

大数据与云计算—震撼开班

教材作者 实践+理论 全面讲解

智能硬件大数据

免费托管平台

阅读新闻

背景: □□□□□□□□

2015深度学习回顾：ConvNet、Caffe、Torch及其他

[日期：2015-08-06]

来源：ComputerVisionBlog 作者：

[字体：大 中 小]

CVPR可谓计算机视觉领域的奥运会，这是vision.ai的Co-Founder，前MIT研究人员T. Malisiewicz针对CVPR'15尤其是Deep Learning的综述文章，谈到了ConvNet的Baseline，Caffe和Torch之间的分歧，ArXiv论文热，以及百度的ImageNet违规事件等。原文标题为：Deep down the rabbit hole: CVPR 2015 and beyond。

CVPR是主要的计算机视觉会议，可以把它看作是计算机视觉研究的奥林匹克。今年，CVPR将在我的后院举行，离可爱的剑桥不到一英里。我很多麻省理工的同事都参加了，如果Google在这次CVPR 2015上有着最好的表现，我也不会感到惊讶。从2004年开始，我几乎参加了每年的CVPR，那么，让我们来一场计算机视觉研究世界新而短暂的精彩旅行吧。

灰白阴影下的兔子洞艺术

多年来，发生了很多变化，但本质上还是没变。学者曾经是最顶端的，捍卫他们自己的大学以及发生在他们非商业研究实验室的令人惊叹的事情。如今，学者仍然处于顶端，但是现在维护的是他们的Google、Facebook、Amazon以及X公司的某个附属子公司。随着招聘的最佳预算和以出版为导向的沉重企业文化，如果接下来连续几年的大量学术外流，我也不会感到惊讶。由于CVPR只有两周，因此Google一直在忙于制作ConvNet(卷积神经网络)艺术，同时向世界展示了如果你想做最好的深度学习研究，他们就是(合适的)王国。

仅仅是博士生和博士后的部队并不能打败软件工程师和科学研究者部队。在以前，学生在获得计算机视觉的博士后之后通常会选择离开学校(通常会受到一些视觉研究工作和华尔街的工作的诱惑)。现在，之前的博士生在大公司运作研究实验室已经紧锣密鼓的进入我们的视野。似乎还没有足够的深度专家来填补这个大需求。

数据集通常是一件大事——请下载我们的数据!数据集依旧是件大事——但是我们抱歉告诉你，你所在大

Digg排行

181 云计算入门必读(图)

93 轻解云裳：图解云计算（轻松有趣）

86 云计算的概念和内涵

69 不懂这显然您不专业！云计算术语

36 云计算Cloud Computing简介

31 华为桌面云深度体验

30 云计算的概念和原理

23 什么是云计算？

21 【视频】认识云计算（王翔）

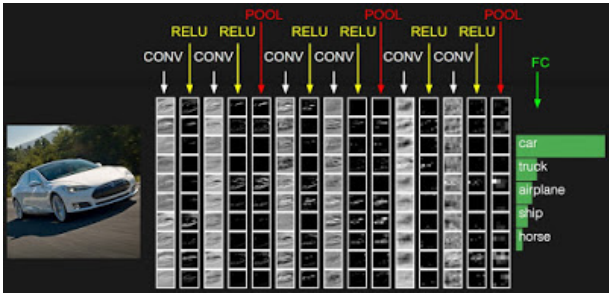
19 Gleasy：云端运行的操作系统

学的计算资源达不到要求(但幸运的是，我们 X 公司总在招聘，所以来加入我们吧，让我们一起推动研究的向前发展)。

如果你想要查看个人文献，我建议Andrej Karpathy的 CVPR 2015文献在线导航工具或者看看在计算机视觉基础网站上的CVPR 2015文献 。还有就是一个麻省理工学院的博士候选人Zoya Bylinskii，也列出了一些有趣的CVPR 2015 的文献。

ConvNet革命：一个网络预训练

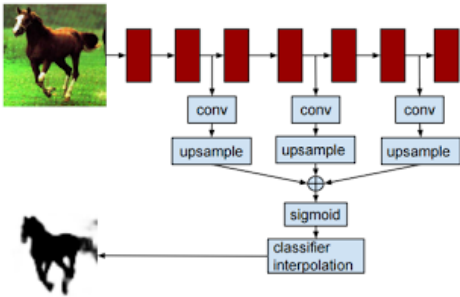
机器学习过去是女王，现在是国王。机器学习过去是浅显的，但如今的学习方法如此之深，以致于图表在单个滑动下很难拟合。研究生过去常常避开了关于Yann LeCun的笑话，Yann LeCun坚持认为机器学习在将来的某一天会做特征设计阶段的工作。现在似乎是这种情况，当你坚持认为“手工特征设计”将省掉一天时间的时候，整个视觉社区都会忽略你。Yann LeCun做了一个主题报告，并给了它一个有趣的标题：“深度学习怎么了？”，这表明了卷积神经网络(也叫CNNs或ConvNets)存在于CVPR的任何角落。



