2011夏天在Google Research 实习，当时深度学习主要集中在无监督学习，主要起源于3点：

1,人类学习是无监督式的

2.学术界大多都在研究无监督学习（RBMS,autoencoders）

3.无标签数据更易获得

2012 在斯坦福攻读phd

3D无监督式学习（但他个人不满意）

随后他加入了李菲菲实验室并寻找一个计算机视觉的研究方向，此时深度学习在计算机视觉仍不受待见

CNN在ImageNet Classification challenge 中夺冠，并且在其他数据集上也有很好的表现，这一现象震惊了所有人(监督式的)，而非监督式的主要目的就是为了学习未被标记的数据的表示。那么这两者有什么联系呢？ 或许我们应该以多人任务目标函数训练监督式数据，而不是对未被标记的数据进行无监督式学习。

2013年夏天在Google Research实习

此时，Google Research已经将研究方向集中于深度学习

2014 参加ImageNet Large Scale Visual Recognition Challenge ,错误率仅为6.7%

发表了Deep Fragment Embeddings for Bidirectional image sentence mapping,并公布了matlab源码

{

输入：N张图片以及相应的句子描述

}{

2015 Deep Visual-Semantic Alignments for Generating Image Description(importance)

输入：N张图片以及相应的句子描述

输出：图片的整体描述以及各区域描述

phd:

输入图片以及相应的句子描述

输出测试图片的句子描述，以及局部特征描述