

文本复制检测报告单(全文对照)

№:BC201905262133172308388482

检测时间:2019-05-26 21:33:17

检测文献: 一种融合注意力机制的多模态可控文本生成

作者: 周会全

检测范围: 中国学术期刊网络出版总库

中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库

中国重要会议论文全文数据库

中国重要报纸全文数据库

中国专利全文数据库

图书资源

优先出版文献库

大学生论文联合比对库

互联网资源(包含贴吧等论坛资源)

英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)

港澳台学术文献库

互联网文档资源

CNKI大成编客-原创作品库

时间范围: 1900-01-01至2019-05-26

检测结果

总文字复制比: 7%

跨语言检测结果: -

去除引用文献复制比: 7%

去除本人已发表文献复制比: 7%

单篇最大文字复制比: 2.2% (基于深度学习的中文命名实体识别研究)

重复字数: [1789]

总段落数: [7]

总字数: [25634]

疑似段落数: [5]

单篇最大重复字数: [563]

前部重合字数: [220]

疑似段落最大重合字数: [552]

后部重合字数: [1569]

疑似段落最小重合字数: [220]



指标: ☐ 疑似剽窃观点 ☒ 疑似剽窃文字表述 ☐ 疑似自我剽窃 ☐ 一稿多投 ☐ 疑似整体剽窃 ☐ 过度引用 ☐ 重复发表

表格: 0 脚注与尾注: 0

0% (0) 中英文摘要等 (总2545字)

9.4% (220) 第1章绪论 (总2346字)

8.9% (267) 第2章 (总2988字)

8.6% (305) 第3章理论模型 (总3550字)

15.5% (552) 第4章实施方案 (总3559字)

4.6% (445) 第5章实验与分析 (总9570字)

0% (0) 第6章总结与展望 (总1076字)

(注释: 无问题部分 文字复制比部分 引用部分)

1. 中英文摘要等

总字数: 2545

相似文献列表 文字复制比: 0%(0) 疑似剽窃观点: (0)

2. 第1章绪论

总字数: 2346

相似文献列表 文字复制比: 9.4%(220) 疑似剽窃观点: (0)

1 041301224_严杰

4.6% (108)

严杰 - 《大学生论文联合比对库》- 2017-05-31

是否引证: 否

2	基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20	4.5% (105) 是否引证：否
3	三维可视化装配工艺发布系统关键技术研究 谷秋实(导师：闫崇京) - 《南京航空航天大学硕士论文》 - 2018-03-01	2.5% (59) 是否引证：否
4	高速接口物理设计与布局算法研究 裴秉玺(导师：郭阳) - 《国防科学技术大学硕士论文》 - 2016-12-01	2.0% (46) 是否引证：否

原文内容		相似内容来源
1	<p>此处有 44 字相似</p> <p>010年，图像文本生成开始活跃在TACL、ACL和EMNLP等知名会议和期刊；在2013年，包括IEEE TPAMI (模式识别与人工智能领域顶级国际期刊)、IJCV (计算机视觉领域顶级国际期刊) 在内的期刊也开始收录相关内容，在2015年，在 CVPR (计算机视觉领域的知名国际会议) 发表了[1-3]等10篇相关论文，同时在IC</p>	<p>041301224_严杰 严杰 - 《大学生论文联合比对库》 - 2017-05-31 (是否引证：否)</p> <p>1.论文在自然语言处理界的知名国际会议和相应的期刊 ACL、TACL 与 EMNLP 中都有发表；而自 2013 年起，模式识别与人工智能领域顶级国际期刊 IEEE TPAMI 以及计算机视觉领域顶级国际期刊 IJCV 也逐渐刊登图像标注工作的研究进展，至 2015 年，更是有近 10 篇相关工作的论文在计算机视觉领域的知名国际会议</p>
2	<p>此处有 65 字相似</p> <p>智能领域顶级国际期刊)、IJCV (计算机视觉领域顶级国际期刊) 在内的期刊也开始收录相关内容，在2015年，在 CVPR (计算机视觉领域的知名国际会议) 发表了[1-3]等10篇相关论文，同时在ICML (机器学习领域知名国际会议) 中也有2篇相关论文发表。</p> <p>在不断的发展中，图像到文本的自动生成任务被认为是人工智能领域中的一项基本挑战[4-6]。而可控文本生成作为近年来研究发展</p>	<p>041301224_严杰 严杰 - 《大学生论文联合比对库》 - 2017-05-31 (是否引证：否)</p> <p>1.觉领域顶级国际期刊 IJCV 也逐渐刊登图像标注工作的研究进展，至 2015 年，更是有近 10 篇相关工作的论文在计算机视觉领域的知名国际会议 CVPR 中发表，同时机器学习领域知名国际会议 ICML 中也有 2 篇相关论文发表。由此可见，图像标注任务越来越受人关注，其已被认为是人工智能领域中的一项重要挑战。毕业设计 (论文) 报告纸2图像</p>
3	<p>此处有 60 字相似</p> <p>用。</p> <p>3) 构建面向多模态的可控文本生成的神经网络模型。</p> <p>4) 在COCO数据上验证其有效性</p> <p>5) 在真实数据集上验证验证其实用性。</p> <p>1.4 论文的组织结构</p> <p>全文分为6章：</p> <p>第一章绪论，主要介绍了课题的来源，选题背景及研究意义，分析说明当前研究现状以及明确该课题的主要任务是研究构建一套面向多模态数据的可控文本生成模型体系框架。</p> <p>第二章相关研究，</p>	<p>三维可视化装配工艺发布系统关键技术研究 谷秋实 - 《南京航空航天大学硕士论文》 - 2018-03-01 (是否引证：否)</p> <p>1.维可视化装配工艺发布原型系统，并以轨道车辆转向架的装配工艺发布为实例，对三维可视化装配工艺的发布的应用流程进行实例验证。1.4 论文的组织结构全文共分为六章，每章具体内容如下：第一章 绪论本章介绍了课题的研究背景和意义，对三维装配工艺数字化设计技术、装配工艺信息的管理与发布技术的相关国内外研究现状进行了分析，最后简要介绍了论文研究内</p> <p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20 (是否引证：否)</p> <p>1.别在传统文本和大数据量互联网文本上实验BLSTM-NER，验证其有效性；5)将实验验证过的模型应用于真实系统，进一步验证其实用性；1.4论文的组织结构全文分为6章：第1章绪论，主要说明了本课题的背景、基本概念和研究意义，介绍了命名实体识别的研究现状以及现阶段的困难。并且对深度学习做了简单的介绍，最后提出了本文的研究内容</p> <p>高速接口物理设计与布局算法研究 裴秉玺 - 《国防科学技术大学硕士论文》 - 2016-12-01 (是否引证：否)</p> <p>1.成了算法主体代码的编写和调试并做了对比。与知名布局器和 EDA 工具的对比结果表明该算法具有有效性</p>

