大连理工大学软件学院毕业设计（论文）开题报告

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **姓名** | 王丹丹 | **学号** | 201692354 | **题目类型** | ☑校内工程 □校内研究  □校外工程 □校外研究 | | |
| **班级** | 软1607 | **专业** | 数字媒体技术 | **开题日期** | 2020 年 1 月 1 日 | | |
| **联系电话** | | 18742585730 | | **E-mail** | dandan3029@mail.dlut.edu.cn | | |
| **中文题目** | | 兴趣实例检测模型的研究与实现 | | | | | |
| **英文题目** | | Research and implementation on instance of interest detection model | | | | | |
| **外文翻译题目** | | Panoptic Segmentation | | | | | |
| **外文翻译期刊名称、出版年月** | | IEEE Conference on Computer Vision and Pattern Recognition. Apr 2019 | | | | **是否近五年期刊论文** | ☑是  □否 |
| **校内导师姓名** | | 谢玲 | | **职称** | 工程师 | | |
| **题目来源** | | ☑科研项目 □工程模拟 □实际应用 □自拟课题 □其他 | | | | | |
| **校外导师信息（校外题目填写，校内题目以下三行不必填写）** | | | | | | | |
| **校外导师姓名** | |  | | **职务/职称** |  | | |
| **所在单位** | |  | | | | | |
| **联系电话** | |  | | **Email** |  | | |

