

# 计算机视觉 | 纽约读博五年总结（二）谈谈写paper和做报告

Mike Shou CVer 2020-12-23

点击上方“CVer”，选择加“星标”置顶  
重磅干货，第一时间送达

本文作者：Mike Shou | 编辑：Amusi (CVer)

<https://zhuanlan.zhihu.com/p/338191470>

本文已由原作者授权，不得擅自二次转载

还没看过（一）的同学，先看一下这个帖子！学习一下：

- 计算机视觉 | 纽约读博五年总结

**Disclaimer:** 因为篇幅限制，有的问题难以讲得完善全面，或跟您意见不合，全当看小说，寻个乐子罢了。

## 3. 谈谈Presentation

Presentation分为作报告，还有就是写paper

### 3.1 谈谈做报告

14年，刚来哥大那会儿，每周五是我们组会，导师让我在组会上present RCNN，这是我第一次给老板作报告，而且是在全组面前报告，自然想要好好表现。我对object detection之前完全没了解过，于是paper读了一遍又一遍，文中用到的前人技术不懂，便找到前人的文章去学怎么回事。感觉自己学到好多，自我感动，觉得花了这么大工夫，一定到时候会让导师刮目相看。

结果就是啪啪打脸：导师极其严谨，当我解释了A，解释了B，问我已经有A了为啥还要B；我大脑一片空白，尝试着解释了半天，导师表示听不懂，这是可想而知的，因为其实我自己也并没有搞清楚为啥；当时自己读paper的思维模式其实只是，memorize怎么做怎么做，但没有去搞清楚要这么做背后的原理；另外当时slides准备的也不够好。总之结果是，这次报告搞砸了。还好事后，师兄们继续跟我讨论，让我对技术原理有了更深刻的认识；有的师兄更用亲身经历安慰我，说当年他刚来的时候，第一次汇报工作连slides都没做，干讲，导师自然也是没能听懂。

之后几年，从导师身上学到了很多presentation的技巧：

- 如果可能的话，**事先了解你的听众背景**，是跟你做同一个topic的，还是同一个大领域但不同topic的，还是完全其他专业背景的。需要根据听众背景，定制和调整：比如，需不需要多介绍些背景？需不需要更深入技术细节？等等
- **一页slide尽可能focus在一个点上**，不要信息量过大，否则听众很容易lost
- 尽可能多用图片表达，不要大段大段的列文字，**A picture is worth a thousand words**
- 上面这两点，其实principle都是尽量让要讲的内容简单明了，因为很多时候我们在听talk，这样被动接受的时候，接受新知识的能力是比主动接受时候（比如看paper）低的。
- 当听众问问题的时候，**If you don' t know the answer, just say don' t know.**
- 如果是跟mentor日常讨论的slides，因为会讨论到很细节的东西，有些图PPT画起来，很花时间，而且通常这样细节的图还挺多，所以可以就ipad上面手画一画，截个图放到PPT里就好了；如果是正式一点的presentation，写slides跟写paper的principle有点像，不要太focus在细节上，更重要的是讲清楚motivation，为什么这样设计，细枝末节的不关键的内容，放在backup slides里面。

19年CVPR，Doctoral Consortium有幸mentor是斯坦福的一位大牛教授，她也提到了presentation的重要性，说她们lab有个开玩笑的说法，一份slides交给她去改，no pixel left.....为了分享如何能让报告听起来有兴趣，她画了下面这张图，**让听众情感（亦是兴趣高低，注意力程度）随着时间的变化，有三个高潮**：首先，介绍你的问题，通常这时候大家都会引发兴趣；但听着听着大家注意力就不集中了，这时候就到了图中第一个低谷，这时候需要指出来这个问题有哪些challenge，大家的兴趣就又被激发了；等大家兴趣来了，精力集中的时候，介绍你的一部分工作work 1；等介绍完第一个工作，大家又疲劳了，这时候指出来，即使有这个work 1，问题还不能被解决，因为有remaining challenge；接着大家又被调动了兴致，可以开始介绍work 2。

## 3.2 谈谈写paper

在2.2里面讲了对某一篇paper，如何选题和做规划。那真的到了写paper的时候，我自己有几点如何让文章写的更好的体会：

- **先给一个Talk。**写paper最难的是构思storyline，而最好的完成这一步的方法就是先对你的工作做一个slides，给周围的人present一遍。这个过程中，你会梳理好自己的思路，画好文中的figure，准备好实验结果的table，周围的人还可以给你提意见，帮助你完善，等这个talk给完了，后面写paper就会顺畅自然了。其实我现在，如果准备投一个paper，当做了一段时间后，就会按照最终presentation的思路，准备slides，用在每周给老板们report时。开头先快速review一下做的task和提出的方法，remind一下context，然后重点focus在那周做的新东西上，所以每周汇报的slides可能80%都是跟上一周一样的，然后新的方法和实验结果的那几页slides是新的，有比较多的细节。
- **用Google doc做语法检查。**刚写好的paper有typo和语法错误是很难避免的，但常常会被reviewer揪着不放。大家写paper如今大都在overleaf上，但overleaf的查错还是不够好，建议可以写完paper后，贴到Google doc里面。几年前开始，估计是由于deep learning对Google NLP的改进很大，感觉Google自动改的质量已经非常高了。
- **Rationale很重要。**不光是要讲清楚你怎么做的，更要justify你问什么这么做；不光要讲你的结果比baseline好，更要解释为什么好；读者看到的不应是一个“使用手册”。有时候我们写paper，花了很多篇幅写了很多实现细节，但是更重要的是，解释“为什么”，这个背后的逻辑和insights。
- 大部分paper都是提出一个新的方法，**这类方法型paper**似乎都可以套下面这个框架：
  - Introduction：可以分为以下几个部分：
    - Problem definition
    - Previous methods and their limits