		和实用性。1.3 论文组织结构全文共分为 6 章：第一章是绪论，引入并介绍本文的研究背景和目的以及本文的主要工作。第二章是物理设计的介绍，主要介绍了物理设计各个阶段的工作以及每个阶段所常用的一些方
4	<p>此处有 51 字相似</p> <p>文本生成的实施方案，本章主要介绍本文所用的模型，以及实现的思路 and 过程</p> <p>第五章一种融合注意力机制的多模态可控文本生成的实验与分析，本章详细介绍了本文所做实验的步骤及细节，以及实验结果的深度分析。</p> <p>第六章总结与展望，总结分析成果和不足，并对以后的多模态可控文本生成任务进行建议。</p>	<p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》- 2017-05-20 (是否引证：否)</p> <p>1.4 章中文命名实体识别的实施方案，本章介绍了本文所用的模型，用来处理命名实体的思路及具体过程。第 5 章中文命名实体识别的实验与分析，本章详细介绍了本文所做实验的步骤及详细细节，以及实验结果的深度分析。第 6 章总结与展望，主要总结了本文研究的成果以及存在的不足，并对以后的命名实体识别任务进行建议。第 2 章相关研究 2.1 国际相关研究</p>

指 标		
疑似剽窃文字表述		
<div>1. 验证验证其实用性。</div> <div>1.4 论文的组织结构</div> <div>全文分为6章：</div> <div>第一章绪论，主要介绍了课题的来源，选题背景及研究意义，</div> <div>2. 实验与分析，本章详细介绍了本文所做实验的步骤及详细细节，以及实验结果的深度分析。</div> <div>第六章总结与展望，</div>		
3. 第2章		总字数：2988
相似文献列表 文字复制比：8.9%(267) 疑似剽窃观点：(0)		
1	53130720_邝孝伟_计算机科学与技术_图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》- 2017-05-31	8.8% (264) 是否引证：否
2	2_邝孝伟_图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》- 2017-05-18	8.8% (264) 是否引证：否

原文内容		相似内容来源
1	<p>此处有 60 字相似</p> <p>的模板匹配识别到的结果，这类似于一个填槽过程，当槽位填充完毕，我们根据模板生成结果。Feng与Lapata[9, 10]</p> <p>使用概率图模型同时建模文本信息和图像信息，并从新闻图像所在的文本报告中选择适当的关键词作为体现图像内容的关键词，语言模型</p> <p>将所选内容关键字和必要的功能词汇链接到基本上符合语法的图像标题中。还有一些工作[11-15]能够识别出图像中物体，并找到</p>	<p>2_邝孝伟_图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》- 2017-05-18 (是否引证：否)</p> <p>1.) 具有逻辑结构关系 (即物体的空间布局)，所以自动图像解析的研究是有用的。2.3 基于概率图模型的图像文本标注算法使用概率图模型同时对文本信息和图像信息进行建模，并从新闻图片所在的文本报告中选择合适的关键字作为体现图像内容的关键字，然后使用语言模型来选择所选择的内容关键字以及图像标题的基本语法规则所需的功能词汇连接，最后生成所需要的文本。图像文本标注为的是生成句</p> <p>53130720_邝孝伟_计算机科学与技术_图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》- 2017-05-31 (是否引证：否)</p> <p>1.) 具有逻辑结构关系 (即物体的空间布局)，所以自动图像解析的研究是有用的。2.3 基于概率图模型的图像文本标注算法使用概率图模型同时对文本信息和图像信息进行建模，并从新闻图片所在的文本报告中选择合适的关键字作为体现图像内容的关键字，然后使用语言</p>

