

武汉纺织大学计算机与人工智能学院

# 深度学习基础

## 5. 深度学习用于计算机视觉

吴晓堃

xkun.wu at gmail dot com

2021/05/10

2 / 27

Outline

## 本章内容

卷积神经网络简介。在小型数据集上从头开始训练一个卷积神经网络。使用预训练的卷积神经网络。卷积神经网络的可视化。

**重点：**卷积和池化层的工作原理、数据增强、使用预训练的卷积神经网络；

**难点：**卷积神经网络的可视化方法。

3

## 学习目标

- 理解卷积神经网络的基本构成单元的工作原理：卷积层、池化层；
- 掌握在小型数据集上从头开始训练一个卷积神经网络的方法，并能应用数据增强来降低过拟合；
- 掌握使用预训练的卷积神经网络的两种基本方法：特征提取、微调模型；
- 了解卷积神经网络的可视化方法：可视化中间激活、可视化卷积神经网络的过滤器、可视化类激活的热力图。

4

## 图像多分类卷积网络：样例框架

```
from keras import models, layers

model = models.Sequential()
model.add(layers.SeparableConv2D(32, 3, activation='relu',
                                input_shape=(height, width, channels)))
model.add(layers.MaxPooling2D(2))

model.add(layers.SeparableConv2D(64, 3, activation='relu'))
model.add(layers.MaxPooling2D(2))

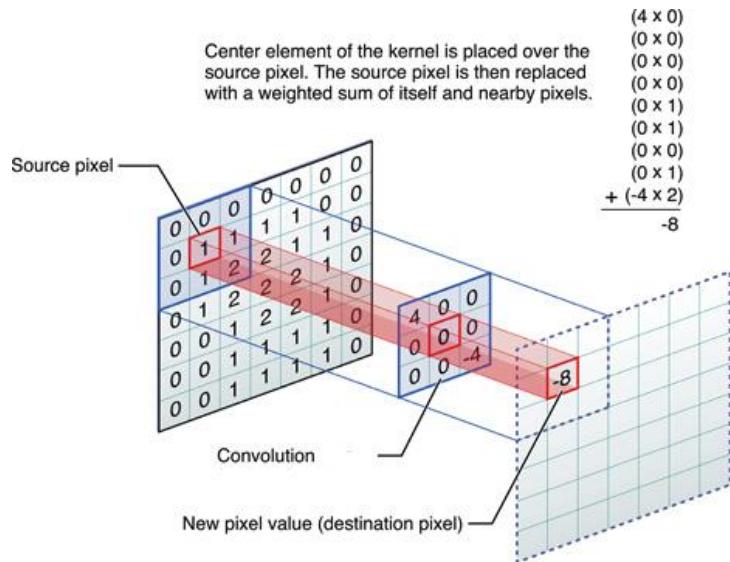
model.add(layers.SeparableConv2D(128, 3, activation='relu'))
model.add(layers.GlobalAveragePooling2D())

model.add(layers.Dense(32, activation='relu'))
model.add(layers.Dense(num_classes, activation='softmax'))

# model.summary()
```

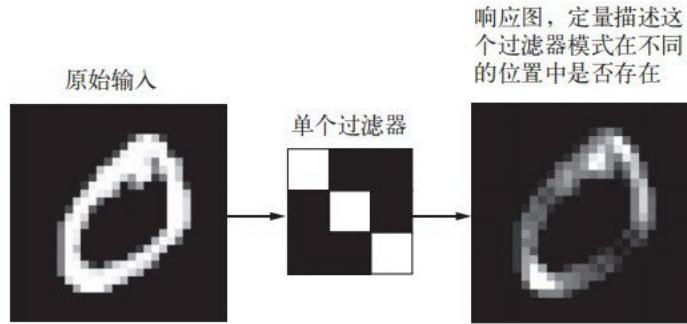
5

## 卷积运算



6

## 特征图、过滤器、响应图



7

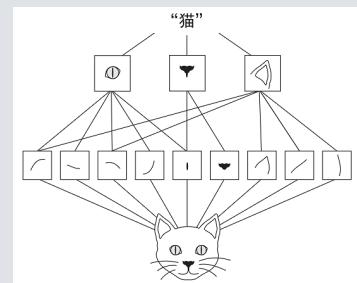
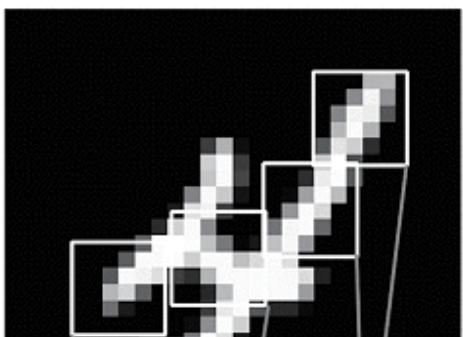
## 卷积层与卷积网络

