

2022 全球算法精英大赛初赛参赛手册

题目：车道渲染数据智能质检

1. 简介

在地图业务中，导航过程中所看到的背景道路界面是基于地图数据渲染生成。渲染过程中，部分数据会存在不同程度的问题，导致该部分数据渲染生成的道路存在缺陷，诸如缺边少角，异形道路等。为了更高效的检测这部分数据，降低人工成本，质检模型需要达到更高的准确度。根据问题数据所具有的共同模式建模，提高模型精度是降低成本的关键。希望通过本次比赛，挖掘计算机视觉方向人才，推动该领域发展。

2. 赛题说明

本赛题为选手提供地图渲染数据及其部分标注。训练集包括车道渲染数据的图片集及部分标注，选手使用训练数据进行模型训练。测试集分为 A/B 两个测试集。测试集仅提供车道渲染数据图片集，选手使用模型预测测试图片是否存在问题。

3. 数据说明

3.1. 车道渲染数据

本赛题给出的车道渲染数据为图片格式，包含两个大类，即问题图片和无问题图片，问题图片的问题范围包括：中心线问题、停止线问题、引导面问题、路肩问题、路面问题、箭头问题、车道线问题。请注意存在问题的图片可能有多个问题。图 1~7 分别给出了中心线、停止线、引导面、路肩、路面、箭头、车道线问题图片的例子。



图 1. 中心线存在问题的图片（中心线长出路口）



图 2. 停止线问题（停止线位于路口面内）

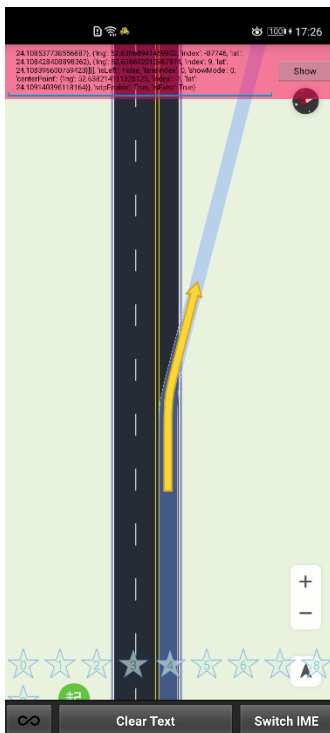


图 3. 引导面问题（引导面超出路面，与底图不匹配）

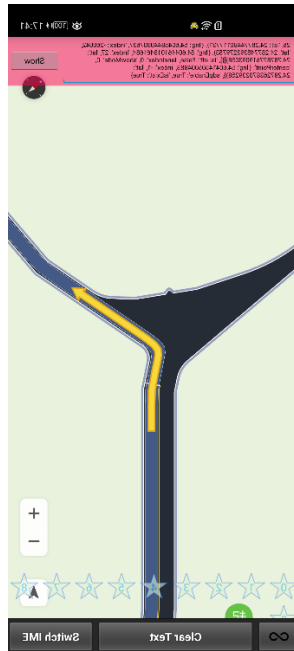


图 4. 路肩问题（路肩凸起，凹凸不平）



图 5. 路面问题（路面断裂）



图 6. 箭头问题（车信箭头布局有压盖）

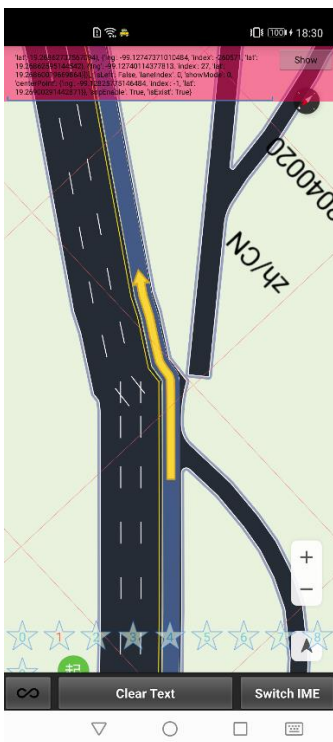


图 7. 车道线问题（车道线交叉重叠压盖）

3.2. 标注数据

标注文件以 csv 文件形式给出。每行包括图片名及该图片是否存在问题标注(若存在问题, 会给到具体的问题编号)。格式如下:

imagename, defect_type

image_0.png, 1

image_1.png, 0

3.3. 数据文件说明

本次比赛给出的数据文件及相关信息如表 1。

表 1. 数据集文件说明

文件名	格式	说明
train_image.rar	图像	训练集图片
train_label.rar	csv	训练集标注
testA_image.rar	图像	测试集 A 图片

4. 评估方式

选手根据测试集图片, 预测图片是否存在问题(无需给出具体问题类别)。比赛使用 AUC(Area Under Curve)作为评价指标, AUC 越高, 代表结果越优, 排名越靠前。

5. 提交方式

选手提交 csv 文件, 编码采用无 BOM 的 UTF-8。格式如下:

imagename, defect_prob

其中 imagename 对应测试图片的图片名, defect_prob 表示测试图片存在问题的概率。imagename, defect_prob 间采用英文逗号分隔。