		模型来选择所选择的内容关键字以及图像标题的基本语法规则所需的功能词汇连接，最后生成所需要的文本[6]。图像文本标注为的是
2	<p>此处有 39 字相似</p> <p>合语法的图像标题中。还有一些工作[11-15]能够识别出图像中物体，并找到物体获得它的上下位置关系，然后依靠概率图模型和</p> <p>语言模型来选择适当的描述顺序，将这些对象概念和介词短语块连接成一个完整的句子。</p> <p>Hodosh 等人[16]则利用核函数的典型关联分析 (Kernel Canonical Correlation Anal</p>	<p>2 邝孝伟_图像文本标注自动生成系统设计与实现 邝孝伟 -《大学生论文联合比对库》- 2017-05-18 (是否引证：否)</p> <p>1.像中提取对象 (包括对象，动物，花，汽车，表和其他常见对象类型)，并定位它们以获得它们之间的位置关系，然后依靠概率图模型和语言模型来选择适当的描述，利用这些对象的顺序，将介词短语块变成一个完整的句子。还有一类基于内核函数的典型关联分析来查找文本和图像之间的关系，并根据候选句子排序的图像信息，来获得句子的最佳描述。随</p> <p>53130720 邝孝伟_计算机科学与技术_图像文本标注自动生成系统设计与实现 邝孝伟 -《大学生论文联合比对库》- 2017-05-31 (是否引证：否)</p> <p>1.像中提取对象 (包括对象，动物，花，汽车，表和其他常见对象类型)，并定位它们以获得它们之间的位置关系，然后依靠概率图模型和语言模型来选择适当的描述，利用这些对象的顺序，将介词短语块变成一个完整的句子。还有一类基于内核函数的典型关联分析来查找文本和图像之间的关系，并根据候选句子排序的图像信息，来获得句子的最佳描述。随</p>
3	<p>此处有 54 字相似</p> <p>像标注的信息并没有在文本端进行使用。</p> <p>图 2 - 2 多模态 m-RNN 模型</p> <p>在2015年，谷歌通过借鉴文本生成的最新研究成果推进了图像到文本自动生成的进展[22, 23]。Google利用深层卷积神经网络DCNN，将图像特征经过Encoder神经网络层编码后，在经过 Decoder神经网络层 (该神经网络层由LSTM (Long-Short T</p>	<p>2 邝孝伟_图像文本标注自动生成系统设计与实现 邝孝伟 -《大学生论文联合比对库》- 2017-05-18 (是否引证：否)</p> <p>1.化为自然语言句子生成的序列过程，取得了良好的效果。谷歌和蒙特利尔大学以及多伦多大学的研究人员正在利用统计机器翻译领域最新的研究进展促进图像到文本生成的联合建模。前者使用深度卷积神经网络DCNN对图像进行建模和“编码”图像信息，然后直接从另一个长短时记忆递归神经网络 (LSTM) 生成语言文本，无需图像对齐，排序和其他传统</p> <p>53130720 邝孝伟_计算机科学与技术_图像文本标注自动生成系统设计与实现 邝孝伟 -《大学生论文联合比对库》- 2017-05-31 (是否引证：否)</p> <p>1.化为自然语言句子生成的序列过程，取得了良好的效果。谷歌和蒙特利尔大学以及多伦多大学的研究人员正在利用统计机器翻译领域最新的研究进展促进图像到文本生成的联合建模。前者使用深度卷积神经网络DCNN对图像进行建模和“编码”图像信息，然后直接从另一个长短时记忆递归神经网络 (LSTM) 生成语言文本，无需图像对齐，排序和其他传统</p>
4	<p>此处有 74 字相似</p> <p>Decoder网络结构与传统模型相比，不再需要图像-词对齐、调序等繁琐步骤。与此同时，加拿大蒙特利尔大学和多伦多大学通过</p> <p>机器翻译领域的新技术开创性的提出在借鉴计算机视觉领域中的“注意力”(Attention)机制来获取图像块和词语之间的关联，从而使文本段在句子生成时，能够有效的参考图像标注信息，进而模仿人类视觉到句子生成的过程，使生成的句子更符合人的表述习惯。</p>	<p>2 邝孝伟_图像文本标注自动生成系统设计与实现 邝孝伟 -《大学生论文联合比对库》- 2017-05-18 (是否引证：否)</p> <p>1.接从另一个长短时记忆递归神经网络 (LSTM) 生成语言文本，无需图像对齐，排序和其他传统步骤的子步骤。后者在基于神经网络的机器翻译的框架下，提出在计算机视觉领域使用“注意”机制来促进单词和图像块之间的对齐。在句子生成过程中，可以通过单词序列的生成过程来促进传递过程，使得生成的句子更符合人的表达。此外，微软的研究人员使用卷积神经网络CNN和</p>

	图 2 - 3 视觉“	53130720 邝孝伟 计算机科学与技术 图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》 - 2017-05-31 (是否引证 : 否) 1. 接从另一个长短时记忆递归神经网络 (LSTM) 生成语言文本, 无需图像对齐, 排序和其他传统步骤的子步骤。后者在基于神经网络的机器翻译的框架下, 提出在计算机视觉领域使用“注意”机制来促进单词和图像块之间的对齐。在句子生成过程中, 可以通过单词序列的生成过程来促进传递过程, 使得生成的句子更符合人的表达。此外, 微软的研究人员使用卷积神经网络CNN和
5	此处有 40 字相似 微软研究院[23]利用CNN卷积神经网络和多示例学习 (Multiple Instance Learning , MIL) 对图像建模, 并使用判别语言模型生成候选句子, 使用机器翻译模型中的经典最小错误率训练 (Minimum Error Rate Training , MERT) 来发掘文本和图像级别的特征以及对候选句子进行排序。	2 邝孝伟 图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》 - 2017-05-18 (是否引证 : 否) 1. 程来促进传递过程, 使得生成的句子更符合人的表达。此外, 微软的研究人员使用卷积神经网络CNN和多重实例学习(MIL)对图像进行建模, 并使用判别语言模型生成候选句子, 并使用经典最小误差率训练 (MERT) 中的统计机器翻译研究来探索候选句子的文本和图像级特征的排序。由此, 从图像到文本的生成技术需要整合模式识别 53130720 邝孝伟 计算机科学与技术 图像文本标注自动生成系统设计与实现 邝孝伟 - 《大学生论文联合比对库》 - 2017-05-31 (是否引证 : 否) 1. 程来促进传递过程, 使得生成的句子更符合人的表达。此外, 微软的研究人员使用卷积神经网络CNN和多重实例学习(MIL)对图像进行建模, 并使用判别语言模型生成候选句子, 并使用经典最小误差率训练 (MERT) 中的统计机器翻译研究来探索候选句子的文本和图像级特征的排序。由此, 从图像到文本的生成技术需要整合模式识别

