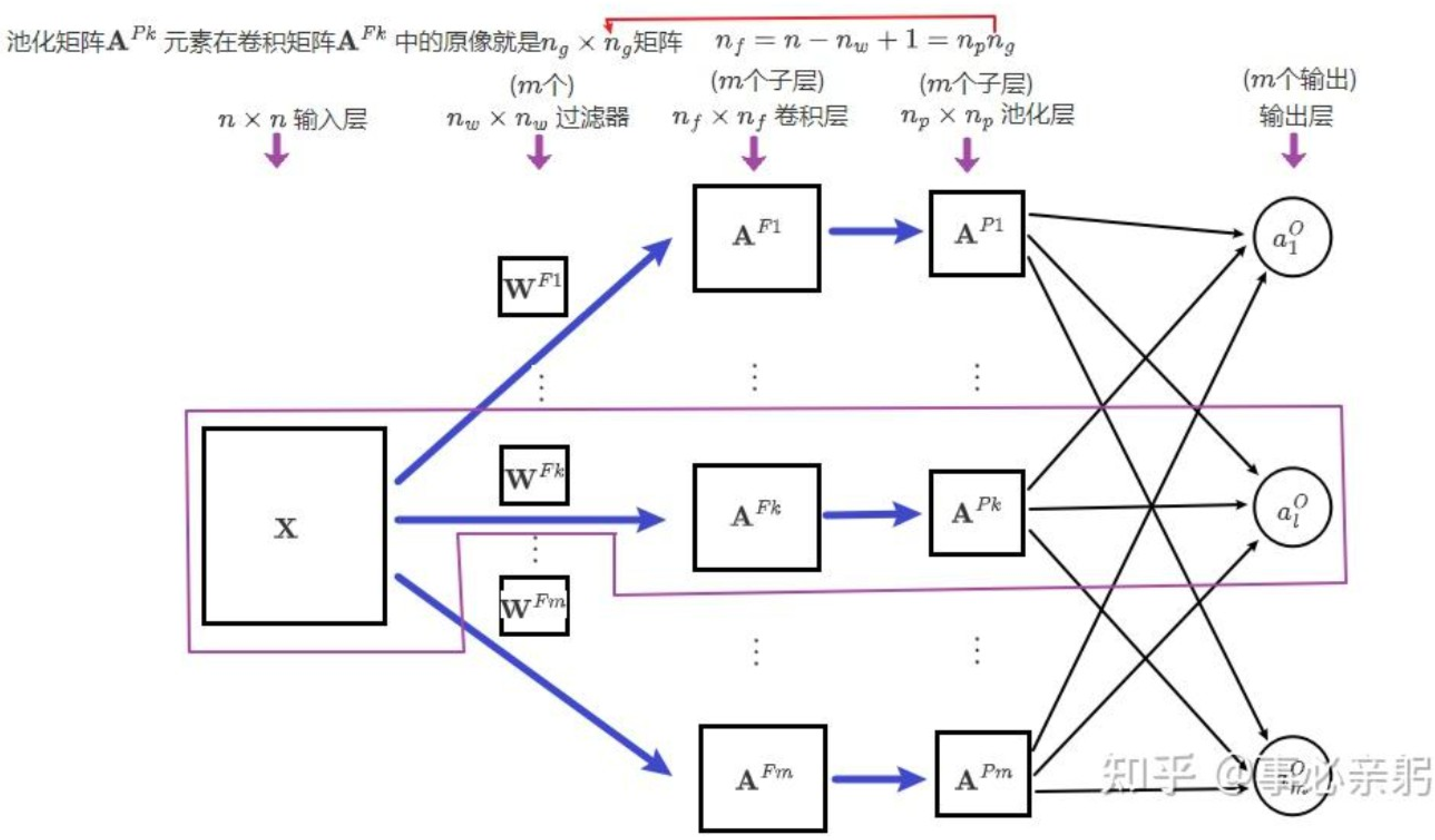
卷积神经网络（简洁的推导）

事必亲躬

模型变量和参数



卷积神经网络总体结构图

- 模型变量

【输入层变量】 输入层的输入变量 $\mathbf{X}$ ，作为一个 $n \times n$ 矩阵，其分量 x_{ij} 。每个分量作为一个神经元其加权输入和输出相等： $a_{ij}^I = z_{ij}^I = x_{ij}$ ， 或其矩阵形式： $\mathbf{A}^I = \mathbf{Z}^I = \mathbf{X}$ 。

