

第十二届服创大赛A类命题企业答疑1.0

(截止至1月6日)

【A02】基于非特定类别的图像前景主体分割算法【万兴科技】

1.问题：关于最后算法结果的评定

最后关于算法结果评定是光看那20多张图的MIOU还是会有自己的所有的测试图的MIOU来评定，是否需要加入其他指标？因为光是评价主体边缘细节清晰也太主观了。

答：

那23张图只是样例图而非测试图，出题方会有专用的测试集进行评估（300+图片）评价会包含客观评价(MIOU,PA,MSE等)和主观评价两个部分，包括但不限于上述指标。

2.问题：关于题目要求的半透明物体而测试图里并没有这类图这件事是不是最后会有关于半透明图的分割测试？

答：

最后会进行半透明分割测试，具体针对半透明物体的分割要求可详见视频讲解PPT。

注意：本算法为通用前景抠图，适用于任意场景的主体抠图，不限于23张示例图片的主体。

3.问题：关于网络模型工程化，包含全部依赖的 C++ 例程

最后的网络模型工程化是指c++用推理框架（如openvino、ncnn、tensorrt、tvm）对训练出来的如Pytorch模型进行网络模型部署嘛，能否稍微介绍一下方法。

答：

是的，用上述OpenVINO/MNN/TensorRT/TVM等任何一种推理框架部署，具体可以去OpenVINO等推理框架的github/官网上去搜。

本题中要求统一用C++进行部署，方便横向对比性能。

4.问题：测试图片大小差别大不大

最后需要进行测试的图片大小差距大不大，是不是就是按照常规情况进行设计。

答:

测试图片分辨率不做限制，任意分辨率。

5.问题：关于训练集是小样本还是大量数据

出题方最后想要得到的是一个基于小样本而训练出来的符合这20多张图情况的模型，还是基于大量数据的，general的一个模型

答:

需要一个比较general的模型，也就是泛化性能好鲁棒性强的模型。

6.问题：图片压缩

是否允许对给的测试集图片先进行压缩，之后再行分割，这样时间会更短

答:

可自行选择是否压缩，但最终输出结果分辨率应与原始分辨率相同。

图像可以压缩inference，但是最后的边缘锯齿等问题需要解决，因为最后的评定标准包括了边缘的抠图效果。

7.问题：可以使用python嘛，需不需要进行前端展示

可以使用python嘛，需不需要进行前端展示

答:

可以使用python，开发语言不限，但是部署语言要求C/C++，并且要能进行前端展示。

【A03】数据资产估值定价分析【浪潮卓数】

1.问题：关于算法的要求

我注意到我们算法要求使用RNN，意思是只能使用循环神经网络吗？不可以使用其他的手段包括机器学习（lightgbm, xgboost）或者其他的深度学习手段（lstm），因为京东万象的数据集十分有限，差不多一千左右，其实用深度学习方法效果达到90%基本上不具备什么泛化能力

答:

可以使用其他算法，我们也希望学生用多样的算法来实现，当时审题老师觉得不限定算法比较难，我们才加了限定

2.问题：关于展示的选择

请问咱们展示只能是web端嘛？ 是否可以移动端？

答：

展示也可以用移动端。

【A04】 航空旅客付费选座意愿识别 【东软】

1.问题：竞赛用数据集中emd_label的含义

数据集中出现了emd_label和emd_label2，其中后者是目标列，前者给的说明是emd,我们团队对emd的含义存在疑惑。

答：

业务上有两种规则用来确定是否为付费选座行为。用emd_label1确认的付费选座范围较大，emd_label2被包含在其中。

本次竞赛中采用emd_label2为付费选座业务准则。

2.问题：“用这些特征描述付费选座旅客的画像”是指眼睛鼻子那些的还是抽象的行为模型？（5.项目说明中的【问题说明】中，提到，“用这些特征描述付费选座旅客的画像”。请问这个画像是指眼睛鼻子那些的还是抽象的行为模型？）

答：

“付费选座旅客的画像”在这里特指航空旅客的业务行为特点。用旅客已经发生的业务行为的期间内统计结果，描述旅客在航空业务中的共性和特性。

3.问题：数据集中某特征的用处

数据集中字段名为emd_label2的特征为目标列，但还有一个字段名为emd_label的一列数据，请问这一列数据是什么用呢？

答：

同1.

4.问题：提供的数据有问题

按照赛题视频的介绍，数据项里面6个月的数据是包括3个月的数据，但是为什么有的列y1即第一年有数据，y2即第二年数据为0，y3即第三年有数据呢？如果按照说明，第二年是包括第一年的数据的，不应该出现这样的情况。这样的列包括JL到JN列、JO到JQ列、RE到RI列。

答：

你的指摘是正确的，这部分y2的统计结果为空，系样本数据预处理时的错误。

5.问题：航班日期

数据特征列的 航班日期 指的是截止数据统计结束时的最后一次坐飞机时间还是什么意思？

与 最近一次出行日期 这一特征列有什么区别？

答：

数据列”航班日期“指，旅客搭乘的航班的起飞时间。如果同一个旅客搭乘多次飞机，会有多条乘机记录和相应的航班日期。

“最近一次出行日期”指，同一旅客在统计期间内的最大的航班日期，Max(航班日期，统计期间)。此列数据部分为空，系样本数据预处理时的错误。

【A06】基于AI学习白流量过滤【深信服】

1.问题：关于数据集的问题

请问本题(白流量过滤)是否会提供数据？

答：

企业不提供数据，学生可以在自己的网络出口抓取数据包

【A07】高级计划与排产平台【雪浪云】

1.问题：关于作品项目范围的具体问题

题目是“高级计划与排产平台”，但是任务要求部分提及的却是具体调度方面要求。因此我们的问题是此次任务是仅仅聚焦于APS调度算法部分还是基于APS算法的ERP平台

答：

算法和平台均有考虑 更偏向于算法 因为算法有很多约束 会描述的比较丰富 平台的话属于完成整删改查以及报表和插单的设计 这属于比较开放和需要同学们去考虑的东西 不做具体扩展功能的限制 故提及的较少。

