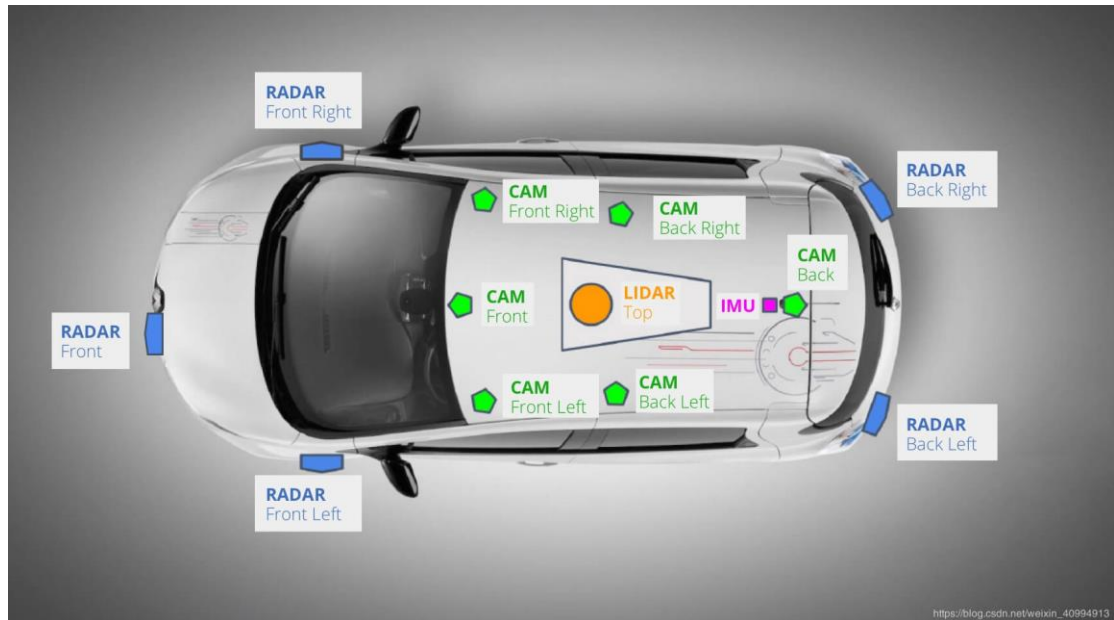


## 概述

nuScenes 是一个类似于 KITTI 的用于无人驾驶的大型公共数据集，但它提供了更多种的传感器数据，更大的数据量，更多样的场景和更多的类别。

**传感器：**它使用了 6 个相机，1 个 32 线 LIDAR，5 个 RADAR，及 1 个 IMU 采集数据。由于 KITTI 中相机覆盖了车辆前方区域，因此标注也只是局限于前方。而 nuScenes 则使用了 6 个相机覆盖了 360 度的范围，且在全部范围都有标注。



**数据量：**KITTI 的 trainval 集共有 7k 多个样本。而 NuScenes 共公布了 1000 个场景，其中 trainval 集包括 850 个场景，每个场景为约 20s 长的视频。为了便于同步，每秒选取 2 帧作为关键帧进行标注。因此 trainval 集共有约 3 万个样本。此外，nusenes 还提供了其他非关键帧的数据。

**场景：**nuScenes 数据集中包含了更多样的场景。包括不同地点（波士顿、新加坡，包括城市场景、居民区、野外和工业区），不同天气（晴天，阴天，雨天，），夜晚场景和左右手交通等。

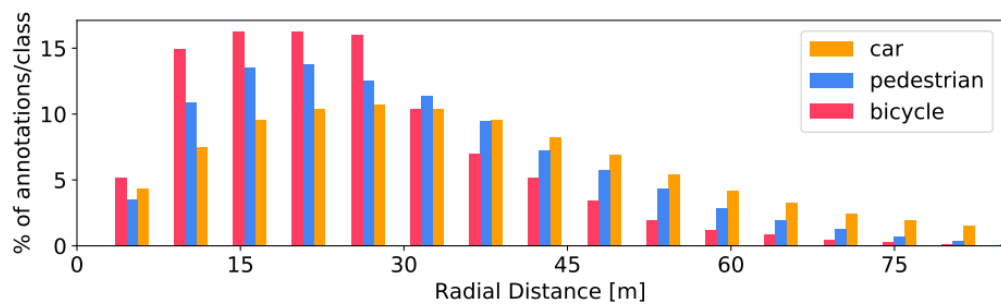
**类别：**nuScenes 提供了 35 类目标，对一些类别做了更细致的分类。如行人目标被分为了成年人，小孩，警察，建筑工人等。目前我们只是将这些类目标统称为行人，以后可做特殊需求的扩展。此外，nuScenes 还对某些类目标有特殊的状态标注，如车辆的移动、停止和停车，行人的移动或伫立等。