【卷积层加权输入】 卷积层的第 k 个卷积子层加权输入变量 $\mathbf{Z}^{Fk}$ ， 作为一个 $n_f \times n_f$ 矩阵，其分量 z_{ij}^{Fk} 代表第 k 个卷积子层的第 i 行 j 列神经元的加权输入。

【卷积层输出】 卷积层的第 k 个卷积子层输出变量 $\mathbf{A}^{Fk}$ ， 作为一个 $n_f \times n_f$ 矩阵，其分量 a_{ij}^{Fk} 代表第 k 个卷积子层的第 i 行 j 列神经元的输出（激活函数值）。

【池化层输入输出】 池化层的第 k 个池化子层输入输出相等（ $\mathbf{A}^{Pk} = \mathbf{Z}^{Pk}$ ），第 k 个池化子层输入变量 $\mathbf{Z}^{Pk}$ ， 作为一个 $n_p \times n_p$ 矩阵，其分量 z_{ij}^{Pk} 代表第 k 个池化子层的第 i 行 j 列神经元的输入。

【输出层加权输入】 输出层，用 z_l^O 代表第 l 个神经元的加权输入。

【输出层输出】 输出层，用 a_l^O 代表第 l 个神经元的输出（激活函数值）。

- 模型参数

【过滤器】第 k 个过滤器 $\mathbf{W}^{Fk}$ ，作为一个 $n_w \times n_w$ 矩阵，其分量 w_{ij}^{Fk} ，代表第 k 个特征映射的过滤器 i 行 j 列的值。

【卷积层偏置】卷积层，用 b^{Fk} 代表第 k 个卷积子层的所有神经元的共同偏置。

【池化层到输出层权重】从池化层第 k 个子层指向输出层第 l 个输出神经元的权重 $\mathbf{W}^{Olk}$ ， 作为一个 $n_p \times n_p$ 矩阵， 其分量为 w_{ij}^{Olk} 则代表从池化层第 k 个子层的 i 行 j 列神经元指向输出层第 l 个输出神经元的权重 。