4. 第3章理论模型

总字数 : 3550

相似文献列表 文字复制比 : 8.6%(305) 疑似剽窃观点 : (0)

1	基于机器学习的扫视路径估计方法研究 高帅(导师 : 齐飞) - 《西安电子科技大学硕士论文》 - 2018-06-01	3.8% (134) 是否引证 : 否
2	基于神经网络的软件缺陷严重性预测和修复分派模型 周影(导师 : 彭黎) - 《湖南大学硕士论文》 - 2018-05-10	2.4% (84) 是否引证 : 否
3	文新禹-201400301192-电商产品的流行度预测 文新禹 - 《大学生论文联合比对库》 - 2018-05-23	2.2% (77) 是否引证 : 否
4	基于深度学习的RNA打分函数的研究 王帅(导师 : 吕强) - 《苏州大学硕士论文》 - 2017-05-01	1.6% (57) 是否引证 : 否
5	医学文献中疾病与病症关系抽取研究与应用 李梦叶(导师 : 杨志豪) - 《大连理工大学硕士论文》 - 2018-04-28	1.3% (47) 是否引证 : 否
6	基于深度学习的语音识别及其交互应用研究 汪优升(导师 : 王镇道;龚岳平) - 《湖南大学硕士论文》 - 2017-05-20	0.8% (29) 是否引证 : 否

原文内容		相似内容来源
1	此处有 33 字相似 编码器被用来卷积图片特征, 模型使用DCNN[26]来提取L个D维向量, 其中每个向量都对应图像的一个块 : , CNN与DCNN网络不同, 我们通过多层网络卷积得到的特征会比单层得到的特征更准确, 而且因为多层的缘故, 我们可以在多层网络卷积过程中	基于深度学习的语音识别及其交互应用研究 汪优升 - 《湖南大学硕士论文》 - 2017-05-20 (是否引证 : 否) 1. 输入原始语音的MFCC特征, 将特征输入到6层的扩张因果卷积网络中, 经过网络处理结束后, 将隐藏层信息输入到卷积网络中, 通过Softmax网络层生成特征, 将生成的特征通过CTC层处理, 且CTC使用字母级别的标签, 通过训练, 使得网络的误差值收敛。模型的主要代码如下 : #残差模块

	，让Decoder得到正确的图像特征。这好比自然语言处理中的Seq2Se	
2	<p>此处有 80 字相似</p> <p>LSTM (Long Short Term) 网络作为RNN网络的改进，其主要优点在于它有效的解决了普通RNN网络无法解决长距离依赖这一问题。LSTM由Hochreiter & Schmidhuber (1997) 提出，近几年由Alex Graves改良了新版本，这使得LSTM网络在包括文本生成，文本分类等任务获得了广泛的应用。</p> <p>LSTM由于其独特的神经结构，使得它在解决长距离依赖问题上并不需要耗</p>	<p>文新禹-201400301192-电商产品的流行度预测 文新禹 - 《大学生论文联合比对库》 - 2018-05-23 (是否引证：否)</p> <p>1.LSTM) 是一种特殊的时间循环神经网络 (Recurrent Neural Network , RNN) ，它克服了RNN的长期依赖问题。1997年Hochreiter & Schmidhuber提出LSTM，并在近期由Alex Graves进行改良和推广。LSTM在很多实际问题的解决上都取得了良好的成果，并被广泛应用。</p> <p>3.3.1 长短期记忆人工神经网络 (LSTM) 时间循环神经</p> <p>基于深度学习的RNA打分函数的研究 王帅 - 《苏州大学硕士论文》 - 2017-05-01 (是否引证：否)</p> <p>1.经网络(Long Short-term Memory)LSTM是一种特殊的RNN类型，它最重要的能力是可以学习长期依赖的信息。LSTM由Hochreiter & Schmidhuber于1997提出，并在Alex Graves的改良和推广下[37]，在很多问题上，LSTM都取得相当巨大的成功，并得到了广泛的使用。LSTM是通过</p> <p>医学文献中疾病与病症关系抽取研究与应用 李梦叶 - 《大连理工大学硕士论文》 - 2018-04-28 (是否引证：否)</p> <p>1.twork, LSTM)，一般称之为 LSTM，是一种改进的循环神经网络，可以学习到之前的神经网络模型所不能学到的长距离依赖信息。LSTM 由 Hochreiter & Schmidhuber (1997)提出，并被 Alex Graves 进行了改良和推广[32]。它成功地改善了 RNN 对于长距离依赖信息无能为力的缺点，</p>
3	<p>此处有 116 字相似</p> <p>能够预测当前时间步的输出。</p> <p>首先我们需要知道，Attention机制通常作用于Decoder端，用以解码不是时间步长下不同的图像区域，进而可以生成更合理的词。那么，在Attention中就有两个比较关键的量，一个是和时刻t相关，对应于解码时刻；另一个是输入序列的区域，对应图像的一个区域。</p> <p>实现这种机制的方式就是在时刻t，为输入序列的各个区域i计算出一个权重。因为需要满足输入序列的各个区域的权重的加和是唯一的，我们使用Softmax来实现这一点。至于Softmax需</p>	<p>基于神经网络的软件缺陷严重性预测和修复分派模型 周影 - 《湖南大学硕士论文》 - 2018-05-10 (是否引证：否)</p> <p>1.ntext vector，是捕捉了特定区域信息的上下文向量。Attention 要实现的是在解码的不同时刻可以关注不同的文本区域，进而预测更合理的词。那么，在 Attention 中有两个关键量，一个与时刻t 相关，对应于解码时刻；另一个是输入序列的区域ia，对应于文本的一个区域i。即在时刻 t，为输入序列的各个区域 i 计算出一个权重ti?。??????</p> <p>基于机器学习的扫视路径估计方法研究 高帅 - 《西安电子科技大学硕士论文》 - 2018-06-01 (是否引证：否)</p> <p>1.解码时，在不同时刻，网络需要关注到图像不同的区域，进而可以生成更合理的坐标。其中有两个重要的参数，一个是时间参数，对应于解码时刻t；另一个是输入序列的区域????4)，对应图像的一个区域。具体的算法实现如下：在时刻t，对每个输入序列包含的区域 i 计算一个权重tia，同时保证各个区域的权重和为 1，使图像的各个区域都对生成注视点起到贡献，使用 Softmax 函</p>
4	<p>此处有 76 字相似</p> <p>种机制的方式就是在时刻t，为输入序列的各个区域i计算出一个权重。因为需要满足输入序列的各个区域的权重的加和是唯一的，我们使用Softmax来实现这一点。至于Softmax需要输入的信</p>	<p>基于机器学习的扫视路径估计方法研究 高帅 - 《西安电子科技大学硕士论文》 - 2018-06-01 (是否引证：否)</p> <p>1.序列包含的区域 i 计算一个权重tia，同时保证各个区域的权重和为 1，使图像的各个区域都对生成注视点起到贡献，使用 Softmax 函数来实现这一点；至于</p>

