

标题/关键词...



登录

注册

宏观经济

A股

港股

美股

基金

理财

黄金

期货

能源

互联网金融

科技

更多

# 入门 | 一文概览深度学习中的卷积结构

2017年09月27日 13:45:07 机器之心

选自Medium

作者: Paul-Louis Prove

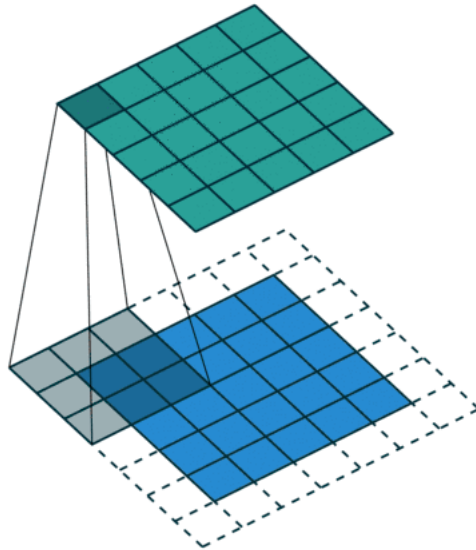
机器之心编译

参与: 路雪、李亚洲

本文对三种不同的卷积进行了介绍,同时讲解了各自的优点,对初学者而言,是理解卷积的一篇好文章。

卷积

首先,我们需要定义卷积层的几个参数。



kernel 为 3、stride 为 1,使用 padding 的 2D 卷积

卷积核大小: 卷积核决定卷积的视野。2D 卷积的常见卷积核为 3,即 3x3 像素。

stride: stride 决定卷积核遍历图像时的步子大小。默认值通常为 1,我们可以将 stride 设置成 2,对图像进行类似最大池化的下采样



机器之心

专业的人工智能媒体与产业服务平台。

## 热文排行

日榜

周榜

月榜

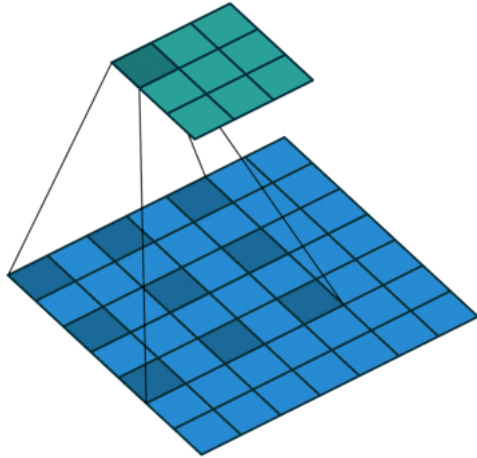
- 1 中国省会,三个最“窝里横”,三个最“...
- 2 又一个国家倒下!特朗普做空亚洲却...
- 3 西方倒逼 中国接招造就又一世界级超...
- 4 马云:我们的教育方式会让孩子在未...
- 5 印度制造业崛起为何走得越来越艰难?
- 6 诺贝尔奖得主:不在乎几百万奖金!...
- 7 连续暴涨!人民币,真成了最强货币...
- 8 富二代奢华婚礼:郭广昌夫妇主婚 张...
- 9 日本变态的“厕所文化”,细节武装到了...
- 10 这政策一出,很多买了房的人肠子悔青

**Padding:** padding 决定处理样本时的边界。(半)填充的卷积使输出空间维度等于输入,而未填充的卷积会裁剪部分边界,如果卷积核大于1的话。

**输入输出通道:** 卷积层通常需要一定数量的输入通道(I),计算一定数量的输出通道(O)。所需参数可以通过  $IOK$  来计算,  $K$  就是卷积核的值。

机器之心曾介绍过用于语义分割中的各种卷积: 从全连接层到大型卷积核: 深度学习语义分割全指南

扩张卷积(又叫空洞卷积)



kernel 为 3、扩张率为 2、没有 padding 的 2D 卷积

扩张卷积向卷积层引入另一个参数「扩张率」。它决定了卷积核中值之间的空间。 $3 \times 3$  卷积核、扩张率为 2 的卷积视野和  $5 \times 5$  卷积核的视野相同,并且前者仅使用了 9 个参数。想象一个  $5 \times 5$  的卷积核,每个都删去第二行和第二列。

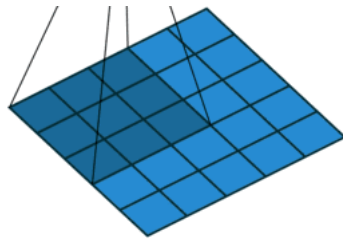
这种卷积用同样的计算成本生成了更大的视野。扩张卷积在实时分隔领域中尤为流行。如果你需要一个宽阔视野,但无力使用多个卷积或更大的卷积核,那么你可以使用这种卷积。