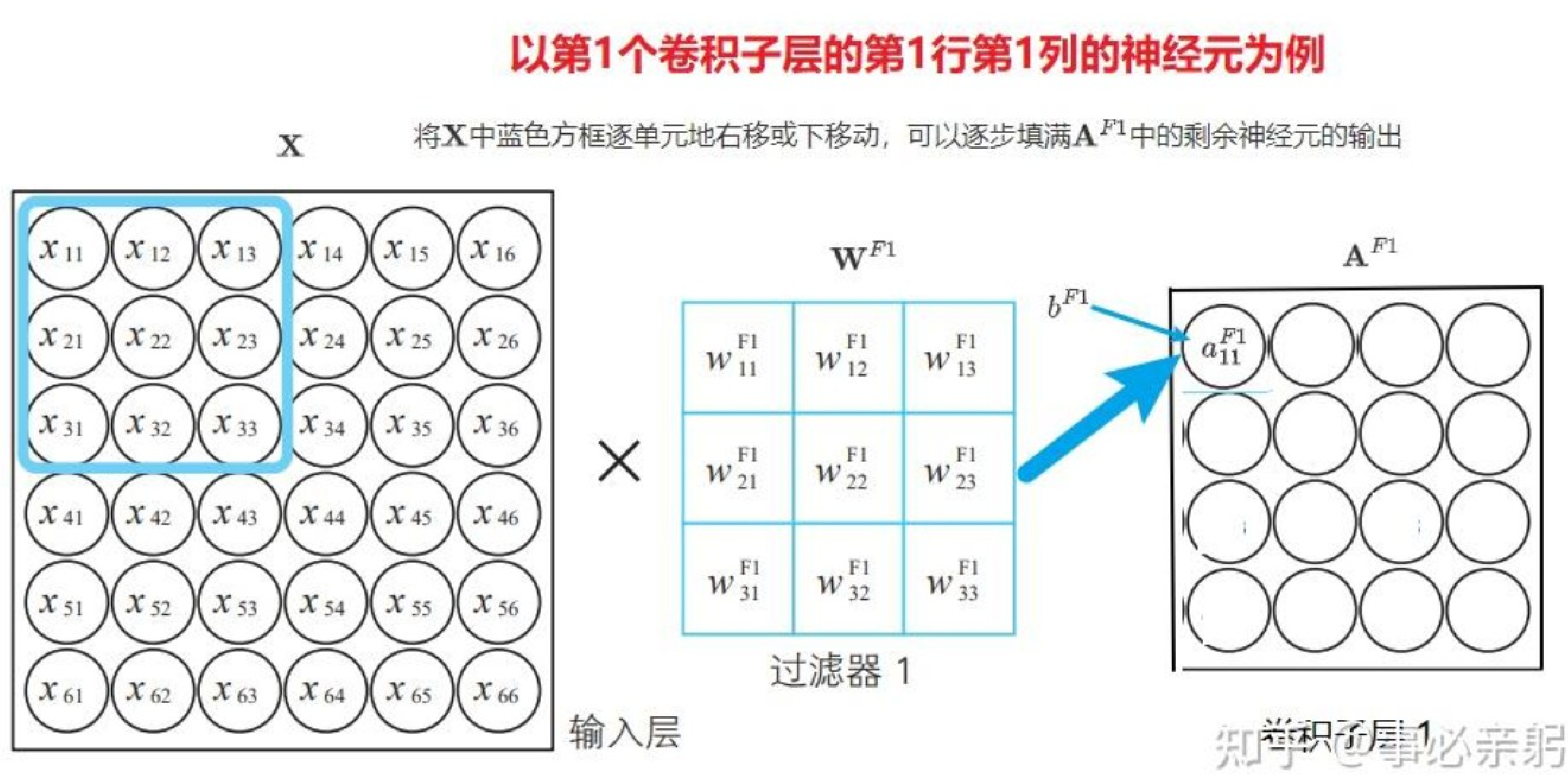
【输出层偏置】输出层， b_l^O 代表第 l 个输出神经元的偏置；

- 学习数据

准备大量学习数据，逐个将某条学习数据从输入层 $\mathbf{X}$ （分量 x_{ij} ）输入， 卷积神经网络将输出一个结果 a_l^O ， 而该条学习数据提供的结果是 t_l 。

