



卷积与卷积神经网络

(Convolution and Convolutional Neural Network)

刘远超

哈尔滨工业大学

计算机科学与技术学院

什么是卷积?



输入图像

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

卷积核

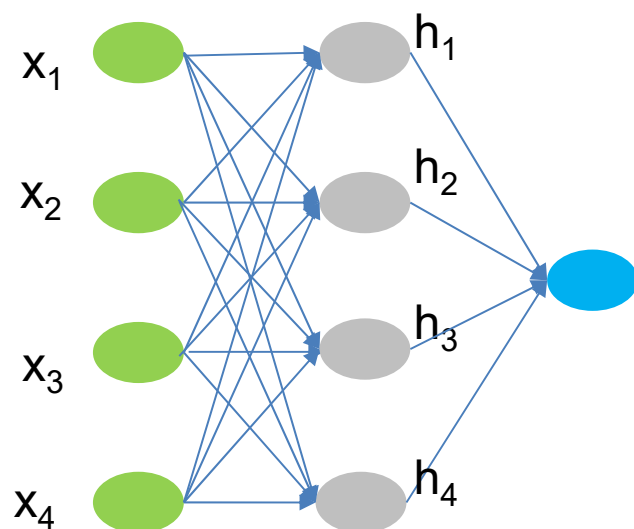


特征图

什么是卷积神经网络?

- 卷积神经网络（Convolutional Neural Networks, CNN）是一类包含卷积计算且具有深度结构的前馈神经网络，是深度学习的代表算法之一。卷积神经网络仿造生物的视知觉（visual perception）机制构建，能够进行平移不变分类。对卷积神经网络的研究始于二十世纪80至90年代，LeNet-5是较早出现的卷积神经网络。
- 卷积神经网络具有以下特点：
 - 局部卷积
 - 参数共享
 - 多卷积核
 - 池化操作
 - 多层处理

假如采用全连接网络进行图像分类

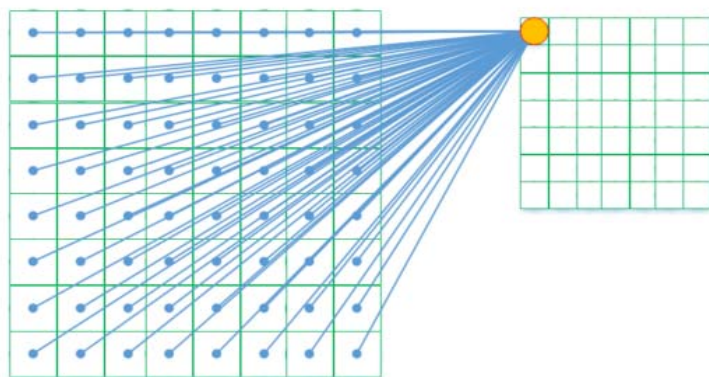


如果采用传统全连接神经网络，则

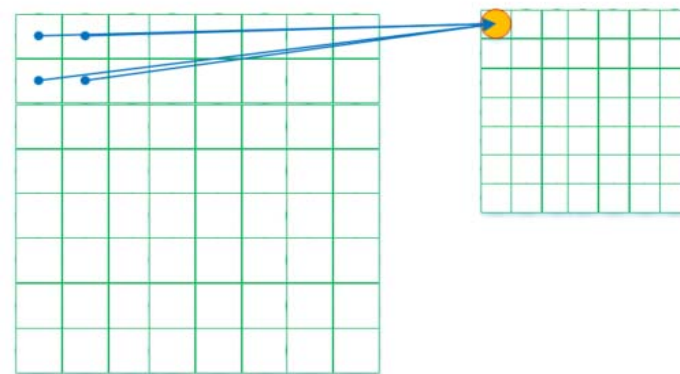
↓

$$\begin{aligned} h_1 &= f(W_{11}x_1 + W_{12}x_2 + W_{13}x_3 + W_{14}x_4) \\ h_2 &= f(W_{21}x_1 + W_{22}x_2 + W_{23}x_3 + W_{24}x_4) \\ h_3 &= f(W_{31}x_1 + W_{32}x_2 + W_{33}x_3 + W_{34}x_4) \\ h_4 &= f(W_{41}x_1 + W_{42}x_2 + W_{43}x_3 + W_{44}x_4) \end{aligned}$$

