22-S2-Q4

(a)Q: $x_{ijk} \rightarrow y_{ijk}$? parameters?

Solution ① understand XIA SONG assistant

A fully connected layer generally squeezes
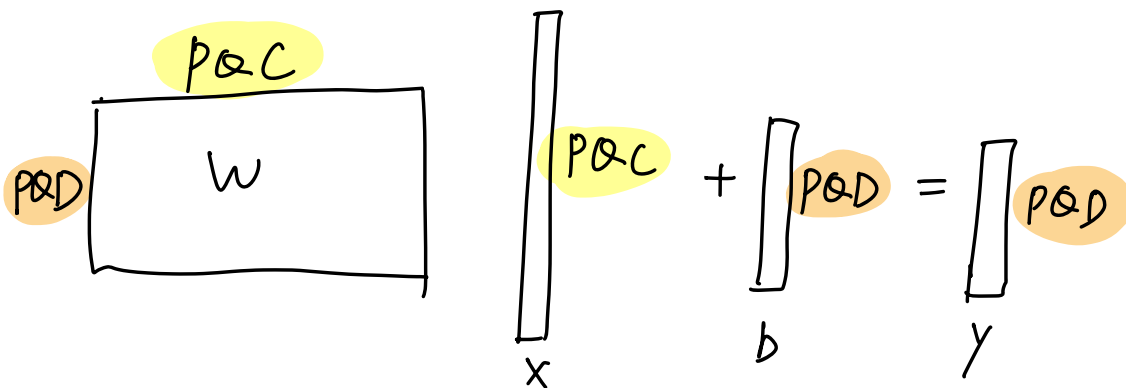
[any] input into a [one-dimensional vector]

So, an image with $P \times Q \times D$ is turned into

a $1 \times PQD$ vector, Thus I think the

solution PQC PQD + PQD is correct

② 逐行扫描

③ W矩阵 · X列向量

So parameters = $PQD \times PQC + PQD$

④ scalar form

$$y_{ijk} = \sum_{l=1}^{C} \sum_{m=1}^{P} \sum_{n=1}^{Q} w_{ijkmnl} \, X_{mnl} + b_{ijk} \quad , \quad 1 \le k \le D$$

有 $PQC \times PQD$ 个   有 PQD 个

(b) CNN ?

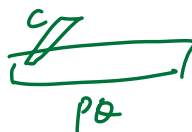$$y_{i,j,k} = \sum_{u=-1}^{1} \sum_{v=-1}^{1} \sum_{l=1}^{C} w_{u,v,l,k} \, X_{i-u, \, j-v, \, l} + b_k$$

$$3 \times 3 \times C \times D + D = 9 \times C \times D + D$$

(c)
$$y_{ijk} = \sum_{l=1}^{C} w_{l,k} \, X_{i,j,l} + b_k$$

$$C \times D + D \text{ bias}$$

RGB   D个特征图

(d)
$$y_{i,k} = \sum_{l=1}^{C} w_{l,k} \, X_{i,l}$$

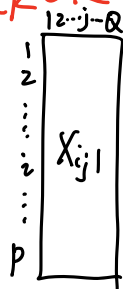$$C \times D$$

(e) $Y = XW$

① Both architectures apply matrix multiplications

to process input

② $Y = XW$ shows the CNN perform a linear transformation on the input features at each spatial position, identical across all position.

It is similar to the linear layers used in Transformer, where input are transformed via weight matrices
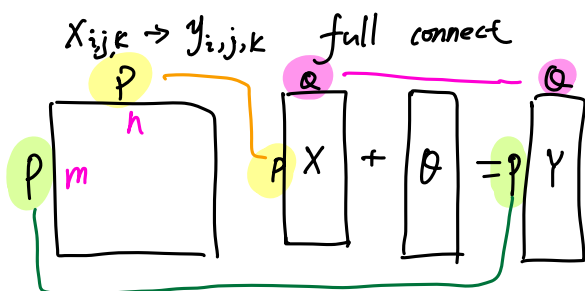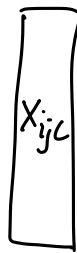
ERROR Understand

channel 1   channel 2   ...k... channel C

$X_{ijk} \to y_{i,j,k}$ full connect

$$y_{ijk} = \sum_{l=1}^{C} \sum_{m=1}^{P} \sum_{n=1}^{P} W_{mnlk} X_{ijl} + b_{ijk} \quad (1 \le k \le D)$$

parameters $P \times P \times C \times D + P \times Q \times D$

$-$ 张 W

13张 bias

C张 W 钱 13张 Y

D张 bias

生成 D张 Y