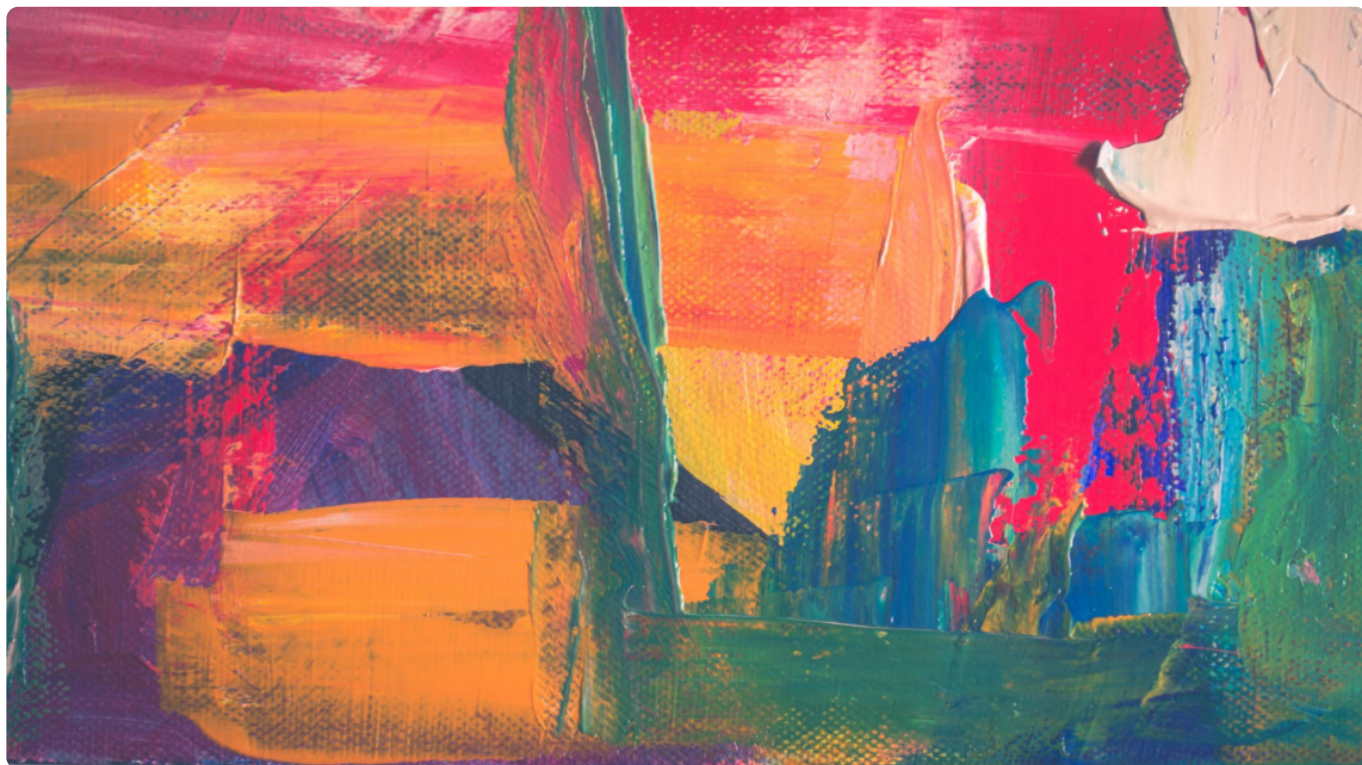


128 | CVPR 2018论文精读：如何研究计算机视觉任务之间的关系？

2018-07-27 洪亮劫

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 07:50 大小 3.59M



今年 6 月 18 日~22 日，计算机视觉和模式识别大会 CVPR（Conference on Computer Vision and Pattern Recognition），在美国的盐湖城举行。CVPR 大会从 1985 年开始举办，已经有 30 多年的历史，是计算机视觉领域的顶级会议。

