1

from formula (1) and (2)

$$20 = \frac{(20 - 1 + 2P)}{S} + 1$$

assume it is a zero padding layer

$$19 = 19/S$$

$$S = 1$$

$P = 0$, $S = 1$ and $K = 10$.

zero padding $= 0$ # of channels $= 10$

Stride $= 1$

Scanned with OKEN Scanner

b) filter 3×3

from equation ① and ② we get

$$20 = \frac{(20-3+2P)}{S} + 1$$

here the padding is 1 to ensure that the spatial dimensions of the future maps are maintained

$$20 = \frac{20-3+2}{S} + 1$$

$S = 1$

$\therefore$ $P = 1$, $S = 1$ and $K = 10$.

zero padding = 1

Stride = 1

channels = 10 = k

c) filter 1×1

Similar to 1st question.

$$\Rightarrow \quad 20 = \frac{(20-1+2P)}{S} + 1$$

$$19 = \frac{19+2P}{S} \qquad P=0,\ S=1$$

$\Rightarrow$ $19 = \frac{19}{S}$ $\Rightarrow$ stride = 1

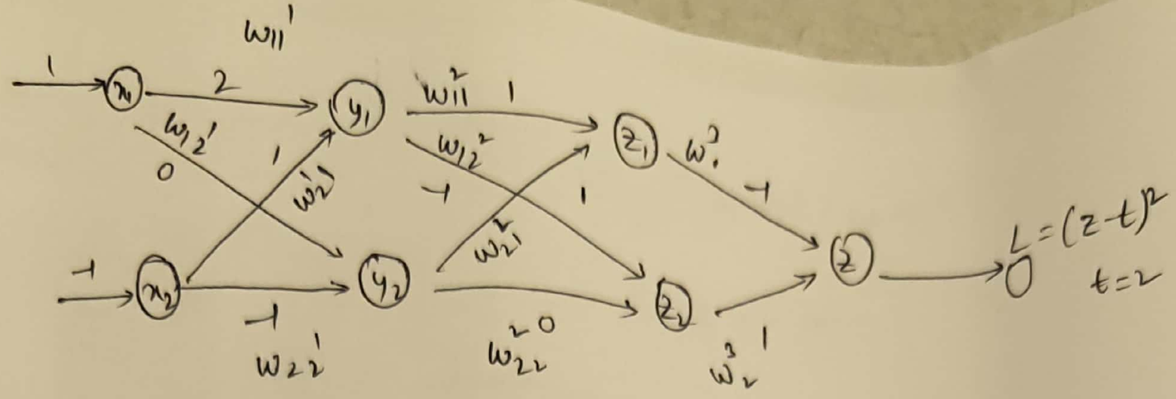
padding = 0

Stride = 1

no of channels = 10.

Scanned with OKEN Scanner

3.



(a) Forward Propagation.

Given $x_1 = 1$ and $x_2 = -1$

$y_1 = x_1 w_{11}^1 + x_2 w_{21}^1 = 1 \times 2 + (-1) \times (1) = 1$

$y_2 = x_1 w_{12}^1 + x_2 w_{22}^1 = 1 \times 0 - 1 \times -1 = 1$

$z_1 = y_1 w_{11}^2 + y_2 w_{21}^2 = 1 \times 1 + 1 \times 1 = 2$

$z_2 = y_1 w_{12}^2 + y_2 w_{22}^2 = 1 \times -1 + 1 \times 0 = -1$

$z = z_1 w_1^3 + z_2 w_2^3 = 2 \times -1 + (-1) \times 1 = -3.$

loss $= (z-t)^2 = (-3-2)^2 = 25.$

(b)

$$\frac{\partial L}{\partial z} = \frac{\partial}{\partial z}(z-t)^2 = 2(z-t) = -10.$$

$$\frac{\partial L}{\partial z_1} = \frac{\partial L}{\partial z} \times \frac{\partial z}{\partial z_1} = -10 \cdot \frac{\partial}{\partial z_1}(z_1 w_1^3 + z_2 w_2^3)$$

$$= -10 \times w_1^3 = 10.$$

$$\frac{\partial L}{\partial z_2} = \frac{\partial L}{\partial z} \cdot \times \frac{\partial z}{\partial z_2} = -10 \cdot \frac{\partial}{\partial z_2}(z_1 w_1^3 + z_2 w_2^3)$$

$$= -10 \times w_2^3 = -10 \times 1 = -10$$

$$\frac{\partial L}{\partial w_1^3} = \frac{\partial L}{\partial z} \times \frac{\partial z}{\partial w_1^3} = 0 - 10 \times \frac{\partial}{\partial w_1^3} \left(z_1 w_1^3 + z_2 w_2^3 \right)$$

