

如何快速用脚本下载 Nuscenes 数据集



自动化企鹅

互联网行业 从业人员

7 人赞同了该文章

最近换了一台服务器，可以独享4张v100 :)，但是需要重新下载Nuscenes数据集，在网上找了一些解决办法，最终选择用脚本的方式无痛下载Nuscenes

首先笔者参考的脚本是 github 上的一个批量下载的脚本：

GitHub - li-xl/nuscenes-download: script for downloading nuscenes

这是一个python 脚本文件，打开来如下图所示：

```
Code Blame 136 lines (109 loc) · 4.68 KB Code 55% faster with GitHub Copilot
1 import requests
2 import os
3 import hashlib
4 from tqdm import tqdm
5 import tarfile
6 import gzip
7
8 # replace with your API url and Bearer Token
9 bearer_token = 'eyJraWQ1OjI0aUk1Z2ZgZGh0UnRlVGRVl1hpZm9pa2U0bVJGaVlKN1le'
10
11 output_dir = "/path/to/save"
12 region = 'asia' # 'us' or 'asia'
13
14
15 download_files = {
16     "v1.0-trainval_meta.tgz": "3eee698806fc52330faa2e682b9f3a1",
17     "v1.0-trainval01_blobs.tgz": "8b5eaece969aea173a5317be153ca63",
18     "v1.0-trainval02_blobs.tgz": "116085f49ec4c60958f9d49b2d6b6fdd",
19     "v1.0-trainval03_blobs.tgz": "9de7f2a72864d6f9ef5ce0b74e84d311",
20     "v1.0-trainval04_blobs.tgz": "4d0bec5cc581672bb557c777cd0f0556",
21     "v1.0-trainval05_blobs.tgz": "3747bb98cdf6b0f29b236a61b95d66a",
22     "v1.0-trainval06_blobs.tgz": "9f6948a19b1104385c30ad58ab64dab",
23     "v1.0-trainval07_blobs.tgz": "d92529729f5506f5f0cc15cc82070c1b",
24     "v1.0-trainval08_blobs.tgz": "90897e7b58ea38634555c2b9583f4ada",
25     "v1.0-trainval09_blobs.tgz": "7cf0ac8b8d9925edbb6f23b96c0cd1cb",
26     "v1.0-trainval10_blobs.tgz": "fedf0df4e82630abb2d3d517be12ef9d",
27     "v1.0-test_meta.tgz": "f473fa9bb4d91e44ace5989d91419a46",
28     "v1.0-test_blobs.tgz": "3e1b78da1e8eed076ab3df082a54366",
29 }
30
31 # set request header
32 headers = {
33     'Authorization': f'Bearer {bearer_token}',
34     'Content-Type': 'application/json',
35 }
36
```

知乎 @minushorno

使用方法就是将 bear_token 替换成自己的token即可，要获取自己的token，需要先在Nuscenes注册并登录，登录上自己的账号之后再点击 nuscenes 进入下载页面，如下图所示：

Full dataset (v1.0) ^

In March 2019 we released the full nuscenes dataset with 1000 scenes. Due to the huge size of the dataset, we provide the mini, trainval and test splits separately. Mini (10 scenes) is a subset of trainval used to explore the data without downloading the entire dataset. Trainval (700+150 scenes) is packaged into 10 different archives that each contain 85 scenes. Test (150 scenes) is used for challenges and does not come with object annotations. Alternatively, it is also possible to download only selected modalities (camera, lidar, radar) or only keyframes. The meta data is provided separately and includes the annotations, ego vehicle poses, calibration, maps and log information.

Mini ^

Subset of trainval, 10 scenes, used to explore the data without downloading the whole dataset.

Metadata and sensor file blobs [US, Asia] 3.88 GB (4167696325 Bytes) md5: 791dd9ced556cfa1b425682f177b5d9b

Trainval ^

850 scenes, 700 train, 150 val. Metadata is for all 850 scenes, sensor file blobs are split into 10 subsets that each contains 85 scenes. It download only a subset of the data, you can choose between different sensor modalities (lidar, radar and camera) or keyframes only.