最近几年，CVPR 大会发展成为了人工智能领域的盛会。受人工智能浪潮的影响，大会的投稿数量和参会人数都有了明显增加。大会今年共收到了 3300 份论文投稿，录取了 979 篇，录取率将近 30%。最终，选出了 70 篇论文做口头报告，224 篇论文做快速汇报。近两年的参会人数都保持着近 1 千人的增长势头，而今年更是达到了 6 千多人，是 2014 年参会人数的 3 倍多。同时，大会的审稿人也达到了惊人的 1 万人。

除了主会议以外，CVPR 大会还组织了 21 个讲座，48 个研讨班和博士论坛，有超过 115 家公司的赞助。

想要在这么多论文里找到最有价值、最有影响力的信息，可以说是大海捞针。我在这里为你精选了三篇今年 CVPR 的论文，希望能够起到抛砖引玉的作用。

今天，我们来分享大会的最佳论文，题目是——Taskonomy: Disentangling Task Transfer Learning。

我先来简单介绍下论文的作者群。

第一作者阿米尔·扎米尔 (Amir R. Zamir) 目前是斯坦福大学和加州大学伯克利分校的博士后研究员，已经在计算机视觉领域发表了 30 多篇论文，还获得过 CVPR 2016 的最佳学生论文奖。

第二作者亚历山大·萨克斯 (Alexander Sax) 刚刚从斯坦福大学计算机系硕士毕业，即将前往加州大学伯克利分校攻读博士，已经以硕士生的身份发表了两篇 CVPR 论文。

第三作者沈博魁刚从斯坦福大学计算机系本科毕业，即将在本校继续攻读博士。尽管是本科刚刚毕业，他已经发表了 2 篇 CVPR 论文和 1 篇 ICCV 论文。

第四作者利昂奈达·圭巴斯 (Leonidas Guibas) 是斯坦福大学计算机系教授，也是 ACM 和 IEEE 院士，还是美国工程院和科学院院士。他的博士导师是图灵奖获得者高德纳 (Donald Knuth) 。

第五作者吉腾德拉·马立克 (Jitendra Malik) 是加州大学伯克利分校计算机系教授，也是 ACM 和 IEEE 院士，并且是美国工程院以及科学院院士。马立克是计算机视觉方向的学术权威。

最后一位作者西尔维奥·萨瓦瑞斯 (Silvio Savarese) 是斯坦福大学计算机系的教授。他的研究方向是计算机视觉和计算机图形学。我们对华人学者李飞飞都很熟悉，萨瓦瑞斯是李飞飞的丈夫。

论文的主要贡献

概括来说，这篇论文主要是研究了计算机视觉任务之间的关系，并且提出了一个计算框架，能够定量地学习到这些任务之间的相似度。同时，这些相似的任务可以帮助数据较少的任务达到比较好的效果。这其实就是**迁移学习**（Transfer Learning）的核心思想：如何从已经学习到的任务或者领域迁移到数据较少、学习更加困难的任务或者领域。

很多研究人员在平时的研究过程中可能都会有这样的感觉，一些计算机视觉任务之间有某种逻辑或者直觉上的联系。例如，在计算机视觉界，像物体识别（Object Recognition）、景深估计（Depth Estimation）、边界发掘（Edge Detection）以及姿势估计（Pose Estimation）这些任务，大家都普遍认为它们是有关系的一系列任务。但是，有一些视觉任务之间的关系则显得没有那么直观，比如，边界发掘和光影（Shading）如何帮助姿势估计，就不得而知了。

如果我们单独来解决每一类任务，必然会有很大的挑战。这篇论文其实展示了，很多任务之间是有关联性的，而利用这些任务的关联性其实可以带来数据上的巨大便利。也就是说，我们可以利用更少的数据来学习到更多的任务。从这个角度来看，迁移学习也为新任务带来了希望，当我们没有大量的人工标注的数据时，依然能够在新任务上获得有效的结果。