<p>息，则如上所讲，需要包含两个方面：一个是被计算的区域，另一个就是上一时刻t-1的信息。</p> <p>图 3 - 4 Attention示例图</p> <p>3.2 本文理论模型论述</p> <p>通过对Encoder-Decoder网络结构的分</p>	<p>Softmax 需要输入的信息，需要包含两个方面：一个是被计算的区域i?，另一个就是上一时刻 t-1 的信息 t1h? : ()1,ti att i te f a h? = (3-11) () ()1exp</p>
--	---

指 标
疑似剽窃文字表述
<p>1. 不同的图像区域，进而可以生成更合理的词。那么，在Attention中就有两个比较关键的量，一个是和时刻t相关，对应于解码时刻；另一个是输入序列的区域，对应图像的一个区域。</p> <p>实现这种机制的方式就是在时刻t，为输入序列的各个区域i计算</p> <p>2. 使用Softmax来实现这一点。至于Softmax需要输入的信息，则如上所讲，需要包含两个方面：一个是被计算的区域，另一个就是上一时刻t-1的信息。</p>

5. 第4章实施方案	总字数：3559
相似文献列表 文字复制比：15.5%(552) 疑似剽窃观点：(0)	
1 基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-29	8.1% (287) 是否引证：否
2 基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20	7.9% (280) 是否引证：否
3 基于深度学习的图像流自动摘要生成 李嘉成 - 《大学生论文联合比对库》 - 2018-06-07	4.1% (145) 是否引证：否
4 详解循环神经网络(Recurrent Neural Network) - aliceyangxi1987的博客 - CSDN博客 - 《网络 (http://blog.csdn.net) 》 - 2017	1.0% (35) 是否引证：否

原文内容	相似内容来源
<p>此处有 54 字相似</p> <p>试阶段使用DCNN卷积图片特征，送入训练好的GAN网络，指定生成的标签类型，生成指定主题</p> <p>上述过程是图像受控文本生成的</p> <p>基本过程，接下来就是对模型识别准确率以及BLEU的分析判断，然后利用对比实验来验证模型的可用性程度。</p> <p>4.</p> <p>2 WordEmbedder</p> <p>首先我们要做的是将词转化为词向量，常见的作法是在神经网络中加入embedding层[28</p>	<p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20 (是否引证：否)</p> <p>1.2vec进行特征抽取。5、 进行LSTM模型的训练及测试。6、 抽取测试结果并输出。上述过程是命名实体识别的基本过程，接下来就是对模型识别准确率的分析判断，然后利用对比实验来验证模型的可用性程度。</p> <p>4.2词向量表示我们要使用LSTM对命名实体进行预测，对于一个整句的输入，每输入一个词，LSTM网络就输出到目前为止的下</p> <p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-29 (是否引证：否)</p> <p>1.2vec进行特征抽取。5、 进行LSTM模型的训练及测试。6、 抽取测试结果并输出。上述过程是命名实体识别的基本过程，接下来就是对模型识别准确率的分析判断，然后利用对比实验来验证模型的可用性程度。</p> <p>4.2词向量表示我们要使用LSTM对命名实体进行预测，对于一个整句的输入，每输入一个词，LSTM网络就输出到目前为止的下</p>
<p>此处有 75 字相似</p> <p>..93199320... 9359</p>	<p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20 (是否引证：否)</p> <p>1. 对于训练集建立一个字典或词典，每个字或词在字典或词典里具有唯一的编号。2) 字典或词典中的任意一</p>