数学模型

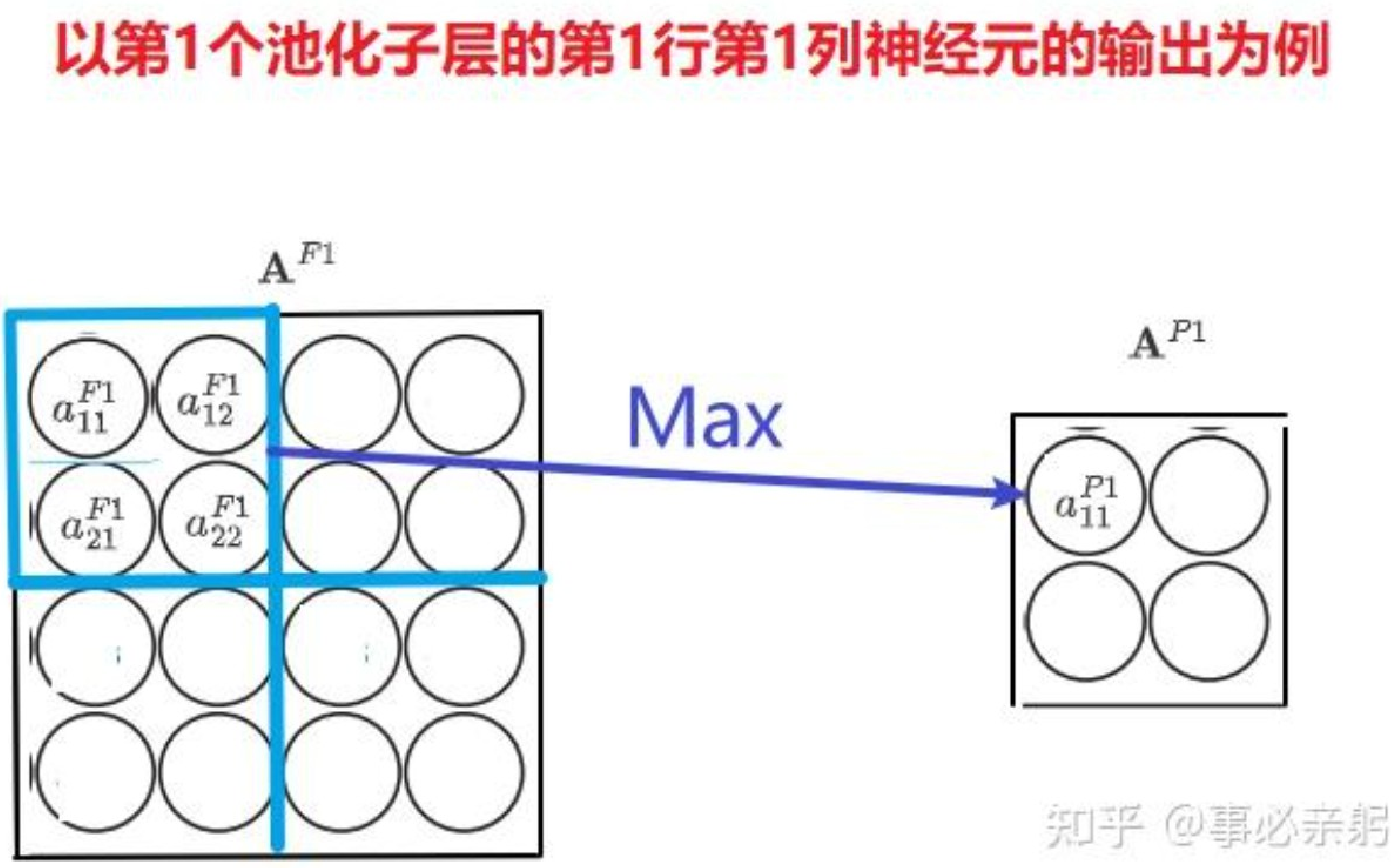
- 卷积层的加权输入和输出



卷积神经网络的卷积层

$$\left.\begin{aligned} z_{ij}^{Fk} &= \left(\sum_{r,s=1}^{n_w} x_{r+i-1,\ s+j-1} w_{rs}^{Fk}\right) + b^{Fk} \\ a_{ij}^{Fk} &= a(z_{ij}^{Fk}) \end{aligned}\right\} \tag{1}$$

- 池化层的输入输出（以最大池化法为例）



卷积神经网络的池化层

$$\left.\begin{aligned} z_{ij}^{Pk} &= \max_{1\le r,s\le n_g} a_{r+n_g(i-1),\ s+n_g(j-1)}^{Fk} \\ a_{ij}^{Pk} &= z_{ij}^{Pk} \end{aligned}\right\} \tag{2}$$

- 输出层的加权输入输出

$$\left.\begin{aligned} z_l^O &= \left(\sum_{k=1}^{m}\sum_{i,j=1}^{n_p} w_{ij}^{Olk} a_{ij}^{Pk}\right) + b_l^O \\ a_l^O &= a(z_l^O) \end{aligned}\right\} \tag{3}$$

