



卷积神经网络

Convolutional Neural Network

冯禹豪 2020/10/23



一个场景

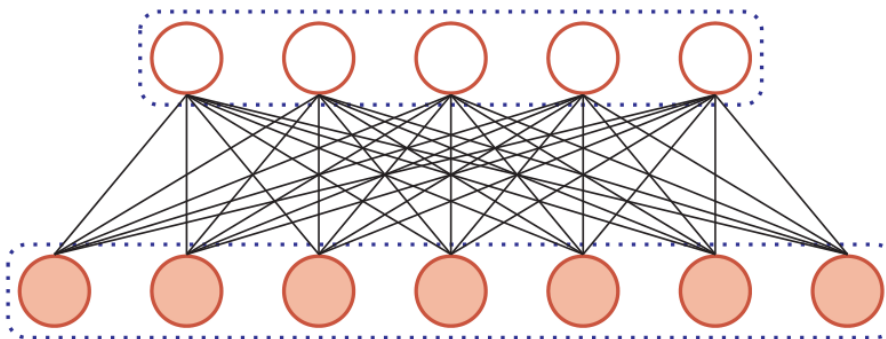


给你一个图片分类的数据集，要怎么做？



一个场景

假设一张图片是 226×226 ，然后有RGB三通道，那么拼成一个向量就是的维度就是153228。假设第一层全连接层的节点数是500，那么权重矩阵大小就是 $[153228, 500]$ ，这个参数量是比较惊人的。





一个场景

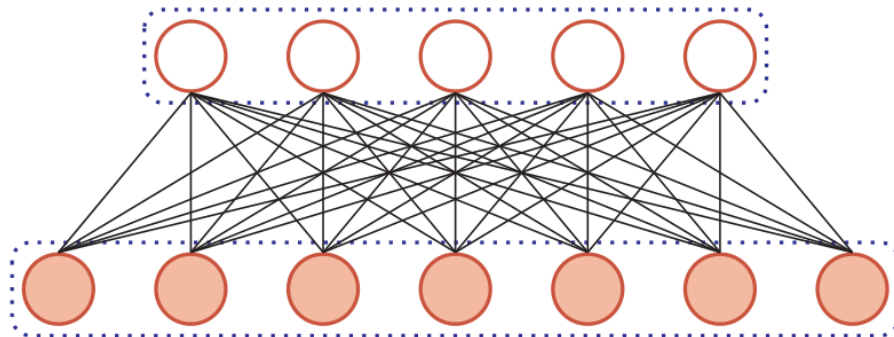
而且图像还具有很多不同的特性，例如局部不变性。尺度缩放、平移、旋转等操作不影响其语义信息。所以单纯的拼接成向量就会破坏掉图像的这种特性，例如破坏掉局部的轮廓、关联等，所以往往效果不好。





全连接神经网络的缺点

- 权重矩阵的参数非常多



- 无法提取局部不变性特征
 - 尺度缩放、平移、旋转等操作不影响其语义信息



卷积神经网络 (CNN)

- 基于生物学上感受野 (Receptive Field) 的机制

感受野 (Receptive Field) 主要是指听觉、视觉等神经系统中一些神经元的特性，即神经元只接受其所支配的刺激区域内的信号。在视觉神经系统中，一个神经元的感受野是指视网膜上的特定区域，只有这个区域内的刺激才能够激活该神经元。

- 卷积神经网络有三个结构上的特性

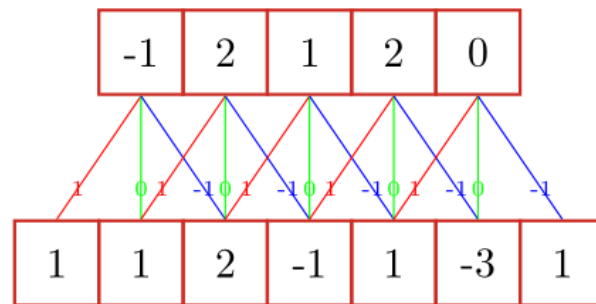
- 局部连接
- 权重共享
- 空间上的下采样



卷积

- 一维卷积经常用在信号处理中，用于计算信号的延迟累积。
- 假设一个信号发生器每个时刻 t 产生一个信号 x^t ， k 个时间后的信息衰减率为 w^k ，则当前信号输出 y_t 为：

$$y_t = \sum_{k=1}^m w_k \times x_{t-k+1}$$



Filter: [-1,0,1]