【A09】轨道交通智慧客流分析预测【八维通】

1.请问能否告知所给数据的具体来源呢？在企业方所给的任务清单中提到了增加温度、天气的影响因素，也提到了制作热力图用于展示，然而这些都依赖于具体的城市及地铁结构信息。如果不能给出具体城市的话，我们是无法找到与数据相匹配的天气信息的，“为使模型跟贴近实际，团队可以自己模拟加上相关的影响因子例如：温度、天气等等”这一点故而不可能达成。同样的，如果不能得知具体的线路拓扑结构，仅靠现有数据推算的话，是不能保证得到的拓扑结构的正确性的，线路结构不正确，模型准确性也就无从谈起；而且如果由我们自己根据数据推算线路结构的话，我们只能知道站点间的连接情况，而无法得知真实的地图，那么企业方要求的热力图也是不可能准确地画出来的。这些问题，希望企业方予以考虑。

答：

关于数据来源以及城市的话，涉及到保密不便透露。因此给出的用户ID/站点名称都是修改过的，但是所有行程数据都是实际的。因此，线路拓扑图各个团队可以自己设计，这个不影响热力图的展现。如示例图

居住



2、有老师反映官给的数据有误，如钱不对劲（有非整数的钱），trips里面有几个站在station里面没有（5,155号）

答：

test_trips 表里面有近80W行数据，1. 其中里面有77条金额不是整数的，可以去除，或者自行修改，并不会影响最终计算结果。2. 有32583条行程里面的 in/out_station_name 不在station基础表里面，这个是由于原数据的差异导致的。按照比例看来，异常数据只占全部数据的3.7% 因此不影响结果，可以不予理会或者抛弃。同时作为大数据处理，数据的清洗和容错是考虑的一方面。

【A11】 Web3D 短视频学习教育平台 WSVTP 【文思海辉】

1、问题：教学课程设计内容是否直接和提供接口数据相关？

答：

部分功能需要用到数据接口

2.问题：是否有固定指向 譬如 有关软件设计 java php 等教学内容？还是自己定任意领域教学内容？

答：

有 JAVA做服务。达梦数据库存储数据 数值分析和计算。可以用Python 但最后要转为 JAVA模型

【A12】 基于手势识别的会议控制系统 【长安计算】

1.问题：控制手势和功能应用

1.是否可以自己设计手势实现五种基本的功能。

答：

五种基本动作手势为题目要求必选项，选题者可在这五种基本手势动作的基础上进行扩展。

2.对于五种基本功能是否需要考虑其具体应用场景。例如，抓取功能在日常会商会议中具体会用在什么地方。

答：
需要考虑。

【A13】 聚焦人脸识别的大数据测试系统 【虹软】

1.问题： 是否有baseline?

是否有baseline?

答：
没有

2.问题： 数据集是否有指定?

数据集是否有指定? 用什么?

答：
数据集没有指定的，可以从网上下载，展示时最好能说明一下数据集构成

3.问题： 以服务端形式是在服务器环境上吗?

以服务端形式是在服务器环境上吗?

答：
是的

【A19】 智能化共享租赁平台 【中信银行信用卡中心】

1.问题： 关于淘宝速卖通平台接口接入使用的问题

以下出自速卖通品平台的接入要求的文档：

速卖通自研型开发者要求：

必须为企业，需上传开发者企业营销执照。仅限于具有开发能力的速卖通卖家授权接入仅适合于自己企业的店铺（以店铺企业认证信息为准）申请商家接入必须入驻聚石塔

申请入驻需要在产品说明书中提供自研发系统详细功能模块信息（调用API信息，系统架构，功能截图，自研发软件著作权，自研发源代码）

请问我们参赛者如何处理？

答：

速卖通主营业务为跨境电商，对接相对复杂，题目中列出主要是做为参考，本次选题的主要创新环节不一定非要体现在支付环节，而且参加竞赛也并不需要完全复制企业实现业务全流程，所以建议参赛同学，可以尽量简化支付环节，网上可以找到一些供个人账号对接的支付api，或者采用充值虚拟货币到个人账户进行管理的方式实现模拟扣费，着重体现对整个租赁流程的设计实现

2.问题：关于支付，物流，用户信用的第三方接口问题

命题文档中数据接口部分提供的网址需要企业账户才能获取对应的接口服务，这个企业账户贵公司是否提供？或者能否提供更详细的第三方支付，物流的API接口文档呢？项目考虑增加用户信用评估功能，请问中信银行提不提供查询用户信用信息的第三方接口API？

答：

目前无法对外提供企业账户，如实现困难，建议同学们到网上自行查找相关服务接口，支付宝，以及各个物流公司都能找到详细对接api，如果因为资质问题，无法接入，建议按实际api格式进行模拟实现。由于客户信息比较敏感，有强监管，无法对外提供相应的api。

3.问题：是否提供金额在银行中转的API

我们考虑将收取的定金等金额存入平台绑定的中信银行卡中暂存，请问是否提供一个平台的银行卡，以及存入和取出金额的API？

答：

目前无法提供相应的api，建议同学自行调研相应第三方支付平台的实现。

4.问题：商品要与中信银行有关吗？

您好，这里的商品指的是类似天猫那样的商品，还是指定是中信银行相关的商品？

答：

商品可以是任意具备租赁属性的物品，同学们可以自行设计。