图来自Karpathy的卷积神经网络教程

过去在CVPR很难发表ConvNet(卷积神经网络)研究论文，而现在如果你没有对ConvNet做一个基本的比较的话，很难得到一篇CVPR文献。得到一个很酷的新问题了么?哦，你没有试一下基于ConvNet的基本方法么?很好，这说明为什么没人关心了。

尽管如此，但这并不是机器接管了视觉科学家的工作。如今的视觉科学家更是一个应用型机器学习黑客，而且由于强大的CNN主题，理解和重新实现如今的视觉系统变得更加容易。我们在CVPR上看到的在本质上是一个类似分割和运动的经典问题的回访，只是使用的是这种新的机器方法。正如Samson Timoner在本地Boston Vision Meetup网站上概括的那样，当互信息变得流行，社区也随之变得时尚——这次围绕的是ConvNets。但这不仅是一种趋势，非CNN(卷积神经网络)的竞争正在被摧毁。



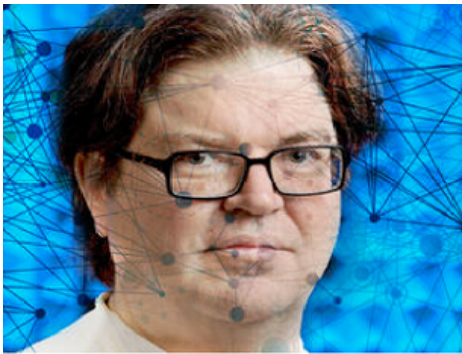
来自Bharath HariharanCVPR 2015的文章-在切割上使用卷积神经网络

还有很多事情等着视觉科学家去做，一个扎实的数学上的正式教育比其他任何都要重要。我们过去是使用梯度下降来训练，现在也是如此。就好比我们过去喝咖啡，现在也喝咖啡一样。其本质上，其实还是数学。

内心深处的兔子洞

CVPR2015让人想起了物理学上牛顿之前的日子。许多聪明的科学家能够使用数学来预测物体的运动，曾经聪明的笛卡尔教会我们如何将我们的物理思维考虑到坐标系统中。非常清楚的是：通过卷积神经网络语言来铸造你的计算机视觉问题，你将打败所有用手做计算机视觉的人。我认为Yann LeCun(深度学习之父其中之一)就是一个现代的笛卡尔，只是因为我认为开创性的工作是指日可待。他的ConvNets思想框架就像是一个必备的坐标系统——我们可能不知道目的地像什么，但我们现在知道如何来绘制一张地图。

深度网络每个月都表现的很出色，但我仍然等待着艾萨克(牛顿)的到来，让我们的生活更轻松。我想要一个简化，但我并不悲观——一个很好的原因就是会有一系列的ConvNet空间活动(以防你没能参加CVPR 2015)，所以我直言不讳：ConvNets这该死的工作!其实我只想要深度学习的F=ma公式。



Yann LeCun: 1960 - Current
Father of Convolutional Neural Networks



René Descartes: 1596 - 1650
Father of Cartesian Geometry

计算机视觉的开源深度学习：Torch VS Caffe

CVPR2015开始的第一天是一些优秀的软件教程。外面有许多优秀的非- α 深度学习软件，并且让每个人的生活变得容易。在CVPR上，我们既有Torch教程也有Caffe教程。我参加了DIY深度学习的教程，这是一个充满Caffe的屋子——在会议开始的5分钟前，站在房子里的参会人员只有像我这样的懒虫。相比之下，Caffe是更受欢迎的，但对于Torch，当谈及到一些深度学习的权威人员时(如+Andrej Karpathy和其他深度思维科学家)，一些特定的专家小组似乎从Caffe转移到了Torch。



Caffe是在Berkeley发展起来的，有一些充满活力的社区，与Python结合并且在大学生中似乎相当流行。Trevor Darrell教授甚至找了一些博士后来帮助Caffe发展。如果我再年轻几岁并且刚获得博士学位，那么我一定会申请的。

Torch并没有跟随Python的潮流，它是基于Lua的。对于解释器没有必要像Matlab或者Python那样，Lua会给你神奇的控制台。Torch被Facebook人工智能研究实验室和位于伦敦的谷歌DeepMind大量使用。对于那些害怕类似于Lua这样新语言的人，在此不用担心——如果你已经涉足Python，JavaScript或者Matlab的话，那么你会感到Lua语言学起来特别“轻松”。