### 卷积层

学到的是局部模式。

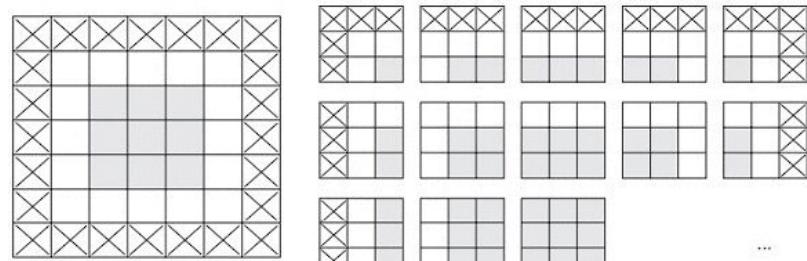
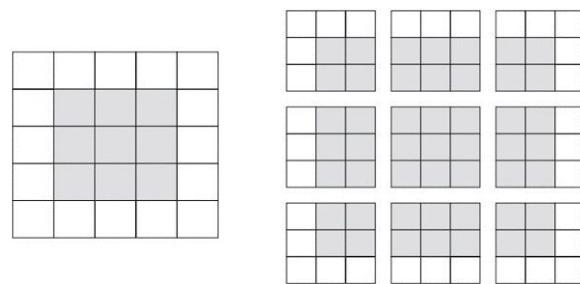
### 推论：卷积网络的两大性质