animal		787
human.pedestrian.adult		208240
human.pedestrian.child		2066
human.pedestrian.construction_worker	Pedestrian	9161
human.pedestrian.personal_mobility		395
human.pedestrian.police_officer		727
human.pedestrian.stroller		1072
human.pedestrian.wheelchair		503
movable_object.barrier		152087
movable_object.debris		3016
movable_object.pushable_pullable		24605
movable_object.trafficcone		97959
vehicle.bicycle	cyclist	11859
vehicle.bus.bendy		1820
vehicle.bus.rigid		14501
vehicle.car	sedans, hatch-backs, wagons, vans, mini-vans, SUVs and jeeps	493322
vehicle.construction		14671
vehicle.emergency.ambulance	Car	49
vehicle.emergency.police		638
vehicle.motorcycle	cyclist	12617
vehicle.trailer		24860
vehicle.truck		88519
static_object.bicycle_rack *		2713
<b>Total</b>		<b>1166187</b>

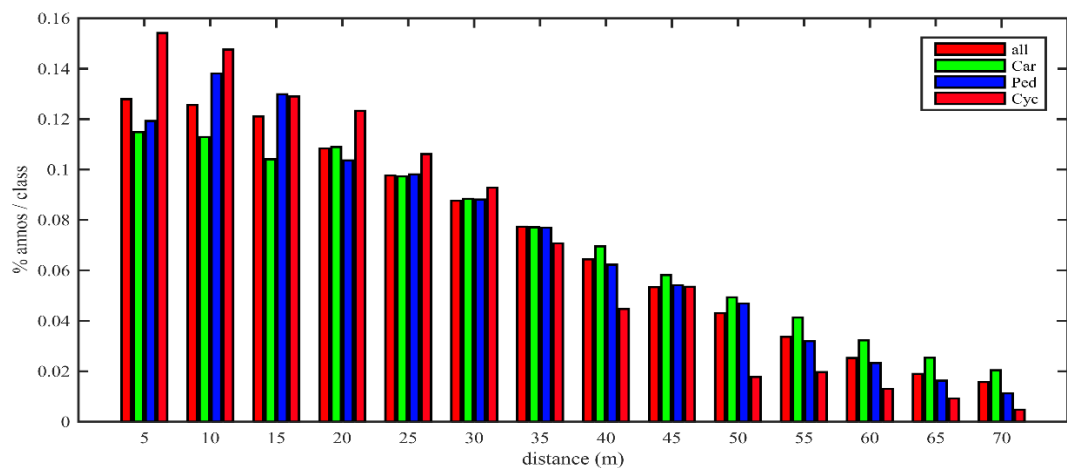
## 样本分布

### 1. 距离分布

#### a) 全部 gt 的径向距离分布

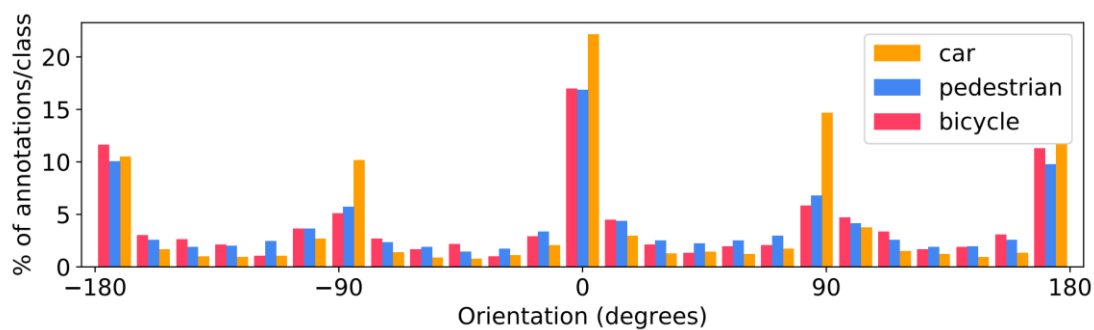


a) 前向[-40,40,0,70]范围内 gt 的距离分布