- 简单描述你是提出了什么技术来overcome上面的limits
- 一个图，非常high-level的解释前人工作的limits和你的工作怎么解决了这些limits，**最好让人30秒内完全看懂**
- 最后一段如今大都是，In summary, this paper makes three contributions:
  - First work to解决什么limits
  - 提出了什么novel的技术
  - outperforms了state-of-the-art多少
- Related Work：一般三五个subsection，分别review下相关的topics，同样不光讲previous work做了啥，更要讲自己的方法跟前人工作有啥不同
- Method
  - 这是文章的主体，按照你觉得最容易让别人看懂的方式来讲
  - 可以第一个subsection是overview，formulate一下你的problem给出notation，配一个整体framework的图，图里面的字体不能太大或者太小看不清，要有些细节，让人光看图就能明白你的方法是怎么回事，但不要过于复杂，让人在不超过2分钟的时间看完这张图
  - 然后几个subsection具体介绍你的方法或者模型；如果testing跟training不太一样，最后一个subsection介绍inference时候的不同，通常是一些post-processing操作
- Experiment
  - Datasets
  - Implementation details such as pre-processing process, training recipe
  - Evaluation metrics
  - Comparisons with state-of-the-art
  - Detailed analysis
    - Alternative design choice exploration
    - Ablation studies
    - Visualization examples
- Conclusion (and Future Work)
- Abstract：是全文的精简版，建议在paper写完第一稿差不多成型了，有定下来的成熟的storyline了，再去写abstract；大概就是用一两句话分别概括paper里面每个section，然后串起来

另外paper提交时候，可以交supplementary materials，虽然reviewer并不被要求强制看这个，但其实给我们机会，去include更多文章技术细节、实验结果的好地方；在后面rebuttal阶段，通常篇幅有限制，但如果你已经在**supp里面未雨绸缪**，可以省很多空间，refer reviewer去看你supp里面的内容就好了。

说到rebuttal，我还是比较幸运的，从导师那学到很多。导师已经是功成名就，业界泰斗那种，起初我以为他这个级别会对我们是放养；但我在哥大投自己第一篇一作paper的时候，导师可以说是手把手带我入门了。还记得16年CVPR review出来后，导师找我讨论rebuttal，我那会儿对写rebuttal并没有什么经验，也不知道可以用R1代表review 1等等。那天meeting开始已经晚上7点了，估计导师还没吃饭，我两就挨着坐在他办公室里，对着他的电脑，讨论

reviewer提的一个一个问题。因为很多时候其实reviewer表达问题并不准确，他教我分析每个问题背后reviewer真正关注的点是什么。边讨论，导师边直接敲下我们讨论的notes，meeting完后，我看这notes基本上就可以算是个rebuttal的初稿了，比我meet前准备的draft强多了.....

---End---

上述是【计算机视觉 | 纽约读博五年总结】系列的第二部分，剩余部分即将更新！**如果期待后续内容，请点个赞和在看到，支持一下！谢谢！**

## 目标检测综述下载

后台回复：**目标检测二十年**，即可下载39页的目标检测最全综述，共计411篇参考文献。

## 下载2

后台回复：**CVPR2020**，即可下载代码开源的论文合集

后台回复：**ECCV2020**，即可下载代码开源的论文合集

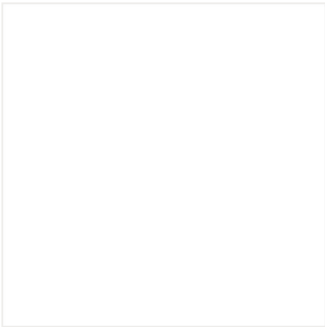
后台回复：**YOLO**，即可下载YOLOv4论文和代码

## 重磅！CVer-论文写作与投稿交流群成立

扫码添加CVer助手，可申请加入**CVer-论文写作与投稿** 微信交流群，目前已满2400+人，旨在交流顶会（CVPR/ICCV/ECCV/NIPS/ICML/ICLR/AAAI等）、顶刊（IJCV/TPAMI/TIP等）、SCI、EI、中文核心等写作与投稿事宜。

同时也可申请加入CVer大群和细分方向技术群，细分方向已涵盖：**目标检测、图像分割、目标跟踪、人脸检测&识别、OCR、姿态估计、超分辨率、SLAM、医疗影像、Re-ID、GAN、NAS、深度估计、自动驾驶、强化学习、车道线检测、模型剪枝&压缩、去噪、去雾、去雨、风格迁移、遥感图像、行为识别、视频理解、图像融合、图像检索、论文投稿&交流、PyTorch和TensorFlow等群。**

**一定要备注：研究方向+地点+学校/公司+昵称（如论文写作+上海+上交+卡卡）**，根据格式备注，可更快被通过且邀请进群



▲长按加微信群



▲长按关注CVer公众号

整理不易，请给CVer点赞和在看！

阅读原文

喜欢此内容的人还喜欢

年薪201万！华中科大CV博士生入选华为"天才少年"计划！  
CVer

多校划片暴雷后，学区房还能买吗？  
樱桃大房子

有冰川「擦痕」的人  
人物