	<p>(2) One-hot编码</p> <p>词典中的任何一个词，都有其对应的二进制编码，我们称为</p> <p>“One-hot”编码，这种方式可以表示成一个N 维的向量</p> <p>字典或词典中的任意一个字或词，都可以用One-hot方式表示成一个N 维的向量。</p> <p>表</p> <p>2 One-hot编码</p> <p>单词单词向量</p> <p>...foothilsCosmetics...briskshark... .</p>	<p>个字或词，都可以用“One-hot”方式表示成一个N维的向量。表 4-1 词典字或词编号.....我 1981..... 很 1982..... 吃 19</p> <p>2.词表示为向量的形式：1) 对于训练集建立一个字典或词典，每个字或词在字典或词典里具有唯一的编号。2) 字典或词典中的任意一个字或词，都可以用“One-hot”方式表示成一个N维的向量。表 4-1 词典字或词编号.....我 1981..... 很 1982.....</p> <p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》- 2017-05-29 (是否引证：否)</p> <p>1. 对于训练集建立一个字典或词典，每个字或词在字典或词典里具有唯一的编号。2) 字典或词典中的任意一个字或词，都可以用“One-hot”方式表示成一个N维的向量。表 4-1 词典字或词编号总长度我很.....吃苹果..... 19811982.....19841985..... 467</p> <p>2.词表示为向量的形式：1) 对于训练集建立一个字典或词典，每个字或词在字典或词典里具有唯一的编号。2) 字典或词典中的任意一个字或词，都可以用“One-hot”方式表示成一个N维的向量。表 4-1 词典字或词编号总长度我很.....吃苹果..... 19811982.....19841985.....</p>
3	<p>此处有 77 字相似</p> <p>hilsCosmetics...briskshark... ...92289229...93199320... .. [0,0,...,1,...,0,0][0,1,...,0,...,0,0]...[0,0,...,0,...,1,0][0,0,...,0,...,0,0,...,0,0,1]... 1]...</p> <p>(3) 词嵌入</p> <p>RNN的输入是文本序列，因此我们需要找到一种简单的方式来表达它。因为计算机更擅长数值处理，</p>	<p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》- 2017-05-20 (是否引证：否)</p> <p>1.字或词向量.....我 [0,0,0,...,1,...,0,0,0]..... 很 [0,0,0,...,1,...,0,0,0]..... 吃 [0,0,0,...,1,...,0,0,0]苹果 [0,0,0,,1,...,0,...,0,0]..... 用“One-hot”方式表示的向量，具有高维、稀疏的特点，所以在输入之前我们要使用</p>
4	<p>此处有 164 字相似</p> <p>示效率高，因此我们使用矩阵运算把高维、稀疏的向量表示成低维、稠密的向量。</p> <p>表 3 WordEmbedder表示方式</p> <p>One-hot 词嵌入</p> <p>...[0,0,...,1,...,0,0][0,1,...,0,...,0,0]...[0,0,...,0,...,1,0][0,0,...,0,...,0,1]... [-4.42496777e-01 2.49782610e...][1.42921007 0.4299607824281...]...[-3.424963293795 2.43842345...][2.415727685 0.3627612462</p>	<p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》- 2017-05-29 (是否引证：否)</p> <p>1.高维、稀疏的向量表示成低维、稠密的向量。例如对于表4-2中One-Hot向量通过词嵌入得到表：表 4-3 词嵌入One-Hot向量词嵌入后 [0,0,0,...,1,...,0,0,0][0,0,0,...,1,0,0,0][0,0,0,,1,...,0,...,0] [-4.42496777e-01 7.78621614e-01 2.49782610e+00 ...][1.429210070.42996076 -0.82428557 ...] [-4.51417826e-027.34614357e-02 3.04548368e-02 ...]4.3 Bi-</p> <p>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》- 2017-05-20 (是否引证：否)</p> <p>1.字或词向量.....我 [0,0,0,...,1,...,0,0,0]..... 很 [0,0,0,...,1,...,0,0,0]..... 吃 [0,0,0,...,1,...,0,0,0]苹果 [0,0,0,,1,...,0,...,0,0]..... 用“One-hot”方式表示的向量，具有高维、稀疏的特点，所以在输入之前我们要使</p>

