

卷积层和池化层前向 / 后向传播的实现原理

所有卷积及池化过程都利用 `im2col` 实现，没有调用 `numpy` 和 `scipy` 的卷积函数，特此说明。

`im2col` 及 `col2im` 函数参考了开源代码，地址已在代码中注明。

在以下推导中，矩阵间 \times 符号表示矩阵乘法。

1. 卷积

前馈过程：

参考 Caffe 中卷积的实现方式，对原始图片矩阵进行变换，将每组与卷积核相乘的元素作为一列，卷积核矩阵也进行相应变换之后，直接通过矩阵相乘的方式得到结果。

原始图片矩阵尺寸： $N \times c_{in} \times h_{in} \times w_{in}$

转换后尺寸： $(c_{in} \times kernel_size \times kernel_size) \times (h_{out} \times w_{out} \times N)$

卷积核矩阵输入： $c_{out} \times c_{in} \times kernel_size \times kernel_size$

变换后： $c_{out} \times (c_{in} \times kernel_size \times kernel_size)$

矩阵相乘后结果的尺寸： $c_{out} \times (h_{out} \times w_{out} \times N)$

之后变换为 $N \times c_{out} \times h_{out} \times w_{out}$ 的尺寸即为所求结果。

例：

输入尺寸 $1 \times 3 \times 3 \times 3$

$$input = \begin{bmatrix} \begin{bmatrix} [0,1,2], [0,1,3], [0,2,3] \end{bmatrix}, \\ \begin{bmatrix} [2,1,0], [3,1,0], [3,2,0] \end{bmatrix}, \\ \begin{bmatrix} [1,2,0], [1,3,0], [2,3,0] \end{bmatrix} \end{bmatrix}$$

卷积核矩阵： $2 \times 3 \times 2 \times 2$

$$W = \begin{bmatrix} \begin{bmatrix} [1,1], [1,1] \end{bmatrix}, \\ \begin{bmatrix} [2,2], [2,2] \end{bmatrix}, \\ \begin{bmatrix} [3,3], [3,3] \end{bmatrix} \\ \begin{bmatrix} [0,1], [0,1] \end{bmatrix}, \\ \begin{bmatrix} [1,2], [1,2] \end{bmatrix}, \\ \begin{bmatrix} [2,3], [2,3] \end{bmatrix} \end{bmatrix}$$

pad: 0, stride: 1

`im2col` 输出尺寸： 12×4

$$input_cols = \begin{bmatrix} [0,1,0,1], \\ [1,2,1,3], \\ [0,1,0,2], \\ [1,3,2,3], \\ [2,1,3,1], \\ [1,0,1,0], \\ [3,1,3,2], \\ [1,0,2,0], \\ [1,2,1,3], \\ [2,0,3,0], \\ [1,3,2,3], \\ [3,0,3,0] \end{bmatrix}$$

卷积核矩阵输出： 2×12

$$w_transposed = \begin{bmatrix} [1,1,1,1,2,2,2,2,3,3,3,3], \\ [0,1,0,1,1,2,1,2,2,3,2,3] \end{bmatrix}$$

令：

$$output_conv_transposed = w_transposed \times input_cols$$

之后变换为 $N \times c_out \times h_out \times w_out$ 的尺寸即可。（此为未加偏置的结果）

后向传播过程：

先做 transpose 的逆过程，将 grad_output 转换为 output_transposed 的尺寸。

由于前馈过程做了一个简单的矩阵相乘，所以类比线性层后向传播的实现，有：

$$\frac{\partial E}{\partial w_transposed} = grad_output_transposed \times input_cols.T,$$

$$\frac{\partial E}{\partial input_cols} = w_transposed.T \times grad_output_transposed$$

要得到 $\frac{\partial E}{\partial w}$ ，只需做 transpose 的逆过程，将 w_transposed 变换回原来的形状。由运算

input_cols 的过程易知，要得到 $\frac{\partial E}{\partial input}$ 中的每一个元素，只需将 $\frac{\partial E}{\partial input_cols}$ 中与其对应的各个值

相加即可，这个过程由 col2img 完成。

2.平均池化

前馈：

与卷积类似，先将 input 变换为 $(N \times c_in) \times h_in \times w_in$ 的尺寸，依然用 im2col 得到 input_cols（尺寸为 $(kernel_size \times kernel_size) \times (h_out \times w_out \times (N \times c_in))$ ），之后对每列做平均，再变换为 $N \times c_in \times h_out \times w_out$ 的尺寸。

后向传播：

做平均的过程也可以看作矩阵乘法，所以后向传播也与卷积类似。

其余部分较为简单，并且是按照课件上的公式实现的，在此略过。