现在，越来越清晰的是：深度学习的未来主要是来自像Caffe或Torch那样有自给自足的软件包，而不是像OpenCV或Matlab那样的处在生死边缘上的全能工具。当你在OpenCV上分享创作，你最终会分享源代码，但有了深度学习工具包，你最终提供的是你的网络预训练。对于你的计算机视觉管道，你不必再考虑20个“小”算法的组合——你只要考虑你想要的流行网络架构，然后就是数据集。如果你有GPU和庞大的数据，你可以做完整的端到端的训练。如果你的数据集或者中等数据集，你可以微调最后几层。你甚至可以在最后一层训练的顶部训练一个线性分类器，如果你怕麻烦——那么你要做的只是超越SIFT(尺度不变特征变换算法)，HOG(方向梯度直方图)，GIST(通用搜索树)以及所有在计算机视觉过去二十年里庆祝的算法。

在CVPR 2015上使用ConvNets的方式，使我感觉到我们正在接近某个大的东西。但是在我们捡到黄金之前，ConvNets仍然像是一个微积分的影子，只是“希望”会得到更大，更深层次的东西和更有意义的事。我认为，对于ConvNets的调查可视化算法的研究表明，即使网络建筑师也不能完全确定幕后发生的事情。

嵌入大脑的视频游戏引擎：面向机器智能的不同路径

在CVPR 2015会议的最后一天的现场分析研讨会上，Josh Tenenbaum 给嵌入大脑的视频游戏引擎赋予了一个诱人的讨论标题。你可以在一篇短篇《科学美国人》文章中读到他观点的概括。尽管他的说话风格看起来似乎不符合CVPR的标准，但它却是典型的Tenenbaum。在他的世界里，没有基准可以超越，没有曲线可以拟合阴影，如果你让我把LeCun和笛卡尔作比较，那么LeCun在某种意义上算是教授吧。Tenenbaum可能是现代的亚里士多德。正如Jianxiong Xiao教授全面的介绍Josh那样，他或许是对的——这是你能找到的最聪明的扬声器之一。它在一秒内能说100个单词，你会觉得你的大脑在放大你听到的。

Josh的一个主要研究主题是基于图像识别阴影的超越。Josh的所有的工作都是关于在这个世界上建立思维模型，而他的工作确实可以被看作是综合分析。在他模型的内部主要是一些类似于视频游戏引擎的东西，而他展示了许多令人信服的相关实例，这些实例对于人类来说很是方便，但对于今天数据驱动的ConvNets而言，几乎是不可能的。在这个夏天，如果他的学生在谷歌DeepMind工作，那么也不必感到惊讶。

几年前，概率图模型(图论和概率方法的合并)那是风靡一时。Josh给了我们概率编程的味道，虽然我们还没有看到新的方法来主导计算机视觉研究世界，但请睁大你的眼睛。他提到了一篇最近发表在Nature上的



科技头条

科技爱好者的情报站

集成10个科技门户网站(含本站)
和微信热点

科技头条——科技爱好者的情报站，是一款挖掘互联网科技类新闻、微信公众号热点文章的阅读客户端。酷炫科技，远见未来。

精选微信朋友圈100000+的热点文章、知名微信公众号，每日智能推送您关心的内容，帮助您更加高效地获取科技信息、追踪科技热点。

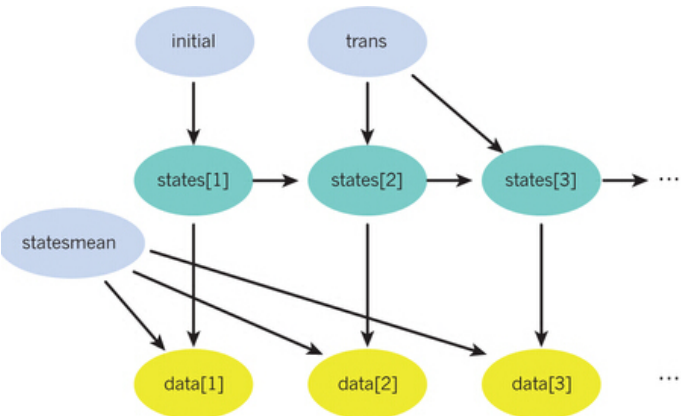
(目前仅支持安卓客户端)

文章(下面引用的), 来自另一个受人尊敬的机器学习研究, 这应该会引领潮流兴奋相当长一段时日。请看看下面Julia代码中尾部的一部分:

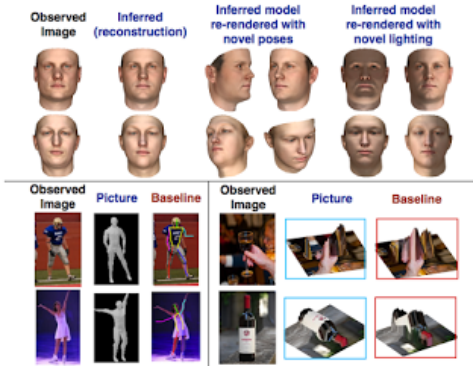
Probabilistic machine learning and artificial intelligence.Zoubin Ghahramani. Nature 521, 452–459 (28 May 2015) doi:10.1038/nature14541

```
statesmean = [-1, 1, 0] # Emission parameters.
initial     = Categorical([1.0/3, 1.0/3, 1.0/3]) # Prob distr of state[1].
trans      = [Categorical([0.1, 0.5, 0.4]), Categorical([0.2, 0.2, 0.6]),
              Categorical([0.15, 0.15, 0.7])] # Trans distr for each state.
data       = [Nil, 0.9, 0.8, 0.7, 0, -0.025, -5, -2, -0.1, 0, 0.13]

@model hmm begin # Define a model hmm.
  states = Array{Int, length(data)}
  @assume(states[1] ~ initial)
  for i = 2:length(data)
    @assume(states[i] ~ trans[states[i-1]])
    @observe(data[i] ~ Normal(statesmean[states[i]], 0.4))
  end
  @predict states
end
```



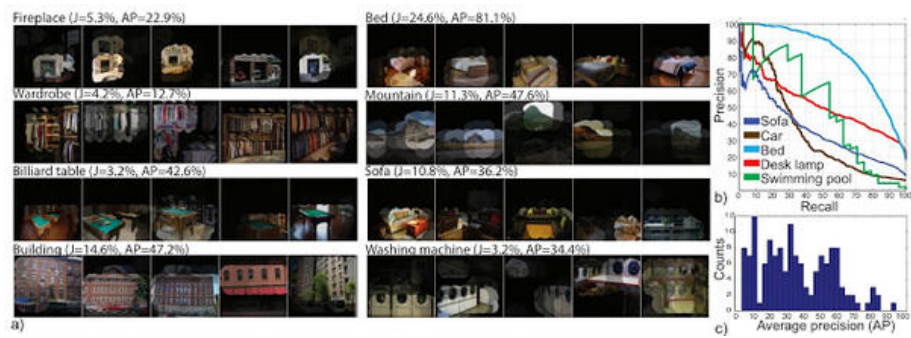
下面来看一看一些教授。Tenenbaum的想法正在行动中, 看看下面的这篇CVPR 2015的文章, 标题为: 一种面向场景感知的概率编程语言。恭喜Tejas D. Kulkarni, 第一作者, 一个MIT的学生, 因为这个令人兴奋的新作而获得最佳论文荣誉奖。有了谷歌DeepMind, 你会有一个有趣的夏天。



Picture: A Probabilistic Programming Language for Scene Perception

深度场景CNNs物体检测器

在现场分析研讨会上有着大量的好的演讲报告, 在另一次演讲中, 真正突出的是一个新的大型数据集(麻省理工学院的地方)和当你使用场景与物体来训练发生什么时的一个彻底调查。



来自于麻省理工学院的Antonio Torralba做了一个关于地点数据库的讨论，以及当你做以物体为中心的数据库(如ImageNet)与以场景为重的数据库(如麻省理工学院的地点)训练时能学到什么的一个深层分析。你可以查看“目标探测器浮现”幻灯片或arXiv论文来学习更多的指示。这些都来自即将到来的研究员Bolei Zhou的出色工作!

