

超全的 3D 视觉数据集汇总

hello_uncle · 发表

于 2020-01-04

21:28:27 文章

来源: 人脸识别

数据集

分享

KITTI数据集

KITTI数据集由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办，是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集。该数据集用于评测立体图像 (stereo)，光流(optical flow)，视觉测距(visual odometry)，3D物体检测(object detection)和3D跟踪(tracking)、语义分割等计算机视觉技术在车载环境下的性能。KITTI包含市区、乡村和高速公路等场景采集的真实图像数据，每张图像中最多达15辆车和30个行人，还有各种程度的遮挡与截断。整个数据集由389对立体图像和光流图，39.2 km 视觉测距序列以及超过200k 3D标注物体的图像组成。

数据集链接：



hello_uncle
北邮

45

关注

146

获赞

9

精选文章

近期动态

人脸识别方向

文章专栏

人脸识别

关注

作者文章

更多

- 在注意力中重新思考 Softmax：分解非线性，这个线性 transformer 变体实现多项 SOTA

50
- 从原理和代码详解 FAIR 去年的惊艳之作：全新的纯卷积模型 ConvNeXt

19
- 打破静态融合操作常规！腾讯提出新的 MLP 架构 DynaMixer

17

https://www.cvlibs.net/datasets/kitti/raw_data.php



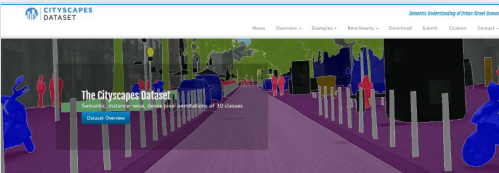
2、Cityscapes

Cityscapes是一个较为新的大规模数据集，它包含50个不同城市的街道场景中记录的各种立体视频序列，除了一组较大的20 000弱注释帧外，还具有5 000帧的高质量像素级注释。因此，数据集比以前的类似尝试要大一个数量级。

Cityscapes数据集旨在评价视觉算法在城市场景语义理解中的性能：像素级、实例级和全景语义标注；支持旨在开发大量（弱）注释数据的研究，例如用于训练深层神经网络包含城市场景下双目图像及像素级语义分割标注。

数据集链接：

<https://www.cityscapes-dataset.com/>



3、牛津数据集

91.3%！首个将 Transformer 解码器应用于多标签图像分类的方法
Query2Label

26

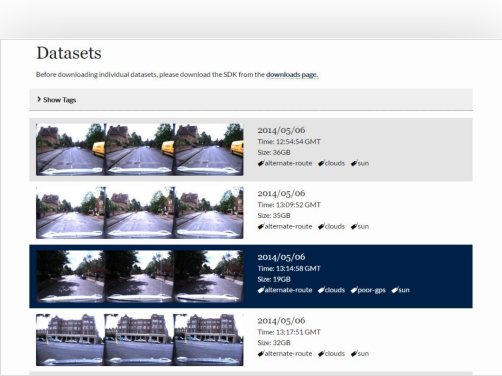
Sampling-Argmax：用重参数技巧优化 Soft-Argmax

145

对牛津的一部分连续的道路进行了上百次数据采集，收集到了多种天气、行人和交通情况下的数据，也有建筑和道路施工时的数据。包含全景图像、激光雷达点云、导航信息。

数据集链接：

<https://robotcar-dataset.robots.ox.ac.uk/datasets/>

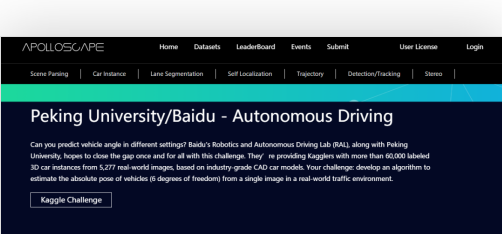


4、ApolloScape

百度Apollo开源的数据集，包含3D目标检测、语义分割、目标跟踪、立体视觉、场景识别等各类信息，数据量非常大！

数据集链接：

<https://apolloscape.auto/>



5、BDD100K

主要包括视频数据、道路目标检测、实例分割、可驾驶区域等相关

数据。

其中：

视频数据：在一天中的许多不同时间、天气条件和驾驶场景中，探索超过1100小时驾驶体验的100000高清视频序列。我们的视频序列还包括GPS位置、IMU数据和时间戳。

