

“中信银行杯”第三届中国研究生人工智能创新大赛赛题

发布时间：2021-05-19

来源：中国研究生人工智能创新大赛

阅读次数：23912

一、大赛赛题

(一) **技术创新**：开放题

(二) **应用创意**：开放题

(三) **企业赛题**：华为赛题共4个，均为华为技术有限公司从实际需求出发拟定的与智能技术的题目。属于此种选题方式的作品还可以参与华为专项奖的评奖。

1. 华为赛题一

1.1 题目名称：提升AI模型在工业质检中“通用性”

1.2 题目描述

计算机视觉在智能制造工业检测中发挥着检测识别和定位分析的重要作用，为提高工业检测的准确率和智能化程度做出了巨大的贡献。然在应用过程中会存在如下“通用性”差的问题

(1) 易受光照影响。目前的AOI工业质检设备，都需要加补光装置，而通常在某一个光照条的AI模型，很难“泛化”到其他光照条件下。

(2) 样本数据少。基于工业质检场景对高精度要求，目前AI模型，通常采用有监督学习方式集大量缺陷样本进行训练，当在实际生产中，缺陷产品占比很小，很难收集大量样本。

(3) 无法跨域“迁移”。比如训练了一个PC外观缺陷检测的模型，但是却无法直接用于手机、洗衣机甚至不同型号的PC的外观缺陷检测。

1.3 具体要求

(1) 所提方法，必须具有“通用性”，不可只能解决某一个小问题。

(2) 可在一些公开数据集上实现算法，进行验证，并通过多组对比实验，证明方法的有效性。

(3) 在保证模型“通用性”条件下，尽量保证模型精度不下降。

1.4 华为赛题一咨询专家及联系邮箱：

OCR及工业质检通用性--温雨金 --wenyujin@huawei.com (<mailto:wenyujin@huawei.com>)

2. 华为赛题二

2.1 题目名称：自然场景下的小样本光学字符识别算法

2.2 题目描述

当前光学字符识别算法，主要采用深度学习的方法进行识别，目的是识别出图像中出现的文字。传统的深度学习方法需要基于大量样本进行有监督训练，在训练样本不充足的情况下，难以得到泛模型。这导致在新增样本识别任务中，出现误判或者错判。

在制造、供应仓储等场景下，对于元器件、电路板的字符图片，由于元器件供应商多，器件和字符的样式多种多样，且针对一家供应商无法大量搜集电路板、元器件上的字符图片样例，同样对运输的货车，车牌样式，字体均不一致，且数据难以搜集。需要一种小样本训练的算法，在提供少以内）图片的情况下，即可训练出一个识别模型，识别待测物体的字符。

2.3 具体要求

(1) 基于小样本学习的光学字符识别算法，训练样本数量小于等于20张图片。对采集到的图片进行标注、学习，生成识别模型。可实现图像的自动采集、在线处理和检测及识别。

(2) 模型以一张图片作为数据，以Json格式输出图片中的文字的位置和内容。

2.4 华为赛题二咨询专家及联系邮箱

自然场景下的小样本光学字符识别算法--温雨金--wenyujin@huawei.com (<mailto:wenyujin@huawei.com>)

3. 华为赛题三

3.1 题目名称：视觉-语义模式匹配

3.2 题目描述

当代基于深度神经网络的视觉模型已经取得了极大的成功。但许多实验都表明，常规视觉模型的形式实现各类视觉任务，如检测、分割、识别等等。一旦这些任务中涉及抽象的语义，视觉模型会受到困扰，并且容易发生过拟合。一种可能的方式是通过自然语言描述与视觉信息的模式匹配。视觉模型可以区分不同语义的信息，完成更高鲁棒性的特征提取过程，并且可以利用这一点直接支持各类任务。比如，在仓储、物流等场景中，模型可以直接根据“卸货车辆后方未着工作服的行人”提示潜在货物风险或行人危险；在商业文档识别等场景中，模型可以根据“价格栏下方的数字”的物品价格等等可能。请设计一种视觉-语义模式匹配的方式，使得至少可以通过自然语言输入完成检索，或者对给定的图像完成正确的自然语言描述输出。

3.3 具体要求

(1) 图像检索任务可以是从小图库中检索单张或多张图片，也可以从图片中进一步检测出给定区域，或是两者结合。输入越接近自然语言、结果越是准确、检索的内容越是精细的检索方式得分限定自然语言输入的格式要求，允许设计带有限制性的语言格式（比如可以设计一个JSON列表但JSON的值仍然需要尽可能地包含自然语言）。

(2) 语言输出任务是对选定图片进行语言描述，描述的准确性越高、信息越丰富，得分越高。流畅度不作为主要得分项（所以也可以一定程度上结构化输出），但至少应该可以被人类较方便地理解。

(3) 任务数据限定为自然场景数据和OCR类文档数据。参考文献仅提供参考，可以使用其他数据集，但所使用数据集需要本身为公开数据集或可以被公开展示。

(4) 提供与现有基线方法的比较和差异化创新点说明。

3.4 参考文献与数据集

Modeling Context in Referring Expressions: <https://arxiv.org/pdf/1608.00272> (<https://arxiv.org/pdf/1608.00272>)

