第六讲
卷积神经网络
Lecture 6
Convolutional Neural Network

明玉瑞 Yurui Ming

yrming@gmail.com

声明
Disclaimer

- 本讲义在准备过程中由于时间所限，所用材料来源并未规范标示引用来源。所引材料仅用于教学所用，作者无意侵犯原著者之知识版权，所引材料之知识版权均归原著者所有；若原著者介意之，请联系作者更正及删除。

  The time limit during the preparation of these slides incurs the situation that not all the sources of the used materials (texts or images) are properly referenced or clearly manifested. However, all materials in these slides are solely for teaching and the author is with no intention to infringe the copyright bestowed on the original authors or manufacturers. All credits go to corresponding IP holders. Please address the author for any concern for remedy including deletion.

回忆：交叉熵
Recall: Cross-Entropy

- 特征脸：一张是脸的图片之所以会被认成脸，可以认为特定位置的像素的值导致的结果，这些位置像素值可假设是满足一定概率分布的。设$x_{i,j}$表示位置为(i,j)的随机变量，则人脸的这些位置的像素值的联合概率分布可记为$P(\{x_{i,j}\})$。在简化情况下，我们可以认为对这个分布采样，就会得到一张人脸。

  Eigenface: a face of a human being is just being identified as it is, is due to the pixels in specific positions. And these pixel values are hypothesized to satisfy a certain probability distribution. Let $x_{i,j}$ denote the rando of pixel at (i,j), then we have a joint distribution of the pixel values $P(\{x_{i,j}\})$. And by sampling such a distribution, we can get a face.

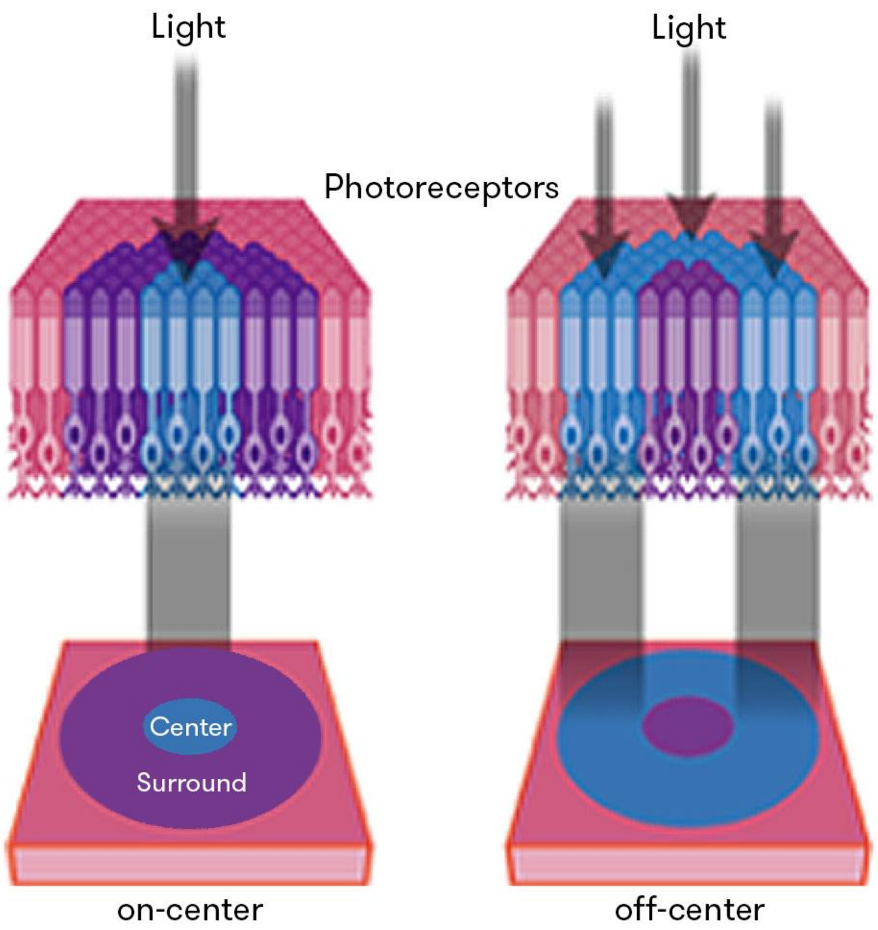
回顾
Retrospect

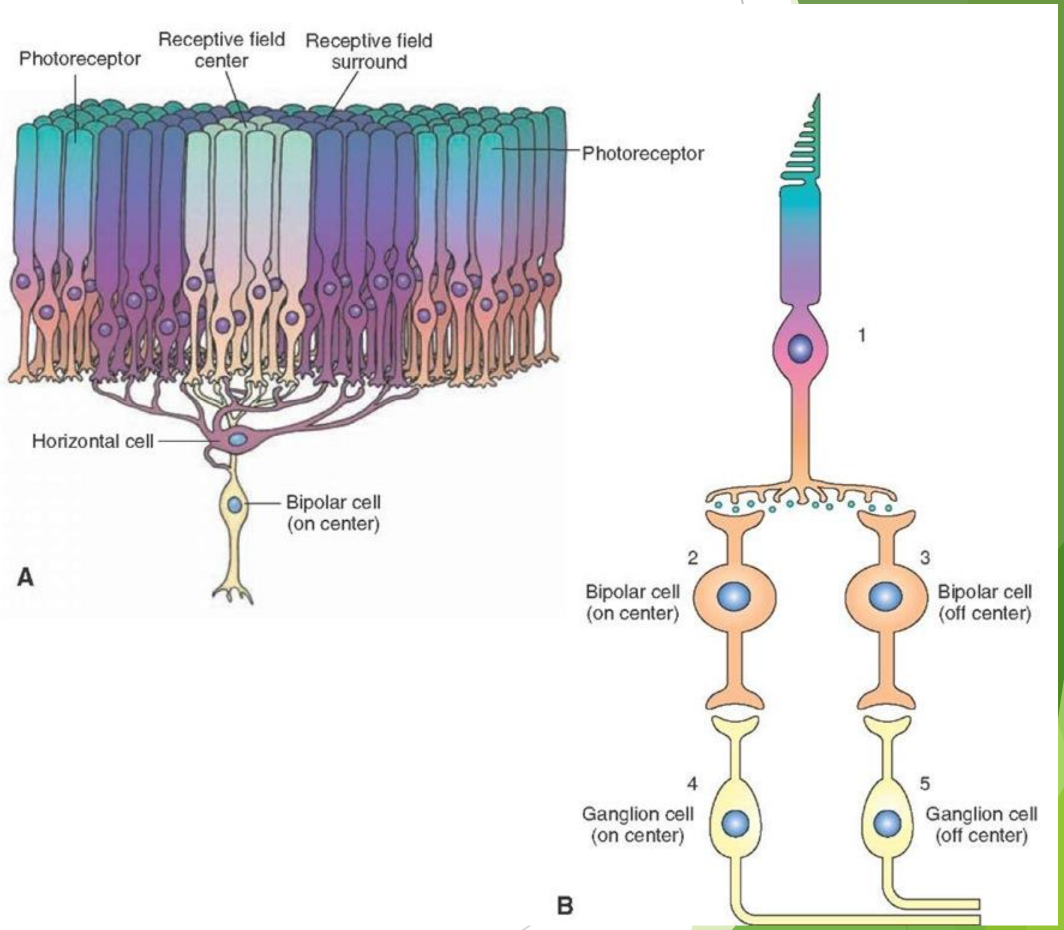
- 为什么简单的拒绝采样，采出的结果很吓人？原因在将每个像素视为独立的随机变量是不行的，因为其和周围像素是相关的