CNN特点之一: 局部卷积



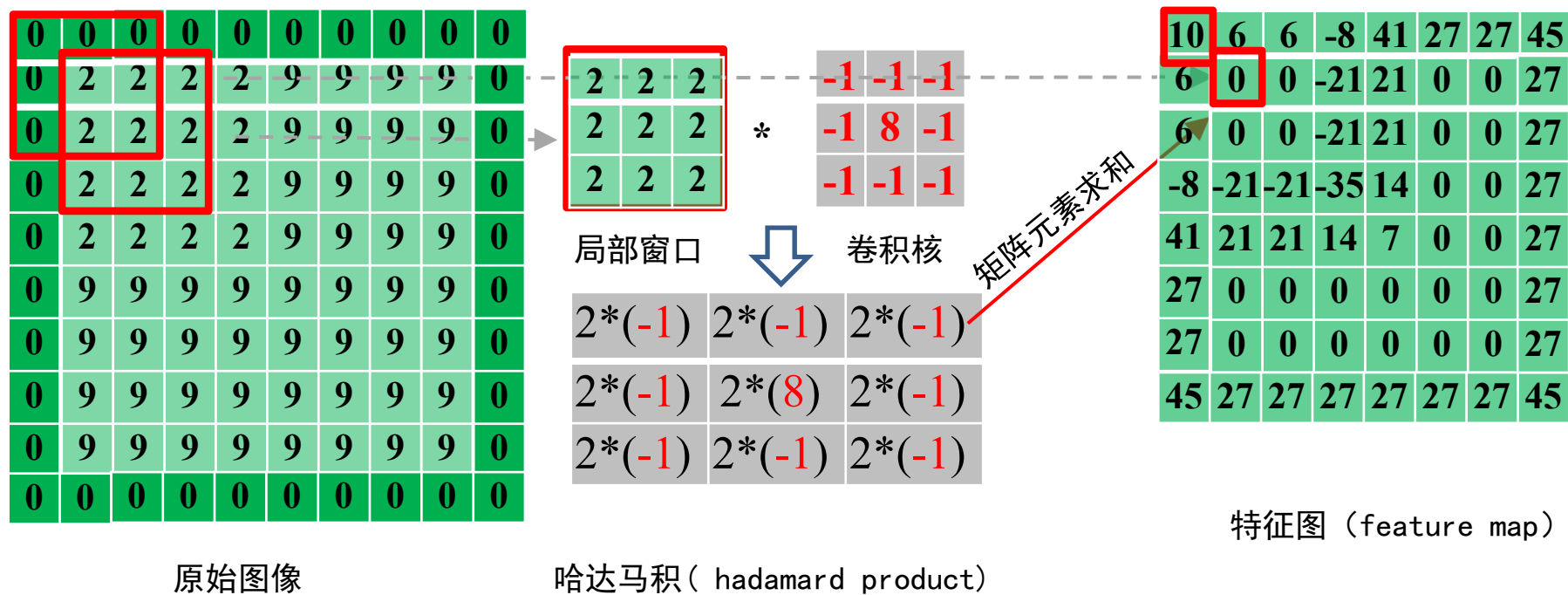
全局感知



局部感知

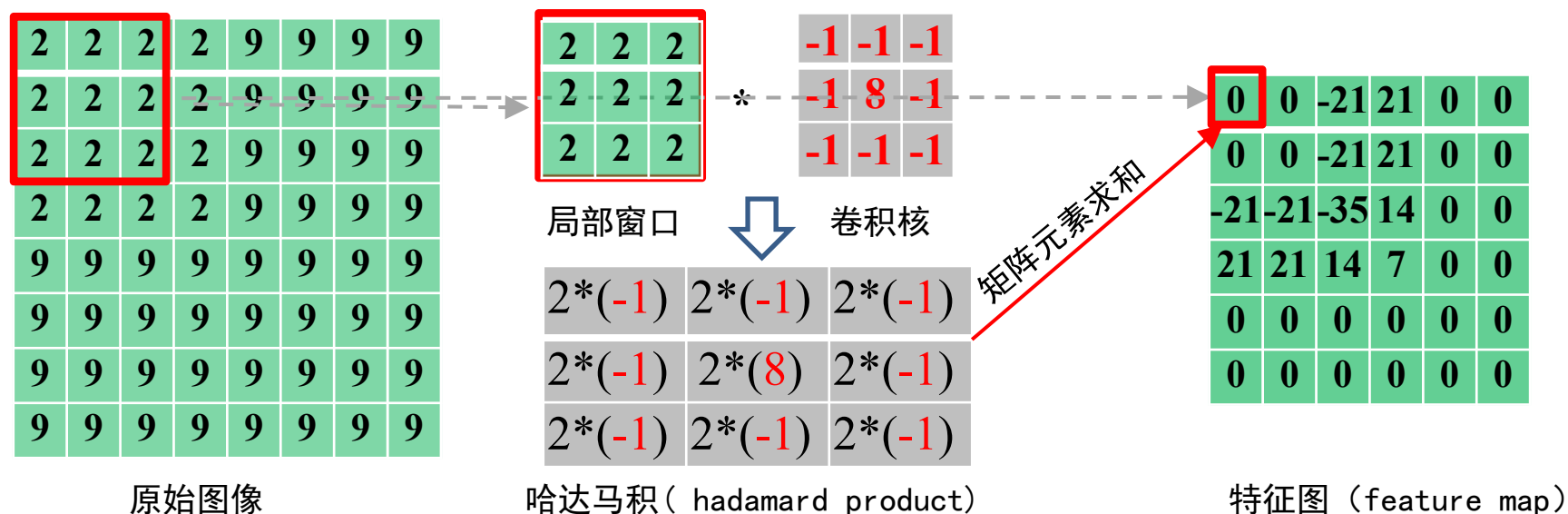
- 可以把卷积想象成作用于矩阵的一个滑动窗口函数。滑动窗口又称作**卷积核**、滤波器或是特征检测器。
- 对于给定的输入图像，输出特征图中每个像素实际上是输入图像中局部区域中像素的加权平均，其权值由**卷积核**定义。

卷积具体实现举例



- 卷积的步长 (stride size)
- 宽卷积与窄卷积 (Wide vs. Narrow Convolution)

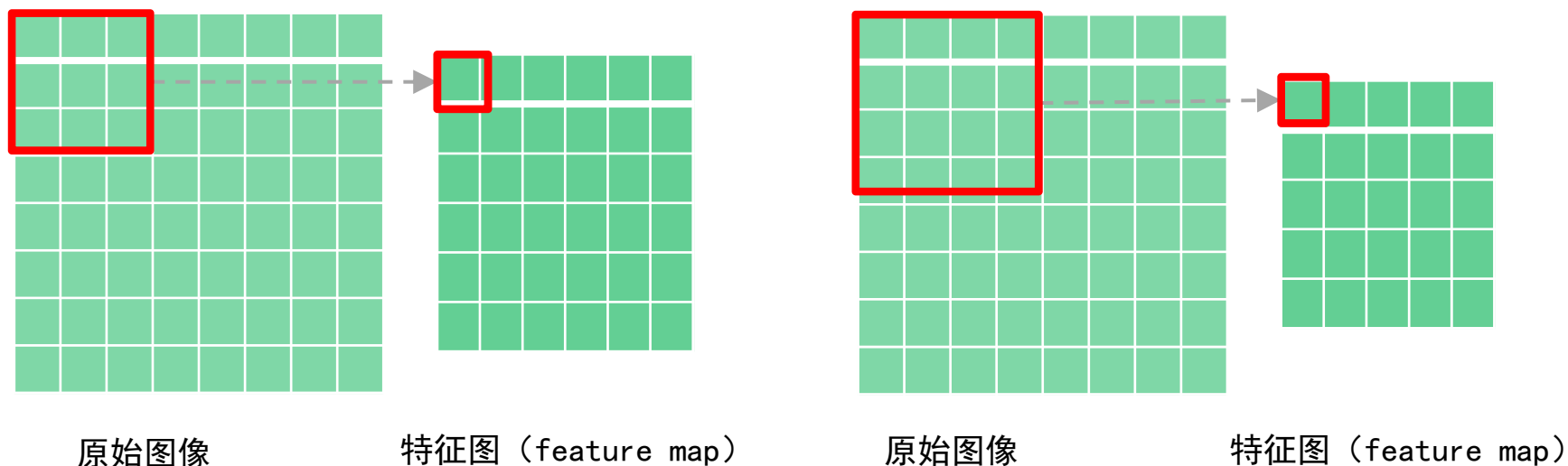
CNN特点之二:参数共享



- 在上面的局部连接中，右边每个神经元都对应 $3*3=9$ 个参数（卷积核），这9个参数就是共享的。
- 可以将这9个参数看成是卷积提取特征的方式，该方式与位置无关。

