



Let's solve this question step-by-step.我们来一步步解答这个问题。

**(a) Express the outputs  $y_j^k$  in terms of the inputs  $z_i^k$ .**

**(a) 表达输出  $y_j^k$  就投入而言  $z_i^k$  。**

Since each  $z^k$  is a  $10 \times 1$  column vector that includes nine pixel values from a  $3 \times 3$  patch of the original image  $X$  and a bias term (1), the output for each neuron  $j$  in the fully-connected layer is given by:

由于每个  $z^k$  是一个  $10 \times 1$  列向量，包括来自  $a$  的九个像素值  $3 \times 3$  原始图像的补丁  $X$  和偏置项 (1)，每个神经元的输出  $j$  在全连接层中由下式给出：

$$y_j^k = \sum_{i=1}^9 w_{ij} z_i^k + w_{0j} \cdot 1 + b_j$$

This can be simplified to:这可以简化为：

$$y_j^k = \sum_{i=1}^9 w_{ij} z_i^k + b_j$$

where  $w_{ij}$  are the weights associated with each input  $z_i^k$  and  $b_j$  is the bias for the output neuron  $j$ .在哪里  $w_{ij}$  是与每个输入相关的权重  $z_i^k$  和  $b_j$  是输出神经元的偏差  $j$  。

**(b) Construct the matrix  $W$  that contains all network parameters, and express the output  $y^k$  in terms of the input  $z^k$ .**

**(b) 构造矩阵  $W$  包含所有网络参数，并表达输出  $y^k$  就输入而言  $z^k$  。**

Define a weight matrix  $W$  of size  $6 \times 10$  where each row corresponds to a specific output  $y_j^k$  and each column corresponds to an input  $z_i^k$  (including the bias term).

定义权重矩阵  $W$  尺寸的  $6 \times 10$  其中每一行对应一个特定的输出  $y_j^k$  每列对应一个输入  $z_i^k$  （包括偏差项）。

So, the matrix  $W$  can be represented as:所以，矩阵  $W$  可以表示为：

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{19} & b_1 \\ w_{21} & w_{22} & \dots & w_{29} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ w_{61} & w_{62} & \dots & w_{69} & b_6 \end{bmatrix}$$

The output vector  $y^k$  can then be expressed in terms of the input  $z^k$  as:

输出向量  $y^k$  然后可以用输入来表示  $z^k$  作为：

$$y^k = W \cdot z^k$$

where  $W$  is the  $6 \times 10$  weight matrix, and  $z^k$  is the  $10 \times 1$  input vector.

在哪里  $W$  是  $6 \times 10$  权重矩阵，以及  $z^k$  是  $10 \times 1$  输入向量。

**(c) Six images  $Y_j$  are constructed by the 10,000 outputs  $y^k$ ,  $k = 1, \dots, 10,000$ . Express the output images  $Y_j$  in terms of the input image  $X$ .**

**(c) 六幅图像  $Y_j$  由 10,000 个输出构成  $y^k$  ,  $k = 1, \dots, 10,000$  。****表达输出图像  $Y_j$  就输入图像而言  $X$  。**

Each output  $y_j^k$  corresponds to a specific  $3 \times 3$  patch in the original image  $X$ . Therefore, each output image  $Y_j$  is constructed by mapping each  $y_j^k$  back to the spatial position in the original  $100 \times 100$  grid.

每个输出  $y_j^k$  对应于特定的  $3 \times 3$  原始图像中的补丁  $X$  。因此，每个输出图像  $Y_j$  是通过映射每个  $y_j^k$  回到原来的空间位置  $100 \times 100$  网格。

Thus, each output image  $Y_j$  can be seen as a filtered version of the input image  $X$ , where each pixel in  $Y_j$  is the result of applying the weights  $w_{ij}$  and bias  $b_j$  to the corresponding  $3 \times 3$  patch in  $X$ . Mathematically, this can be expressed as a convolution operation over  $X$  with the filter  $w_{ij}$  for each output image  $Y_j$ , followed by the addition of bias  $b_j$ :

因此，每个输出图像  $Y_j$  可以看作是输入图像的过滤版本  $X$  ，其中每个像素  $Y_j$  是应用权重的结果  $w_{ij}$  和偏见  $b_j$  到对应的  $3 \times 3$  修补  $X$  。在数学上，这可以表示为卷积运算  $X$  带过滤器  $w_{ij}$  对于每个输出图像  $Y_j$  ，然后添加偏差  $b_j$  ：

$$Y_j = X * w_j + b_j$$

where  $*$  denotes the convolution operation.在哪里  $*$  表示卷积运算。

**(d) Suppose this network is trained by 100 images of size  $100 \times 100$ . What is the number of training samples used to train the network parameters  $W$  or  $w_{ij}$  and  $b_j$ ?**

**(d) 假设该网络由 100 张大小为  $100 \times 100$  。****用于训练网络参数的训练样本数量是多少  $W$  或者  $w_{ij}$  和  $b_j$  ？**

Each  $100 \times 100$  image generates 10,000  $3 \times 3$  patches, which correspond to 10,000 training samples per image.

每个  $100 \times 100$  图像生成 10,000  $3 \times 3$  patch，对应于每张图像 10,000 个训练样本。

With 100 images, the total number of training samples is:100张图像，训练样本总数为：

$$100 \times 10,000 = 1,000,000$$

So, there are **1,000,000 training samples** used to train the network parameters  $W$  (the weights  $w_{ij}$  and biases  $b_j$ ).

因此，有**1,000,000 个训练样本**用于训练网络参数  $W$  （权重  $w_{ij}$  和偏见  $b_j$  ）。