  Why the outcome of reject sampling is so awkward? This is due to the assumption that each pixel is independent of each other. Actually, it correlates to its neighbours.

```
% Load the eigenface as pixel distributions
data = imread('eigenface-gray.PNG','png');

% Convert to single values for each pixel position
dist = rgb2gray(data);
dist = im2double(dist);

% Generate random numbers for reject sampling
pixels = rand(size(dist));

% Perform reject sampling
image = pixels;
image(pixels > dist) = 0;
```

回忆：视野域
Recall: Receptive Field

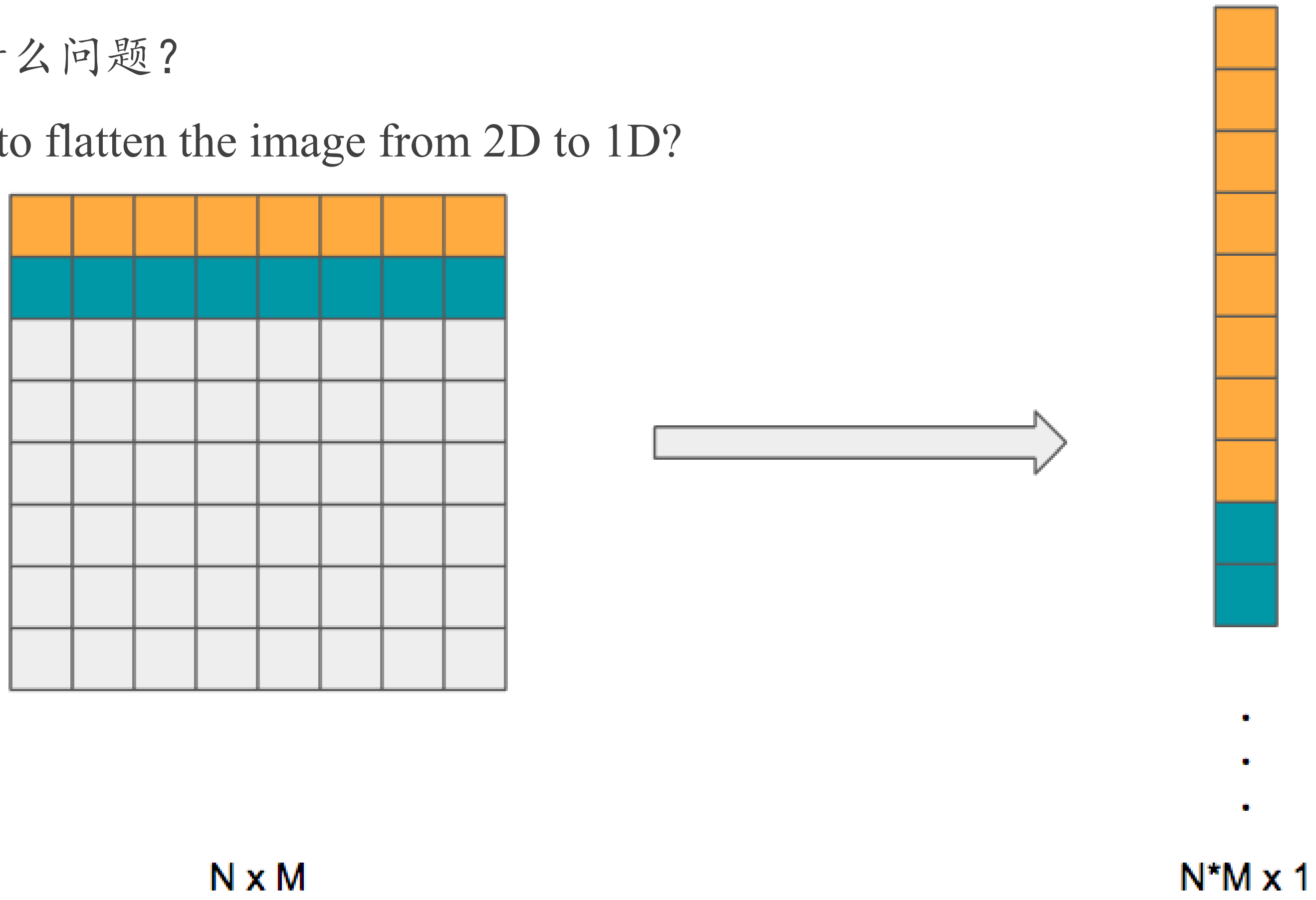




回忆：视野域
Recall: Receptive Field

- 将2D拉成1D会有什么问题？

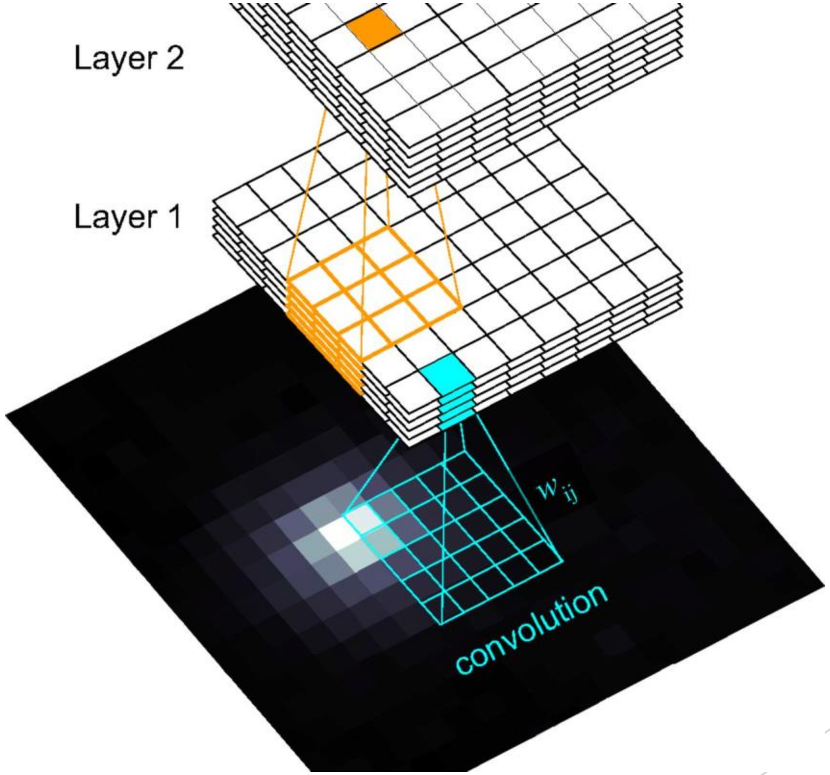
  What’s the problem to flatten the image from 2D to 1D?



卷积
Convolution

- 卷积似乎很自然地解决了上述问题

  The convolution seems solve the above problem naturally.



卷积
Convolution

- 傅里叶定理告诉我们，好多波动函数可由正弦函数合成。

  We draw from Fourier theorem that many wave functions can be synthesized by sinusoidal functions.

- 振动越快，或者变化越快，频率越高；噪声是由高频成分组成的。

  The quicker the vibration or the swifter the change, the higher frequencies. Noise are mainly composed of high frequency components.

