

日期: /

一. 神经网络特征图计算

1. 卷积层

输入图像大小: $H_{in} \times W_{in} \times n_c$

每个滤波器(卷积核)大小: $k \times k \times n_c$

滤波器(卷积核)个数: K

加边填充 padding: P

卷积核滑动步幅(stride): S

输出特征图像大小:

$$\underbrace{\left(\frac{H_{in} - k + 2P}{S} + 1 \right)}_{\text{高}} \times \underbrace{\left(\frac{W_{in} - k + 2P}{S} + 1 \right)}_{\text{宽}} \times \underbrace{K}_{\text{通道数}}$$

计算卷积层输出特征图大小, 当除不尽时, 一般向**下**取整。

2. 池化层

- 输入图像大小: $H_{in} \times W_{in} \times n_c$

- 每个滤波器(卷积核)大小: $k \times k \times n_c$

- 滤波器(卷积核)个数: K

- 卷积滑动步幅(stride): S

- 加边填充 padding: P

□ 输出图像大小:

$$\left(\frac{H_{in} - k + 2P}{S} + 1 \right) \times \left(\frac{W_{in} - k + 2P}{S} + 1 \right) \times K$$

计算池化层输出特征图大小, 当除不尽时, 通常向**上**取整。

日期: /

二. 参数计算: (lenet5 为例)

1. Input

输入图像统一归一化为 32×32 。

2. C1卷积层

经过 $(5 \times 5 \times 1) \times 6$ 卷积核, $\text{stride}=1$, $\text{pad}=0$, 生成feature map为 $28 \times 28 \times 6$ 。

3. S2池化层

经过 (2×2) 池化核, 平均池化, $\text{stride}=2$, $\text{pad}=0$, 生成feature map为 $14 \times 14 \times 6$ 。

4. C3卷积层

经过 $(5 \times 5 \times 6) \times 16$ 卷积核, $\text{stride}=1$, $\text{pad}=0$, 生成feature map为 $10 \times 10 \times 16$ 。

5. S4池化层

经过 (2×2) 池化核, 平均池化, $\text{stride}=2$, $\text{pad}=0$, 生成feature map为 $5 \times 5 \times 16$ 。

6. C5卷积层

经过 $(5 \times 5 \times 16) \times 120$ 卷积核, $\text{stride}=1$, $\text{pad}=0$, 生成feature map为 $1 \times 1 \times 120$ 。

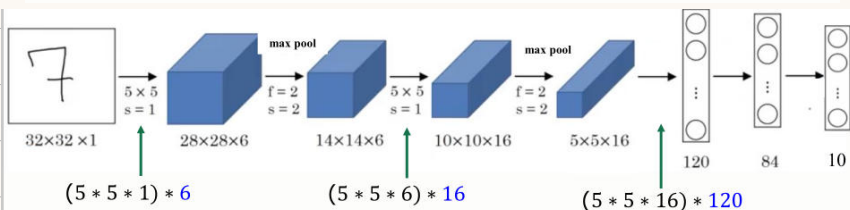
7. F6全连接层

输入为 $1 \times 1 \times 120$, 输出为 $1 \times 1 \times 84$, 总参数量为 120×84 。

8. F7 全连接层 (输出层)

输入为 $1 \times 1 \times 84$, 输出为 $1 \times 1 \times 10$, 总参数量为 84×10 。10就是分类的类别数。输出层激活函数是 softmax。

中间隐层激活函数是ReLU



待估计的权重参数量:

$$5 \times 5 \times 1 \times 6 + 5 \times 5 \times 6 \times 16 + 5 \times 5 \times 16 \times 120 + 120 \times 84 + 84 \times 10$$

待估计全部参数量 (权重+偏置):

$$\begin{aligned} & 5 \times 5 \times 1 \times 6 (\text{卷积核}) + 6 (\text{偏置}) + 5 \times 5 \times 6 \times 16 (\text{卷积核}) + 16 (\text{偏置}) + \\ & 5 \times 5 \times 16 \times 120 (\text{卷积核}) + 120 (\text{偏置}) \\ & + 120 \times 84 (\text{全连接权重}) + 84 (\text{偏置}) + 84 \times 10 (\text{全连接权重}) + 10 (\text{偏置}) \end{aligned}$$

卷积层: 有参数. 池化层: 无参数.

日期: /

卷积核

- 一个卷积核的参数: $2 \times 2 \times 3 = 12$; ($\text{in_channels}=3$, $\text{kernel_size}=2$)
- 16 个卷积核的参数: $16 \times 12 = 192$; ($\text{out_channels}=16$)
- 加上偏置: $192 + 16 = 208$; ($\text{weights} \times x + \text{bias}$)

全连接层:

$$16 \times 133 + \underbrace{133}_{\text{bias}} = 2261$$

三 时间复杂度 FLOPs

① 卷积核

■ 衡量卷积计算量的指标是FLOPs (Floating Point Operations, 浮点运算次数)

■ 一次乘法或一次加法表示一个浮点运算次数

■ CNN 中单个卷积层的乘法和加法浮点运算次数:

$$[(k \times k \times n_c) + (k \times k \times n_c - 1) + 1] \times H_{out} \times W_{out} \times K$$

- 卷积核每滑动一次的乘法浮点计算量: $k \times k \times n_c$
- 卷积核每滑动一次的加法浮点计算量: $k \times k \times n_c - 1$
- 输出单个特征图的卷积乘法浮点计算量: $(k \times k \times n_c) \times H_{out} \times W_{out}$
- 输出单个特征图的卷积加法浮点计算量: $(k \times k \times n_c - 1) \times H_{out} \times W_{out}$
- 输出 K 个特征图的卷积乘法浮点计算量: $(k \times k \times n_c) \times H_{out} \times W_{out} \times K$
- 输出 K 个特征图的卷积加法浮点计算量: $(k \times k \times n_c - 1) \times H_{out} \times W_{out} \times K$
- 输出单个特征图的偏置浮点加法计算量: $H_{out} \times W_{out}$
- 输出 K 个特征图的偏置浮点加法计算量: $H_{out} \times W_{out} \times K$

□ 单个卷积层的乘法和加法的浮点计算量:

$$FLOPs = 2 \times k \times k \times n_c \times H_{out} \times W_{out} \times K$$

- 上式是乘法和加法运算的总和，将一次乘运算或加运算都视作一次浮点运算
- 在计算机视觉论文中，常常将一个‘乘-加’组合视为一次浮点运算，英文表述为‘Multi-Add’，运算量正好是上面的算法减半，此时的运算量为：

$$FLOPs = k \times k \times n_c \times H_{out} \times W_{out} \times K$$

② 全连接层的浮点计算量

□ 单个全连接层的乘法和加法浮点计算量（权重+偏置）：

$$FLOPs = [N_{in} + (N_{in} - 1) + 1] \times N_{out} = 2 \times N_{in} \times N_{out}$$

- 其中 N_{in} 表示输入层神经元个数， N_{out} 表示输出层神经元个数。上述式子中第一个 N_{in} 表示乘法运算量， $N_{in} - 1$ 表示加法运算量，+1表示 N_{out} 个偏置项计算量， $\times N_{out}$ 表示计算 N_{out} 个神经元的值。

□ 单层全连接层的网络模型参数量（权重+偏置）：

$$parameters = (N_{in} + 1) \times N_{out}$$

□ 如果将一个‘乘-加’组合视为一次浮点运算，则此时单个全连接层的浮点运算量为：

$$FLOPs = N_{in} \times N_{out}$$

- 其中 N_{in} 表示输入层神经元个数， N_{out} 表示输出层神经元个数。

日期: /