

## 143 | 基于深度学习的计算机视觉技术（一）：深度神经网络入门

2018-08-31 洪亮劼

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 07:07 大小 3.26M



在最近几年的人工智能发展中，深度学习技术成为了一个强劲的推动力。对于计算机视觉来讲，深度学习在过去几年重新改写了这个领域的核心方法论。时至今日，深度学习已经深入到了计算机视觉技术的方方面面，成为解决各类视觉问题的有力工具。

从今天开始，我们将介绍一系列以深度学习为背景的计算机视觉技术。那么在这个环节的第一篇分享中，我们首先来了解一下什么是深度学习。

### 为什么是深度学习

在了解一些深度学习技术细节之前，我们首先要来看一下为什么需要深度学习技术。

初学者经常会有一个误区，那就是认为和“深度学习”相对的就是“浅层学习”（Shallow Learning）。这种看法也对，也不对。

“对”的地方在于“深度学习”的确强调从数据或者说是特征（Feature）中构造多层或深度的变换，从而能够得到非线性的表征（Representation）。显然，这种效果是线性模型所达不到的。

“不对”的地方是，在所谓的“深度学习”，或者准确地讲是深度神经网络技术发展之前，就已经有了很多构造复杂非线性表征的尝试和技术。这些技术在机器学习和人工智能的发展中都起到了举足轻重的作用。

说到这里，我们就要从**线性模型**聊起了。从线性模型发展到非线性模型，这一步貌似理所当然，但其实这里面有一个非常重要的思路，那就是线性模型并不是不能处理数据中的非线性关系，这一点很容易被忽视。很多时候，我们其实是可以构造非线性的特征，然后利用线性模型来把所有的非线性特征给串联起来。

举个例子，在网页搜索中，我们经常利用类似 PageRank 来表征一个网页的重要性。这个模型本身就是非线性的对网页图（Graph）的一种表征。所以，即便在此之上构建线性模型，整个模型其实也是包含了非线性的表征转换。

其实，对于很多深度学习模型而言，即便进行了复杂的表征转换，在最后一层对最终的输出进行建模的时候依然是一个线性模型。所以，线性模型在非线性特征的帮助下，依然能够满足整体的非线性建模的需求。

那么，这种非线性特征外加线性模型的方法有什么问题呢？

这类方法的主要问题是，如何才能系统性地找到这些非线性特征呢？其实，传统的机器学习中的特征工程（Feature Engineering）主要就是在做特征寻找这个事情，也就是要消耗人力和物力去寻找这些特征。有时候，找到一个好的特征可能还需要灵感和其他领域特定的知识。

那么，有没有办法让模型自身就能从现有的数据中发现这些非线性关系，从而不需要额外的特征工程呢？

其实在机器学习发展的早期，研发人员就意识到了这个方向的重要性，这样就发展出了各种各样的非线性模型。

这里面比较有代表性的模型是“**决策树**”（Decision Tree），以及在此基础上发展出来的一系列“**树模型**”（Tree Models）。我们在专栏里介绍过，树模型在搜索和推荐的一些场景中都有不错的性能表现。在很大程度上，树模型可以表达非线性的关系，但是它的困难之处在于无法表达过于深层次的数据关系。一般来说，3 到 4 层的树已经算是比较深的结构了，如果一个树模型有特别多层次还能被训练成功，这样的例子是比较少见的。

除了树模型之外，还有一类模型，用来挖掘数据中间的隐含关系，特别是非线性关系，这就是“**概率图模型**”（Probabilistic Graphical Model）。例如在文本挖掘领域非常流行的 LDA（Latent Dirichlet Allocation），以及在推荐领域流行的矩阵分解（Matrix Factorization），都可以被看作是概率图模型中的佼佼者。

概率图模型的一大优势是可以融入众多对于数据以及所需要处理问题的直觉，从而能够让模型具有一定的可解释性，甚至是“因果性”（Causality）。但是，概率图模型的最大挑战就是每一个模型都需要单独计算训练算法。也就是说，算法无法做到普适通用。这就极大地限制了概率图模型在实际问题，特别是大数据问题中的应用。

综合来看，**我们急需一种方法，能够自动挖掘数据中的非线性关系，而且最好能够找到数据中的隐含规律，这种隐含的规律可能是非常多层次的非线性转换；并且，这种方法还需要在计算上直接通用，不同模型可以共用一个计算框架。**

所有这些因素如何囊括在一个方法里呢？答案就是深度学习技术。

## 深度学习的特点

深度学习技术慢慢成为了主流的非线性模型。接下来我们来看一看深度学习技术的一些特点。

首先，深度学习技术是一个非常大的外延，这里面包含了很多不同的**模型**和模型的**计算框架**技术。这两者都是深度学习成功必不可少的组成部分。

深度学习有一种最简单也是最基础的模型，就是“**深度神经网络**”（Deep Neural Networks）。这种模型其实很早就已经被提出了。

从形式上来说，深度神经网络就是把多层简单的非线性操作叠加起来，希望能够发现更加复杂的非线性关系。实际上，有理论研究表明，**在有足够的内部隐含变量的情况下，深度神经网络可以表达任意复杂的函数关系**。也就是说，深度神经网络有希望能够对现实世界中的复杂现象进行建模。这一点对于我们刚才提到的树模型和概率图模型来说，有相当大的难度或者说几乎是不可能的。

另外，回到我们刚才说的特征工程的需求，深度神经网络的确可以减轻这方面的压力。虽然并不如很多人预期的那样，深度神经网络依然需要依赖一定的初始数据，但是已经有实验表明，通过深度神经网络学习的特征在很多时候相比于研发人员手动挖掘的特征要更加健壮。

实际上，**在计算机视觉这样的领域里，利用深度神经网络来挖掘特征基本上已经完全代替了手动的特征挖掘**。

深度神经网络还有一个特点就是**计算的普适性**。刚才我们讲到概率图模型有一个“死穴”，就是计算无法做到模型普适，或者说不牺牲性能的前提下，计算无法做到普适性。那么这一点来说，以深度神经网络为代表的深度学习，依赖简单的梯度下降就能对非常复杂的网络进行计算。而这种计算是可以针对不同的模型的，因此这就极大地降低了深度学习在实际工程应用中的代价。

## 小结

今天我为你讲了基于深度学习的计算机视觉技术的第一篇分享，帮助你对深度学习有一个更加明确的认识。

一起来回顾下要点：第一，深度学习技术能够自动挖掘数据中的非线性关系；第二，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，而且，深度神经网络还具有计算的普适性。

最后，给你留一个思考题，和树模型或者概率图模型相比较，深度学习有什么劣势吗？如果有，你认为它最大的劣势是什么？

欢迎你给我留言，和我一起讨论。

---



# 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 142 | 计算机视觉中的特征提取难在哪里？

下一篇 144 | 基于深度学习的计算机视觉技术（二）：基本的深度学习模型

## 精选留言 (3)

 写留言



寻找自我

2018-08-31

 4

最大的劣势应该是复杂且不可解释

展开



吴曦

2019-01-20

 1

参数太多，计算复杂度高，容易过拟合。

展开





**mortimer**  
2018-10-18



我认为应该是"结果"与常识存在更大概率的"不符",因此在应用上相比"传统方法"需要更多产品化相关的工作