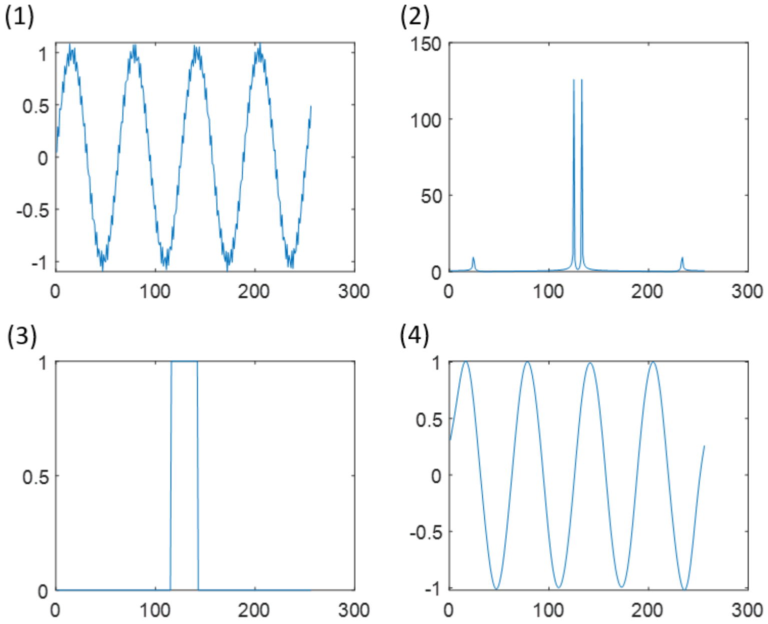
- 去噪将高频成分滤掉就好；具体做法可以是在频域乘以一个窗口函数。

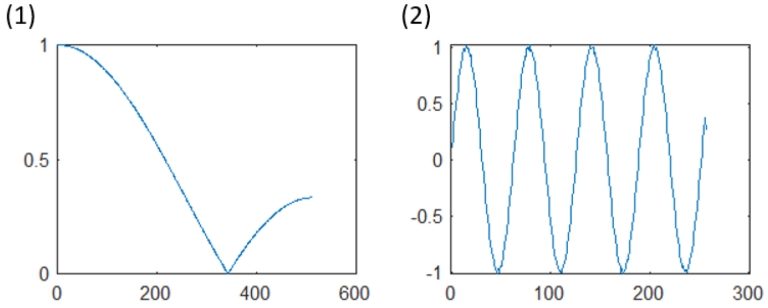
  Denoising can be put into simplicity by removing high frequency components; an implementation could be multiply a window function in the frequency domain.

- 频域上的乘法，等于时域上卷积；与信号所在的维度没有关系。

  Multiplication in the frequency domain is equivalent to the convolution in the time domain, and this is independent of the dimension of the signals.

卷积
Convolution

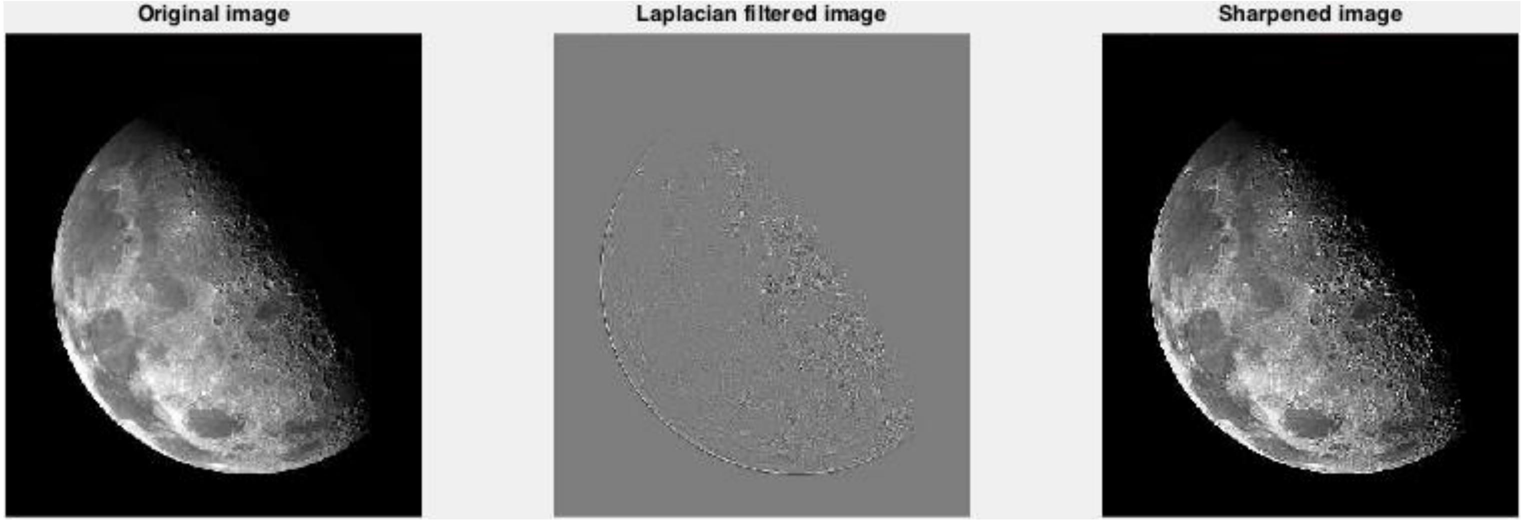




卷积
Convolution

- 图像处理中的许多问题也是藉由卷积来做的，比如图像增强。

  Many operations of digital image processing are also accomplished by convolution, for example, image enhancement.



卷积
Convolution

- 计算机视觉中的许多问题（特别是前置处理）也是藉由卷积来做的，比如图像识别。

  Many operations of computer vision (especially pre-processing) are also accomplished by convolution, for example, image recognition.

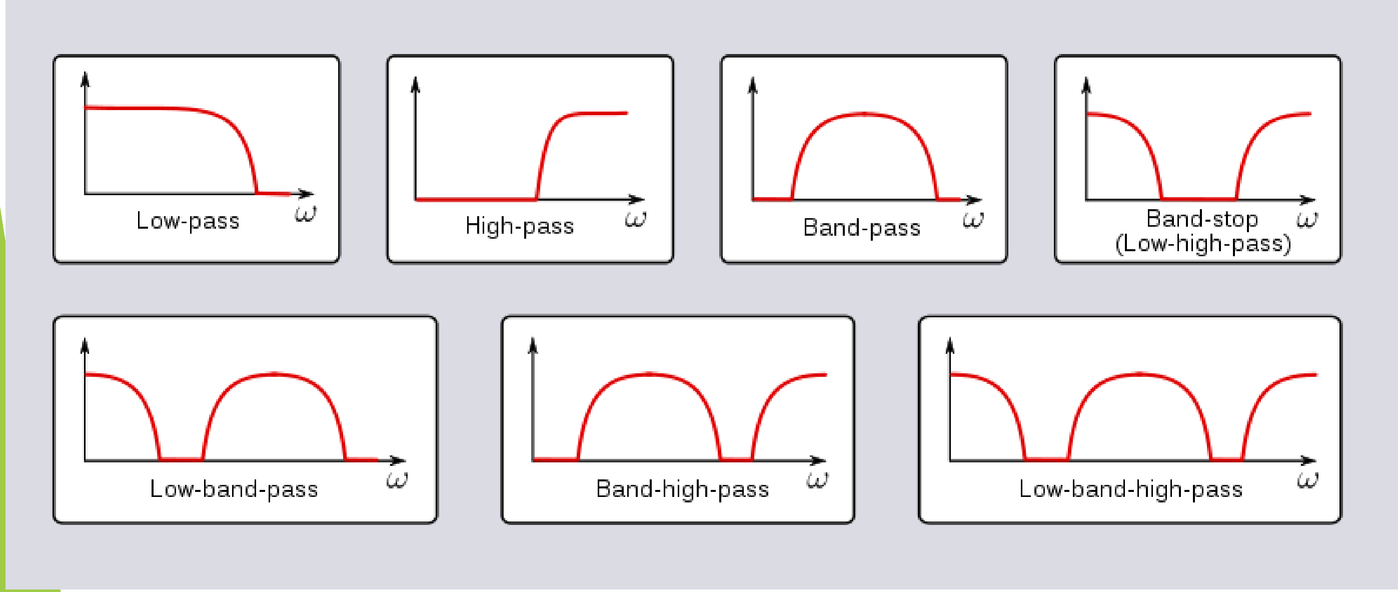
一个大小为1x1x3x1的滤波器:
[[[[0.2989], [0.5870], [0.1140]]]]

