

复盘 5 | 计算机视觉核心技术模块

2018-10-05 洪亮劼

AI技术内参

[进入课程 >](#)



今天我们来对计算机视觉核心技术模块做一个复盘。在这个模块里，我们一起学习了 12 期内容，讨论了四个话题，这些话题主要围绕计算机视觉的基础知识和深度学习技术在这个领域的应用。

之所以这么安排，是因为没有深度学习技术，就不会有现在计算机视觉的发展。我们站得稍微高一点就可以看到，正是因为深度学习技术在计算机视觉中的成功应用，才有了近几年的人工智能浪潮。

提示：点击知识卡跳转到你最想看的那篇文章，温故而知新。

图像技术基础

什么是计算机视觉？

计算机视觉希望模仿人类的视觉系统构架，是一个“跨学科领域”，包含了很多非常深刻的困难问题。



掌握计算机视觉任务的基础模型和操作

像素是对真实世界的采样。在像素表达之上的过滤器，常见的有移动平均和卷积两种方式。我们可以利用函数的梯度计算来进行边界检测。



计算机视觉中的特征提取难在哪里？

在图像理解任务中，高维度的语义理解需要特征提取这个关键步骤，其难点在于图像信息本身的复杂性，一个经典的思路是从局部信息入手，提取局部特征。

深度学习神经网络入门

深度学习技术能够自动挖掘数据中的非线性关系。在计算机视觉领域里，利用深度神经网络来挖掘特征已经基本代替了手动的特征挖掘，深度神经网络还具有计算的普适性。



基本的深度学习模型

前馈神经网络是最基本的神经网络架构。可以引入“激活函数”和“线性整流函数”对线性模型进行非线性的转换。卷积神经网络的核心是用向量来描述一个矩阵的信息。



深度学习模型的优化

模型优化的第一步就是选择目标函数，然后根据目标函数寻找参数的最优解。深度学习模型中，先利用反向传播计算梯度，但不能得到深度模型参数的解析解，只能使用梯度下降来对问题进行近似求解。



AlexNet

AlexNet 模型的四点创新：采用“线性整流函数”作为激活函数；整个模型的训练大量采用 GPU，并且使用多个 GPU 进行计算；局部响应归一化；重叠池化。



VGG & GoogleNet

模型优化的第一步就是选择目标函数，然后根据目标函数寻找参数的最优解。深度学习模型中，先利用反向传播计算梯度，但不能得到深度模型参数的解析解，只能使用梯度下降来对问题进行近似求解。

ResNet

ResNet，残差网络，是一个能够大幅度提升网络层次的深度学习模型。它解决了神经网络结构究竟能够搭建多深这个问题，突破了一个瓶颈。

计算机视觉高级话题



图像物体识别和分割

R-CNN 模型尽可能多地生成选定框来进行物体识别。在此基础上的改进模型是 Fast R-CNN。Faster R-CNN 进一步改进了提出待定选定框的方法。Mask R-CNN 不仅能够识别图像中的物体，还能够做到像素级的抽取。

视觉问答

视觉问答的三大挑战：理解图片中的细节；理解图片的上下文；对图片中的物体进行推理。我们可以利用“关注”机制进行视觉问答建模。



产生式模型

概率分布和概率图模型都是数据产生器。“产生式对抗网络”的思路是定义一个流程，通过这个流程产生数据，从而能够直接模拟真实数据进。

积跬步以至千里

最后，恭喜你学完了这个模块中的内容。今日记一事，明日悟一理，积久而成学。每一个收获都是一个不小的成就。

在计算机视觉领域你有怎样的经验、心得或者疑惑，都欢迎你留言和我一起交流讨论。

AI 技术内参

你的360度人工智能信息助理

洪亮劼

Etsy 数据科学主管
前雅虎研究院资深科学家



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 复盘 4 | 广告系统核心技术模块

下一篇 复盘 6 | 数据科学家与数据科学团队是怎么养成的？

精选留言

 写留言

由作者筛选后的优质留言将会公开显示，欢迎踊跃留言。