CNN特点之三: 多卷积核

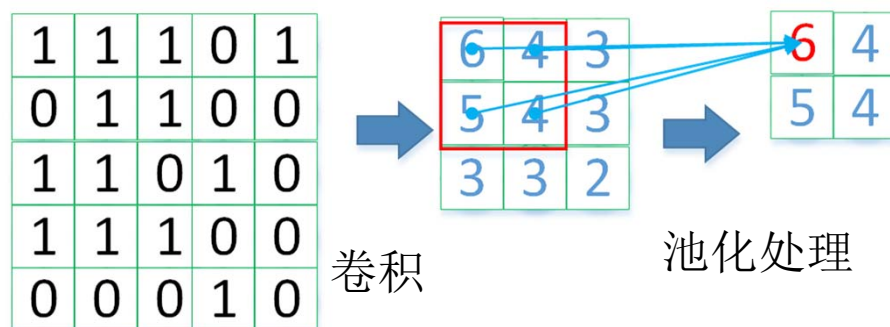
为了充分提取特征, 可以使用多个卷积核:



- 每个卷积核都会对输入图像进行卷积处理, 生成另一幅图像。不同卷积核生成的不同图像可以理解为是该输入图像的不同通道。

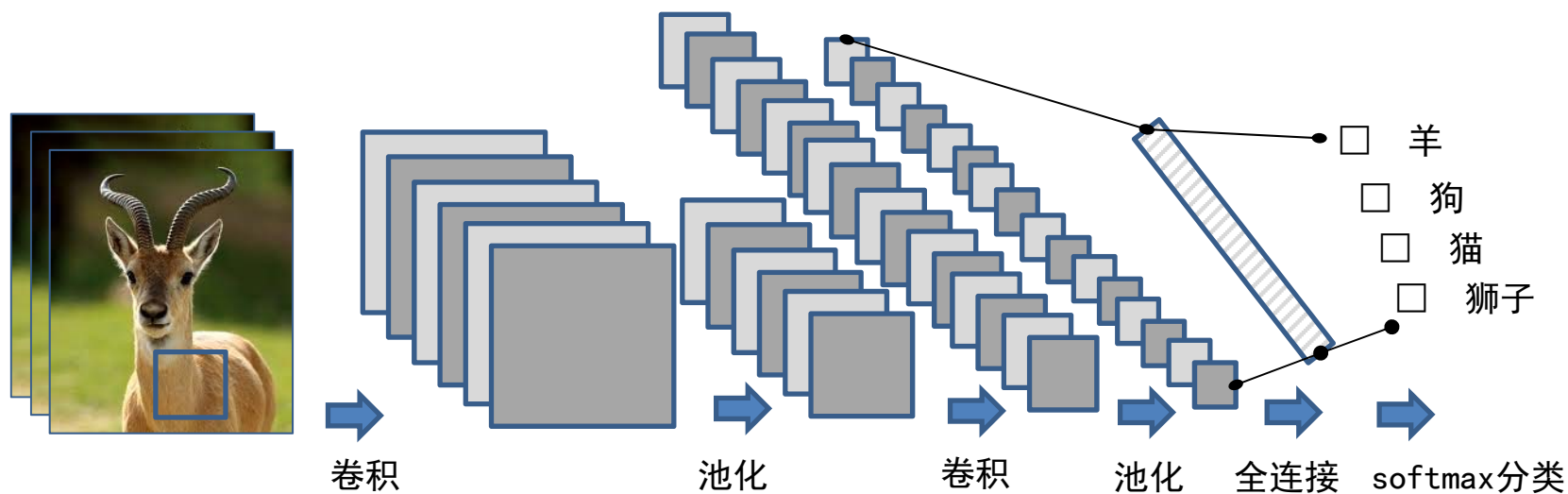
CNN特点之四: 池化处理

- 池化处理也叫作降采样处理（**down-pooling**），是对不同位置的特征进行聚合统计。通常是取对应位置的最大值（最大池化）、平均值（平均池化）等。



- 池化的优点：1) 降维；2) 克服过拟合；3) 在图像识别领域，池化还能提供平移和旋转不变性。

CNN特点之五:多层处理



- 一般而言，在图像处理中，一层卷积及降采样往往只学到了局部的特征。层数越多，学到的特征越全局化。因此通过这样的多层处理，低级的特征组合形成更高级的特征表示。

Thanks!





LeNet-5 模型分析

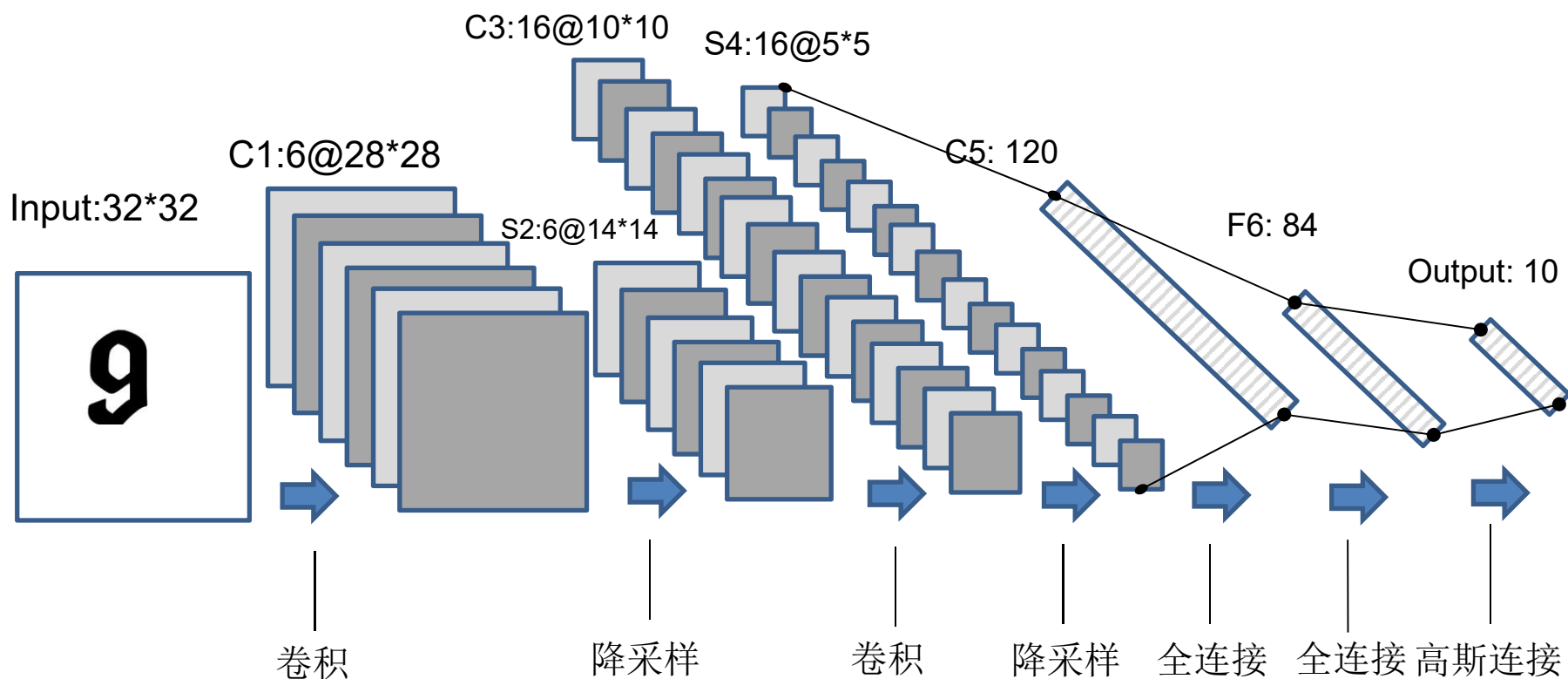
(Analysis of LeNet-5 Model)

