

障碍物轨迹预测任务

本文档描述了百度阿波罗计划中障碍物轨迹预测任务的数据格式和使用说明。

1 数据集介绍

本次障碍物轨迹预测数据集提供了 20,000 个障碍物运动数据，其中包括 10,000 个训练数据和 10,000 个测试数据，每个数据表示一个障碍物的当前运动特征和真实的行为意图。其中训练数据可用于算法模型的训练；测试数据可用于算法的测试。此外，数据集还提供 100 个可供下载的数据，用于算法的调试、数据分析及数据可视化。

2 数据特点

本数据集采样全部来自真实路面场景，数据采集场景丰富，每个数据均为路面上真实的机动车障碍物。其中，包含了该机动车多帧的历史数据信息，并经过感知层一系列处理后，形成相应的障碍物特征和路面特征。数据的标注信息是根据在后续 1s 观测中障碍物是否在预设车道内进行的自动标注，其结果是形成一个二分类问题： >0 为正例，即 1s 后该障碍物在预设车道内； ≤ 0 为负例，即 1s 后该障碍物不在预设车道内。其中，若标注为 1，则为沿当前车道正样本，若标注为 0，则为沿当前车道负样本，若标注为 2，则为变道正样本，若标注为 -1，则为变道负样本。

3 传感器配置

本次数据采集使用了多种传感器综合得到，其中包括激光雷达（型号为 Velodyne HDL-64E S3）、两个视觉摄像头（其中一个相机采用 25mm 焦距镜头，另一个采用 6mm 焦距镜头）、高精度地图和车辆定位系统，见图 1。具体每种传感器的感知结果可参考 Apollo 其它数据开放平台。

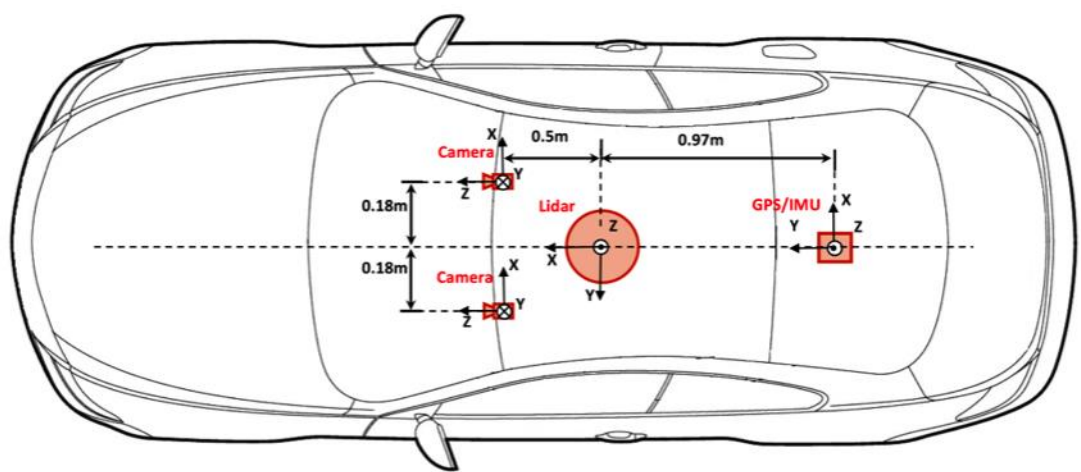


图 1 车辆传感器示意图

4 数据格式

预测模块使用的预测数据和标注数据统一以.h5 文件形式提供，每一组预测数据共有 63 维，前 62 维是相关障碍物和车道信息，最后一维是标注结果（ >0 或 ≤ 0 ）。预测数据具体来说分为两类：障碍物信息和预设车道信息，其中障碍物信息 22 维，车道信息 40 维，具体信息参见表 1。数据中每一个维度的信息均以四字节 float 类型进行存储。本数据集中所采用的坐标系为车道 SL 坐标系，如图 2 所示，每一段车道的中心线为 S 轴，车道垂直方向为 L 轴。

表 1 数据集内容说明

类型	维度	参数说明
障碍物状态信息	1~5	当前时刻 θ 、历史平均 θ 、当前和平均 θ 的差、上时刻与当前 θ 角度差及其导数
	6~10	当前时刻 L、历史平均 L、当前和平均 L 的差、上时刻与当前 L 的差及其导数
	11	车辆观测时间内的平均速度
	12	车辆当前加速度
	13~15	距当前车道左侧线的距离、该距离的平均变化率和当前时刻的变化率
	16~18	距当前车道右侧线的距离、该距离的平均变化率和当前时刻的变化率
	19~22	车道类型，依次为直行、左转、右转、掉头，属于该类型则特征值为 1 否则为 0

预设车道 线信息	23~62	每 4 个特征为一组，共 10 组当前障碍物车辆的预设车道信息 每隔一段距离一个该车道的采样点，每个点有 4 个特征，分别为： 该点与当前车辆位置的角度减去车辆速度方向角度的 \sin 值 该点相对车辆的 L 投影距离 该点的车道朝向 该点车道方向和当前车辆速度方向的夹角
-------------	-------	---

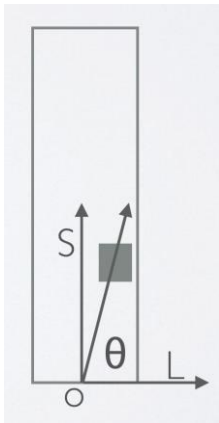


图 2 SL 坐标系

4.1 数据用途

本数据集可以用于障碍物机动车的预测算法的研发和测试。可以构建一个机器学习算法利用提供的训练集进行模型训练并利用测试集进行评价。对于基于规则的算法可以直接利用测试集进行效果测试。

4.2 数据结果格式

用户需要根据提供的障碍物信息和预设道路信息预测出目标 1s 后是否会在预设道路的结果，用二分类结果表示： >0 在预设道路中； ≤ 0 不在预设道路中。

4.3 评测指标

根据用户提供的模型，我们将用测试集对模型的分类准确率（precision）和召回率（recall）进行评价该算法的水平。