Metadata [US, Asia] 0.43 GB (461179800 Bytes) md5: 3eee698806fc52330faa2e682b9f3a1

File blobs of 85 scenes, part 1 [US, Asia] 29.41 GB (31574298075 Bytes) md5: 8b5eaece969aea173a5317be153ca63

File blobs of 85 scenes, part 2 [US, Asia] 28.06 GB (30128325054 Bytes) md5: 116085f49ec4c60958f9d49b2d6b6fdd

File blobs of 85 scenes, part 3 [US, Asia] 27.81 GB (29862707349 Bytes) md5: 9de7f2a72864d6f9ef5ce0b74e84d311

File blobs of 85 scenes, part 4 [US, Asia] 29.87 GB (32070725354 Bytes) md5: 4d0bec5cc581672bb557c777cd0f0556

File blobs of 85 scenes, part 5 [US, Asia] 26.25 GB (28181092231 Bytes) md5: 3747bb98cdf6b0f29b236a61b95d66a

File blobs of 85 scenes, part 6 [US, Asia] 25.61 GB (27503837689 Bytes) md5: 9f6948a19b1104385c30ad58ab64dab

File blobs of 85 scenes, part 7 [US, Asia] 27.50 GB (29523743832 Bytes) md5: 3e1b78da1e8eed076ab3df082a54366

File blobs of 85 scenes, part 8 [US, Asia]

关于作者



自动化企鹅

看书如山倒，写作如抽丝

互联网行业 从业人员

回答

0

文章

31

关注者

237

关注

发私信

大家都在搜

换一换

- 伊朗抗议席卷全国一半省份 373 万 热
- 死了吗App爆火 365 万 热
- 博主杀猪求助网友变百人... 364 万 热
- 秦志戩拟被聘为国乒总教练 363 万 热
- 威海特大暴雪为何这么厉害 363 万 新
- 广东汕头突发山火 363 万 新
- 呆呆杀猪 363 万 热
- 印度卫星运载火箭发射失败 363 万 新
- 河南省考报名延期2天 363 万 热
- 闫学晶 361 万 热

火山引擎 | intel.

2核2G云服务器 低至29元/年

抖音同款服务器,全网超低价

立即购买

广告

登录即可查看 超5亿 专业优质内容

超5千万创作者的优质提问、专业回答、
深度文章和精彩视频尽在知乎。

立即登录/注册

赞同 7

2 条评论

分享

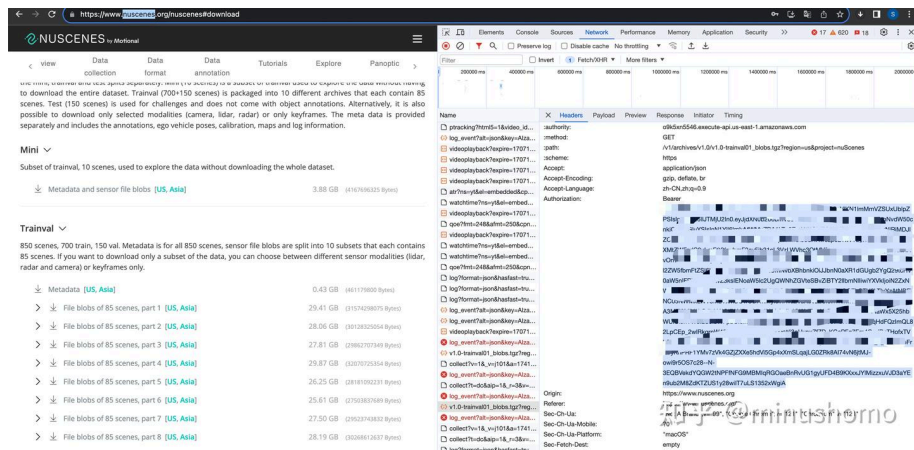
喜欢

收藏

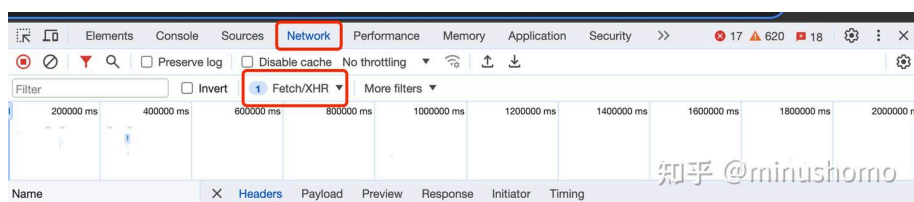
申请转载

...