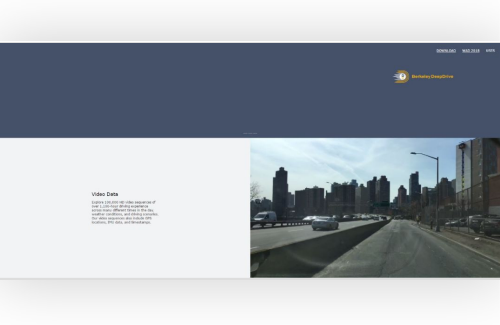
道路目标检测：为公共汽车、红绿灯、交通标志、人、自行车、卡车、汽车、汽车、火车和骑手在100000张图像上标注的二维边框。

实例分割：使用像素级和丰富的实例级注释，浏览超过10000个不同的图像。

可驾驶区域：从100000张图片中学习复杂的驾驶决策。

数据集链接：

<https://bdd-data.berkeley.edu/>



6、Waymo Open Dataset

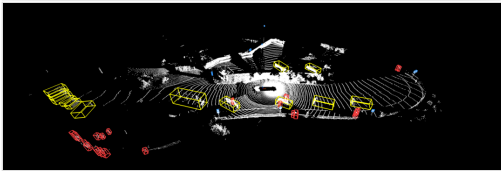
Waymo 数据集包含 3000 段驾驶记录，时长共 16.7 小时，平均每段长度约为 20 秒。整个数据集一共包

含 60 万帧，共有大约 2500 万 3D 边界框、2200 万 2D 边界框。

此外，在数据集多样性上，Waymo Open Dataset 也有很大的提升，该数据集涵盖不同的天气条件，白天、夜晚不同的时间段，市中心、郊区不同地点，行人、自行车等不同道路对象，等等。

数据集链接：

<https://github.com/waymo-research/waymo-open-dataset>



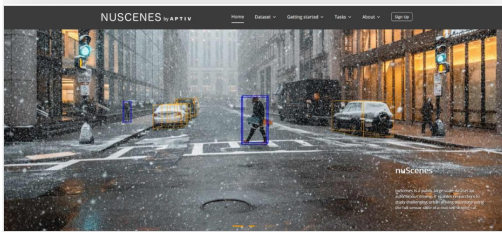
7、nuScenes数据集

nuScenes数据集 是自动驾驶公司 nuTonomy建立的大规模自动驾驶数据集，该数据集不仅包含了Camera和Lidar，还记录了雷达数据。这个数据集由1000个场景组成（即scenes，这就是该数据集名字的由来），每个scenes长度为20秒，包含了各种各样的情景。在每一个scenes中，有40个关键帧（key frames），也就是每秒钟有2个关键帧，其他的帧为sweeps。关键帧经过手工的标注，每一帧中都有了若干个annotation，标注的形式为bounding box。不仅标注了大小、

范围、还有类别、可见程度等等。
这个数据集不久前发布了一个
teaser版本（包含100个
scenes），正式版（1000个
scenes）的数据要2019年发布。这
个数据集在sample的数量上、标注
的形式上都非常好，记录了车的自
身运动轨迹（相对于全局坐标），
包含了非常多的传感器，可以用来
实现更加智慧的识别算法和感知融
合算法。

数据集链接：

<https://www.nuscenes.org/download>



8、3D Photography Dataset

华盛顿大学3D相机标定数据库。

数据集链接：

https://www-cvr.ai.uiuc.edu/ponce_grp/data/mview/

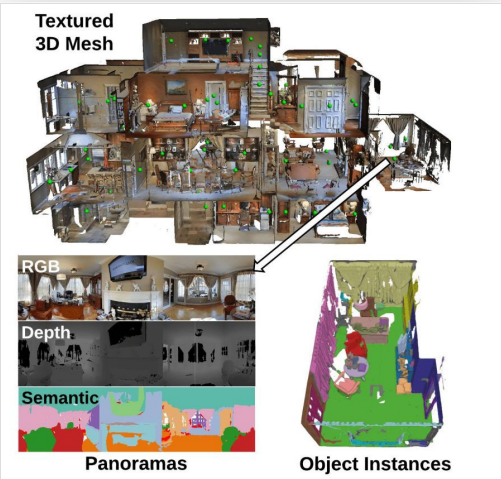
9、Matterport 3D重建数据集

A large-scale RGB-D dataset。该数据集包含10800个对齐的三维全

景视图（RGB+每个像素的深度），来自90个建筑规模场景的194400个RGB+深度图像。

数据集链接：

<https://matterport.com/>



10、NoW Dataset

3D人脸重建相关数据集。该数据集包含用iPhone X拍摄的100名受试者的2054张2D图像，以及每个受试者的单独3D头部扫描。头部扫描是评估的基本依据。受试者的年龄、体重指数和性别（55名女性，45名男性）各不相同。