转置卷积(解卷积或微步卷积)

解卷积(deconvolution)这种叫法不太合适,因为这并不是解卷积。解卷积确实存在,但在深度学习领域中并不常见。真正的解卷积是卷积过程的逆转。想象一下将一个图像输入到单个卷积层上。再把输出放到黑箱中,然后再次输出的是原始输入图像。这个黑箱就叫作解卷积。这是卷积层执行的数学逆运算。

转置卷积与解卷积有一些相似,因为它所输出的空间分辨率反卷积层也能够输出。但是,在这些值上真正进行的数学运算是不同的。转置卷积层使用的是常规的卷积,但仍然能够进行空间分辨率转换。

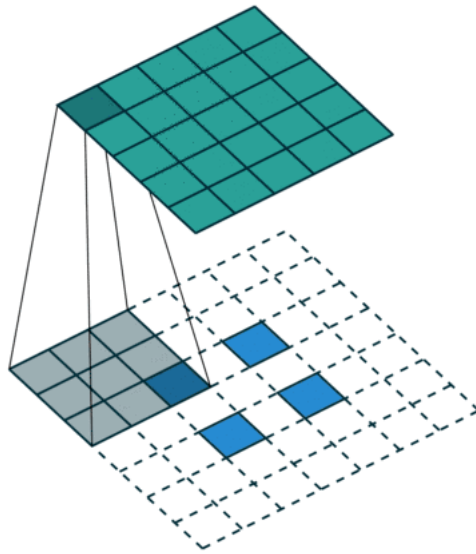




没有 padding、stride 为 2、卷积核为 3 的 2D 卷积

到这里你可能有些疑惑，那么让我们看一下具体的例子吧。一个 5x5 的图像输入到卷积层中，stride 设置为 2，没有 padding，卷积核为 3x3。输出的是 2x2 的图像。

如果我们想逆转该过程，则我们需要进行数学逆运算，以使我们输入的每个像素都能够生成 9 个值。之后，我们用值为 2 的 stride 遍历输出图像。这就是解卷积。



没有 padding、stride 为 2、kernel 为 3 的转置 2D 卷积

转置卷积并不这么做。二者唯一的共同点是输出的都是 5x5 的图像，虽然它执行的仍旧是常规的卷积运算。为了做到这一点，我们需要在输入上执行某种 padding。

如同你能想象的，这一步不会从顶部逆转该流程，至少在数值方面不会实现逆转。

它只不过是前面重建了空间分辨率，且完成一个卷积。这可能不是数学意义上的逆转，但对编码器-解码器架构而言，它仍旧非常有帮助。通过这种方式，我们可以将卷积和图像的 upscaling 结合起来，而不是执行两个独立的流程。

可分离卷积

在可分离卷积中，我们能将卷积核运算分离到多个步骤中。例如，我们可以把一种卷积表

达为  $y = \text{conv}(x, k)$ ，其中  $y$  是输出图像， $x$  是输入图像， $k$  是卷积核。下面假设  $k$  按  $k =$

$k_1 \cdot \text{dot}(k_2)$  进行计算。这样可使其成为可



过用 **k1** 和 **k2** 做两个 **1D** 卷积得到同样的结果。

|    |   |    |
|----|---|----|
| -1 | 0 | +1 |
| -2 | 0 | +2 |
| -1 | 0 | +1 |

x filter

|    |    |    |
|----|----|----|
| +1 | +2 | +1 |
| 0  | 0  | 0  |
| -1 | -2 | -1 |

y filter

#### Sobel X 与 Y 滤波器

拿经常用于图像处理的 **Sobel** 核为例。你可以通过乘以向量 **[1, 0, -1]** 和 **[1,2,1]** 的转置向量获得相同的核。在进行相同操作时，这只需要 **6** 个参数，而无需 **9** 个。

上述实例展示了空间可分离卷积，据我所知它并不用于深度学习。我只是想让大家在看到这个术语时不会感到困惑。在神经网络中，我们通常使用深度可分离卷积（**depthwise separable convolution**）。

这种卷积将执行空间卷积，同时保持通道分离，接着跟从深度卷积。为了便于理解，我们来看一个实例。

假设我们在 **16** 个输入通道和 **32** 个输出通道上有一个 **3x3** 卷积层。每一个输入通道都由 **32** 个 **3x3** 内核遍历，产生 **512**（**16x32**）个特征图。下一步，我们通过叠加每一个输入通道中的特征图，合并形成一个特征图。由于我们这样做了 **32** 次，我们得到了 **32** 个想要的输出通道。