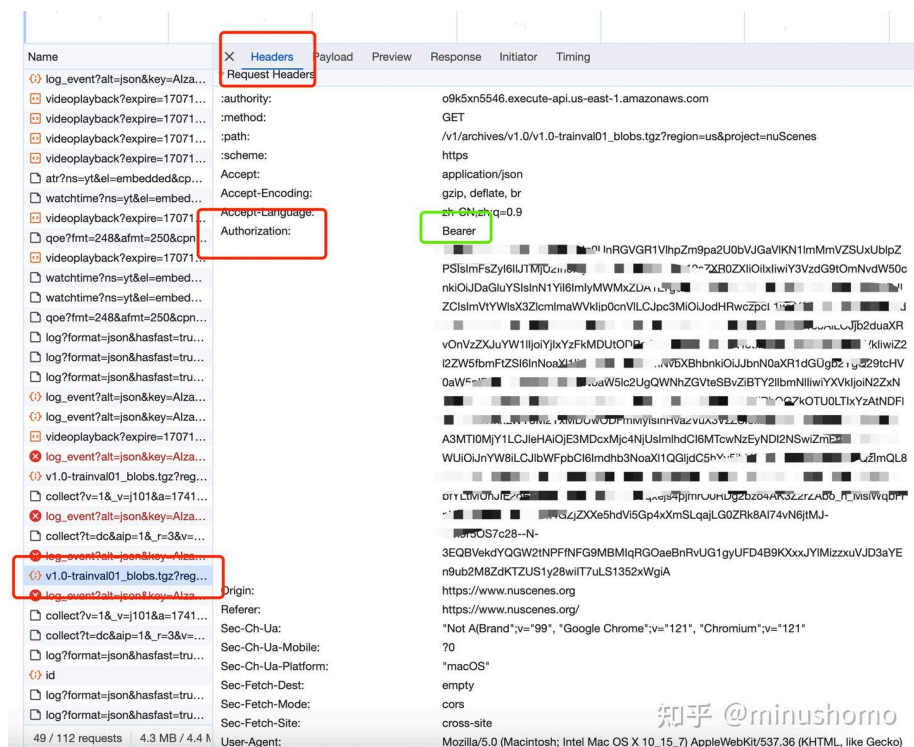
Chrome 浏览器地址下方输入 `https://www.nuscenes.org/nuscenes/download`，按下 `Enter` 键，即可进入下载页面。



选择控制台中的 network -> filter 选择 Fetch/XHR



在下面 Name 那一栏找到刚刚下载的数据集：



找到 Headers -> Authorization -> Bearer 下面那串很长的代号就是我们自己的 bearer_token，复制这一段 token，替换脚本中的 bearer_token

此外脚本中还可以选择是下载 us 还是 asia（参考代码注释）

运行脚本，显示：

登录即可查看 **超5亿** 专业优质内容

超 5 千万创作者的优质提问、专业回答、深度文章和精彩视频尽在知乎。





即表示下载成功

reference:

[GitHub - li-xl/nuscenes-download: script for downloading nuscenes](#)

发布于 2024-02-05 18:06 · 日本

[深度学习 \(Deep Learning\)](#)



未登录用户

2 条评论

默认 最新



白白白向

这样下载的是全部数据吗

2025-11-11 · 北京

回复 喜欢



再成熟一点就好了

我又来下数据集了🐼

2024-05-31 · 北京

回复 喜欢

推荐阅读



深度盘点：机器学习中常见的7种 AutoML 工具库！

大模型爱好... 发表于Pytho...



基础巩固之：xmlns属性梳理

果冻布丁兔 发表于菜鸟架构师...



【AutoML】当前有哪些可用的 AutoML 平台？

龙鹏·笔名... 发表于有三AI学...



AutoML是算法工程师的末日吗？

McGL

登录即可查看 超5亿 专业优质内容

超 5 千万创作者的优质提问、专业回答、深度文章和精彩视频尽在知乎。