		用
5	<p>此处有 45 字相似 761246283...]....</p> <p>4.3 CNN卷积网络</p> <p>图 4 - 2 CNN网络结构</p> <p>在图像特征提取，我们可以利用CNN的空间特性，使用卷积层的Feature Map，给图片的不同位置都提取一个特征。</p> <p>举例来说，用表示我们提取的图片特征，一共L个位置，每个位置的特征为一个D维的向量，对于一个高、宽为14，通道数为256的</p>	<p>基于深度学习的图像流自动摘要生成 李嘉成 - 《大学生论文联合比对库》 - 2018-06-07 (是否引证：否)</p> <p>1.上下文向量，该向量是当前时刻图像的显著区域的特征表达。CNN提取的是固定长度的向量特征，而Attention的思路是使用卷积层的feature map，给图片的不同位置都提取一个特征。例如，将图片划分为L×L的位置，每个位置都是用卷积网络获得维度为D的特征向量，记RNN前一阶段的隐层状态，则上下文向量为</p>
6	<p>此处有 100 字相似</p> <p>下文向量为，RNN前一阶段的隐层状态为。这个上下文向量就是的一个加权平均，具体地，和的关系用式子表达：如公式（1），就是</p> <p>衡量生成第t个单词时，第i个位置的图像特征所占的权重。这个权重实际是前一个隐层状态和第i个位置图像特征的函数。具体的表达式为：，公式（2）由于只和已有的信息，有关，因此这些参数也是可以从数据中进行</p> <p>end-to-ends的自动学习。</p> <p>（1）</p> <p>（2）</p> <p>4.4 UnidirectionalRNNEncoder</p> <p>单向</p>	<p>基于深度学习的图像流自动摘要生成 李嘉成 - 《大学生论文联合比对库》 - 2018-06-07 (是否引证：否)</p> <p>1.卷积网络获得维度为D的特征向量，记RNN前一阶段的隐层状态，则上下文向量为特征向量的加权平均和：(2.6)其中是衡量生成第t个单词时，第i个位置的图像特征所占的权重。这个权重实际是前一个隐层状态和第i个位置图像特征的函数。具体的表达式为：(2.7)(2.8)由于只和已有的，信息有关，因此这些参数也是可以从数据中进行端到端的自动学习的。通过向LSTM输入上下文信息，可以使得模型在生成摘要时去关注应当关注的部分，而不是均匀的使用整个</p>
7	<p>此处有 37 字相似</p> <p>作为输入。</p> <p>图 4 - 3 单向RNN展开图</p> <p>循环神经网络的计算公式如下：</p> <p>（1）</p> <p>（2）</p> <p>公式(1)用以计算</p> <p>输出层，输出层作为一个全连接层，它的每个节点和隐藏层的每个节点都是相连的。</p> <p>V是一个矩阵，被用来计算输出层的权重，g是激活函数（常用的激活函数用RELU、Sinmod等）。公式(2)用以</p>	<p>详解循环神经网络(Recurrent Neural Network) - aliceyangxi1987的博客 - CSDN博客 - 《网络 (http://blog.csdn.net) 》 - (是否引证：否)</p> <p>1.e also known as “simple recurrent networks” (SRN). 其中各变量含义：输出层是一个全连接层，它的每个节点都和隐藏层的每个节点相连，隐藏层是循环层。图来自wiki：https://en.wikipedia.org/wiki/Recurrent_ne</p>

	计算隐藏层	
--	-------	--

指 标
疑似剽窃文字表述
<div>1. 基本过程，接下来就是对模型识别准确率以及BLEU的分析判断，然后利用对比实验来验证模型的可用性程度。</div>

6. 第5章实验与分析	总字数：9570
相似文献列表 文字复制比：4.6%(445) 疑似剽窃观点：(0)	
1 AI Challenger 2018完整获奖名单公布，中科院成最大赢家 雷锋网 - 《网络 (https://www.leiphone) 》 - 2018	2.2% (209) 是否引证：否
2 基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20	1.9% (178) 是否引证：否
3 基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-29	1.9% (178) 是否引证：否
4 5306676_蒋焮夷_基于LSTM-RNN神经网络的聊天机器人 蒋焮夷 - 《大学生论文联合比对库》 - 2018-05-16	0.6% (59) 是否引证：否
5 低传LED照明产品无线电传导骚扰测量方法的研究 程江河;- 《厦门科技》 - 2014-04-25	0.3% (31) 是否引证：否

	原文内容	相似内容来源
1	<div>此处有 178 字相似</div> <div>第5章</div> <div>实验与分析</div> <p>本章主要介绍实验的过程、评价指标，以及在不同参数下的多次对比实验结果分析。为了达到对图像受控文本生成模型效果的客观评价，在实验中，我们在相同的实验环境条件下，用两种模型对测试语料分别进行测试和对比分析。通过对测试结果的分析比对，我们初步得出的结论表明基于GAN网络的深层神经网络更适合于可控文本生成。</p> <div>5.1 实验总体设计</div> <div>本文数据采用</div> <div>COCO2014图像数据集作为图像特征提取的训练语料，使用yelp情感数据集作为可控文本生成的训练语料，使用COCO20</div>	<div>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-20 (是否引证：否)</div> <div>1.析。第5章中文命名实体识别实验与分析为了检测模型的有效性，我们基于上述理论模型针对中文命名实体识别进行了一系列的实验。本章主要介绍实验的过程、评价指标，以及在不同参数下的多次对比实验结果分析。为了达到对实体识别模型效果的客观衡量，在实验中，我们在相同的环境条件下用其他模型对测试语料进行标注和结果统计。通过对大量的测试结果的分析比对，我们初步得出的结论表明基于Bi-LSTM的深层神经网络更适合于中文命名实体识别。5.1实验总体设计本实验数据采用1998年01月份《人民日报》的语料，取80%作为训练，另外20%作为测试，利用Bi-LSTM深层神经网络分别从字向量和</div>
		<div>基于深度学习的中文命名实体识别研究 李广财 - 《大学生论文联合比对库》 - 2017-05-29 (是否引证：否)</div> <div>1.析。第5章中文命名实体识别实验与分析为了检测模型的有效性，我们基于上述理论模型针对中文命名实体识别进行了一系列的实验。本章主要介绍实验的过程、评价指标，以及在不同参数下的多次对比实验结果分析。为了达到对实体识别模型效果的客观衡量，在实验中，我们在相同的环境条件下用其他模型对测试语料进行标注和结果统计。通过对大量的测试结果的分析比对，我们初步得出的结论表明基于Bi-LSTM的深层神经网络更适合于中文命名实体识别。5.1实验总体设计本实验数据采用1998年01月份《人民日报》的语料，取80%作为训练，另外20%作为测试，利用Bi-LSTM深层神经网络分别从字向量和</div>
		<div>低传LED照明产品无线电传导骚扰测量方法的研究 程江河;- 《厦门科技》 - 2014-04-25 (是否引证：否)</div> <div>1.低压的LED照明产品在实际使用中是连接适配器工作的,因此在测试中应该连接适配器一起测试。针对以上两</div>