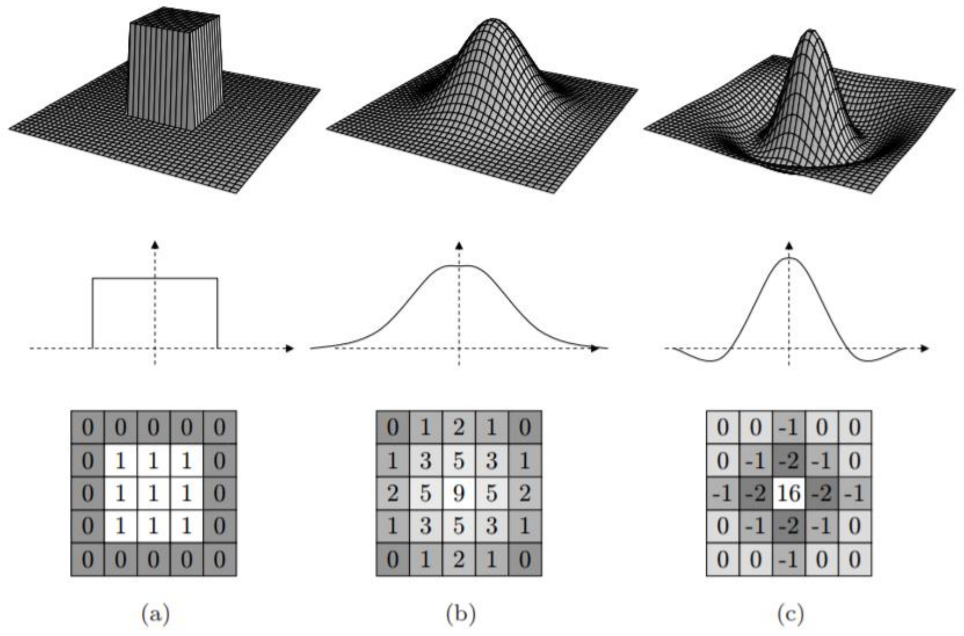
一个大小为3x3x1x1滤波器:
[[[[0.1667]], [[0.6667]], [0.1667]]],
[[[0.1667]], [[-3.3333]], [0.1667]]],
[[[0.1667]], [[0.6667]], [0.1667]]]]

滤波器
Filters

- 和卷积密切相关的一个概念叫滤波器，也叫模板或核

  The concept closed related to convolution is called filter, or mask, or kernel. The convention depends on the underlying domains.

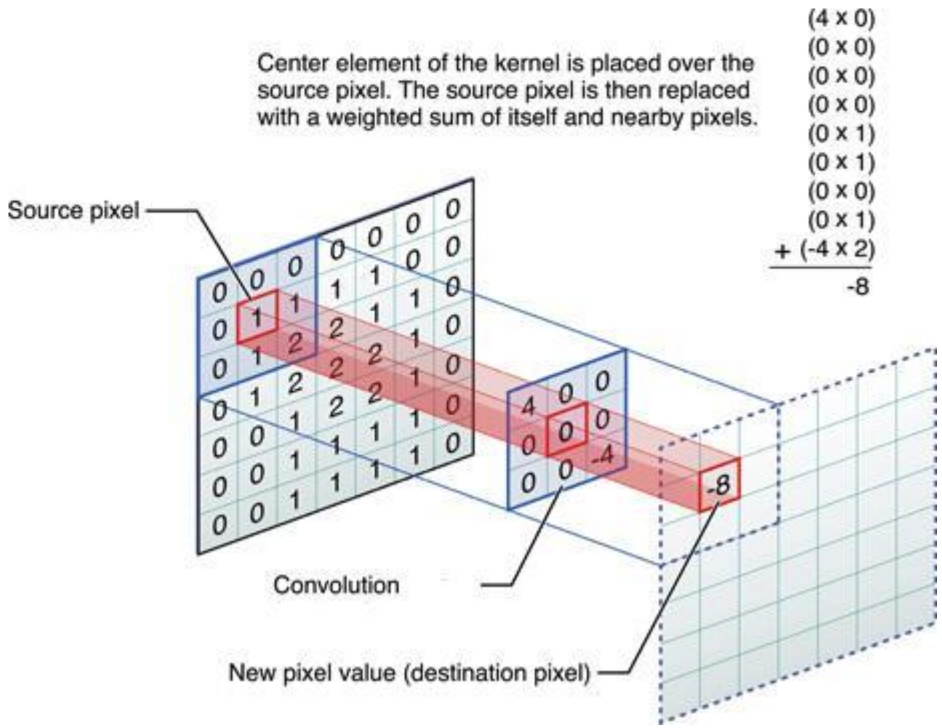


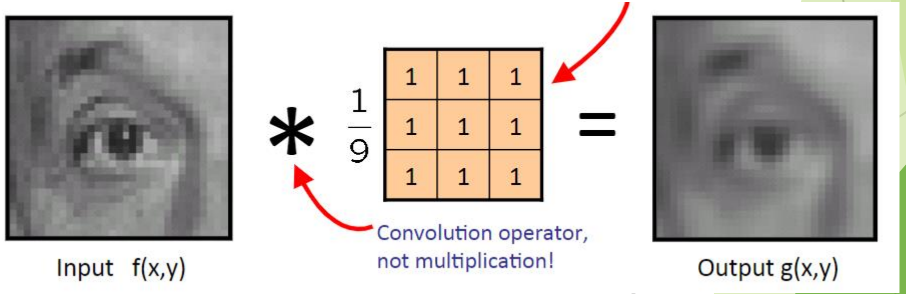


滤波器
Filters

- 在频域上进行滤波很简单，就是乘以滤波器函数；在时域或空域上，就是卷积

  Filtering in the frequency domain is straight-forward, in the time domain, it is the so-called convolution.