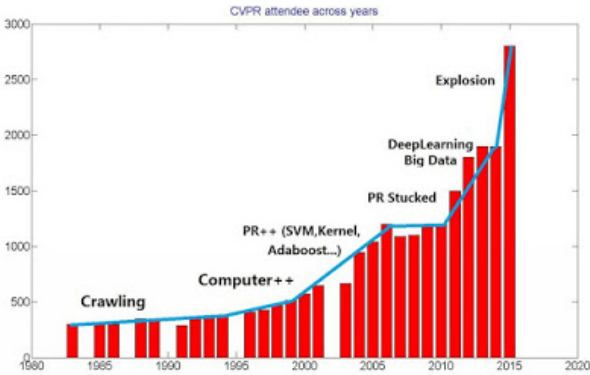
CVPR无意之事：ArXiv出版狂热 & 百度惨败

从长远来看，最近快速推到ArXiv.org的预印本热潮对学者和商业研究者老说贡献都很大。当你有大量的专家以最快的速度探索思想，然后等待6个月直到下一次回忆的最后截止日期，这样是没有意义的。唯一的缺点就是，它使新发布的CVPR论文过时了。这就像在ArXiv发布文章的那一天每个人都已经详细精读了这个好东西。但你得到了你的“想法主张”而不用担心一个不适当的评审影响你。双盲评审，将准备进行一个深度改造。我们现在知道谁在做什么，特别是在发表时间之前。对于学生，发布或退稿仅仅得到的是一个结果。ARXIV狂热是一件好事还是坏事取决于你，可能更多的是你的资历比其它什么都重要。在接下来的日子里，CV的各种声音，肯定会越来越响，并会持续如此。

外行人仅仅是阅读人工智能标题的话，百度作弊丑闻似乎是个大新闻，但在计算机视觉上，在测试集上的过拟合并不是一件新颖的事情。这篇论文最后被驳回了，研究生通常多次在测试集上评估他们的算法，而真相就是人无完人。当它对#1很重要的时候，不要对你的比较变得不合适而感到惊讶。重要的是意识到地面破碎研究和百分比追逐的不同。我们都会犯错，在重大的压力下，我们都会表现出自己的弱点。所以，让我们一笑置之，坦然面对吧。让我们招聘最好的，鼓励真正伟大的研究并且停止追逐百分比。事实上，很多的顶级性能方法都是相似的。

总结

CVPR的参会者在持续的增加。我们现在有博士生，创业者，教授。招聘人员，大公司，甚至是来学生，来展示自己的成果。CVPR会成为新的SIGGRAPH(美国计算机协会)么？



来自Changbo Hu绘制的CVPR参会者图片

ConvNets留在了这里，但是如果我们想要让ConvNets能超越阴影的计算，那么还有大量的工作需要完成。Geoff Hinton的capsules在深夜的讨论中不断出现。“我想使用神经元组来取代非结构化层，我称之为‘capsules’，它更像一个皮层柱”——Geoff Hinton的新闻网站AMA。很多人(像来自CMU的Abhinav Gupta教授)也在讨论非监督学习CNN(卷积神经网络)的训练，我预测，在明年的CVPR中，从没有注释的视频学习大型卷积神经网络应该是一个大的方向。

最重要的是，当深度学习的巨头去提他们最喜欢的方法有错误时，我只期待有最好的研究也会随之而来。快乐的计算吧，并记住，永远不要停止学习。



我的PM2.5
扫二维码下载APP

随时随地告诉你最精确的PM2.5值！走南闯北，随时查看，PM2.5值随你而变！不是城市的平均值，是用离你最近的数个PM2.5传感器，在云计算平台精确计算出来的实时值。还可接入酷炫的“我的PM2.5”，配套智能硬件，实时监测室内空气质量。

5

顶一下

分享 阿九的古道，李鸢gaso... 等15人分享过



收藏 推荐 打印 | 录入：574107552 | 阅读：3997 次

华为企业云：压强式投入自主研发，确保技术领先

那些脱颖而出的云计算认证

同声译



扫二维码下载APP

跟外国人交流有困难？您需要带上同声翻译官！直接对着手机说话，就能翻译并说出来，中译英，英译中。您能听懂老外的话了，老外能听懂你的话了。接待外宾的神器！出国旅行的神器！



中国云计算论坛



云创大数据



中国存储资讯



中国智慧城市资



中国_物联网



大数据资讯



中国PM25资讯



专注云计算



笑看云存储

一键关注

注册微博

相关新闻

本文评论 查看全部评论 (0)

表情: 姓名: ☒ 匿名 字数 0

点评:

☒ 同意评论声明

- 评论声明
- 尊重网上道德，遵守中华人民共和国的各项有关法律法规
 - 承担一切因您的行为而直接或间接导致的民事或刑事法律责任
 - 本站管理人员有权保留或删除其管辖留言中的任意内容
 - 本站有权在网站内转载或引用您的评论
 - 参与本评论即表明您已经阅读并接受上述条款

版权所有:© 2008-2016 南京云创大数据科技股份有限公司（股票代码：835305），保留一切权利。苏ICP备11060547号-9

本站所转载、链接、引用及翻译的内容等，其版权属于各自合法的拥有者。

本站创建、管理、商务合作：zhujianguangjie#163.com(用@代替#)

Powered by iwms 5.4
Processed in 0.03 second(s), 3 queries, Gzip enabled