- 用平方误差总和做代价函数：

$$C = \sum_{\text{data}} \frac{1}{2} \sum_{l=1}^{m} (t_l - a_l^O)^2 \tag{4}$$

其中，第1个求和号，代表逐个数据求和，后续将会省略这个符号，但默认是存在这个求和的。

我们的目标是最小化这个代价函数。

模型求解

我们将采用 **梯度下降法** 求解代价函数的最小化问题：

$$\boxed{\begin{aligned} \Delta w_{ij}^{Fk} &= -\eta \frac{\partial C}{\partial w_{ij}^{Fk}} \qquad & \Delta w_{ij}^{Olk} &= -\eta \frac{\partial C}{\partial w_{ij}^{Olk}} \\ \Delta b^{Fk} &= -\eta \frac{\partial C}{\partial b^{Fk}} \qquad & \Delta b_l^{O} &= -\eta \frac{\partial C}{\partial b_l^{O}} \end{aligned}} \tag{5}$$

如果给出 $w_{ij}^{Fk}, w_{ij}^{Olk}, b^{Fk}, b_l^O$ 的初值，选定合适的学习率 η，通过计算（5）中的四种偏导， 通过有限步的迭代，原则上可以获得满足精度要求的参数值。

我们先计算卷积层和输出层的**神经元误差** $\dfrac{\partial C}{\partial z_{\alpha\beta}^{F\gamma}}, \dfrac{\partial C}{\partial z_{\mu}^{O}}$ ：

$$\begin{aligned} \frac{\partial C}{\partial z_\mu^O} &= \sum_{l=1}^{m}(a_l^O - t_l)\frac{\partial a_l^O}{\partial z_\mu^O} \qquad \text{将（4）代入} \\ &= \sum_{l=1}^{m}(a_l^O - t_l)a'(z_l^O)\frac{\partial z_l^O}{\partial z_\mu^O} \qquad \text{将（3.2）代入} \\ &= (a_\mu^O - t_\mu)a'(z_\mu^O) \end{aligned} \tag{6}$$

$$\begin{aligned} \frac{\partial C}{\partial z_{\alpha\beta}^{F\gamma}} &= \sum_{l=1}^{m}(a_l^O - t_l)\frac{\partial a_l^O}{\partial z_{\alpha\beta}^{F\gamma}} \qquad \text{将（4）代入} \\ &= \sum_{l=1}^{m}(a_l^O - t_l)a'(z_l^O)\frac{\partial z_l^O}{\partial z_{\alpha\beta}^{F\gamma}} \qquad \text{将（3.2）代入} \\ &= \sum_{l=1}^{m}\frac{\partial C}{\partial z_l^O}\frac{\partial z_l^O}{\partial z_{\alpha\beta}^{F\gamma}} \qquad \text{将（6）代入} \\ &= \sum_{l=1}^{m}\frac{\partial C}{\partial z_l^O}\sum_{k=1}^{m}\sum_{i,j=1}^{n_p} w_{ij}^{Olk}\frac{\partial z_{ij}^{Pk}}{\partial z_{\alpha\beta}^{F\gamma}} \qquad \text{将（3.1）（2.2）代入} \\ &= \sum_{l=1}^{m}\frac{\partial C}{\partial z_l^O}\sum_{k=1}^{m}\sum_{i,j=1}^{n_p} w_{ij}^{Olk}\sum_{\mu,\nu=1}^{n_f}\frac{\partial z_{ij}^{Pk}}{\partial a_{\mu\nu}^{Fk}}\frac{\partial a_{\mu\nu}^{Fk}}{\partial z_{\alpha\beta}^{F\gamma}} \\ &= \sum_{l=1}^{m}\frac{\partial C}{\partial z_l^O}\sum_{k=1}^{m}\sum_{i,j=1}^{n_p} w_{ij}^{Olk}\sum_{\mu,\nu=1}^{n_f}\frac{\partial z_{ij}^{Pk}}{\partial a_{\mu\nu}^{Fk}}a'(z_{\mu\nu}^{Fk})\frac{\partial z_{\mu\nu}^{Fk}}{\partial z_{\alpha\beta}^{F\gamma}} \qquad \text{将（1.2）代入} \\ &= \sum_{l=1}^{m}\frac{\partial C}{\partial z_l^O}a'(z_{\alpha\beta}^{F\gamma})\sum_{i,j=1}^{n_p} w_{ij}^{Ol\gamma}\frac{\partial z_{ij}^{P\gamma}}{\partial a_{\alpha\beta}^{F\gamma}} \end{aligned} \tag{7}$$