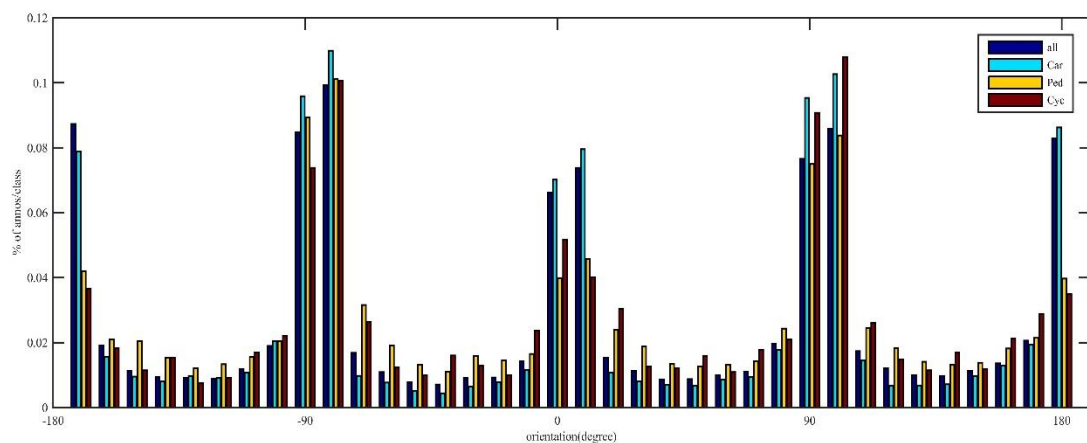


## 2) gt 朝向分布

a) all gt (前向为  $\pi/2$ , 左方为 0)



b) 前向[-40,40,0,70]范围内 gt(前向为 $-\pi/2$ , 左方为 0)



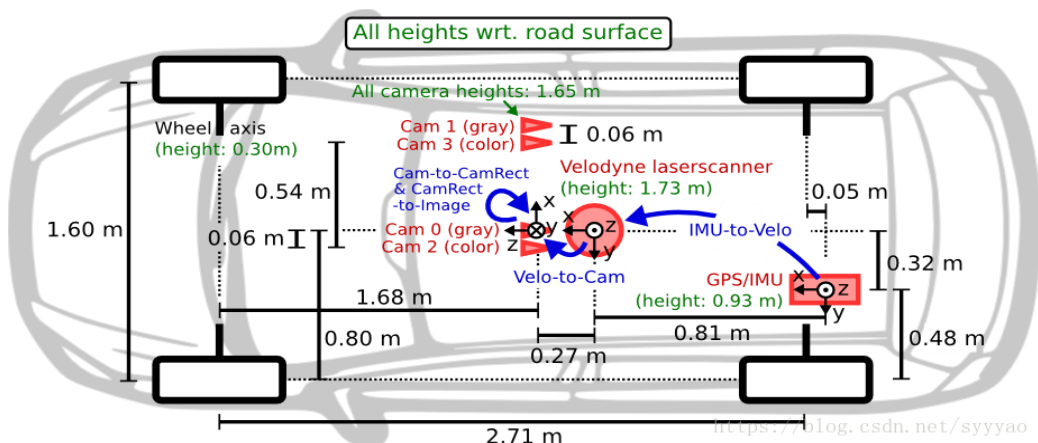
## 生成 infos 文件

### 1. KITTI/Car6 格式的 infos

```
{'image_idx': 0,  
'pointcloud_num_features': 4,  
'velodyne_path':  
'img_path':  
'img_shape':  
'calib/P101':  
'calib/R101':  
'calib/Tr_velo_to_cam':  
'annos': {'name':,  
          'truncated':),  
          'occluded':),  
          'alpha':),  
          'bbox':  
          'dimensions':  
          'location':  
          'rotation_y':  
          'score':  
          'index':  
          'group_ids':  
          'difficulty':  
          'num_points_in_gt':  
        }  
}
```

### 2. 标定

#### 1) KITTI



KITTI 数据集中各坐标系如上图所示。由于在 KITTI 中的标注是定义在相机坐标系下的，因此需要将点云从 lidar 坐标系转换到相机坐标系中。KITTI 的 `calib` 文件中已经为我们提供了以下几个矩阵用于坐标系转换。

`'calib/Tr_velo_to_cam'`: transformation matrix from lidar coordinate to non-rectified camera coordinate