卷积

- 二维卷积经常用在图像处理中。因为图像通常为一个二维结构，所以需要将一维卷积进行扩展。

$$y_{i,j} = \sum_{u=1}^m \sum_{v=1}^n w_{u,v} \times x_{i-u+1,j-v+1}$$

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

4		



卷积

- 实际上二维卷积已经广泛地被用在图像处理中，可以起到特征提取的作用。
- 能不能不通过先验经验构建滤波器，而是让模型自动学习如何提取特征？



原始图像

高斯平滑

$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

=



拉普拉斯算子

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \otimes$$

=



Robinson算子

0	1	1
-1	0	1
-1	-1	0

=



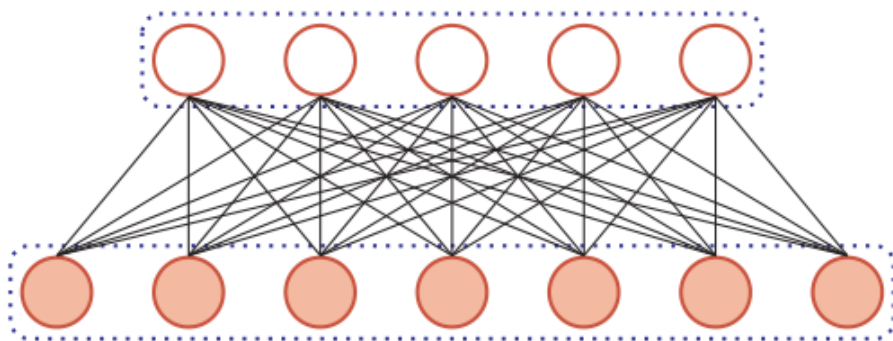
滤波器

输出图像

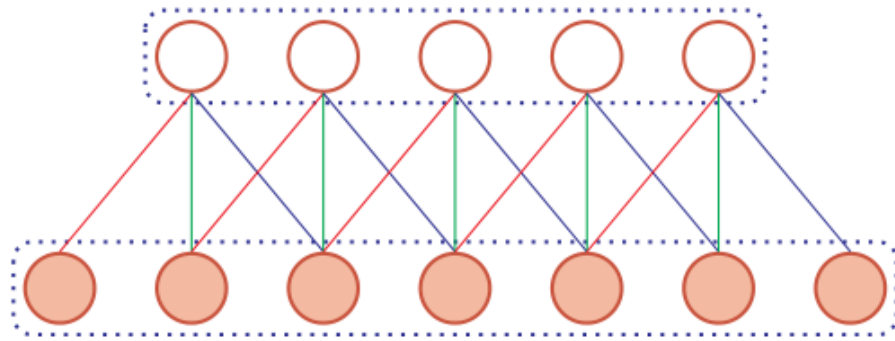


卷积层

- 用卷积层代替原来的全连接层。
 - 局部连接
 - 权重共享



(a) 全连接层



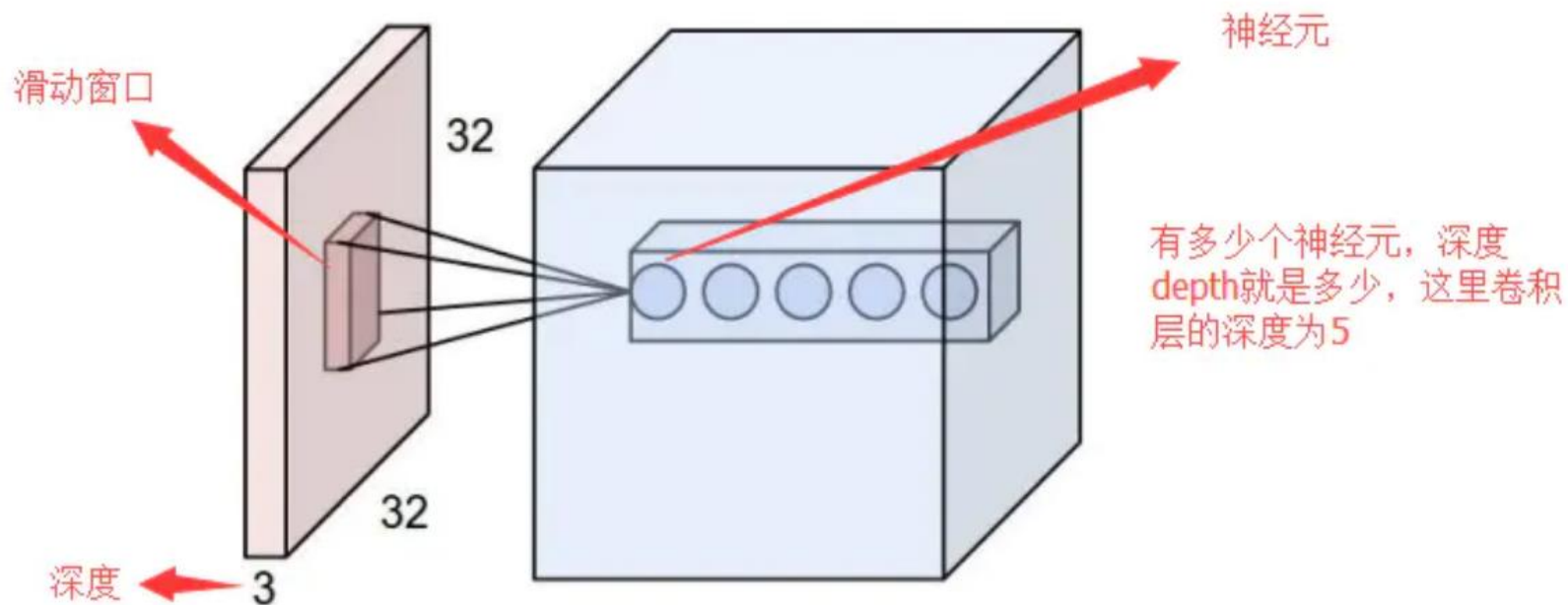
(b) 卷积层



卷积层

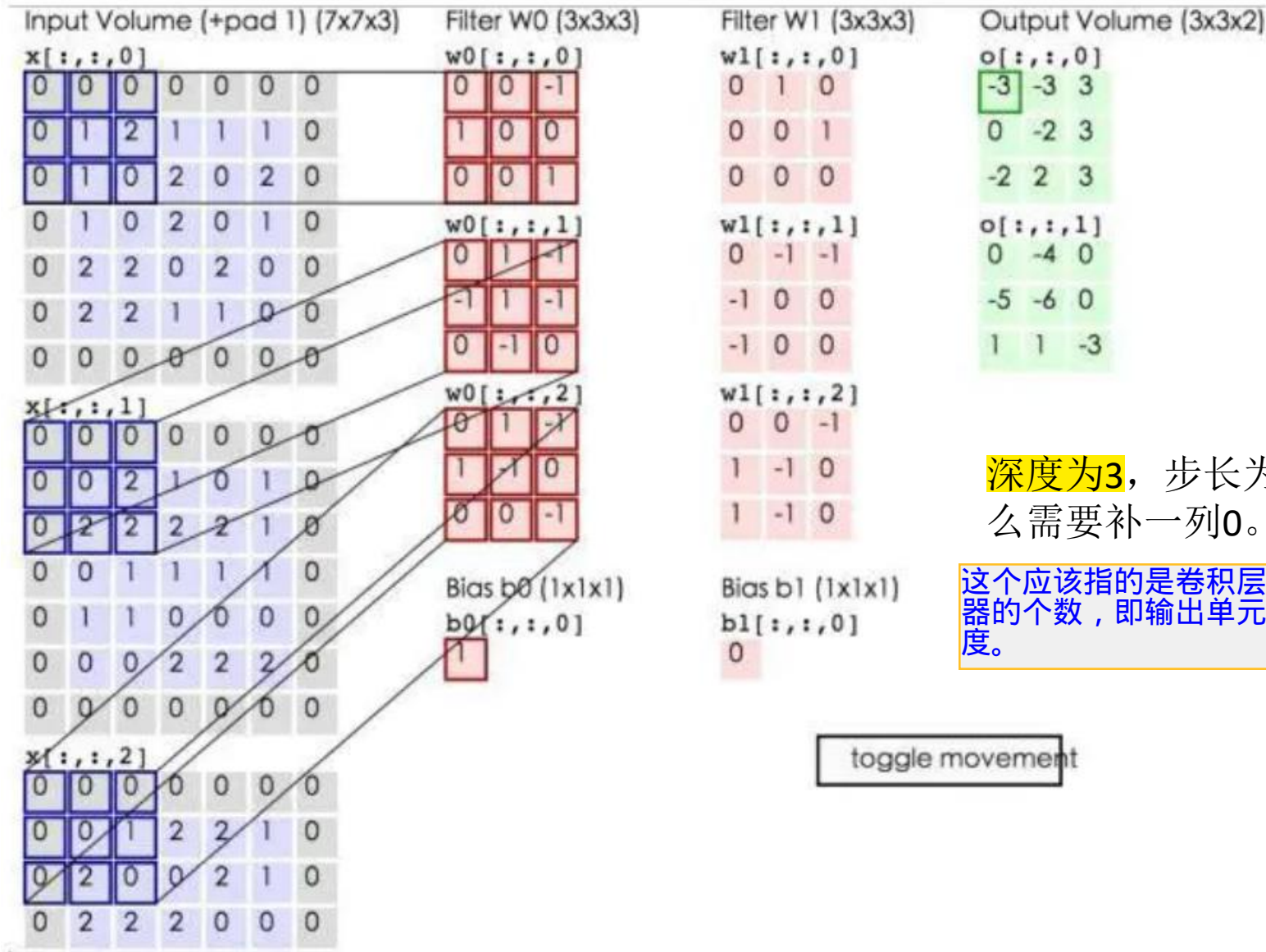
- 卷积层的结构。
 - 深度 depth
 - 步长 stride
 - 填充值 zero-padding