其中：

$$\frac{\partial z_{ij}^{P\gamma}}{\partial a_{\alpha\beta}^{F\gamma}} = \frac{\partial}{\partial a_{\alpha\beta}^{F\gamma}}\left(\max_{1\le r,s\le n_g} a_{r+n_g(i-1),\ s+n_g(j-1)}^{F\gamma}\right) \qquad \text{将（2.1）代入}$$

$$=\begin{cases} 0 & \text{if } i \neq [\frac{\alpha-1}{n_g}]+1 \text{ or } j \neq [\frac{\beta-1}{n_g}]+1 \\ \frac{\partial}{\partial a^{F\gamma}_{\alpha\beta}}(\max\limits_{1\leq r,s\leq n_g} a^{F\gamma}_{r+n_g[\frac{\alpha-1}{n_g}],\ s+n_g[\frac{\beta-1}{n_g}]}) & \text{if } i = [\frac{\alpha-1}{n_g}]+1 \text{ and } j = [\frac{\beta-1}{n_g}]+1 \end{cases} \tag{8}$$

为简单计，引入一个记号 ι_α ，$\Gamma_{\alpha\beta}$：

$$\boxed{\begin{aligned} \iota_\alpha &=[\frac{\alpha-1}{n_g}]+1 \\ \Gamma_{\alpha\beta} &=\frac{\partial}{\partial a^{F\gamma}_{\alpha\beta}}(\max\limits_{1\leq r,s\leq n_g} a^{F\gamma}_{r+n_g[\frac{\alpha-1}{n_g}],\ s+n_g[\frac{\beta-1}{n_g}]}) \\ &=\begin{cases} 1 & \text{if } a^{F\gamma}_{\alpha\beta}\text{是所在区块的最大值} \\ 0 & \text{否则} \end{cases} \end{aligned}} \tag{9}$$

将（8）（9）依次代入（7）可得：

$$\frac{\partial C}{\partial z^{F\gamma}_{\alpha\beta}} =\left(\sum_{l=1}^{m}\frac{\partial C}{\partial z^O_l}w^{Ol\gamma}_{\iota_\alpha\iota_\beta}\right)a'(z^{F\gamma}_{\alpha\beta})\Gamma_{\alpha\beta} \tag{10}$$

将前面的结果总结如下：

$$\boxed{\begin{aligned} \frac{\partial C}{\partial z^O_\mu} &=(a^O_\mu - t_\mu)a'(z^O_\mu) \\ \frac{\partial C}{\partial z^{F\gamma}_{\alpha\beta}} &=\left(\sum_{l=1}^{m}\frac{\partial C}{\partial z^O_l}w^{Ol\gamma}_{\iota_\alpha\iota_\beta}\right)a'(z^{F\gamma}_{\alpha\beta})\Gamma_{\alpha\beta} \end{aligned}} \tag{11}$$