数据集链接：

<https://ringnet.is.tue.mpg.de/challenge>



11、Pix3D

单目图像3D模型匹配数据。

数据集链接：

<https://pix3d.csail.mit.edu/>

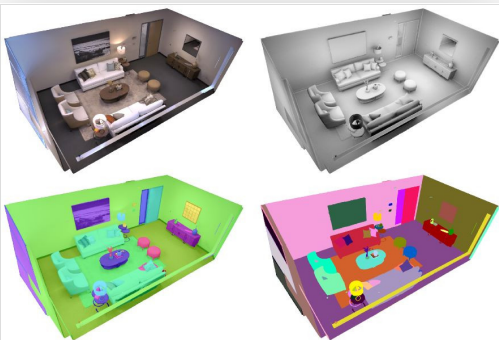


12、Replica Dataset

高质量室内场景三维重建数据。数据集中包含了18个高真实感的室内场景重建数据集Replica。每个场景由一个密集的网格、高分辨率高动态范围（HDR）纹理、每个基本语义类和实例信息以及平面镜和玻璃反射镜组成。副本的目标是使机器学习（ML）研究能够依赖于世界上视觉上、几何上和语义上真实的生成模型。

数据集链接：

<https://github.com/facebookresearch/Replica-Dataset>



13、Scan2CAD

将CAD模型与扫描数据对齐的数据集（适用于3D Object Pose Estimation, 3D Reconstruction）

对于公共数据集，我们为注释提供：

97607扫描模型和CAD模型
之间的关键点对应关系

扫描和计算机辅助设计之间的
14225个对象

1506次扫描

用于Scan2CAD基准测试的附加注释
隐藏测试集包括：

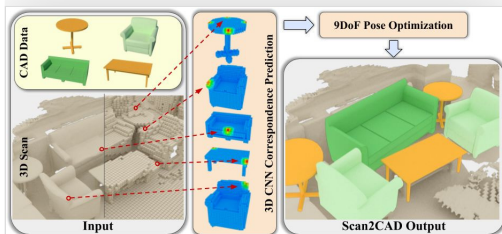
7557扫描模型和CAD模型之
间的关键点对应关系

扫描和CAD之间的1160个对
象

97次扫描

数据集链接：

<https://github.com/skanti/Scan2CAD>



14、ScanNet

ScanNet是一个RGB-D视频数据集，包含1500多个扫描中的250万

个视图，使用3D相机姿势、曲面重建和实例级语义分段进行注释（3D重建相关）。

数据集链接：

<https://www.scan-net.org/>



15、NYC3Dcars

在现实世界中为视觉任务设置的车
辆检测数据库。

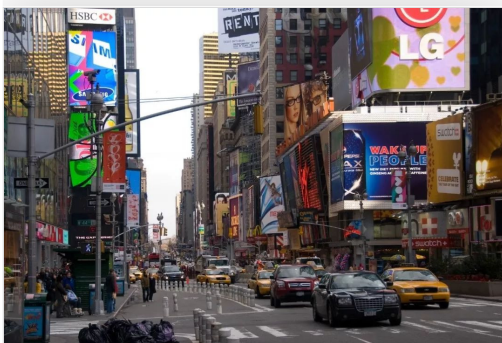
3D重建：NYC3DCAR中的
每张照片都已被GEO注册到
地球，在地球为中心的地球
固定坐标系统中提供完整的
相机内部和外部信息，使得
能够与现有地理空间数据无
缝集成。

地理数据：已集成了诸如
OpenStreetMap和NYC
OpenData提供的配套数据
库，以方便访问道路、人行
道和中间多边形等地理特征
以及道路网络连接。