刘远超

哈尔滨工业大学

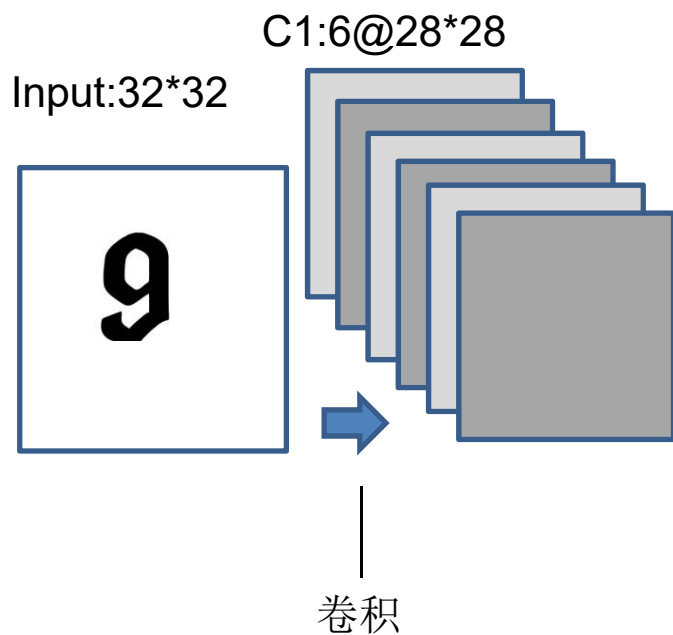
计算机科学与技术学院

LeNet-5



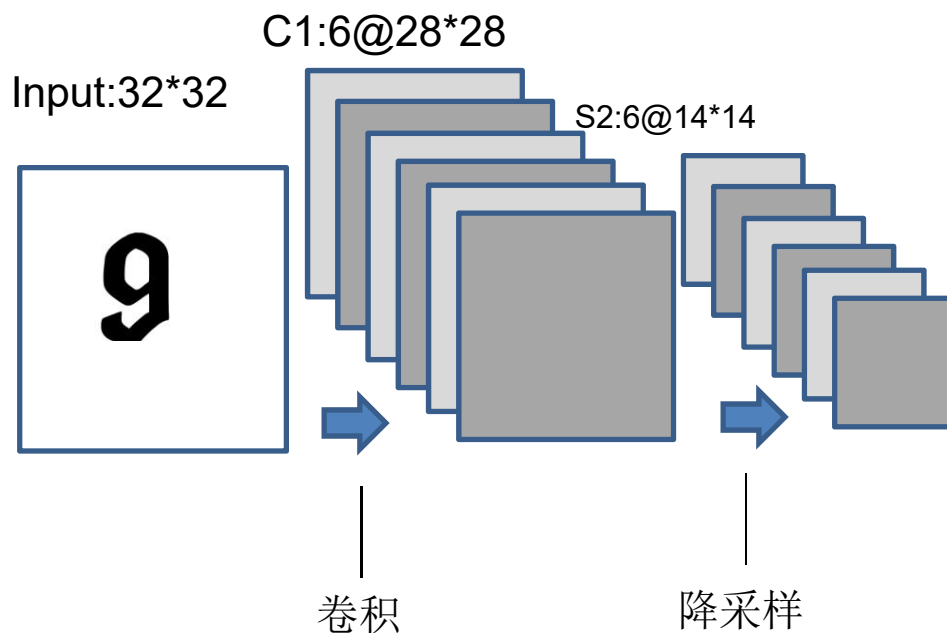
- 输入图像尺寸: 32*32
- 卷积层: 3个
- 降采样层: 2个
- 输出: 输入图像的数字类别 (0-9) 概率