其中用到了（9）中定义的 ι，Γ 。公式（11）就是卷积神经网络的**误差反向传播法**。

进而可计算神经网络代价函数关于参数的偏导：

$$\begin{aligned} \frac{\partial C}{\partial w^{Fk}_{ij}} &=\sum_{\gamma=1}^{m}\sum_{\alpha,\beta=1}^{n_f}\frac{\partial C}{\partial z^{F\gamma}_{\alpha\beta}}\frac{\partial z^{F\gamma}_{\alpha\beta}}{\partial w^{Fk}_{ij}} \\ &=\sum_{\gamma=1}^{m}\sum_{\alpha,\beta=1}^{n_f}\frac{\partial C}{\partial z^{F\gamma}_{\alpha\beta}}\left(\sum_{r,s=1}^{n_w}x_{r+\alpha-1,\ s+\beta-1}\frac{\partial w^{F\gamma}_{rs}}{\partial w^{Fk}_{ij}}\right) \qquad \text{将 (1.1) 代入} \\ &=\sum_{\alpha,\beta=1}^{n_f}\frac{\partial C}{\partial z^{Fk}_{\alpha\beta}}x_{i+\alpha-1,\ j+\beta-1} \end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial C}{\partial w_{ij}^{Olk}} &= \sum_{\mu=1}^{m} \frac{\partial C}{\partial z_{\mu}^{O}} \frac{\partial z_{\mu}^{O}}{\partial w_{ij}^{Olk}} \\ &= \sum_{\mu=1}^{m} \frac{\partial C}{\partial z_{\mu}^{O}} \left(\sum_{\gamma=1}^{m} \sum_{\alpha,\beta=1}^{n_p} \frac{\partial w_{\alpha\beta}^{O\mu\gamma}}{\partial w_{ij}^{Olk}} a_{\alpha\beta}^{P\gamma} \right) \qquad \text{将（3.1）代入} \\ &= \frac{\partial C}{\partial z_{l}^{O}} a_{ij}^{Pk}\end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial C}{\partial b^{Fk}} &= \sum_{\gamma=1}^{m} \sum_{\alpha,\beta=1}^{n_f} \frac{\partial C}{\partial z_{\alpha\beta}^{F\gamma}} \frac{\partial z_{\alpha\beta}^{F\gamma}}{\partial b^{Fk}} \\ &= \sum_{\gamma=1}^{m} \sum_{\alpha,\beta=1}^{n_f} \frac{\partial C}{\partial z_{\alpha\beta}^{F\gamma}} \frac{\partial b^{F\gamma}}{\partial b^{Fk}} \qquad \text{将（1.1）代入} \\ &= \sum_{\alpha,\beta=1}^{n_f} \frac{\partial C}{\partial z_{\alpha\beta}^{Fk}}\end{aligned} \tag{14}$$

$$\begin{aligned}\frac{\partial C}{\partial b_{l}^{O}} &= \sum_{\mu=1}^{m} \frac{\partial C}{\partial z_{\mu}^{O}} \frac{\partial z_{\mu}^{O}}{\partial b_{l}^{O}} \\ &= \sum_{\mu=1}^{m} \frac{\partial C}{\partial z_{\mu}^{O}} \frac{\partial b_{\mu}^{O}}{\partial b_{l}^{O}} \qquad \text{将（3.1）代入} \\ &= \frac{\partial C}{\partial z_{l}^{O}}\end{aligned} \tag{15}$$

总结如下：

$$\boxed{\begin{aligned} &\frac{\partial C}{\partial w_{ij}^{Fk}} = \sum_{\alpha,\beta=1}^{n_f} \frac{\partial C}{\partial z_{\alpha\beta}^{Fk}} x_{i+\alpha-1,\ j+\beta-1} \qquad && \frac{\partial C}{\partial b^{Fk}} = \sum_{\alpha,\beta=1}^{n_f} \frac{\partial C}{\partial z_{\alpha\beta}^{Fk}} \\ &\frac{\partial C}{\partial w_{ij}^{Olk}} = \frac{\partial C}{\partial z_{l}^{O}} a_{ij}^{Pk} && \frac{\partial C}{\partial b_{l}^{O}} = \frac{\partial C}{\partial z_{l}^{O}} \end{aligned}} \tag{16}$$

原则上不难根据（5）（11）（16）再结合引入的记号（9）编程解决本主题的问题。

【完毕】