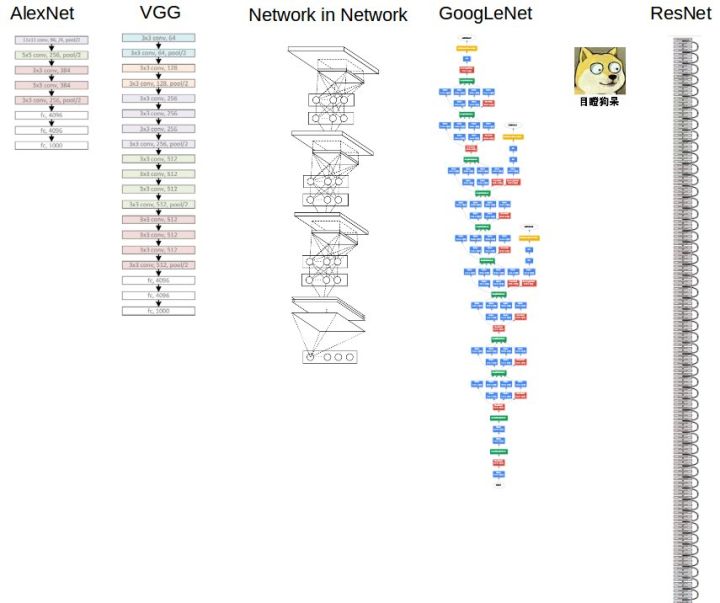
# 为什么现在的CNN模型都是在GoogleNet、VGGNet或者AlexNet上调整的？

https://www.zhihu.com/question/43370067



1. **这些网络都是一些知名的研究机构设计出来的，为领域提供了一个baseline**。

做深度学习的相关研究，一个很重要的点在于将自己的方法和其他人的方法进行对比。对比的时候就要讲究公平性的问题。在ResNet出来之前，多数研究者在基于VGG做相关领域的研究，这样在相同的网络深度下，就检验自身提出的东西的有效性了。还有整个科研界对于VGG、AlexNet、ResNet这些网络有一个基本的认识，你提出的方法在这种baseline下能到什么样的效果大家心里有数，比如在目标检测这个任务中，基于VGG的网络在COCO上边最多能做到27%-28%的mAP，而ResNet就可以做到30以上，这样就便于对比和评价了。再比如一开始用于语义分割的FCN是基于VGG的，ResNet出来之后很容易的变成基于ResNet的，这也侧面说明了方法的有效性。

2. 这个领域只有一小部分人是在研究分类的网络结构，**其他的应用的研究者**，包括视觉、自然语言等**更专注于本领域的内部知识**，在一个base网络的基础之上进行修改，以验证自己方法的有效性。而这些网络正好提供了这样的baseline，所以何乐而不为呢。自己设计的网络结构有可能也会达到很好的效果，但这就偏离了自己本身的研究点。当然如果一个人足够厉害，比如这篇文章([YOLO: Real-Time Object Detection](https://link.zhihu.com/?target=http%3A//pjreddie.com/darknet/yolo/))，作者完全是自己设计、训练的网络结构，这也是完全没有问题的。**研究来讲，最终的准确率只是追求的一个方面**。更重要的，是提供解决问题的一个新途径、思路或者方法。如果大家都去拼准确率、拼机器、搞更深的模型，这个领域就不会发展到现在这个地步了。另外，很多高校，根本就没有这种条件来做这样的事情。

3. **科研与工业应用的区别。**科研界基本上都在这些现有的模型的基础之上做一些改进，而实际的应用中会遇到各种各样的问题（数据标注、计算资源、效率等）。据我所知，一些机构在实际场景中，都在训练适合于自己应用的模型，这些模型可以达到相当好的性能。而这些涉及到商业的东西并没有放出来供大家使用。