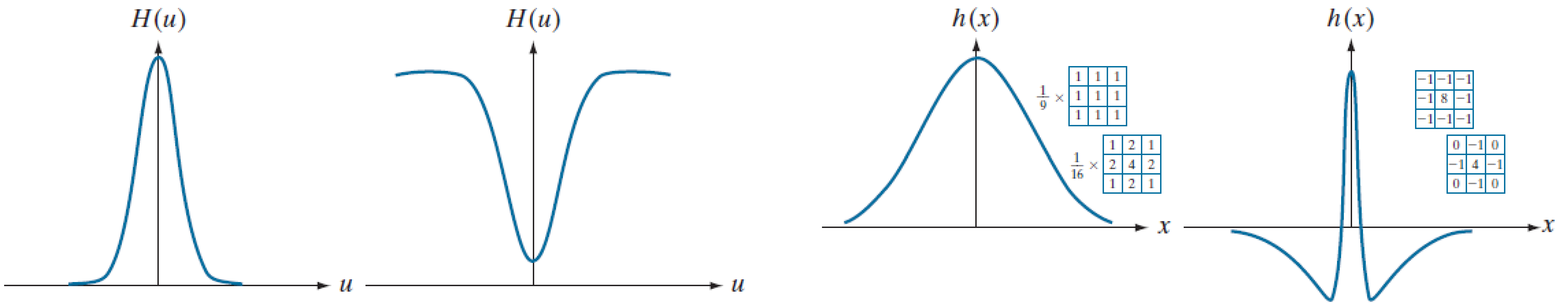




滤波器
Filters

- 怎样得到在时域或空域的滤波器呢？因为在频域上设计滤波器比较直观，所以可以首先在频域上设计，然后变换到空域上进行采样即可。

  How to obtain the filter in the time domain or spatial domain? Since filter design in the frequency domain is rather straight-forward, so it could be firstly designed in the frequency, then transformed to spatial , following by a sampling or discretization.



滤波器
Filters

- 是否可以手工设计滤波器？可以也不可以。手工设计滤波器是比较直观，但费时并且可能也不完善。

  Are the hand-crafted filters ubiquitous? Yes or no. Although manually designed filters are intuitive, but it is time-consuming and imperfect.

−1	−2	−1
0	0	0
1	2	1

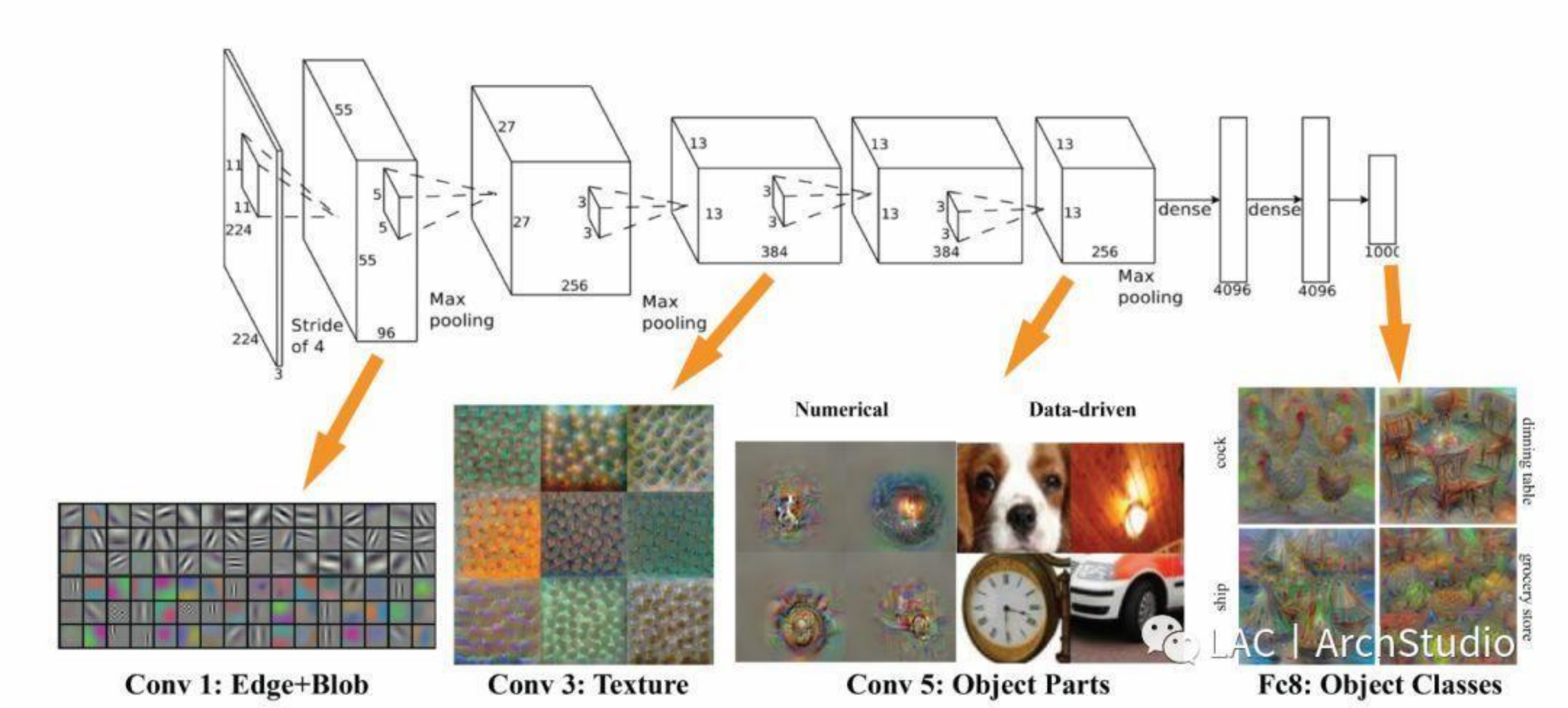
−1	0	1
−2	0	2
−1	0	1

Sobel

滤波器
Filters

- 能否将滤波器设计自动化？可以，这就是卷积神经网络

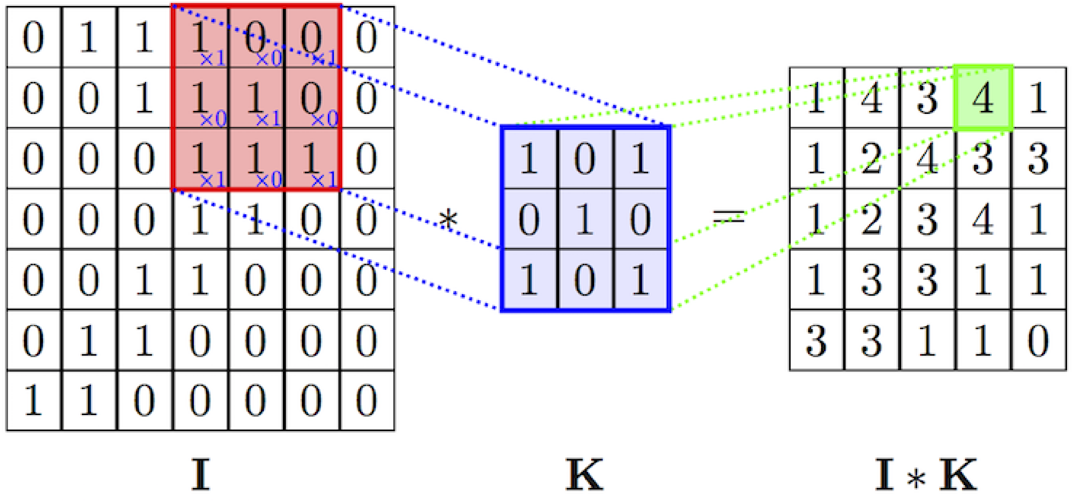
  Can automate the filter design process. Yes, it is intrinsic to the neural network.