对于相同实例上的深度可分离卷积，我们遍历了 **16** 个通道（每个带有一个 **3x3** 内核），得到了 **16** 个特征图。现在，在合并之前，我们遍历了这 **16** 个特征图（每个带有 **32** 个 **1x1** 卷积），然后再把它们叠加在一起。相比于上述的 **4608**（**16x32x3x3**）个参数，这产生了 **656**（**16x3x3 + 16x32x1x1**）个参数。

该实例是深度可分离卷积的特定实现，深度乘数是 **1**，这是目前这类卷积层的最常见设置。

我们这么做是因为假设空间和深度信息可被解耦。看 **Xception** 模型的表现，该理论似乎是成立的。因其对参数的高效使用，深度可分离卷积也可被用于移动设备。

原文地址：<https://medium.com/towards-data-science/types-of-convolutions-in-deep-learning-717013397f4d>

本文为机器之心编译，转载请联系本公众号获得授权。

加入机器之心（全职记者/实习生）：[hr@jiqizhixin.com](mailto:hr@jiqizhixin.com)

投稿或寻求报道：[content@jiqizhixin.com](mailto:content@jiqizhixin.com)



广告商务合作: [bd@jiqizhixin.com](mailto:bd@jiqizhixin.com)

## 作者历史文章

### 专栏 | 深度好奇提出文档解析框架: 面向对象的神经规划



机器之心专栏作者: 深度好奇R&D深度好奇 (DeeplyCurious.ai) 近日在 arXiv 上发布的论文提出了一种基于神经符号智能 (Neural-s[详细]

2017年 10月10日 16:15

### 中科视拓获数千万pre-A轮融资, 机器之心专访山世光



机器之心原创作者: 邱陆陆10月10日, 中科视拓对外公布, 获得安赐资本领投的数千万元 pre-A 轮融资。这是去年秋天中科视拓宣布成立, 并获得线性资本领投的千万[详细]

2017年 10月10日 16:15

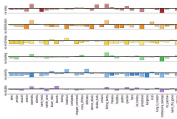
### 业界 | 英特尔研发自我学习芯片Loihi: 我们正在尝试复制大脑!



机器之心发布作者: 英特尔公司全球副总裁兼英特尔实验室研究主任Michael Mayberry 博士能听、能说、能读、能写.....人工智能要掌握如此多项技能必须拥有高[详细]

2017年 10月10日 16:15

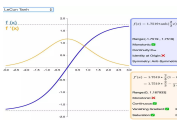
### 学界 | DeepMind提出Rainbow: 整合DQN算法中的六种变体



选自DeepMind机器之心编译参与: 李泽南、蒋思源「AlphaGo 之父」David Sliver 等人最近探索的方向转向了强化学习和深度 Q 网络 (Deep[详细]

2017年 10月10日 16:15

### 资源 | 从ReLU到Sinc, 26种神经网络激活函数可视化



选自GitHub作者: David Sheehan机器之心编译在本文中, 作者对包括 Relu、Sigmoid 在内的 26 种激活函数做了可视化, 并附上了神经网络[详细]

2017年 10月10日 16:15

### 四天速成! 香港科技大学 PyTorch 课件分享



机器之心整理参与: 黄小天、蒋思源前天, 香港科技大学计算机系教授 Sung Kim 在 Google Drive 分享了一个2天速成的TensorFlow课程



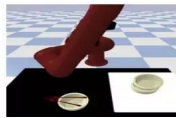
2017年 10月09日 13:45

深度 | 深度学习应该使用复数吗？

$i^2 = -1$  选自Medium机器之心编译参与：Nurhachu Nul、路雪深度学习只能使用实数吗？本文简要介绍了近期一些将复数应用于深度学习的若干研究，并指出使用复数可以[详细]

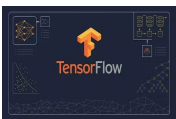
2017年 10月09日 13:45

学界 | 斯坦福提出神经任务编程NTP：让机器人从层级任务中学习

 选自arXiv机器之心编译参与：朱乾树、蒋思源斯坦福视觉与学习实验室与加州大学提出神经任务编程（NTP），它可以指定任务作为输入，并递归地将该任务分解成更精细[详细]


2017年 10月09日 13:45

观点 | TensorFlow sucks，有人吐槽TensorFlow晦涩难用

 选自nicodjimenez机器之心编译参与：李泽南、刘晓坤作为当今最流行的深度学习框架，TensorFlow 已经出现了两年之久。尽管其背后有着谷歌这样的科技[详细]

2017年 10月09日 13:45

教程 | 摄影爱好者玩编程：利用Python和OpenCV打造专业级长时曝光

 选自pyimagesearch机器之心编译参与：乾树、蒋思源在本文中，我们将学习如何使用 OpenCV 和图像处理技术来模拟长时曝光图像。为了模拟长时曝光，我们[详细]

2017年 10月08日 13:15

- 1
- 2
- 3
- 4
- 5
- 
- 