LeNet-5--卷积层C1



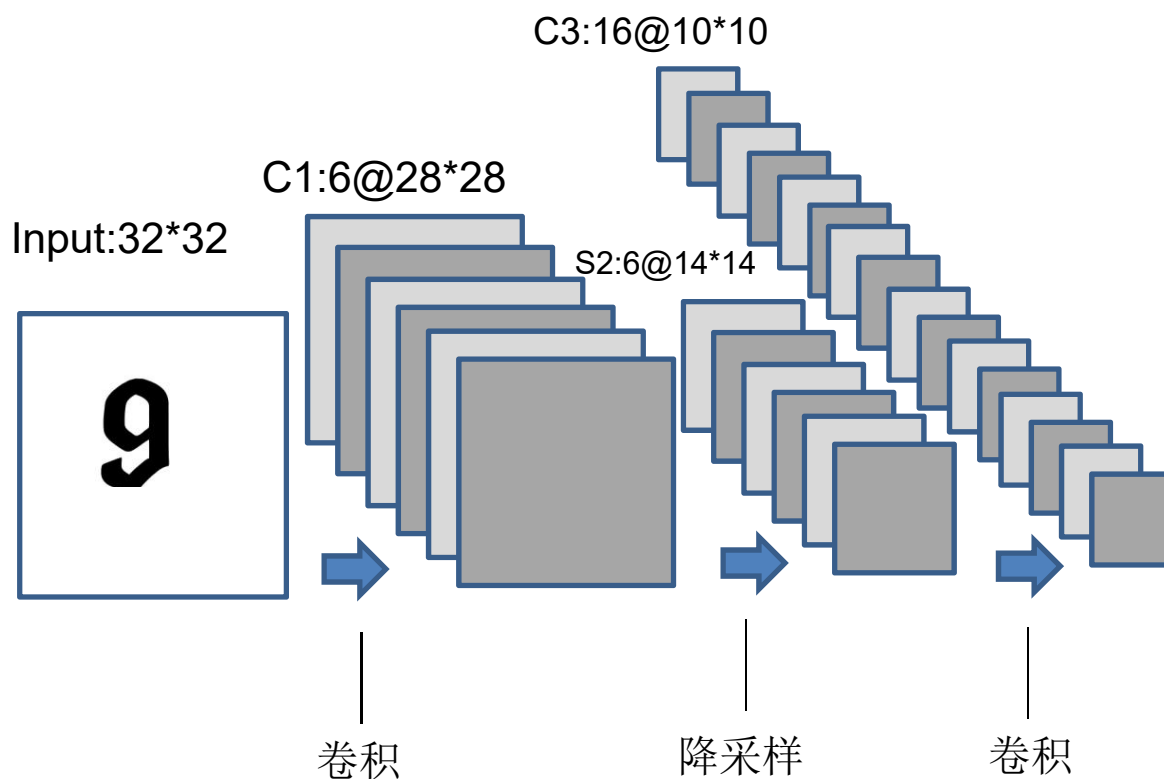
- 卷积核：C1层有6个卷积核（核大小为 $5*5$ ），输出6个大小为 $28*28$ 的特征图，对应6种局部特征。
- 训练参数个数：对于6个卷积核，每个核有 $5*5=25$ 个unit参数和一个bias参数，所以共有 $(5*5+1)*6=156$ 个参数。
- 连接数目：共 $156*(28*28)=122,304$ 个连接。

LeNet-5--降采样层S2



- S2是降采样层（downsampling），目的是降低网络训练参数及克服模型过拟合。通常有两种模式：Max-Pooling或者Mean-Pooling；
- S2层有6个14*14的feature map，其每个单元与上一层的 2*2 领域连接（滑动窗口为2*2）。所以S2层的size是C1层的1/4。

LeNet-5--卷积层C3

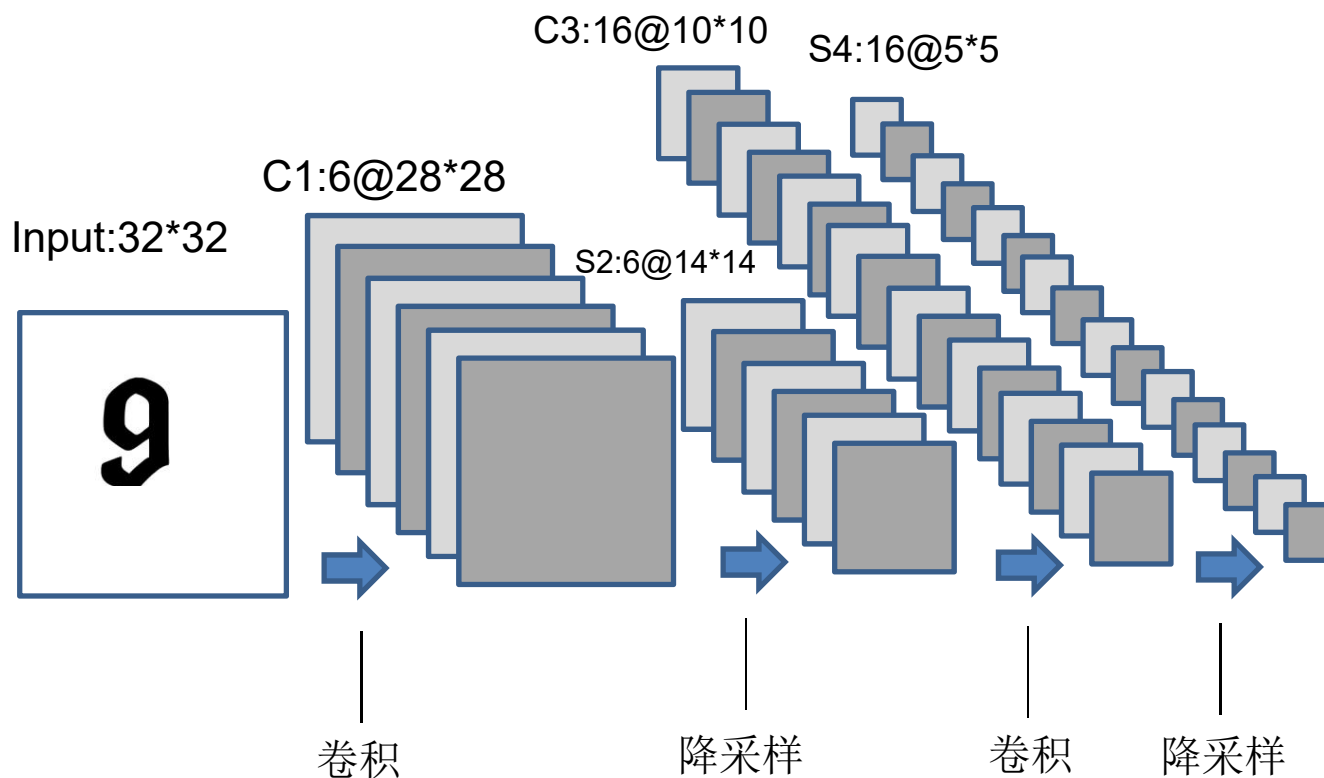


●**C3**卷积层：每个特征图只与上一层S2中部分特征图相连接。

●**C3**卷积层：多通道16核卷积，有16个卷积核，其大小为5*5，输出16个特征图，每个大小为 $(14-5+1) \times (14-5+1) = 10 \times 10$ 。