车辆注释：人工注释器提供了数据库中包含的车辆详细说明。注释包括一个完整的6自由度的车辆姿态，车辆类型，2D车辆包围盒，和大约一天的照片时间。

数据集链接：

<https://nyc3d.cs.cornell.edu/>



16、Expressive Hands and Faces

EHF数据集（丰富姿态的手部和脸部）包含一个受试者穿着最少的衣服的100个精确的帧，执行各种身体姿势，包括自然的手指关节，以及一些面部关节和表情。

每个帧包括以下时间同步模式：

全身RGB图像。

一个JSON文件，包含OpenPose检测到的二维特征（身体关节、手关节、面部特征）。

物体的三维扫描。

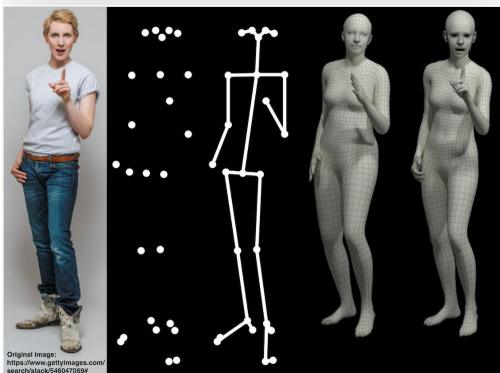
与上述扫描的3D SMPL-X对齐（3D网格），用作伪地面真值。

伪地面真值网格采用顶点到顶点（v2v）误差度量。这是一个比常见的三维关节误差范式更严格的度量标准，它不捕获表面误差和骨骼的旋转。

可以使用SMPL-X模型和SMPLify-X代码从单个RGB图像重建3D人体。

数据集链接：

<https://smpl-x.is.tue.mpg.de/>

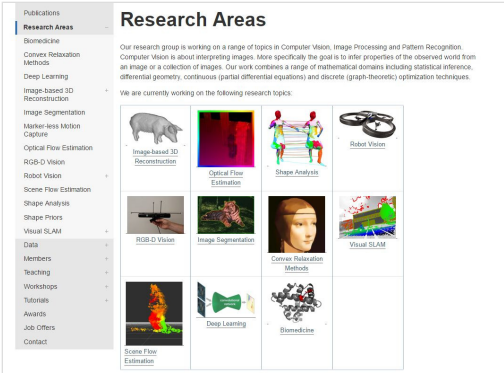


17、TUM数据集

主要包含多视图数据集、3D物体的识别分割、场景识别、3D模型匹配、vSALM等各个方向的数据。

数据集链接：

<https://vision.in.tum.de/>

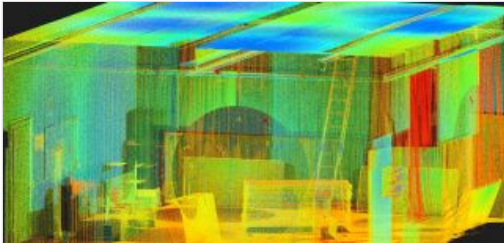


18、EUROC数据集

数据集中主要包含立体图像、同步IMU测量以及精确的运动和真实地面结构。

数据集链接：

<https://projects.asl.ethz.ch/dataset/doku.php?id=knavvisualinertialdatasets>



来源：3D视觉工坊@微信公众号

*延伸阅读

最强战队 | 三维视觉、SLAM
方向全球顶尖实验室汇总

CVPR2019 |AR版“神笔马良”：
从单张2D图片建立3D

人物运动模型，华盛顿大学
与Facebook 3D重建

谷歌重磅推出TensorFlow
Graphics：为3D图像任务打
造的深度学习利器



△ 关注极市平台
获得最新CV干货

👍

2

💬

0

有用

🗑️

☆

收藏

B

I

H

U

S

x²

x₂

“

≡

≡

🔗

🖼️

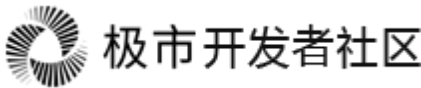
</>

📄

请输入内容


发布评论

暂无评论



极市开发者社区

极市CV社区是人工智能垂直领域计算机视觉技术的开发者社区，致力于为视觉算法开发者提供一个分享创造、结识伙伴、协同互助的平台。

 [极市平台官网](#)

统计信息

社区会员
26026

话题数量
4835

评论数量
5941

其他信息

联系邮箱
developer@cvmart.net

[推荐网站](#)

[社区指南](#)



扫一扫
关注极市平台