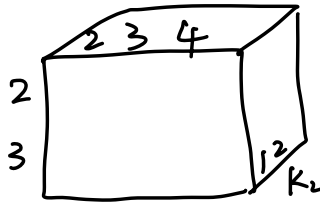
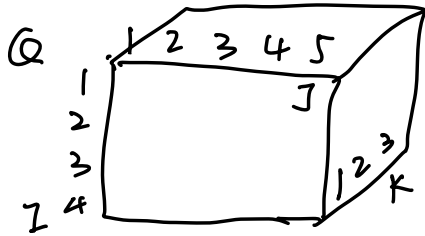


22-52-Q4



(a) fully connect?

Solution

$$y_{ijk} = \sum_{i'=1}^I \sum_{j'=1}^J \sum_{k'=1}^K w_{ij'k'}^{i'j'k'} x_{i'j'k'} + b_{ijk} \quad \begin{array}{l} i = 2 \text{ to } I-1 \\ j = 2 \text{ to } J-1 \\ k = 1 \text{ to } k_2 \end{array}$$

$$N_{\text{params}} = [(I-2)(J-2)k_2] \times (I \times J \times k_1) + (J-2)(I-2) \times k$$

$$N_{\text{multiplications}} = (I-2)(J-2)k_2 \times I \times J \times k_1$$

(b) Q: CNN $3 \times 3 \times k_1$

Solution

$$y_{ijk} = \sum_{m=1}^3 \sum_{n=1}^3 \sum_{k'=1}^{k_1} h_{mnk'} x_{(i+m-2)(j+n-2)k'} + b_{ijk}$$

$$N_{\text{params}} = (3 \times 3 \times k_1) \times k_2 + k_2$$

$$N_{\text{mult}} = (I-2)(J-2)k_2 \times 3 \times 3 \times k_1$$

(c) 2 differences?

Solution

	fully connected	CNN
① parameters	massive \sim	significantly fewer \sim
② training time	longer	fast
③ memory requirement	higher	lower
④ multiplication	a large number	fewer
⑤ capturing local pattern	less effective	effectively