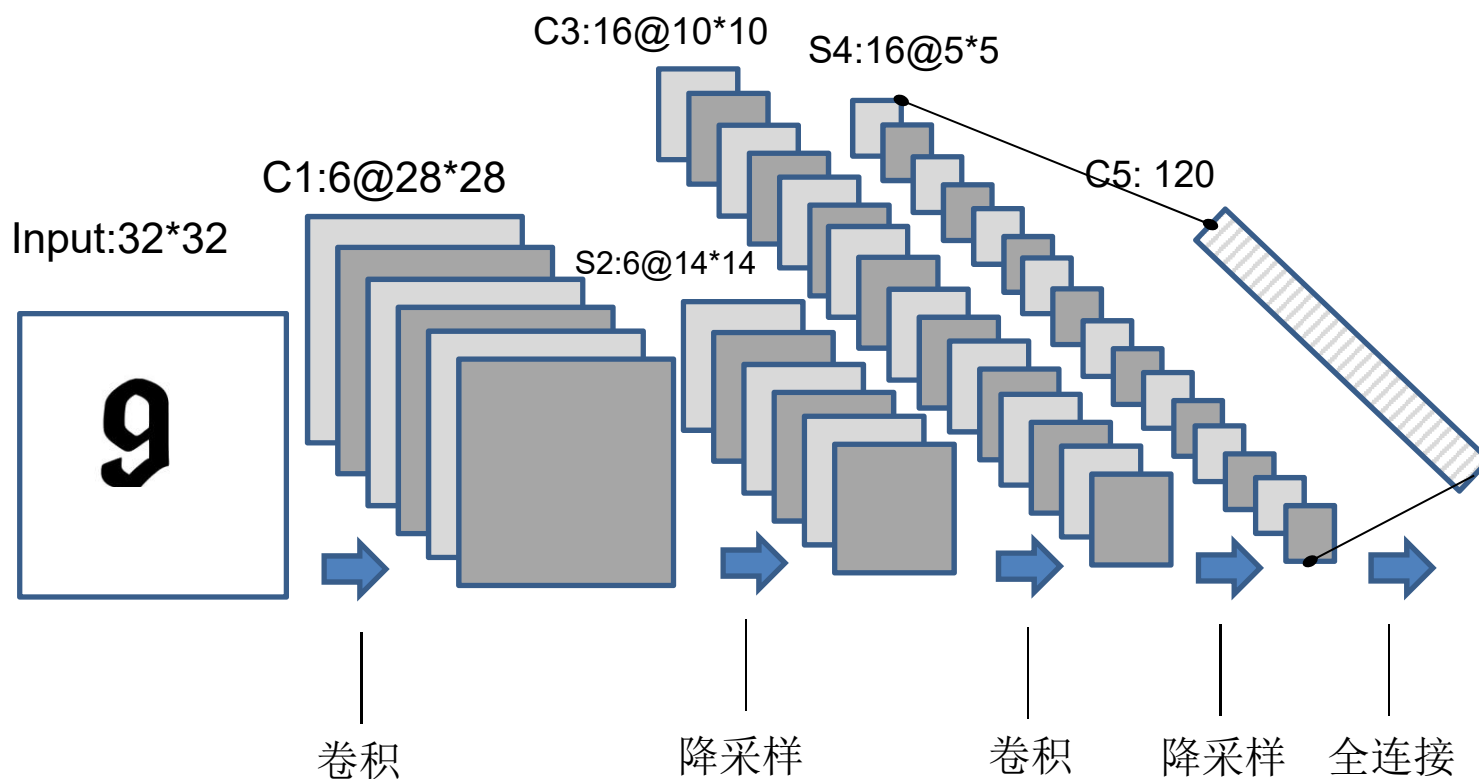
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X			X	X	X	X			X		X	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

LeNet-5--降采样层S4



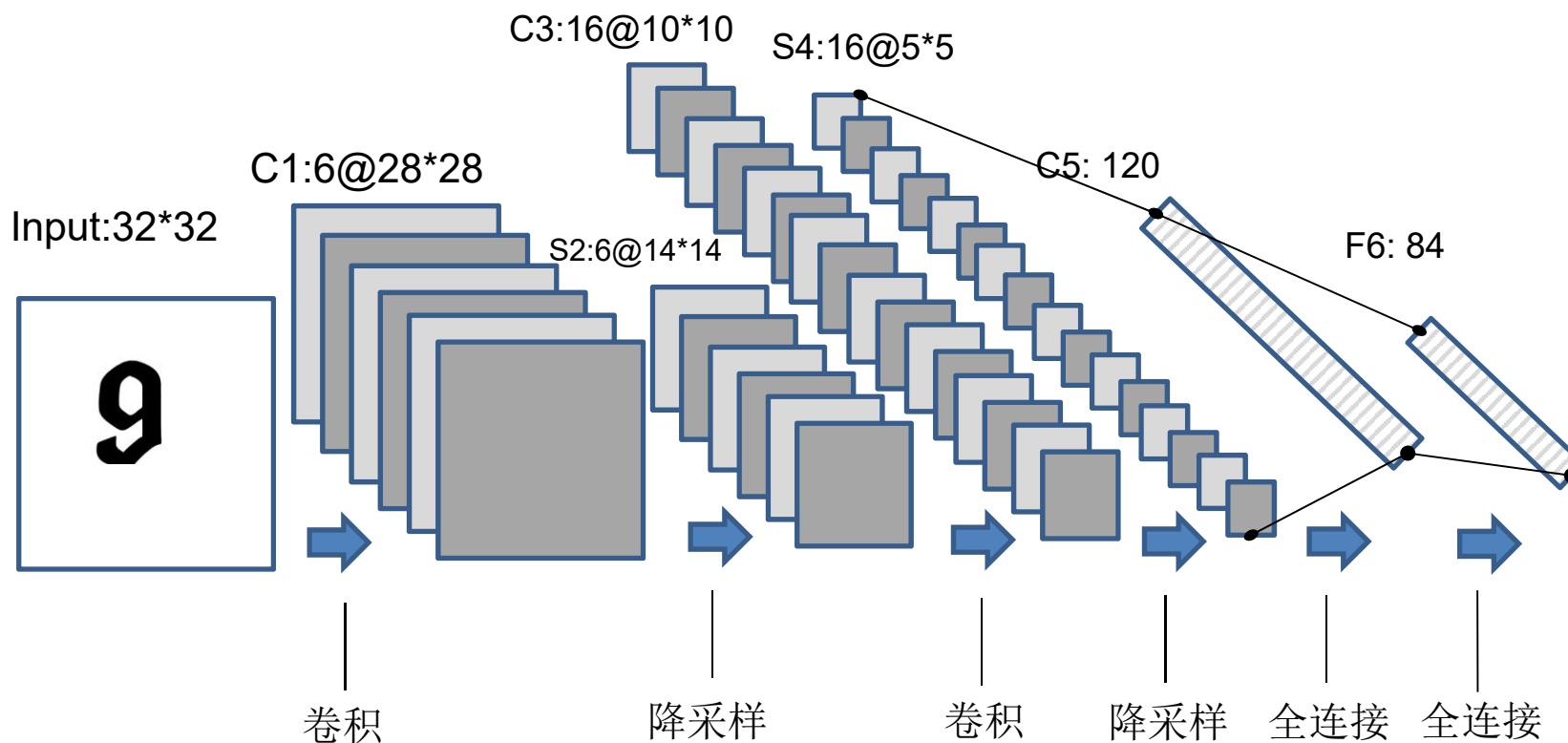
- S4是降采样层(pooling层)。
- S4层有16个5*5的特征图，其每个单元与上一层的 2*2 邻域连接（滑动窗口为2*2）。所以S4层的size是C3层的1/4。