orange	orange	blue	blue	white
orange	orange	blue	blue	white
white	white	white	white	white
white	white	white	white	white
white	white	white	white	white





卷积层



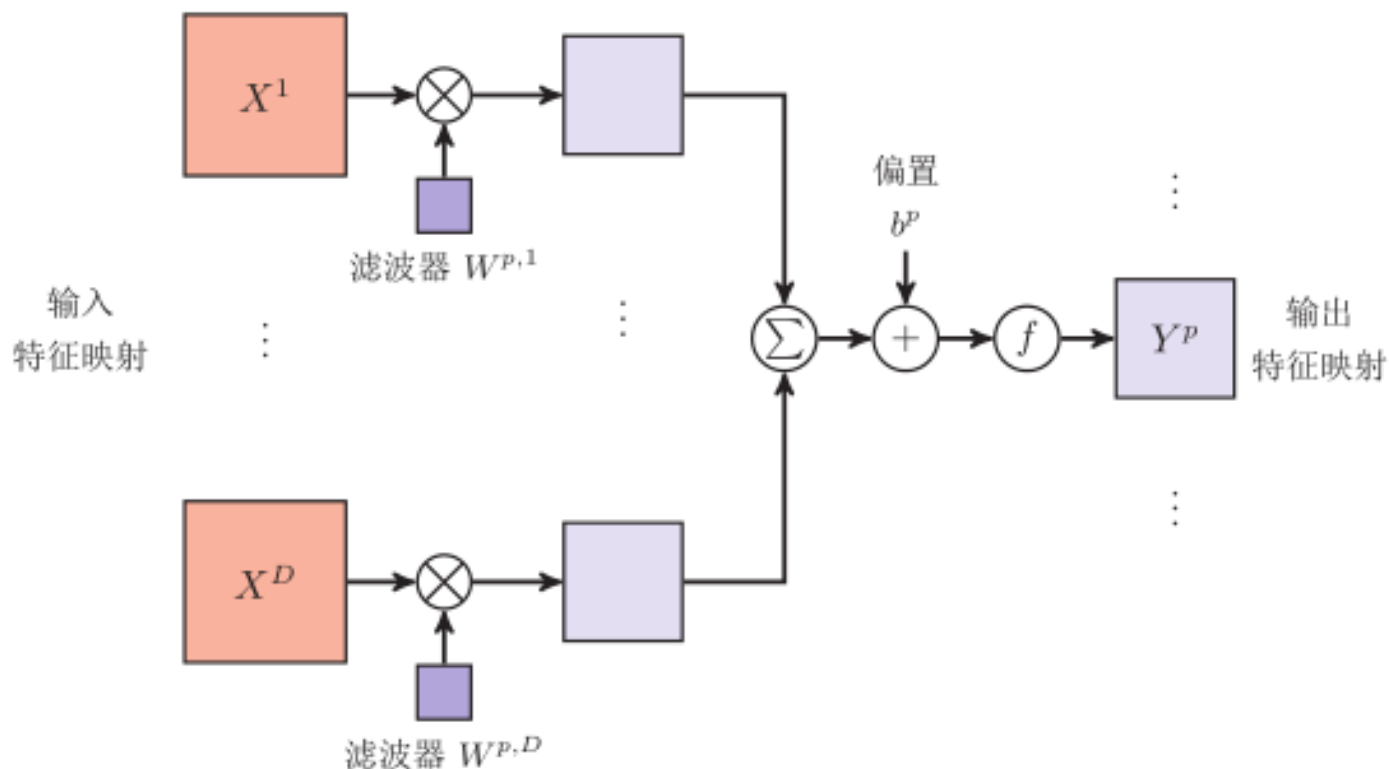
深度为3，步长为2，那么需要补一列0。

这个应该指的是卷积层过滤器的个数，即输出单元深度。



卷积层

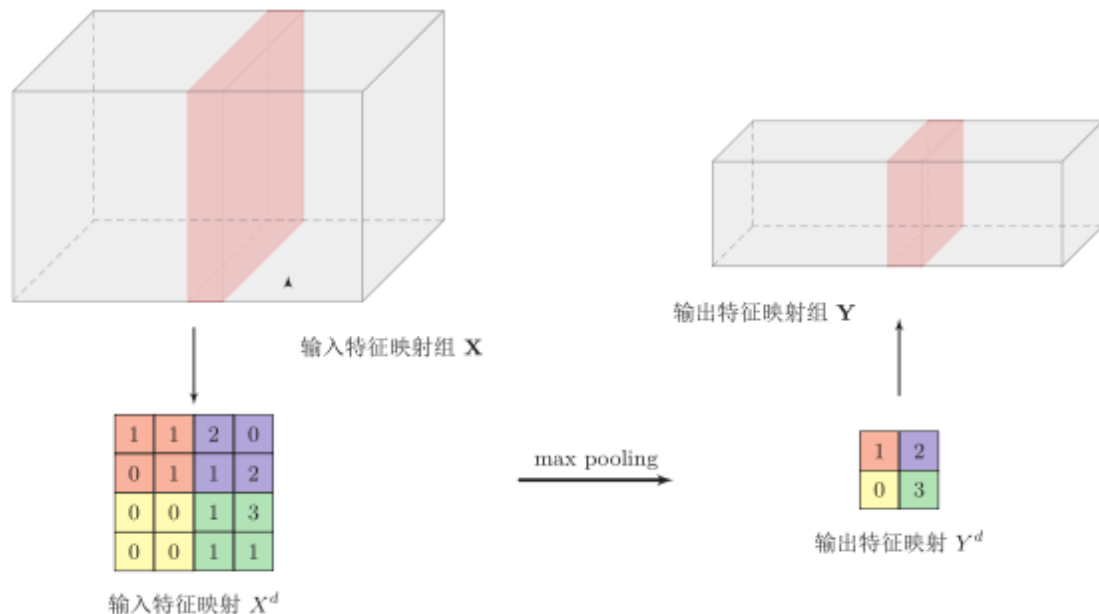
- Y^p 的计算流程：先用卷积核 $W^{p,1}, W^{p,2}, W^{p,3}, \dots, W^{p,D}$ 分别对输入特征 $X^1, X^2, X^3, \dots, X^D$ 卷积，然后将卷积的结果相加，并加上一个标量偏置 b^p 得到 Z^p ，最后通过某个激活函数得到输出 Y^p