'calib/R101': Rotation from non-rectified camera to rectified camera coordinate

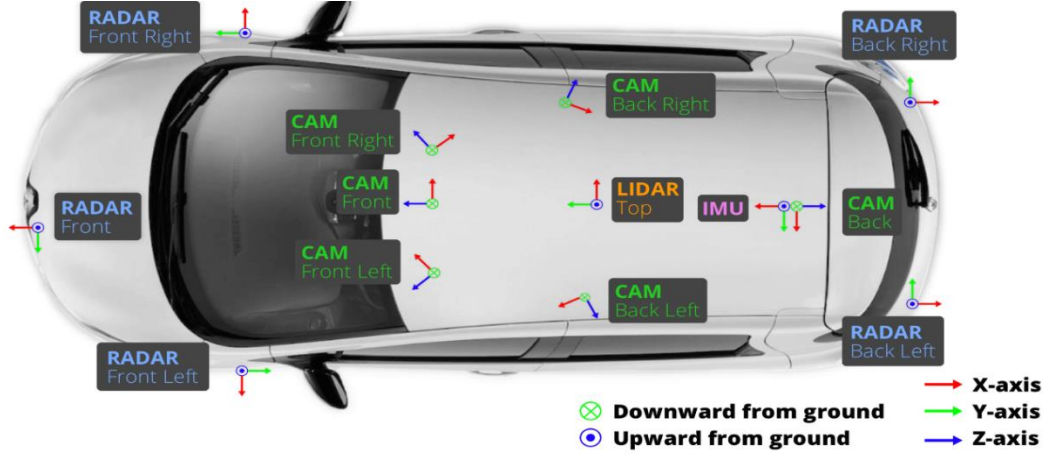
'calib/P101': Projection matrix for camera in rectified coordinate

因此将点 A 从雷达坐标系转换为相机坐标系:

$$A_{cam} = P * R * Tr\_velo\_to\_cam * A_{lidar}$$

## 2) NuScenes

在 nuScenes 中, 各坐标系如下图所示。除此之外还有一个 global 坐标系和一个 ego 坐标系。其中 ego 坐标系原点位于车体后车轴中心。Global 坐标系是对于一个 map 建立的坐标系。



NuScenes 中的标注是定义在 global 坐标系下, 且提供了函数用于将 global 坐标系下的标注转换到任意传感器坐标系下。目前有两个方案:

- 直接使用 lidar 坐标系下的标注, 这样不需要对雷达点做坐标系转换。因此 P、R、Tr 矩阵全部设置为单位矩阵。
- 为了与 KITTI 统一, 同时考虑到以后可能会有 lidar 与相机融合的使用, 还是使用 cam 坐标系下的标注。这样需要计算 P、R、Tr 矩阵:

NuScenes 为我们提供了各个传感器坐标系到 ego 坐标系的平移矩阵 T, 四元数 (w, x, y, z) 以及相机的内参矩阵 cam\_intrinsic。首先将四元数转换为旋转矩阵 R:

$$\begin{bmatrix} 1 - 2y^2 - 2z^2 & 2xy + 2wz & 2xz - 2wy \\ 2xy - 2wz & 1 - 2x^2 - 2z^2 & 2yz + 2wx \\ 2xz + 2wy & 2yz - 2wx & 1 - 2x^2 - 2y^2 \end{bmatrix}$$

将 lidar 坐标系下的点转换到 non-rectified cam 坐标系下需要经过 lidar -> ego -> global -> ego1 -> camera

$$Tr\_velo\_to\_ego: [R_{velo\_ego}^T | T_{velo\_ego}^T]$$

$$Tr\_ego\_to\_global: [R_{ego\_global}^T | T_{ego\_global}^T]$$

$$Tr\_velo\_to\_ego: [(R_{ego\_global}^{-1})^T | (T_{ego\_global}^{-1})^T]$$

$$Tr\_velo\_to\_ego: [(R_{cam\_ego1}^{-1})^T | (T_{cam\_ego1}^{-1})^T]$$

$$Tr\_velo\_to\_cam = Tr\_velo\_to\_ego * Tr\_velo\_to\_ego * Tr\_ego\_to\_global * Tr\_velo\_to\_ego$$

**R** 设为单位矩阵，即不需要 rectify

**P** :

$$\left( \begin{array}{c|c} \text{cam\_intrinsic} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \end{array} \right)$$

### 3. 坐标系旋转

在 KITTI 数据集的雷达坐标系中，将车辆前方设置为 **x** 正方向，左方为 **y** 正方向。

而 NuScenes 的雷达坐标系中，前方为 **y** 正方向，而右方为 **x** 正方向。因此将 nusenes 数据集中的雷达点转换为 kitti 的格式。

### 4. 标注信息

所需的标注格式为: xyz lhw orientation。其中车辆正前方 orientation 为  $-\pi/2$ ，左方为 0。

NuScene 同样是提供了四元数(w,x,y,z)，需要先转换为欧拉角

$$\theta = \arctan 2(2(wz + xy), 1 - 2(y^2 + z^2))$$

但前向为  $\pi/2$ ，左方为 0。

因此  $\theta = -\theta_{nusenes}$

## 生成 reduced point cloud 文件

目前只关注车辆前向的目标，为了节省计算，只保留车辆前方[0,70,-40,40]米范围内的点云。同时需要将此范围外的 gt 也滤除。