LeNet-5--卷积层C5



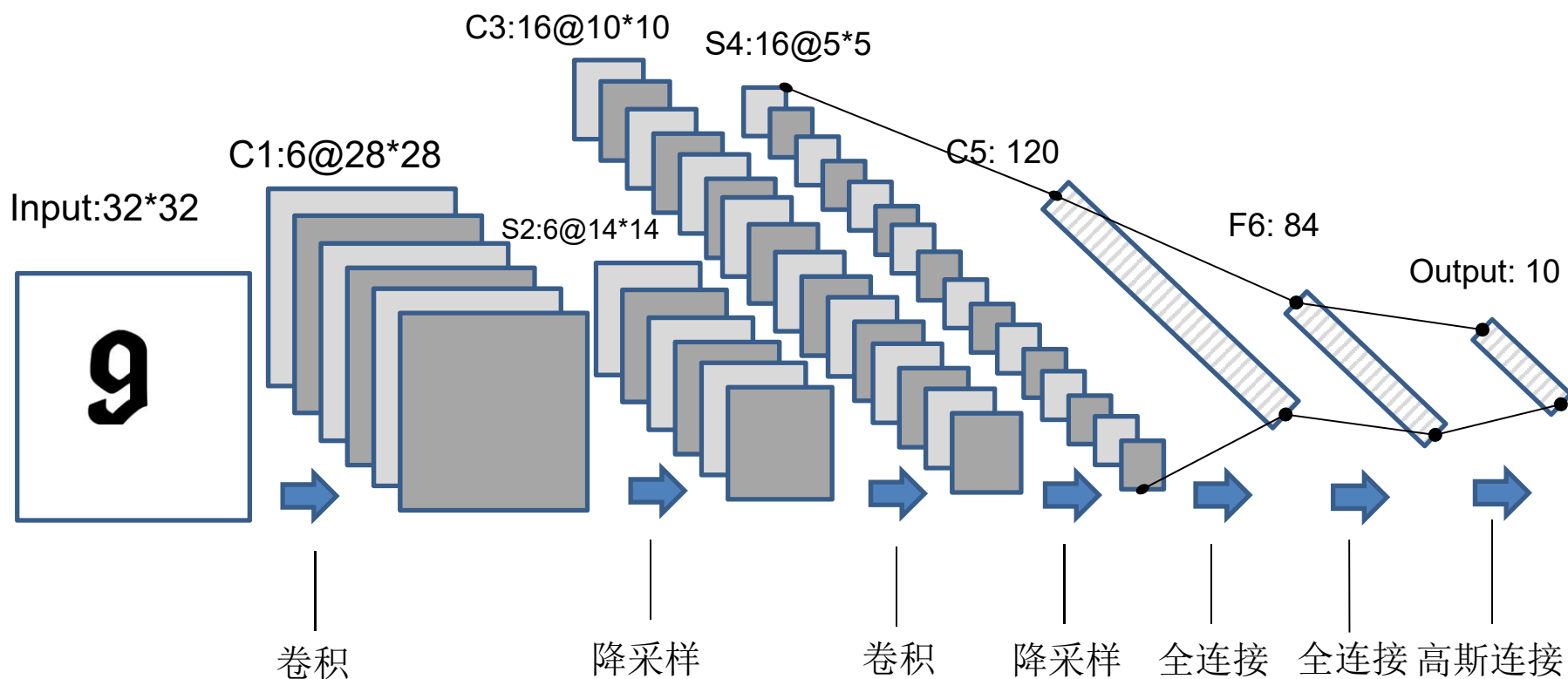
- 卷积层C5有120个卷积核，卷积核大小为 5×5 ，因此生成120个特征图，每个特征图的大小为 1×1 。
- 每个feature map都与上一层S4的所有特征图连接。C5层有 $120 \times (5 \times 5 \times 16 + 1) = 48120$ (16为上一层所有的特征图个数)个连接。

LeNet-5的全连接层



- F6是全连接层，类似MLP多层神经网络中的一个layer，共有84个神经元。这84个神经元与C5层进行全连接，所以需要训练的参数是： $(120+1)*84=10164$ 。
- 如同经典神经网络，F6层计算输入向量和权重向量之间的点积，再加上一个偏置。然后将其传递给sigmoid函数将数值规范化。

LeNet-5的输出层



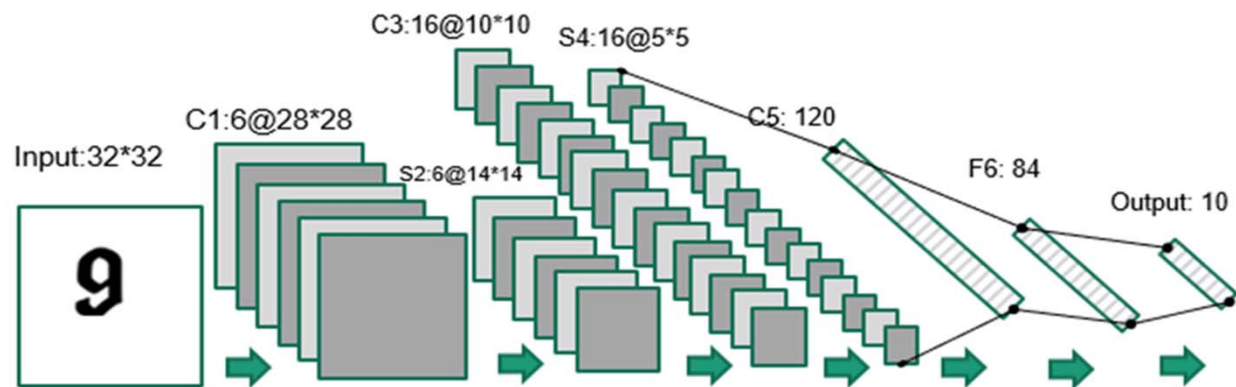
●Output 输出层由欧式径向基函数（Euclidean Radial Basis Function）单元构成。每个输出RBF单元计算输入向量和参数向量之间的欧式距离。