**一、选题的依据（**不少于1000字**）**

|  |
| --- |
| **1.1 选题背景与应用价值（**题目来源、理论意义或工程背景、应用价值等**）**  随着互联网带来的大数据量的传播，如何从海量的图像和视频数据中快速地获取人们感兴趣的重要信息，已经成为[计算机视觉](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E6%9C%BA%E8%A7%86%E8%A7%89/2803351" \t "https://baike.baidu.com/item/%E8%A7%86%E8%A7%89%E6%98%BE%E8%91%97%E6%80%A7%E6%A3%80%E6%B5%8B/_blank)领域一个关键的问题。通过在计算机视觉任务中引入视觉注意机制，可以为视觉信息处理任务带来一系列重大的帮助和改善。  正如俗语所说，一图胜千言。在生活中也的确如此，与冗长的信息描述相比，图片表现了它们的优势。然而图片中丰富的信息往往不全是我们关注的部分，图片中不同的实例拥有不同的重要性。如何使计算机理解图片中的内容，就在于如何找到图片中的比较吸引我们注意力的实例，也就是对表达图片内容比较重要的部分。这成为了我们研究的重要内容。  通过找到图片中引起人们兴趣的实例的应用十分广泛，它可以应用于自动驾驶，面部分割，精准农业等方面，还可以应用在自动生成标题，图像问答等应用中。 |
| **1.2 国内外研究现状分析（**通过文献综述，分析国内外相关研究进展、存在的问题、技术方案选择依据等**）**  为了找到图片中引起人们兴趣的部分，我们首先要做的就是进行物体检测。目前最流行的物体检测方法就是生成大量的区域建议，并预测每个区域的类别概率。比物体检测更进一步的研究还有图像分割。目前比较流行的图像分割方法有超像素，语义分割，实例分割和全景分割。  超像素方法产生的分割是一系列像素的集合，这些像素具有类似的颜色、纹理等特征，距离也比较近。超像素最早的定义来自2003年 Xiaofeng Ren等人的一篇论文《Learning a Classification Model for Segmentation》。超像素中比较常用的一种方法是SLIC（simple linear iterative clustering），是Achanta 等人2010年提出的一种思想简单、实现方便的算法，将彩色图像转化为CIELAB颜色空间和XY坐标下的5维特征向量，然后对5维特征向量构造距离度量标准，对图像像素进行局部聚类的过程。SLIC算法能生成紧凑、近似均匀的超像素，在运算速度，物体轮廓保持、超像素形状方面具有较高的综合评价，比较符合人们期望的分割效果。  语义分割把图像中每个像素赋予一个类别标签（比如汽车、建筑、地面、天空等）。这种分割方式存在一些问题，比如如果一个像素被标记为红色，那就代表这个像素所在的位置是一个人，但是如果有两个都是红色的像素，这种方式无法判断它们是属于同一个人还是不同的人。也就是说语义分割只能判断类别，无法区分个体。语义分割的目标是为每个像素赋予一个类别标签，对于这种任务的一种经典方法就是在一系列的卷积层之后使用全连接层。  实例分割的出现打破了语义分割的瓶颈。实例分割不仅可以判断类别，还可以区分不同的个体。实例分割方式有点类似于物体检测，不过物体检测一般输出的是边界框，实例分割输出的是一个掩模（实例的轮廓）。实例分割和上面的语义分割也不同，它不需要对每个像素进行标记，它只需要找到感兴趣物体的边缘轮廓就行。  全景分割是语义分割和实例分割的结合。每个像素都被分为一类，如果一种类别里有多个实例，会用不同的颜色进行区分，我们可以知道哪个像素属于哪个类中的哪个实例。有很多全景分割方法使用单独的网络来分别处理语义分割和实例分割，也有一些方法采用端到端的方法来共享计算。  另外一个相关的方法是显著性检测(Visual saliency detection)。视觉显著性检测指通过智能算法模拟人的视觉特点，提取图像中的显著区域(即人类感兴趣的区域)。 |
| **1.3 参考文献**（不少于8篇参考文献，格式参照论文模板中参考文献要求）  [1] Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk. 2009. Frequency-tuned Salient Region Detection. In IEEE Conference on Computer Vision and Pattern Recognition.  [2] Sergiu Nedevschi Arthur Daniel Costea, Andra Petrovai. 2018. Fusion Scheme for Semantic and Instance-level Segmentation. In International Conference on Intelligent Transportation Systems.  [3] Liangchieh Chen, George Papandreou, Iasonas Kokkinos, Kevin P Murphy, and Alan L Yuille. 2018. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets and Atrous Convolution and and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence 40, 4 (2018), 843–848.  [4] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross B Girshick. 2017. Mask R-CNN. In IEEE Conference on Computer Vision and Pattern Recognition.  [5] Trevor Darrell Jonathan Long, Evan Shelhamer. 2015. Fully Convolutional Networks for Semantic Segmentation. In IEEE Conference on Computer Vision and Pattern Recognition.  [6] Shaoqing Ren Jian Sun Kaiming He, Xiangyu Zhang. 2016. Deep Residual Learning for Image Recognition. In IEEE Conference on Computer Vision and Pattern Recognition.  [7] Alexander Kirillov, Kaiming He, Ross B Girshick, Carsten Rother, and Piotr Dollar. 2018. Panoptic Segmentation. In IEEE Conference on Computer Vision and Pattern Recognition.  [8] Tsungyi Lin, Piotr Dollar, Ross B Girshick, Kaiming He, Bharath Hariharan, and Serge J Belongie. 2017. Feature Pyramid Networks for Object Detection. In IEEE Conference on Computer Vision and Pattern Recognition.  [9] Tsungyi Lin, Michael Maire, Serge J Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollar, and C Lawrence Zitnick. 2015. Microsoft COCO: Common Objects in Context. In European Conference on Computer Vision.  [10] Alex Graves Koray Kavukcuoglu Volodymyr Mnih, Nicolas Heess. 2014. Recurrent Models of Visual Attention. In Annual Conference on Neural Information Processing Systems. |

