能否设计函数，通过计算词组之间的相似性来获得图像描述任务的好坏而不是由简单的词组匹配。