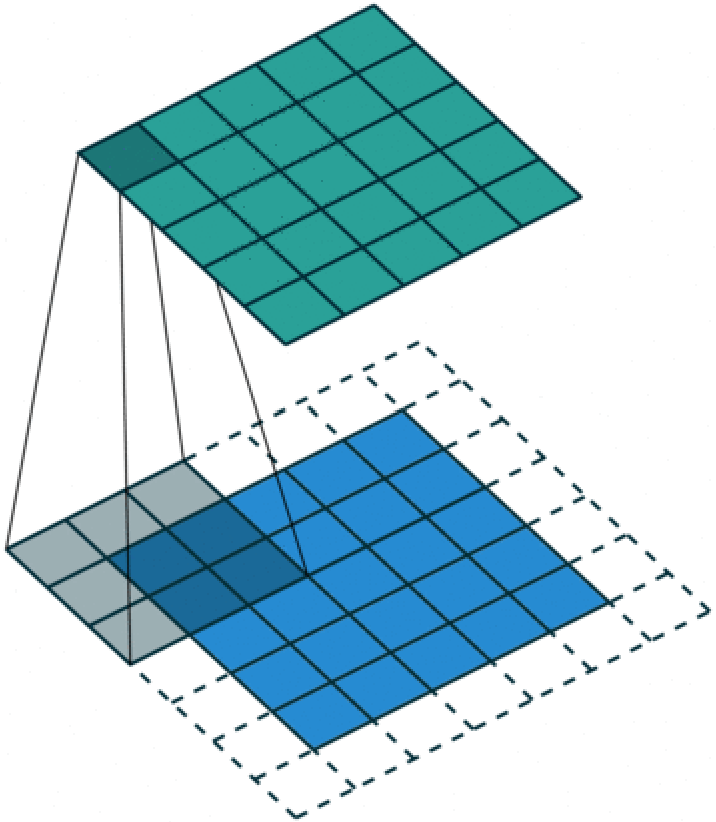


卷积神经网络
Convolutional Neural Network

- 卷积操作

  Convolutions

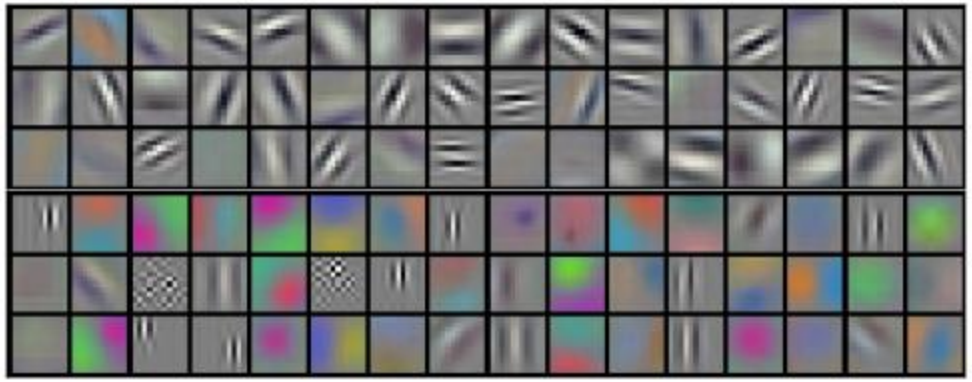




卷积神经网络
Convolutional Neural Network

- 只需指定滤波器的大小，模型训练的过程中，会自动学习最适合的滤波器，而这些滤波器通常很难手工设计

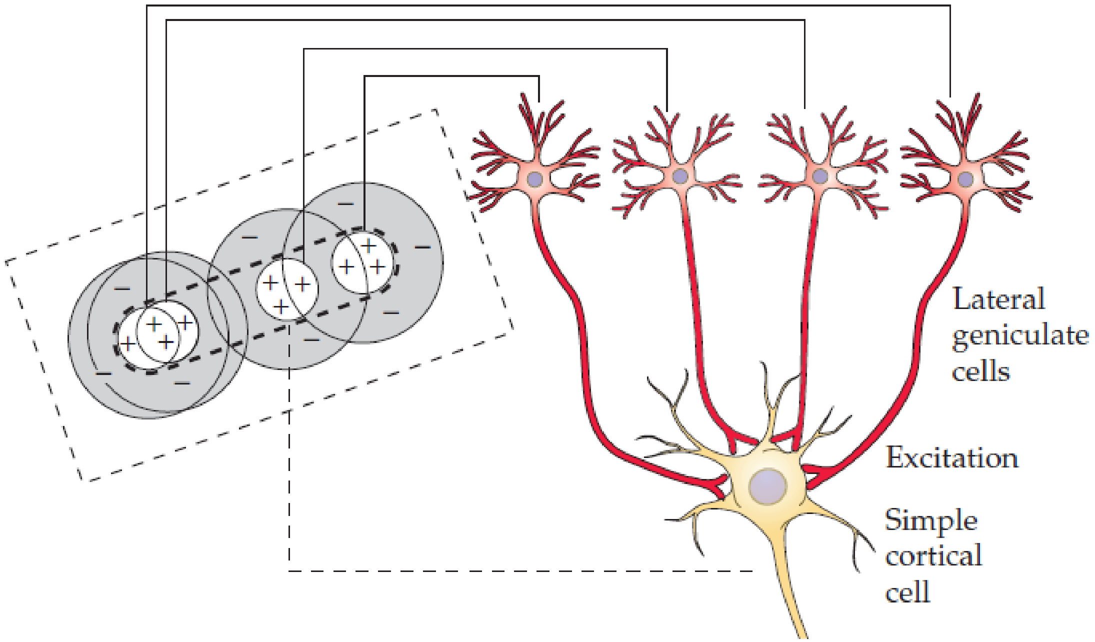
  We only need to specify the size of the filters; the numeric values of the filters will be learned during the training process. And these filters are pretty hard to be designed manually.

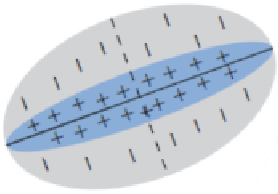


卷积神经网络
Convolutional Neural Network

- 为什么不是很纯的局部视野域（即没有权重共享）？图片中的不同区域的相同模式均会激发高层视觉神经元

  Why not pure local receptive fields (no weight sharing)? The same pattern in different parts of the image can trigger neurons in higher visual cortex indistinguishably.