- 学到的模式具有平移不变性：高效利用数据。
- 可以识别模式的空间层次结构、抽象视觉概念。



8

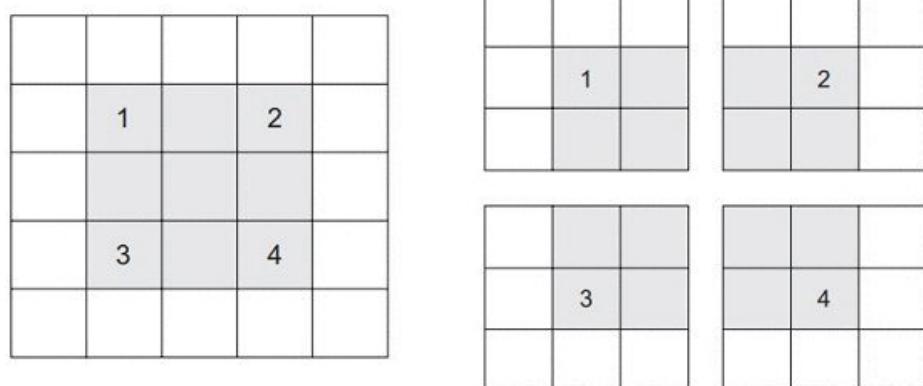
## 边界效应与填充



9

10 / 27

## 卷积步幅



步幅为 2 意味着对特征图做 2 倍下采样（除了边界效应引起的变化）。

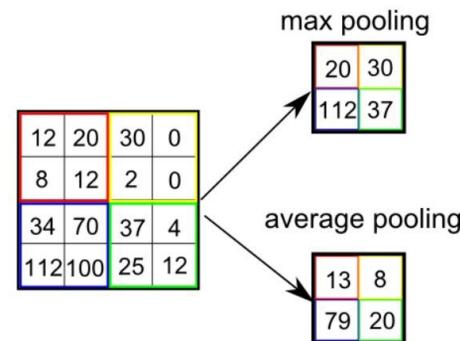
## 最大池化与平均池化

### 池化 (pooling) 的作用：对特征图进行下采样

- 有效减小问题规模，即待处理特征图的元素个数。
- 等价于增大观察窗口（过滤器），从而引入空间层级结构。

一般来说，最大池化的效果最好。

- 步进卷积：可能错过特征；
- 平均池化：特征被淡化。



11

## 小数据问题

### 从头开始训练

- 卷积神经网络可以高效地利用数据：学到的是局部的、平移不变的特征。
- 无须任何自定义的特征工程，但通常使用数据增强。

### 引入预训练的网络

深度学习模型本质上具有高度的可复用性。

- 特征提取 (feature extraction)。
- 微调模型 (fine-tuning)。

12

## 数据增强

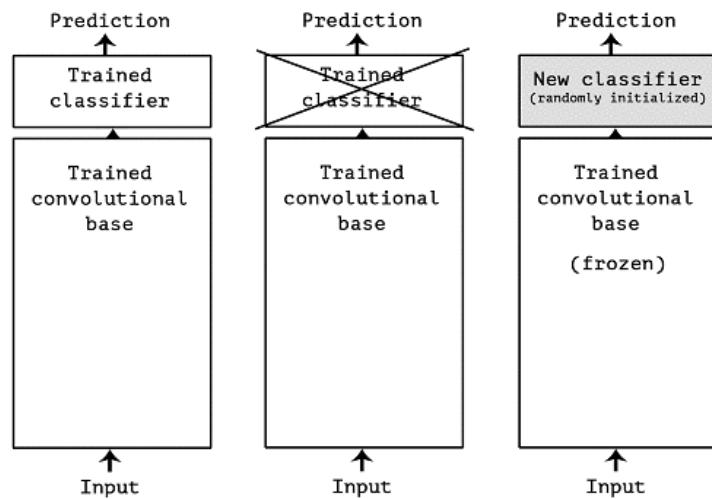


如何训练机器识别倒立的猫？

13

14 / 27

## 使用预训练的卷积神经网络

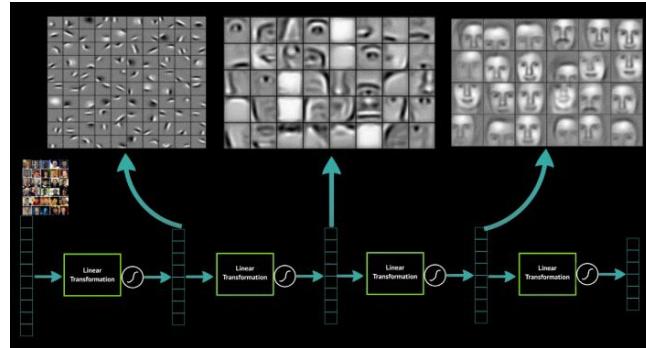


14

## 特征通用性与深度

某个卷积层提取的表示的通用性（以及可复用性）取决于该层在模型中的深度。

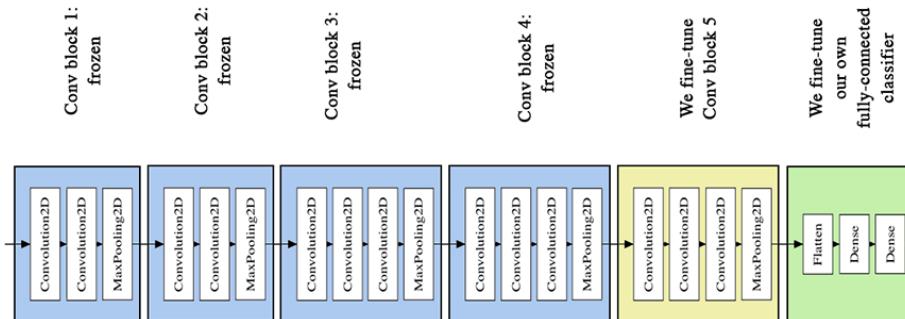
- 模型中更靠近底部的层提取的是局部的、高度通用的特征图；
- 更靠近顶部的层提取的是更加抽象的概念。



15

## 微调模型 (fine-tuning)

将顶部的几层“解冻”，并将这解冻的几层和新增加的部分联合训练：



16

## 卷积神经网络的可视化

### 中间输出（中间激活）

- 有助于理解层如何对输入进行变换。
- 有助于初步了解每个过滤器的含义。

### 过滤器

有助于精确理解每个过滤器容易接受的视觉模式或视觉概念。

### 图像中类激活的热力图

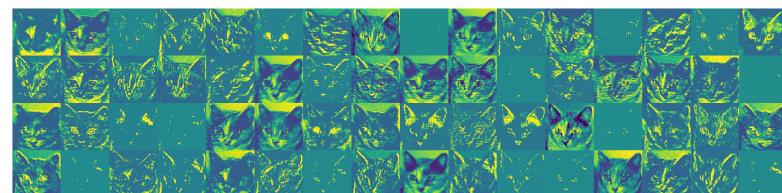
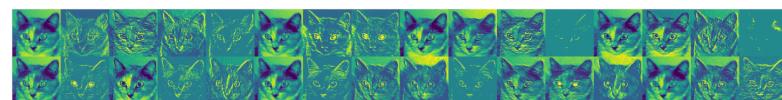
有助于理解图像的哪个部分被识别，从而可以定位图像中的物体。