池化层

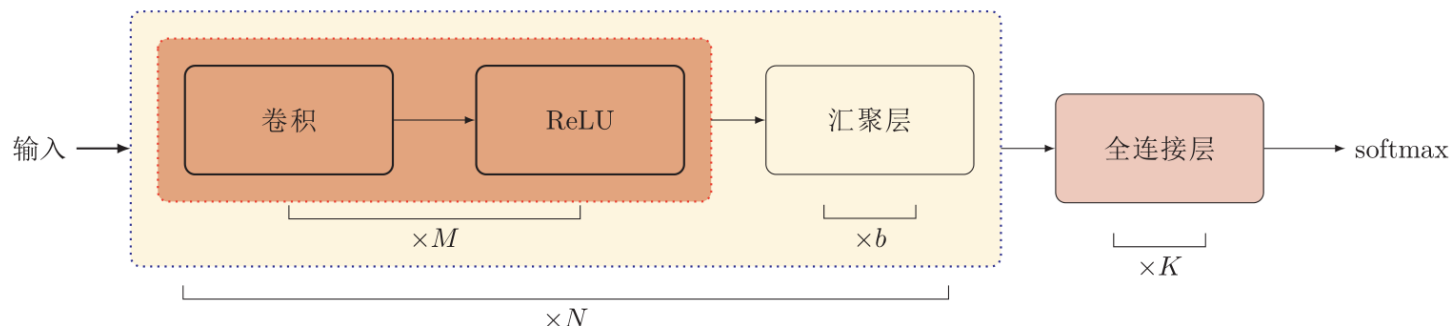
- 卷积层虽然可以显著减少网络中参数的数量，但特征映射输出的维度仍然很高。
- 池化层对输入特征映射组进行下采样，进一步的筛选特征，降低特征数量，并从而减少参数数量，缩减模型的大小，提高计算速度，同时提高所提取特征的鲁棒性。
- 最常用到的池化类型是最大池化（max pooling），大小与步长均设置为2，效果相当于原来的特征高度和宽度缩减一半。