$$y_i = \sum_{j=1}^{84} (x_j - w_{ij})^2$$



使用RBF识别ASCII字符集的初始参数 $\angle U$

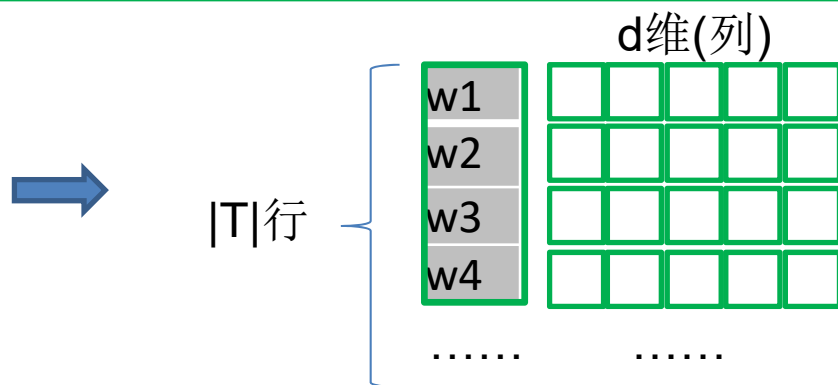
文本信息的矩阵化处理



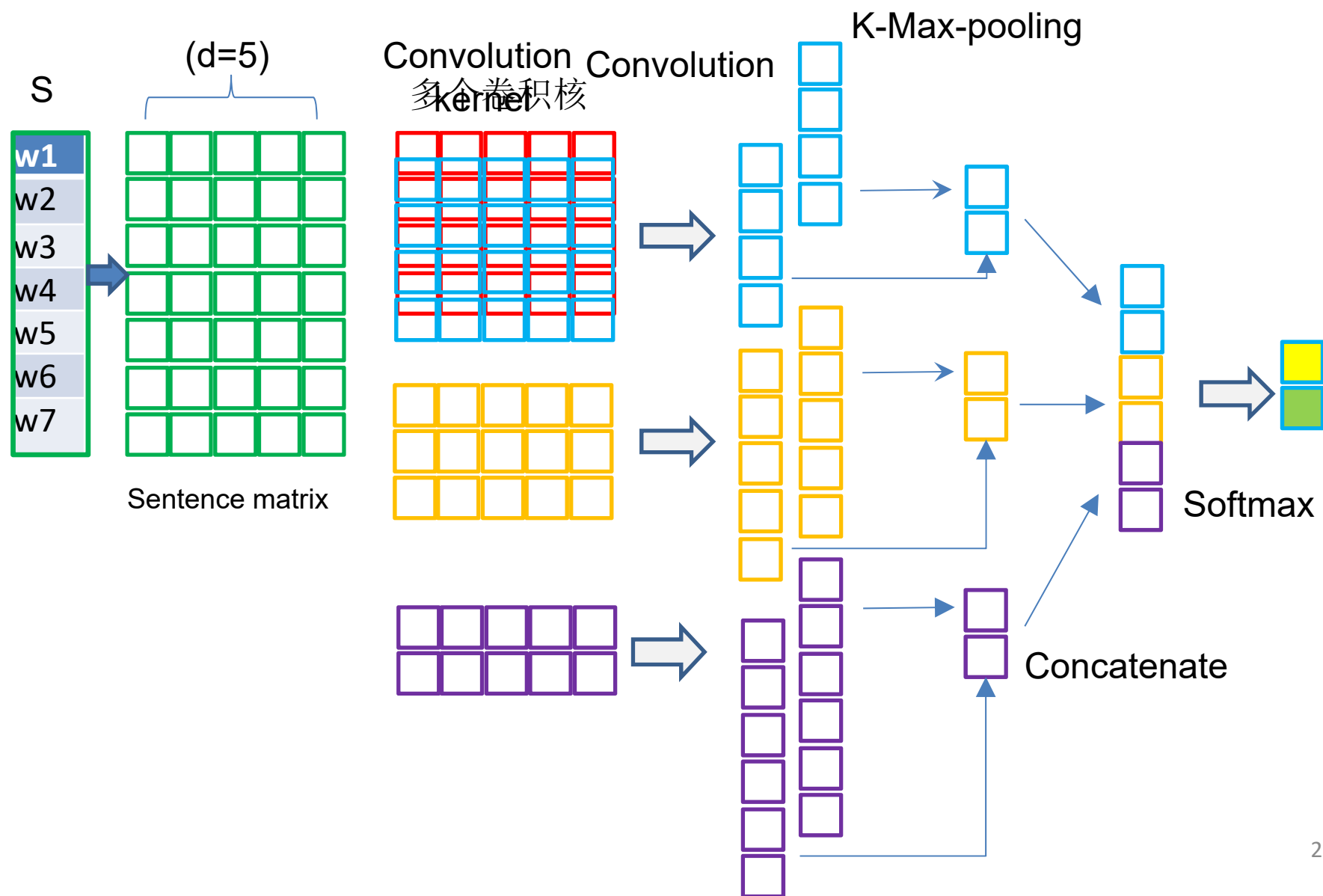
- 如何对无结构的文本信息进行卷积处理？

去过几次，感觉还好吧，也是没有以前好了，现在是吃海鲜的好季节，有时间也是还要去大快朵颐一番的。现在来说这个价格能吃的很丰盛还是不错的。有机会大家都可以去尝尝，环境卫生还有美食都还不错。赞，赞，赞！来来来！

- 假设文本T长度为 $|T|$ ，则文本也可以转化为矩阵形式。



基于CNN的文本分类



CNN代码学习笔记

- 手写数字图像分类

- https://github.com/wepe/MachineLearning/tree/master/DeepLearningTutorials/cnn_LeNet
- 基于tensorflow
- MNIST手写数字数据库

- 文本分类

- https://github.com/yoonkim/cnn_sentence
- 基于Theano
- Yoon Kim. [Convolutional Neural Networks for Sentence Classification](#). New York University. EMNLP 2014. pp. 1746--1751

Thanks!