17

18 / 27

## 可视化中间激活 I.1

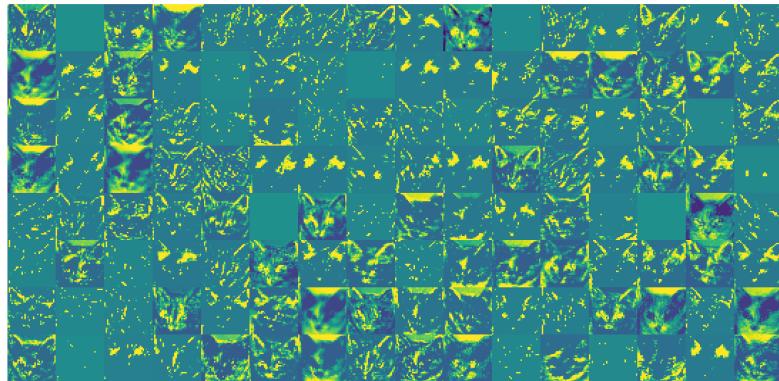
第一层是各种边缘探测器的集合。



18

## 可视化中间激活 II

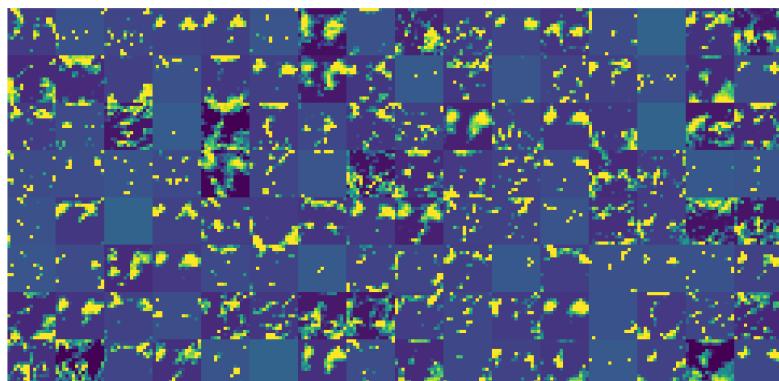
随着层数的加深，激活变得越来越抽象，并且越来越难以直观地理解。



19

## 可视化中间激活 III

层数越深，其表示中关于图像视觉内容的信息就越少，而关于类别的信息就越多。



20

## 具体外观 - 抽象概念

### 信息蒸馏管道

深度神经网络对数据反复变换，将无关的视觉细节过滤掉，总结出高度抽象的概念。

### 人脑感知视觉世界的过程类似



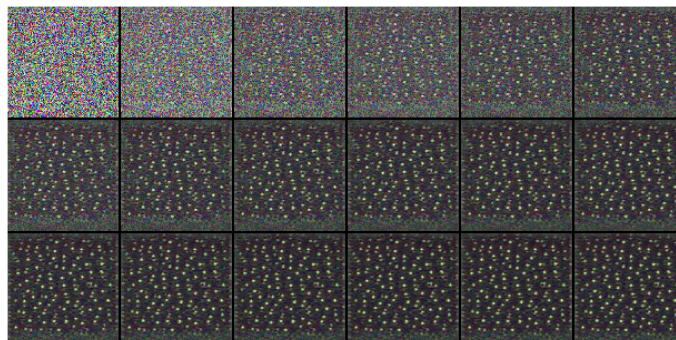
21

## 可视化过滤器：原理

### 梯度上升

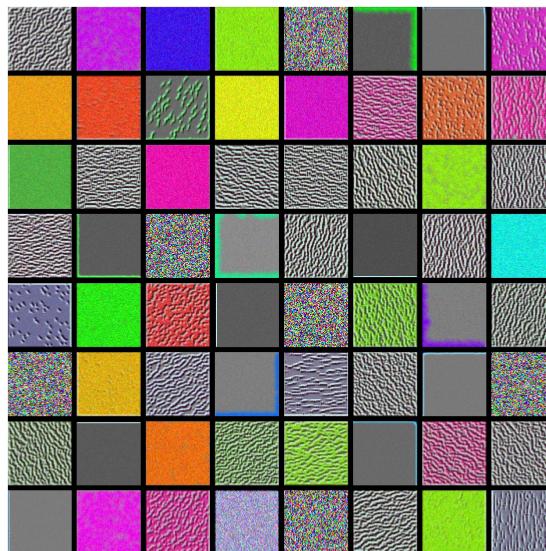