**二、研究内容和方法（**不少于1000字**）**

|  |
| --- |
| **2.1 研究内容与目标（**研究内容、研究目标、技术指标等**）**  **研究内容：**兴趣实例检测模型的研究与实现  **研究目标：**在现有方法的基础上提出更加有效的方法  **技术指标：**准确率，召回率，F值 |
| **2.2 拟采取的研究方案（**需求分析、理论与技术方法、软硬件开发平台参数、技术路线等**）**  **需求分析：**  现有的图像分割算法不能有效地从图片中找到人们感兴趣的实例。对于图像标题生成，视觉问答等应用来说是一个巨大的瓶颈。我们想要提出一种方法能够有效地从图片中找到人们感兴趣的实例。  现有的数据集没有对感兴趣实例的标注。我们需要在现有数据集的基础上对其进行一些处理使其能够满足我们实验的需求。  **理论与技术方法：**  语义分割方法：语义分割的目标是为每个像素赋予一个类别标签，对于这种任务的一种经典方法就是在一系列的卷积层之后使用全连接层。但这种方法有一个弊端就是只能判断类别，无法区分个体。  实例分割方法：实例分割方式有点类似于物体检测，不过物体检测一般输出的是 bounding box，实例分割输出的是一个mask。实例分割和上面的语义分割也不同，它不需要对每个像素进行标记，它只需要找到感兴趣物体的边缘轮廓就行。这种方法可以区分出单个个体。  全景分割方法：它是语义分割和实例分割的结合。每个像素都被分为一类，如果一种类别里有多个实例，会用不同的颜色进行区分，我们可以知道哪个像素属于哪个类中的哪个实例。  视觉显著性检测：视觉显著性（Visual Attention Mechanism，VA，即视觉注意机制）是指面对一个场景时，人类自动地对感兴趣区域进行处理而选择性地忽略不感兴趣区域，这些人们感兴趣区域被称之为显著性区域。视觉显著性包括从下而上和从上往下两种机制。从下而上也可以认为是数据驱动，即图像本身对人的吸引，从上而下则是在人意识控制下对图像进行注意。计算机视觉领域主要做的是从下而上的视觉显著性，而从上而下的视觉显著性由于对人的大脑结构作用不够了解，无法深刻的揭示作用原理，在计算机视觉领域的研究也相应很少。  **软硬件开发平台参数：**  Ubuntu 16.04  Python 3.5.6  Pytorch 0.4.1  Cuda 9.2.148  Cudnn 7.1.4 |
| **2.3 预期成果与创新性（**成果形式、代码量、创新性**）**  **成果形式：**  拟完成论文写作，论文内容为对现有方法的思路的总结，以及在此之上提出的新的方法。  拟完成仿真程序，实现选定的方法，并进行比较，同时实现提出的新的方法，进行对比研究。  **创新性：**  在现有方法的基础上提出效果更好的方法和模型，根据现有数据集创建一个符合该研究方法的数据集。 |
| **2.4 进度计划（**按照“周”展开工作计划，不少于18周**）**  第一周：撰写开题报告  第二周：了解图像分割与显著性分析的理论  第三周：查找并阅读一些综述，并据此查找相应参考文献  第四周：初步阅读参考文献，将其分类汇总  第五~六周：在上一周基础上选择性地精读一些文献，了解各种图像分割和显著性分析，注意力模型算法  第七~八周：在前两周的基础上对经典方法深入研究技术细节  第九~十周：有针对性地选择方法进行实现  第十一~十二周：发现已有方法中存在的问题，并对其进行改进提高  第十三周：实现提出的改进方法  第十四周：对实现的改进方法进行实验与测试  第十五周：开始撰写论文，同时继续调整方法  第十六~十八周：撰写与修改论文，完善改进方法 |

|  |  |
| --- | --- |
| 指导教师意见：  签字： 年 月 日 | |
| 考核成绩 | **□**通过  **□**不通过 |

备注：1、考核组织：由各个系自行组织安排，成绩由指导教师或考核小组给出。

2、考核结论：考核成绩分为通过与不通过两种。

（1）对于不通过者需指出主要问题，导师帮助其分析原因，提出相应的改进措施，待修改完成后再次进行开题。

（2）未提交开题报告及二次考核不合格者，将延期答辩。

3、各个系将开题考核结果及开题报告统一汇总到教务员处。