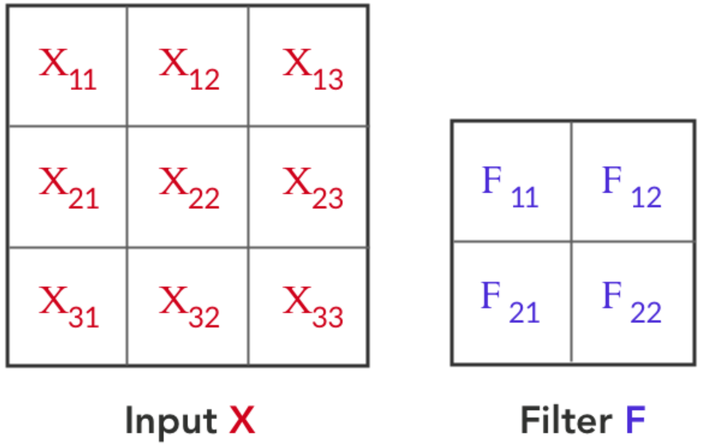


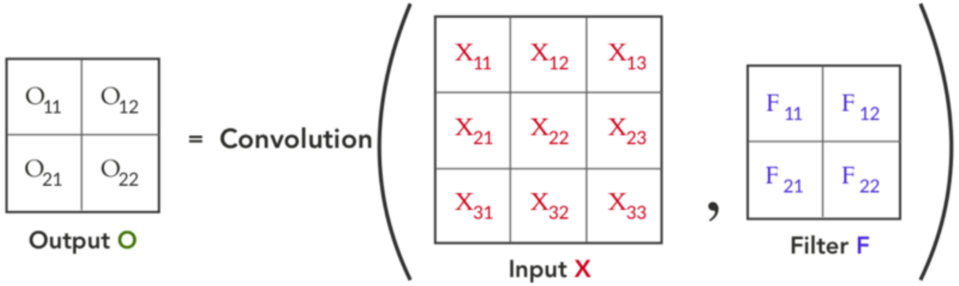


卷积神经网络
Convolutional Neural Network

- 卷积的反向传播算法

  Back-propagation of convolution

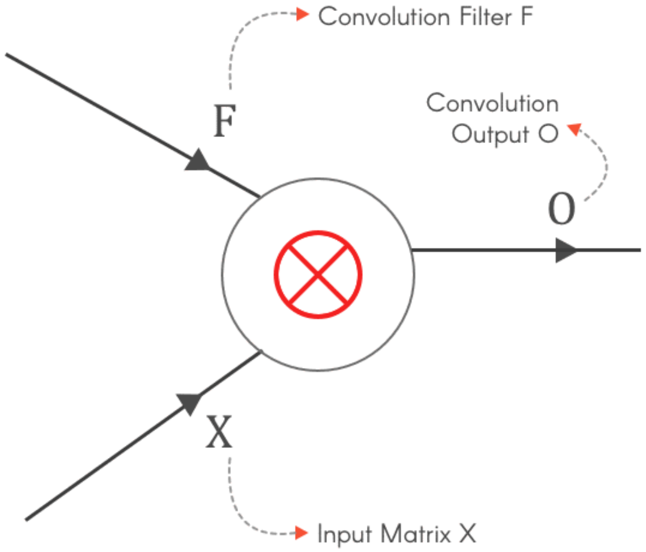




卷积神经网络
Convolutional Neural Network

- 卷积的反向传播算法

  Back-propagation of convolution



Backpropagation in a Convolutional Layer of a CNN

Finding the gradients:

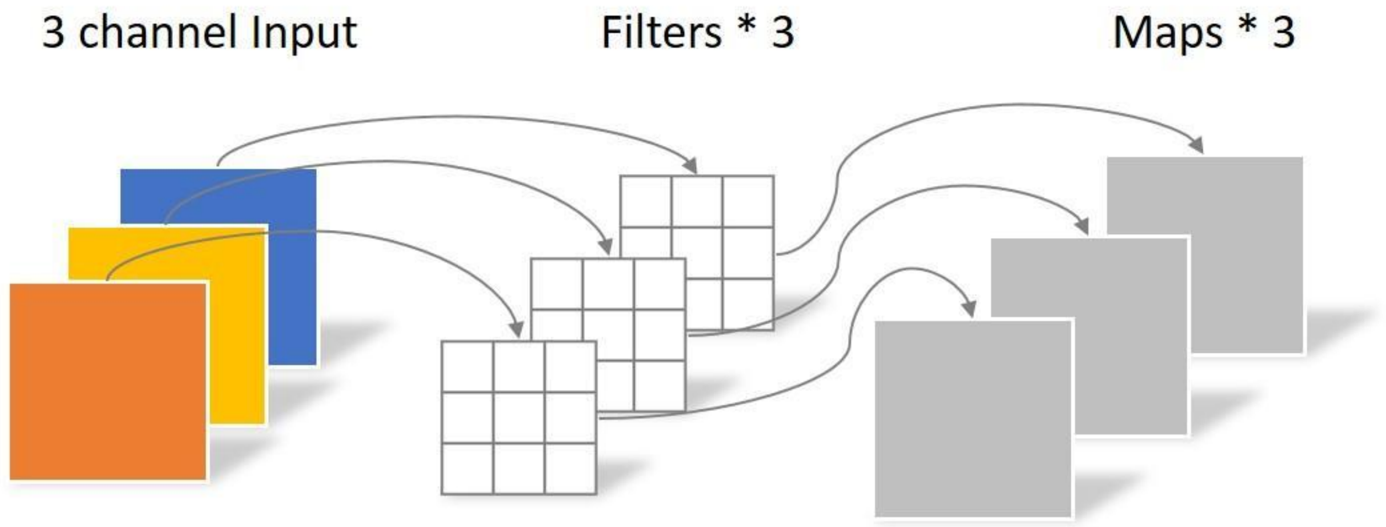
$$\frac{\partial L}{\partial F} = \text{Convolution}\left(\text{Input } X,\ \text{Loss gradient } \frac{\partial L}{\partial O}\right)$$

$$\frac{\partial L}{\partial X} = \begin{matrix}\text{Full}\\ \text{Convolution}\end{matrix}\left(\begin{matrix}180^{\circ}\text{rotated}\\ \text{Filter } F\end{matrix},\ \begin{matrix}\text{Loss}\\ \text{Gradient}\end{matrix}\ \frac{\partial L}{\partial O}\right)$$

卷积神经网络
Convolutional Neural Network

- 实际中，不同的应用可能侧重使用不同的卷积类型，如深度卷积，可分卷积等

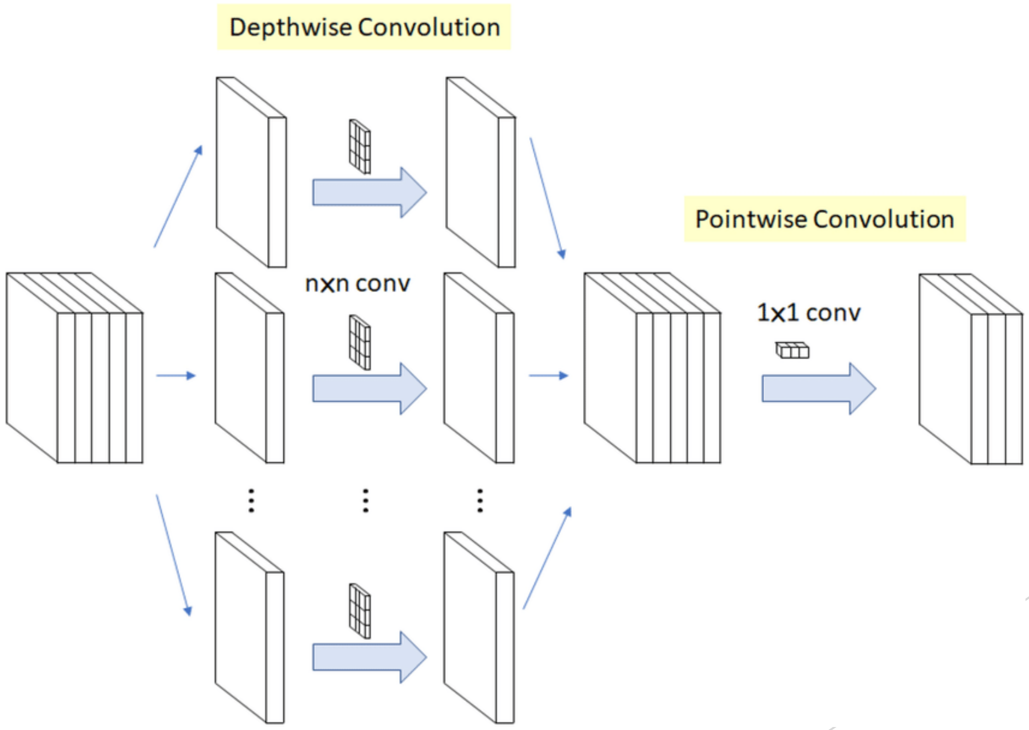
  In practice, different types of convolutions might put into consideration due to the traits of the problems. For example, depth-wise convolutions, separable convolutions.