从带有噪声的灰度图像开始，将梯度上升应用于图像：

- 目的是让目标过滤器的响应最大化，从而得到使其具有最大响应的图像。
- 具体来说：需要构建一个损失函数，目的是让目标过滤器的值最大化。



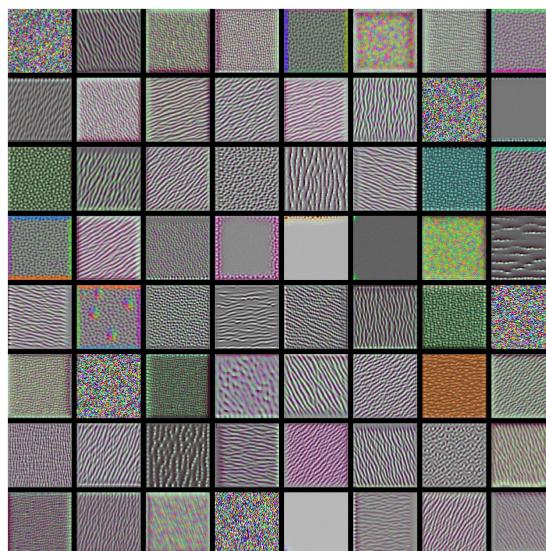
22

## 可视化过滤器 I



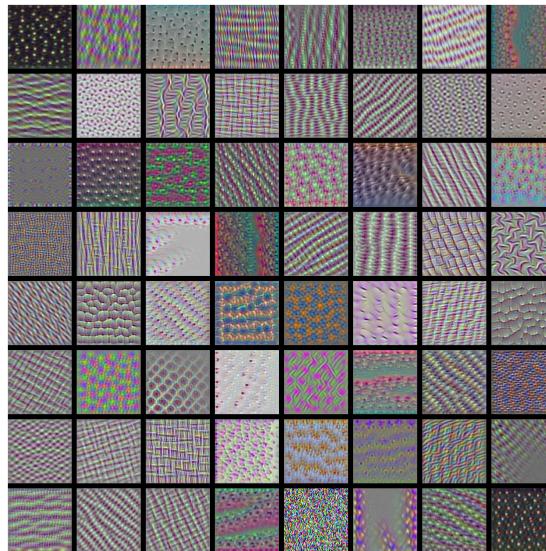
23

## 可视化过滤器 II



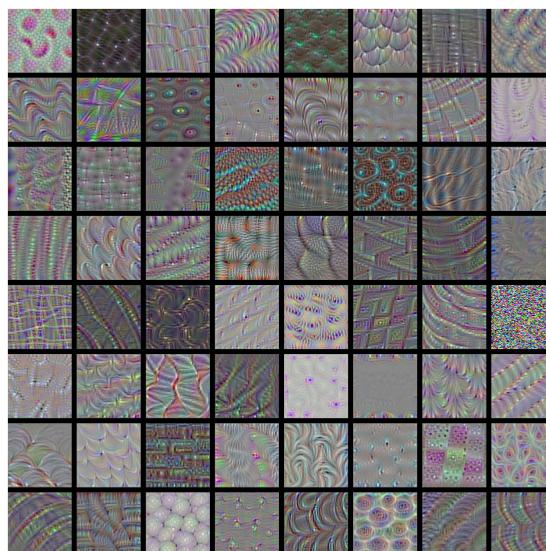
24

## 可视化过滤器 III



25

## 可视化过滤器 IV



26

## 可视化类激活的热力图

用“每个通道对类别的重要程度”对“输入图像对不同通道的激活强度”进行加权，从而得到了“输入图像对类别的激活强度”。



27