卷积神经网络结构

- 由卷积层、池化层、全连接层交叉堆叠而成。
- 典型结构

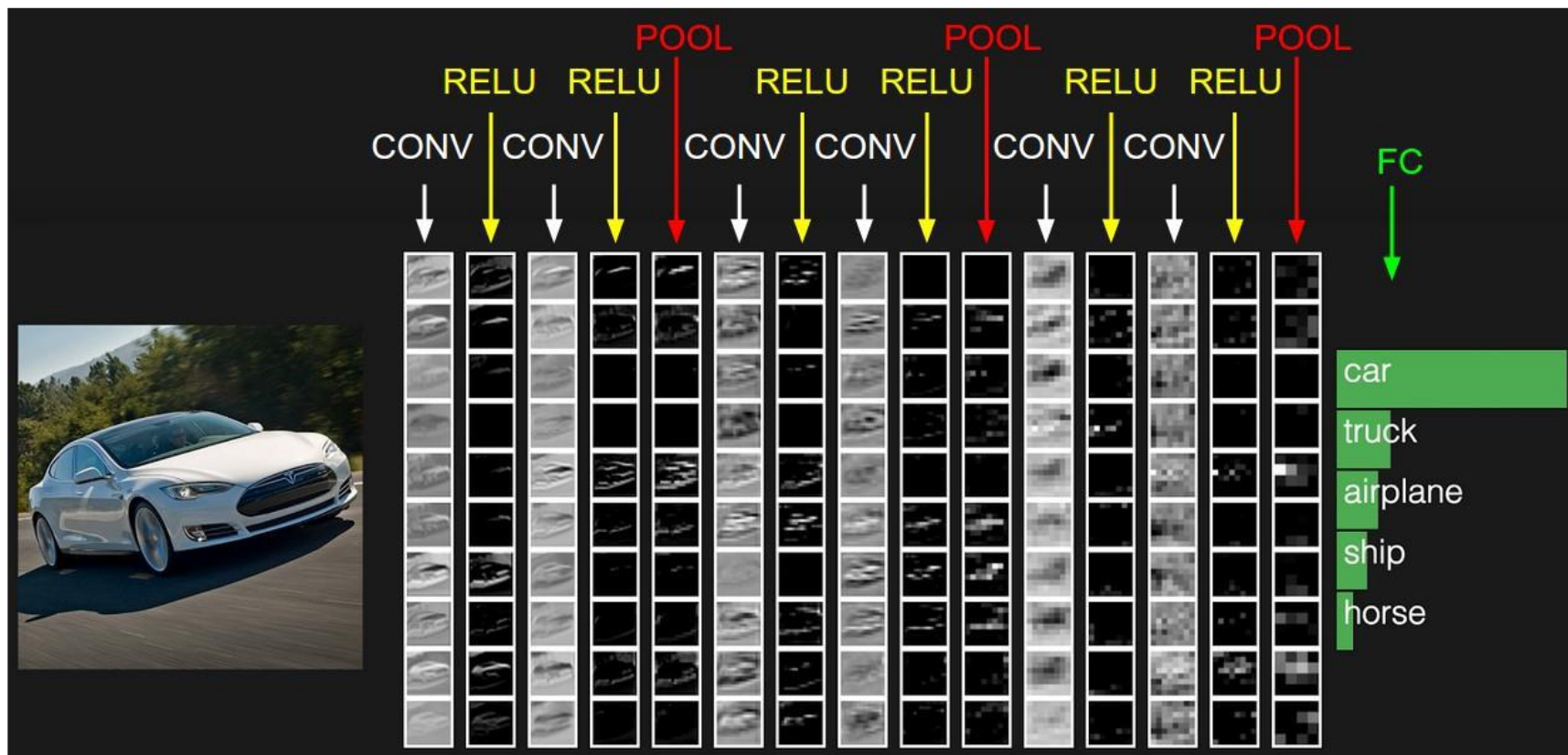


- 定义一个卷积块为连续M个卷积层和b个池化层（M通常设置为2~5，b为0或1）。
- 一个卷积网络中首先堆叠N个连续的卷积块（N的取值区间比较大，比如1~100或者更大）
- 最后再连接K个全连接层（K一般为0~2，相当于最后接一个分类器）



