上一周开始选的论文是北大学生搞得，主要是看重他的神经网络结构，但实际深入阅读后发现，这篇论文主要着重于对联的生成(…………)，然后其模型也主要就是RNN的Encoder和Decoder，且还未我上次看的那篇那么详细，就不赘述了，为了让对联生成更有重点，也用了attention机制（并不创新），唯一有点意思的就是介绍了convolutionary　neural　network，这个模型能更好保持语意的连贯性，且在第二篇论文中也有提及，就稍后再说，关键是我看这篇文章的评估结果也不出彩……

为了拟合图片生成古诗的主题，选了一篇image　inspire　poem的论文（科大的），它主要能从图像中抽取信息生成关键词向量Ｋ= {k1,k2,...,kN}和视觉特征向量V = {v1,v2,...,vB}，然后逐行生成诗句Ｌ＝{l1,l2,..,lL}，这个模型叫做MIPG模型。模型分为两部分，一个叫Image-based　Encoder　(I-Enc)，另一个叫Memory-based　Decoder(M-Dec)，

I－Enc：

从图像中提取视觉特征向量Ｖ= {v1,v2,...,vB},视觉特征提取器使用CNN完成的，即V = CNN(I).

（对CNN不了解的话可以上这个网站看看<http://www.sohu.com/a/277526497_100007727>）

对于诗歌的L生成，用双向的GRU模型实现（估计和我们之前看的方法大差不差），每一个有效信息的传递都靠隐藏状态h的双向传递实现，h是动态的。

M-Dec：

解码的时候，内部状态量用S表示，st = f(st−1,yt−1,ˆ ht, ˆ vt),即s是关于前一刻s,上一句生成的句子，动态的h,和视觉特征向量V,以更好的指示在生成下一个字符时模型应重点关注图像的哪些部分以及前几行中的哪些内容，f函数的计算方法就不贴在这儿了，比较复杂

得到状态量s之后，并不着急生成下一个字符y,文章又给出了一个新的模型Topic Memory Network，此时K向量就派上了用场，如何从图像中生成K向量呢，不用着急，我现在只知道是用了general-v1.3 model算法，clarifai.com网站上有这个算法，但我进去一看好像要注册，就没再看下去，总之先用这个提取出K之后，将每个kj分成两部分，一个输入内存向量qj用于计算yt上kj的importance, 一个输出内存向量mj包含kj的语义信息, 对于每个带有Cj个字符的关键字kj∈K，我们将kj编码为一个语义向量,然后整个过程和GRU的类似，推出m向量后对其加权得到d向量，再加上s向量就得到了能生成语句的o向量，再从字符库里按向量选出字符组成下一句，整个过程就结束了。

评估结果也在各个评价系统里较为突出，

下一步工作应该是把general-v1.3 model算法搞懂，然后可以找诗歌和图片库了，就这样%