$$= -10 \times z_1 = -20$$

$$\frac{\partial L}{\partial w_2^3} = \frac{\partial L}{\partial z} \times \frac{\partial z}{\partial w_2^3} = -10 \times \frac{\partial}{\partial w_2^3} \left(z_1 w_1^3 + z_2 w_2^3 \right)$$

$$= -10 \times z_2 = 10$$

$$\frac{\partial L}{\partial w_{11}^2} = \frac{\partial L}{\partial z_1} \times \frac{\partial z_1}{\partial w_{11}^2} = 10 \times \frac{\partial}{\partial w_{11}^2} \left(y_1 w_{11}^2 + y_2 w_{21}^2 \right)$$

$$= 10 \times y_1 = 10$$

$$\frac{\partial L}{\partial w_{12}^2} = \frac{\partial L}{\partial z_2} \times \frac{\partial z_2}{\partial w_{12}^2} = -10 \times \frac{\partial}{\partial w_{12}^2} \left(y_1 w_{12}^2 + y_2 w_{22}^2 \right)$$

$$= -10 \times y_1 = -10$$

$$\frac{\partial L}{\partial w_{21}^2} = \frac{\partial L}{\partial z_1} \times \frac{\partial z_1}{\partial w_{21}^2} = 10 \times \frac{\partial}{\partial w_{21}^2} \left(y_1 w_{11}^2 + y_2 w_{21}^2 \right)$$

$$= 10 \times y_2 = 10$$

$$\frac{\partial L}{\partial w_{22}^2} = \frac{\partial L}{\partial z_2} \times \frac{\partial z_2}{\partial w_{22}^2} = -10 \times \frac{\partial}{\partial w_{22}^2} \left(y_1 w_{12}^2 + y_2 w_{22}^2 \right)$$

$$= -10 \times y_2 = -10$$

$$\frac{\partial L}{\partial y_1} = \frac{\partial L}{\partial z_1} \times \frac{\partial z_1}{\partial y_1} \times \frac{\partial L}{\partial z_2} \times \frac{\partial z_2}{\partial y_1} = 10 \times \frac{\partial}{\partial y_1} \left(y_1 w_{11}^2 + y_2 w_{21}^2 \right)$$

$$+ -10 \frac{\partial}{\partial y_2} \left(y_1 w_{12}^2 + y_2 w_{22}^3 \right)$$

$$= 10 w_{11}^2 + 10 w_{21}^2$$

$$= 20.$$

$$\frac{\partial L}{\partial y_2} = \frac{\partial L}{\partial z_1} \times \frac{\partial z_1}{\partial y_2} \times \frac{\partial L}{\partial z_2} \times \frac{\partial z_2}{\partial y_2}$$

$$= 10 \times \frac{\partial}{\partial y_2}(y_1 w_{11}^2 + y_2 w_{21}^2) - 10 \frac{\partial}{\partial y_1}(y_1 w_{12}^2 + y_2 w_{21}^3)$$

$$= 10 \times \frac{\partial}{\partial y_1}(w_{21}^2) - 10 w_{22}^2 = 10$$

$$\frac{\partial L}{\partial w_{11}'} = \frac{\partial L}{\partial y_1} \times \frac{\partial y_1}{\partial w_{11}'} = 20 x_1 = 20 \times 1 = 20$$

$$\frac{\partial L}{\partial w_{12}'} = \frac{\partial L}{\partial y_2} \times \frac{\partial y_2}{\partial w_{12}'} = 10 \times \frac{\partial}{\partial w_{12}}(x_2 w_{21}' * x_1 w_{12}')$$

$$= 10 \times x_1 = 10$$

$$\frac{\partial L}{\partial w_{21}'} = \frac{\partial L}{\partial y_1} \times \frac{\partial y_2}{\partial w_{22}'} = 20 \times (-1) = -20$$

$$\frac{\partial L}{\partial w_{22}'} = \frac{\partial L}{\partial y_2} \times \frac{\partial y_2}{\partial w_{22}'} = 10 \times \frac{\partial}{\partial w_{22}}(x_2 w_{22}' + x_1 w_{12}')$$

$$= 10 \times x_2 = -10$$

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial y_1} \times \frac{\partial y_1}{\partial x_1} * \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} = 20 \times 2 + 10 \times 0$$

$$= 40$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial y_2} \times \frac{\partial y_2}{\partial x_2} + \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_2} = 20 - 10 = 10.$$

Scanned with OKEN Scanner