

# 深度学习课程报告

姓名:XX 学号:XX

## Abstract

在计算机视觉领域。。。

关键词：计算机视觉

## 1 一级标题

标题大小、样式可设置，见.tex 源文件中%-标题设置-部分。

正文：人的大脑拥有复杂的结构，大脑中的视觉系统使得人拥有感知环境和描述环境的能力。

### 1.1 二级标题

正文：对于计算机视觉研究领域而言。。。

#### 1.1.1 三级标题

正文：第一段。。第一步，需要从从语义层面理解图像内容。。一种比较常见的图像字幕生成法是基于检索技术，对于给定的查询图像，基于检索的方法首先提取图像特征，后将所提取的特征在数据库中通过检索方法来查询一个或一组与其特征向量相似的图片，最后利用数据库中相似图片的描述直接合成检索图像的描述（或者直接使用相似度最高的图像描述作为查询图像的描述）。

第二段。。。也就是说，生成的查询图像的描述可以是已经存在的句子，也可以是由检索到的句子组成。

## 2 引用

### 2.1 图片引用

图片引用：如图 1 所示，图中。。。引用对应的 label，注意图片的路径即可。latex 会根据页面的篇幅自动调整图片的位置

### 2.2 表格引用

表格引用：表 1 中。。。

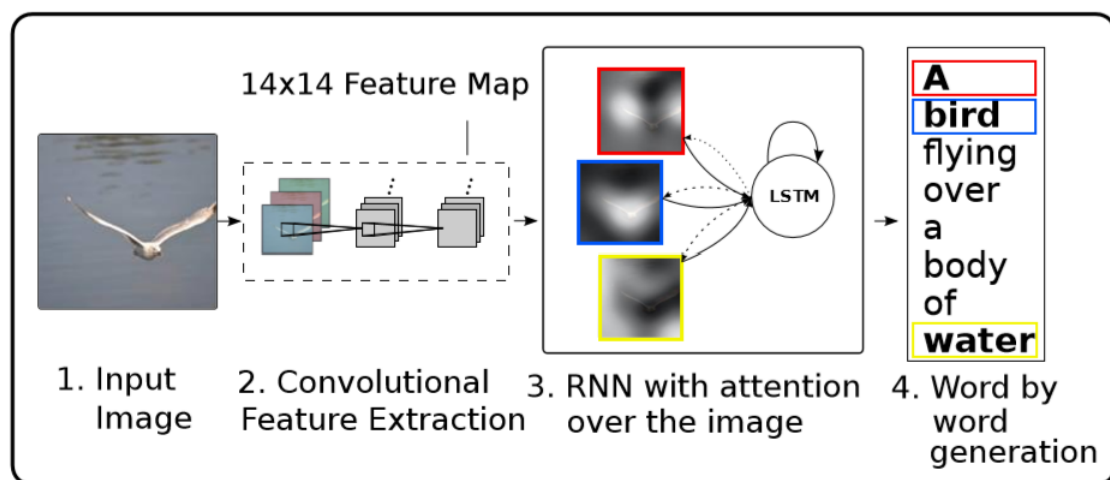


Figure 1: Encoder-decoder 结构

model	BLEU-1	BLUE-4	ROUGE-L	CIDEr
B-up&T-down	77.2	36.2	56.4	113.5
AOAnet	77.4	38.1	58.2	119.8

Table 1: Image caption 其他相关模型

## 2.3 其他引用

章节引用：如 2.3 节所描述。。。

网页引用：百度上。。。

文献引用：正如 Hinton 在文 [1] [2] 中所述。。。注意，文献的引用需要新建 filename.bib 文件，将文献对应的 bitex 格式的引用复制粘贴到文件中然后再引用！谷歌学术或微软学术均可生成文献引用可用的 Bibtex 格式。

## 3 公式

公式等号对齐：

$$f_{att}(c_i, h_i) = \mathbf{v}_a \tanh(\mathbf{W}_a \mathbf{c}_i, \mathbf{W}_b \mathbf{h}_i) \quad (1)$$

$$\alpha_i = f_{att}(\mathbf{c}_i, \mathbf{h}_i) \quad (2)$$

$$\mathbf{s} = \text{softmax}(\boldsymbol{\alpha}) \quad (3)$$

$$\mathbf{z} = \sum s_i \mathbf{c}_i \quad (4)$$

其中  $\mathbf{c}_i$  表示。。。

## References

- [1] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

- [2] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *International conference on machine learning*, pages 2048–2057, 2015.