Contrastive Language-Image Pre-Training: <https://arxiv.org/abs/2103.00020> (<https://arxiv.org/abs/2103.00020>)

From Recognition to Cognition: Visual Commonsense Reasoning:

<https://arxiv.org/abs/1811.10830> (<https://arxiv.org/abs/1811.10830>)

VQA: Visual Question Answering: <https://arxiv.org/abs/1505.00468> (<https://arxiv.org/abs/1505.00468>)

3.5 华为赛题三咨询专家及联系邮箱

视觉语义题目 -- 李成--licheng81@huawei.com (<mailto:licheng81@huawei.com>)

4. 华为赛题四

4.1 题目名称：任务型多轮对话理解

4.2 业务背景

多轮对话系统研发是让机器具备与人交流的能力的人工智能领域的一项关键和极具挑战性的任务，包含自然语言理解、多轮对话管理和自然语言生成，通俗地讲，就是构建能够听懂人话、了解并生成有意义且相关的回复来帮助人类的人工智能系统。现有的多轮对话系统由于需要考虑对话上下文依赖和用户在多领域的意图的动态变化，其可靠性、可扩展性和领域自适应性存在缺陷。问题的细化就是如何在训练语料不足时，利用预训练语言模型和知识表示对意图、目标做出动态的处理；在多轮对话管理和对话文本生成研究中探索包含用户个性、情感、任务和领域的对话特征；多领域适应的目标表示和对话策略连续学习；在用户目的不明确时如何构造引导式的对话管理等。

任务型对话可广泛应用于华为多种业务场景中，典型场景为CBG客服场景（如产品售前售后咨询、退换货等）、12345 IT服务热线（如HR政策咨询、办公软件使用问题等）以及WeLink（预订机票、酒店和会议室等）。

4.3 题目描述

多轮对话系统研发是让机器具备与人交流的能力，是人工智能领域一项关键和极具挑战性的任务，包含自然语言理解、多轮对话管理。通俗地讲，就是构建能够听懂人话、了解用户意图进而响应。

作为下一代人机交互的核心技术，各大公司纷纷入局并推出了自己的产品或服务，例如Apple Siri, Google Assistant, 百度度秘等。同时，各大银行，运营商，电商等也推出了自己的智能导购系统。

请设计针对多轮对话的“自然语言理解”算法来理解（结构化）用户会话(User Utterance)，这里的输入是从 0_0 到 t_t 时刻的用户及系统历史对话 $\langle U_0, M_0, U_1, M_1, \dots, U_{t-1}, M_{t-1} \rangle$ ，其中 U_k 为第 k 轮用户会话， M_k 为第 k 轮系统会话，输出用户会话涉及的领域、意图及槽位（含槽值）。

4.4 具体要求

- (1) 参赛者可自由设计算法方案来比赛，但不可调用第三方外部接口，也不可使用基于规则方案。
- (2) 参赛者可采用业界常见的评估方法，也可根据自己的理解定义评估方法，有合理性即可。
- (3) 需要考虑数据集中可能出现的指代消解问题。

4.5 参考数据集

任务型对话数据集：

MultiWOZ2.3

<https://github.com/lexmen318/MultiWOZ-coref> (<https://github.com/lexmen318/MultiWOZ-coref>).

4.6 参考文献

NLU

[1] Chen Q, Zhuo Z, Wang W. Bert for joint intent classification and slot filling[J]. arXiv:1902.10909, 2019.

MultiWOZ2.3

[3] Han, Ting, et al. "MultiWOZ 2.3: A multi-domain task-oriented dataset enhancement with annotation corrections and co-reference annotation." arXiv preprint arXiv:2010.05594 (2020).
Benchmark

[2] Takanobu, Ryuichi, et al. "Is your goal-oriented dialog model performing really well? An empirical analysis of system-wise evaluation." arXiv preprint arXiv:2005.07362 (2020).

4.7 华为赛题四咨询专家及联系邮箱

中文任务型多轮对话理解 -- 刘喜明--liuximing1@huawei.com (<mailto:liuximing1@huawei.com>).

*华为赛题专项奖

华为公司为选择华为赛题的赛队设立了华为专项奖。选择华为赛题，在初赛评审中成绩排名前100名的赛队将获得华为专项奖，每队奖金1万元人民币。华为专项奖是重复奖，参赛赛队可同时参加竞赛主办方组织的其它竞赛奖项的评审和获奖。

二、报名及提交作品要求

参赛队伍在以上赛题中任选其一进行作品创作参赛。

6月30日前，在大赛官网 <https://cpipc.acge.org.cn/> (<https://cpipc.acge.org.cn/>) 中报名，单位资格审查后，8月31日前，在大赛官网提交参赛作品。（作品提交规范及模板见附件）

附件1：第三届中国研究生人工智能创新大赛初赛作品提交规范
(<http://cpipc.acge.org.cn/sysFile/downloadFile.do?fileId=23aae2e7e2714ef584ef0d4c459e9>)

附件2：第三届中国研究生人工智能创新大赛项目文档模版
(<http://cpipc.acge.org.cn/sysFile/downloadFile.do?fileId=20e4b5bfd3d145df806f103400ffb>)