卷积神经网络
Convolutional Neural Network

- 实际中，不同的应用可能侧重使用不同的卷积类型，如深度卷积，可分卷积等

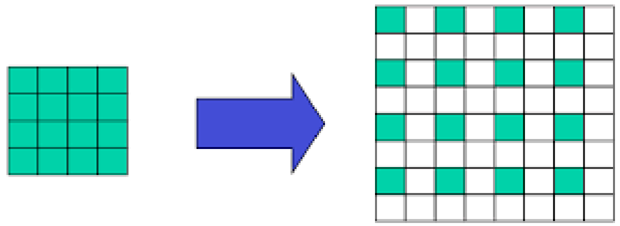
  In practice, different types of convolutions might put into consideration due to the traits of the problems. For example, depth-wise convolutions, separable convolutions.

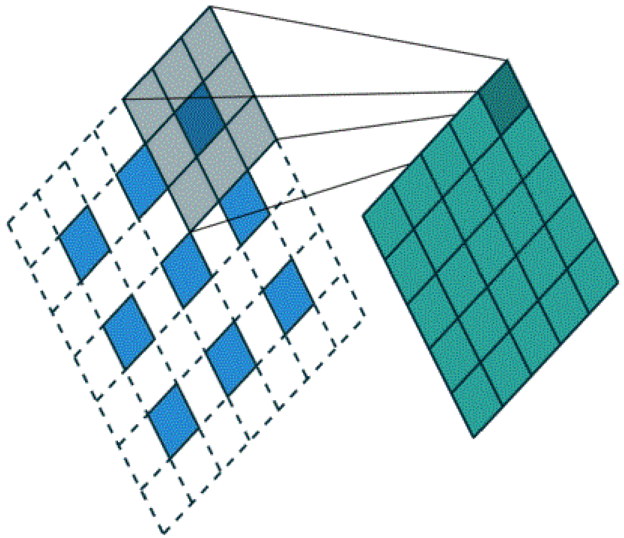


卷积神经网络
Convolutional Neural Network

- 某些应用中，还会用到所谓的反卷积操作，其实还是常规的卷积操作，只不过卷积之前先进行上采样。

  There is a so-called deconvolution in some applications; it is just a normal convolution following up-sampling to the input image or feature-maps.

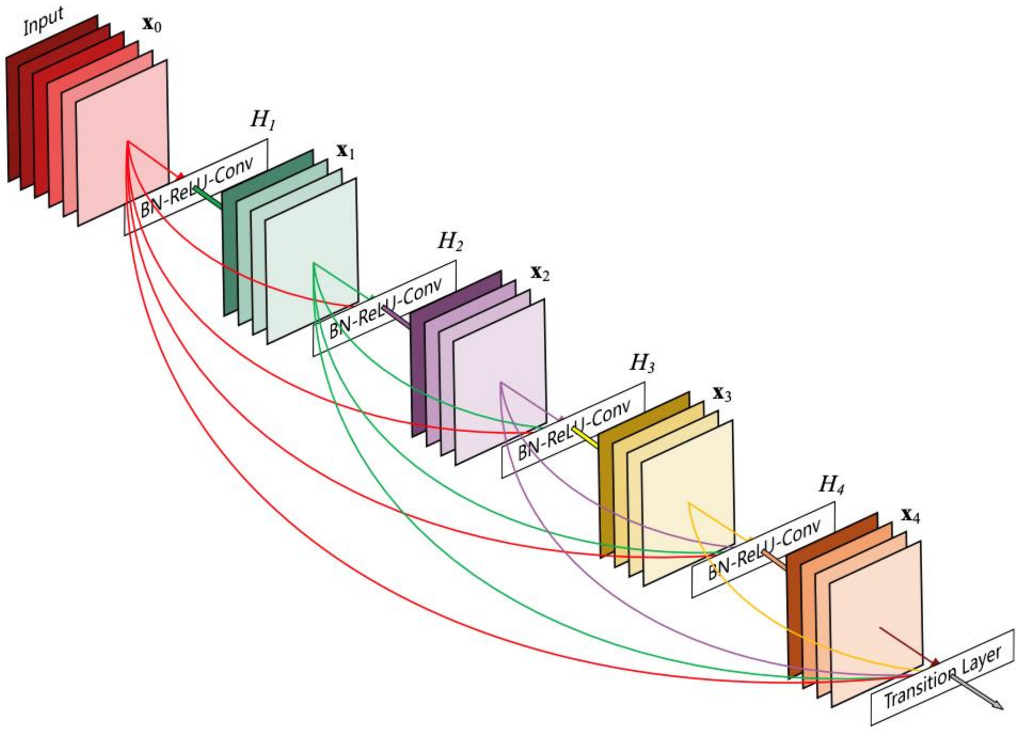




卷积神经网络
Convolutional Neural Network

- 除了卷积类型可以创新，卷积层之间的连接方式也可以创新。

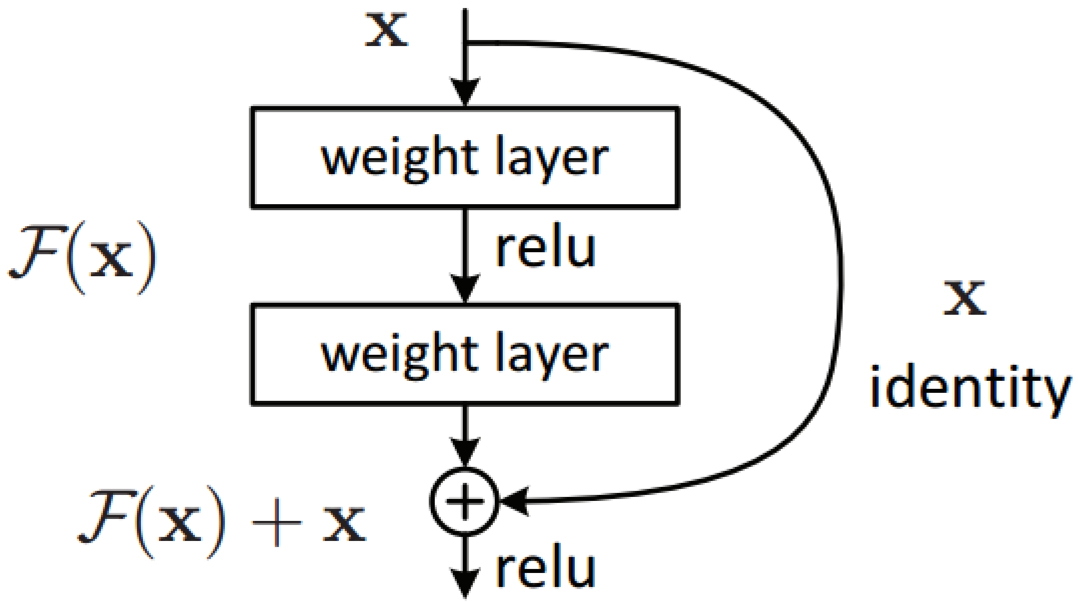
  Besides the types of convolution, there can be variations among links between layers.



卷积神经网络
Convolutional Neural Network

- 除了卷积类型可以创新，卷积层之间的连接方式也可以创新。

  Besides the types of convolution, there can be variations among links between layers.



习题
Problems

1. 简述在时域上手工设计滤波器的思路。

   Describe the idea of designing filters in the time domain.

2. 阅读相关文献，熟悉卷积网络的反向传播算法。

   Understand the backpropagation algorithm for convolutional network by reading related materials.

3. 运用本章内容构建卷积网络，实作MNIST手写数字分类问题。

   Do the hand-written number recognition based on MNIST using convolutional neural network.