卷积神经网络结构

- 由卷积层、池化层、全连接层交叉堆叠而成。

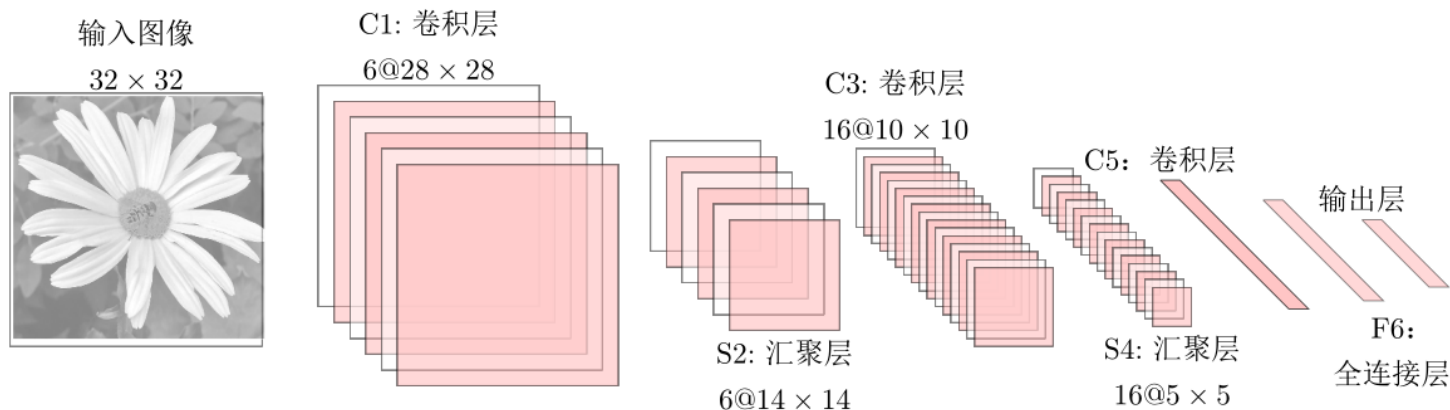




卷积神经网络应用

- LeNet-5

- 基于 LeNet-5 的手写数字识别系统在 90 年代被美国很多银行使用，用来识别支票上面的手写数字。





卷积神经网络应用

- ImageNet

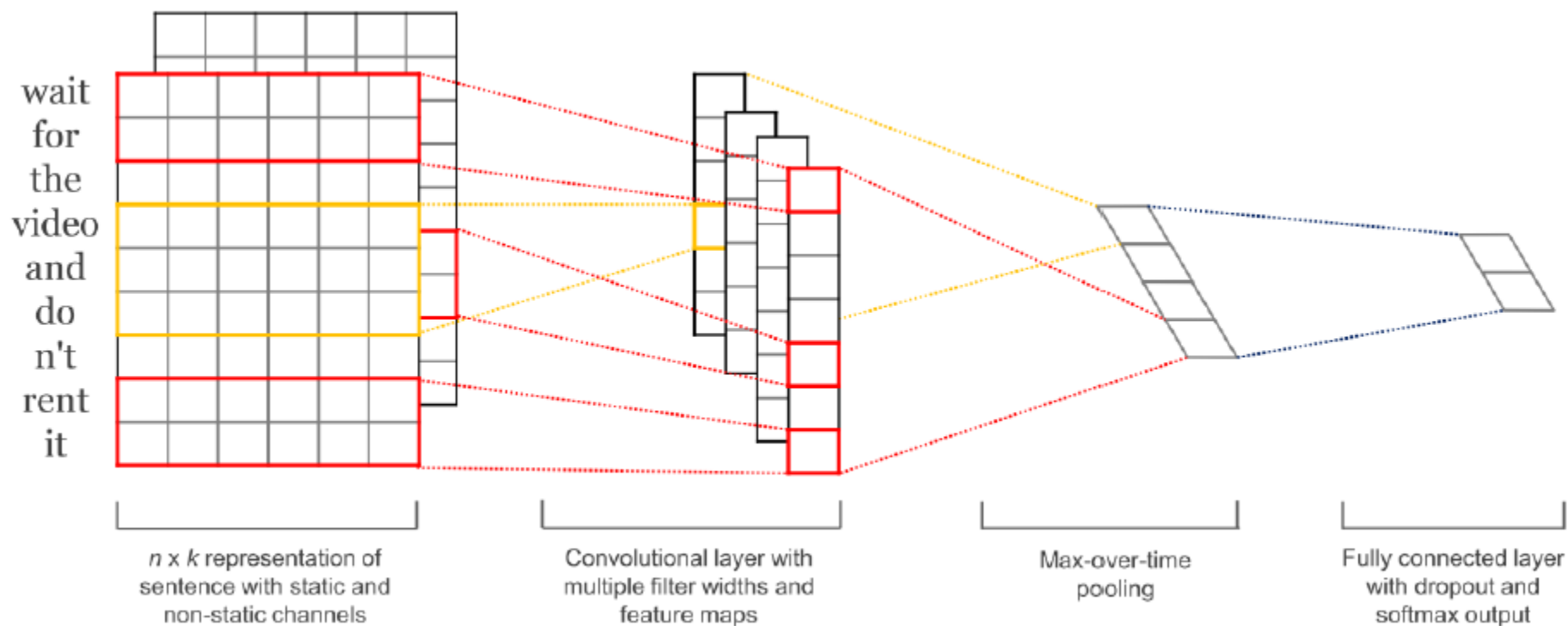
14197122个图像, 总共分为21841个类别(synsets), 大类别包括:
amphibian、animal、appliance、bird、covering、device、fabric、
fish、flower、food、fruit、fungus、furniture、geological formation、
invertebrate、mammal、musical instrument、plant、reptile、sport、
structure、tool、tree、utensil、vegetable、vehicle、person。



卷积神经网络应用

- 文本领域

- 输入为某个句子所有词的词嵌入向量组成的矩阵。





期中project

- 包含两个任务
 - 实现CNN
 - 实现RNN
- 两个数据集
 - CIFAR-10
 - SemEval
- 分组完成，每组两个人



期中project

- 报告评分:

- 每组同学共同完成一份报告。
- 参考的论文要以参考文献呈现出来，参考的代码要有标记来源。
- RNN和CNN各占50分。
- DDL下周公布

评分项	说明	分值
实验原理	总结两种模型的原理	20
网络结构	画出自己模型的网络结构示意图	10
结果分析	展示并分析不同结构下的实验结果	40
创新	可以借鉴现有方法，但需总结原理	20
排版	整体美观性	10
组员分工	总结组员各自做了什么工作	0



附录

源图片下载地址:

<http://www.cs.toronto.edu/~kriz/cifar.html>

参考资料(要善于自己找资料):

<https://www.cnblogs.com/Jerry-Dong/p/8109938.html>

https://blog.csdn.net/mao_hui_fei/article/details/89477938

也可以从深度学习框架里面用api读取, 参考:

<https://www.jianshu.com/p/4ed7f7b15736>