这篇论文的另外一个重要贡献是提出了一个计算框架，这个框架并不需要事先准备的知识，比如人为地决定哪两个任务之间是有关联的，或者说，并不像之前的一些利用概率建模的方法，需要对任务之间的结构加以一个先验概率。**这篇论文提出的框架完全从数据和结果的角度出发，从而避免了这些先验信息的不完整和不准确。**

论文的核心方法

这篇论文提出的方法由四个组成部分，分别是：**任务相关的建模、迁移建模、任务与任务关系归一化以及最后计算任务的关系图谱**。每一个组成部分都有不同的目标。

首先，我们需要建立的是每一个独立任务自己的一个模型。这些模型有两个任务：第一就是尽可能地提高自身任务的精度；第二就是在这个过程中，尽可能提取有代表性的中间表征结果，能够有助于迁移学习。

第二个部分就是迁移建模。这个部分主要是利用第一部分学习到的中间表现层，然后再在目标任务上学习到从原本的表现层到任务目标的迁移。这里面，除了一个原表现层，或者是原任务可以借鉴以外，作者们提出还可以利用多个原任务，来达到提升效果的目的。这样也就把多个任务和一个目标任务关联了起来。

第三个部分是任务关系的归一化。这一部分其实是这篇文章的一个亮点。当我们得到迁移学习的结果以后，我们就可以利用每两个任务之间的关系来获得一个矩阵，这个矩阵就完全表征了所有任务的联系。然而，如果直接利用任务的迁移损失函数的值来刻画两个任务之间的关系，那么每两个任务之间的这个数值其实是没办法直接比较的。如果我们采用机器学习界归一化数据的办法，比如把数据归一到 0 和 1 之间，也是不行的，因为这样就完全没有考虑损失函数变化的速度和目标任务精度之间的关系。

所以，这篇论文的 authors 们提出了一种**按照次序来做归一化的方法**。简单来说，就是不再看两个任务之间的绝对的迁移数值，而是看在测试集上哪一个原任务相比于其他任务能够更多地获取目标任务的精度。这样所有的任务就可比了。总之，**任务关系归一化的目的就是构建了任务与任务之间关系的矩阵。**

最后一个部分的目的就是**从这个关系矩阵中提取出所有任务的一个真正的关系图谱**。也就是说，我们希望从一个完全的全连通图，找到一个最有价值的子图。在这里，authors 们采用了一种叫作“**布尔值整数规划**”（Boolean Integer Programming）的方法，在一些限制条件下，挖掘出了一个有代表性的子图。

实验结果

authors 们提出了一个有 4 百多万张图片的新的数据集。在这个数据集里，有 26 个计算机视觉任务。从实验中，authors 们挖掘出了这样一些情况，例如 3D 的、2D 的任务自然被归类到了一起，而其他的例如上下文分割、景象识别这种高层次的任务则被分在了一起。

为了研究这种挖掘出的结构是否真的能够实现迁移学习的目的，authors 们还把不同的两两任务随机组合在一起，也就是某种随机任务的图谱，按照学习到的结构进行迁移学习，看是不是比随机结果要好。答案是，的确要好很多。在这篇论文里，authors 们展示了学习到的结构不仅能够帮助目标任务提升性能，而且在任务之间关系的解释性上效果也非常不错。

小结

今天我为你们讲了 CVPR 2018 的最佳论文。

一起来回顾下要点：第一，我们详细介绍了这篇文章要解决的问题以及贡献，论文研究了计算机视觉任务之间的关系，并且提出了一个计算框架，能够起到迁移学习的作用；第二，我们简要介绍了文章提出的核心方法，主要有四个组成部分；第三，我们简单介绍了论文的实验结果。

最后，给你留一个思考题，当前挖掘的关系主要是任务的两两关系，能否有一个方法挖掘任务的高维度关系，比如三个任务之间的关系？

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 127 | SIGIR 2018论文精读：如何对搜索页面上的点击行为进行序列建模？

下一篇 129 | CVPR 2018论文精读：如何从整体上对人体进行三维建模？

精选留言 (1)

 写留言



晏

2018-07-27

希望有图片能展示核心方法的设计思路。

展开 ▾

 1

