**中 原 工 学 院**

**毕业设计（论文）选题审批表**

**学院：计算机学院 专业：软件工程 2019年12月25日**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **课题名称** | **基于关系网络的图片描述生成研究** | | | | | |
| **课题来源** | **社会服务（ ） 科研项目（√）** | | | | | |
| **课题类型** | **设计（ ） 研究（ √ ） 其它（ ）** | | | | | |
| **指导教师** | **刘小明** | | **职称** | **讲师** | **学生人数** | **1人** |
| **研究内容：**  图片描述（Image Caption）任务是结合自然语言处理(Natural Language Processing, NLP)和计算机视觉(Computer Vision, CV)两个研究领域的综合性任务，图像的文本描述是构建一个能够根据输入的一幅图自动生成对应的描述性文字的算法模型，这项任务要求模型可以识别图片中的物体、理解物体间的关系，并用一句自然语言表达出来。图像的文本描述作为融合视觉和语言的多模态研究任务，可应用于基于文本内容的图像检索、网络图像分析等众多场景中，从而受到了广泛关注。  本课题主要研究一种将深度学习(Deep Learning, DL)和关系感知的自适应注意力网络模型相结合，并将关系网络与自注意力机制融合，实现在公共隐含空间上的特征分布的差异最小化后，实现对图像内容描述文本的生成。  具体任务如下：   * 1. **利用深度学习算法模型对图像内容的准确理解；**   基于目标检测（Object Detection）算法例如FasterRCNN、MaskRCNN等，标注出图像中对象候选区域(region-proposals)、识别对象的属性(attribute)、对象之间的关系信息（relationship）。   * 1. **研究一种文本生成算法模型实现对图像内容的文本描述自动生成；**   研究一种基于关系感知的自适应注意力网络模型(Relation-Aware Attentive Network)，在 AutoEncoder（WAE)框架下实现图像的文本生成，并将关系网络与自注意力机制融合，实现在公共隐含空间上的特征分布的差异最小化后，实现对图像内容描述文本的生成，并在MSCOCO和 Visual-Genome等公开数据上验证该模型的有效性。 | | | | | | |
|  | | | | | | |
| **专家评审意见**  专家签字： 年 月 日 | | | | | | |
| **院 长**  **审批意见** | | 签字：  年 月 日 | | | | |

注：由指导教师填写，经教研室或课题组论证，学院组织专家评审并经院长审批后，本表仅留存电子档。