		种观点,本文分别按照这两种测试布置方法进行测试,通过对测试结果的比对分析,提出适合于低压LED照明产品的骚扰电压测试方法。电源端子骚扰电压在屏蔽室中进行,人工电源骚扰电压的数值见图4。网
2	<p>此处有 58 字相似</p> <p>中以BLEU指标尤为重要,它决定生成的语句是否流畅通顺,是否符合可控的要求,该评价指标由IBM在2002[31]年提出,常作为机器翻译系统评价指标。其实就是统计生成响应和真实响应中的N-gram词组在整个训练语料中出现次数。公式如下所示</p> <p>:</p> <p>(1)</p> <p>(2)</p> <p>Pn被用来计算多元词组的精度,所谓多元词组(N-gram),就是比较目标语句和人工翻译之间N个</p>	<p>5306676 蒋焱夷_基于LSTM-RNN神经网络的聊天机器人 蒋焱夷 - 《大学生论文联合比对库》 - 2018-05-16 (是否引证:否)</p> <p>1. a Method for Automatic Evaluation of Machine Translation”,常作为机器翻译系统评价指标。其实就是统计生成响应和真实响应中的n-gram词组在整个训练语料中出现次数。公式如图3-13所示:图3-13 BLEU公式图词重叠评价指标基本上都是n-gram方式,去计算生成响应和真是响应之间的重合程度,但</p>
3	<p>此处有 209 字相似</p> <p>网</p> <p>(1) 赛道</p> <p>英中文本机器翻译</p> <p>(2) 任务介绍</p> <p>英中机器文本翻译作为此次比赛的任务之一,目标是评测各个团队机器翻译的能力。本次机器翻译语言方向为英文到中文。测试文本为口语领域数据。参赛队伍需要根据评测方提供的数据训练机器翻译系统,可以自由的选择机器翻译技术。例如,基于规则的翻译技术、统计机器翻译及神经网络机器翻译等。</p> <p>本次竞赛将利用机器翻译的客观考核指标(BLEU、NIST score、TER)进行评分,BLEU得分会作为主要的机器评价指标。组委会将通过客观指标,并结合答辩表现,综合评估参赛者的算法模型。</p> <p>(3)</p> <p>数据说明</p> <p>训练集文件名train.txt,其中每个训练样例包含自左至右4个元素:DocID, SenID, EngSe</p>	<p>AI Challenger 2018完整获奖名单公布,中科院成最大赢家 雷锋网 - 《网络(https://www.leiphone)》 - (是否引证:否)</p> <p>1.技术总监。优秀奖:「new start」团队成员:张礼, Agoda(安可达)公司数据科学家。英中文本机器翻译 本次机器翻译语言方向为英文到中文。测试文本为口语领域数据。参赛队伍需要根据评测方提供的数据训练机器翻译系统,可以自由的选择机器翻译技术。例如,基于规则的翻译技术、统计机器翻译及神经网络机器翻译等。本次竞赛将利用机器翻译的客观考核指标(BLEU、NIST score、TER)进行评分,BLEU得分会作为主要的机器评价指标。组委会将通过客观指标,并结合答辩表现,综合评估参赛者的算法模型。冠军:「蓝色空间」团队成员:唐剑波,金山软件集团算法工程师;李小龙,金山软件集团算法工程师;郭馨泽,金山软件集团算法工</p>

指 标

疑似剽窃文字表述

1. 实验与分析

本章主要介绍实验的过程、评价指标，以及在不同参数下的多次对比实验结果分析。为了达到对图像受控文本生成模型效果的客观评价，在实验中，我们在相同的实验环境条件下，用两种模型对测试语料分别进行测试和对比分析。通过对测试结果的分析比对，我们初步得出的结论表明基于GAN网络的深层神经网络更适合于可控文本生成。

5.1 实验总体设计

本文数据采用

2. 本次机器翻译语言方向为英文到中文。测试文本为口语领域数据。参赛队伍需要根据评测方提供的数据训练机器翻译系统，可以自由的选择机器翻译技术。例如，基于规则的翻译技术、统计机器翻译及神经网络机器翻译等。

7. 第6章总结与展望

总字数：1076

相似文献列表 文字复制比：0%(0) 疑似剽窃观点：(0)

说明：1.总文字复制比：被检测论文总重合字数在总字数中所占的比例

2.去除引用文献复制比：去除系统识别为引用的文献后，计算出来的重合字数在总字数中所占的比例

3.去除本人已发表文献复制比：去除作者本人已发表文献后，计算出来的重合字数在总字数中所占的比例

4.单篇最大文字复制比：被检测文献与所有相似文献比对后，重合字数占总字数的比例最大的那一篇文献的文字复制比

5.指标是由系统根据《学术论文不端行为的界定标准》自动生成的

6.红色文字表示文字复制部分;绿色文字表示引用部分

7.本报告单仅对您所选择比对资源范围内检测结果负责



 amlc@cnki.net

 <http://check.cnki.net/>

 <http://e.weibo.com/u/3194559873/>