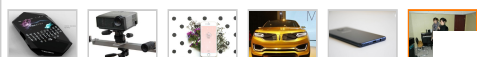




## 建盏大师排名



## 学唱歌



访问：237881次

积分：2816

等级：**BLOG > 5**

排名：第14334名

原创：52篇

转载：0篇

目录视图

摘要视图

RSS 订阅

图灵赠书——程序员11月书单 【思考】Python这么厉害的原因竟然是！ 感恩节赠书：《深度学习》等异  
作译者评选启动！ 每周荐书：京东架构、Linux内核、Python全栈

## YouTube-8M 数据集简介

标签：深度学习 Dataset YouTube-8M 数据集 视频

2016-10-06 11:07

5325人阅读

评论(0)

收藏

分类：

Machine Learning (20)

版权声明：本文为博主原创文章，未经博主允许不得转载。

目录(?)

[+]

## Preface

Google 公布了一个大型的视频数据集：[YouTube-8M 视频数据](#)  
的 Technical Report 在 arXiv 上也挂出来了：[YouTube-8M: A L](#)  
[Benchmark](#)。

关闭



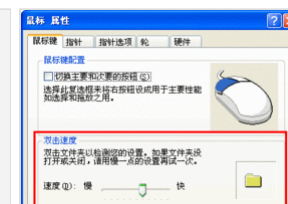
app开发报价单



OA办公系统



短信验证码接口



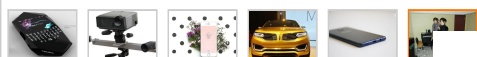
电脑初学入门



建盏大师排名



学唱歌



Developer (11)

Reading and Thinking (2)

NLP (0)

NLP (1)

## 文章存档

2017年03月 (4)

2016年11月 (5)

本文是对 Google 对这个数据集介绍的博文：[Announcing YouTube-8M: A Large and Diverse Labeled Video Dataset for Video Understanding Research](#) 的一个总结归纳，

## Introduction

这个数据集包含 8,000,000 万个 YouTube 视频链接，这些视频集进行了 video-level (视频级) 的标注，标注为 4800 种 Knowledge Graph entities (知识图谱实体)。

这次公布的 YouTube-8M 数据集相比较于之前公布的数据集：YouTube-1M，又是一次突破。之前的 YouTube-1M 是包含 500 种体育项目的 YouTube 视频数据集，包含 1,000,000 万个

## Characteristic

这个的数据集的特点如下：

8 Million  
Video URLs

0.5 Million  
Hours of Video

1.9 Billion  
Frame Features

8 Million video URLs

关闭

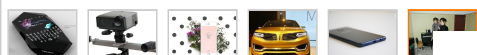




建盏大师排名



学唱歌



SSD: Signle Shot Detector 用于自然场景文...  
Handsome\_1205 : @sunnanaily:设置test\_interval

SSD: Signle Shot Detector 用于自然场景文...  
Handsome\_1205 : 博主你好, 你知道SSD原始论文中Table1的各个类别的检测精度是在哪找的么?

SSD: Signle Shot Detector 用于自然场景文...  
Handsome\_1205 : @yhz5256:自己建

SSD: Signle Shot Detector 用于自然场景文...  
圆橙 : 哇咔咔, 很有用的样子, 先收藏下

为了保证数据集的质量, 在选取视频时, 做了一些限制:

- 每一个视频都是公开的, 且每个视频至少有 1000 帧
- 每一个视频的长度在 120s 到 500s 之间
- 每一个视频至少与一个 Knowledge Graph entities (知识图谱实体) 相联系
- 成人视频由自动分类器移除

## 0.5 Million Hours of video

数据集中视频超过了 5000 个小时, 一般需要 1PB (1PB = 1024TB) 的硬盘来存储, 也需要 50 CPU-years 来处理这个视频。

所以为了减小存储开销与计算时间开销, Google 提供了 pre-computed and compressed features, 这样的话就可以在单台机子上一天内完成模型的训练。

## 1.9 Billion Frame Features

同时, 已经用在 ImageNet 上训练得到的 Inception-V3 image and video features, 提取 frame-level、video-level 特征。

这些特征是从 1.9 Billion 视频帧中, 以每秒 1 帧的时间分辨率处理, 是最后的特征能够存储在一张硬盘中 (小于 1.5T)。

## 4800 Classes

关闭

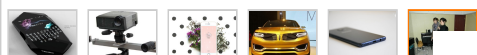




建盏大师排名



学唱歌



- \* 【CSDN】邀请您来GitChat赚钱啦！
- \* 【GitChat】精选——JavaScript进阶指南
- \* 改做人工智能之前，90%的人都没能给自己定位
- \* TensorFlow 人脸识别网络与对抗网络搭建
- \* Vue 移动端项目生产环境优化
- \* 面试必考的计算机网络知识点梳理

这些视频被标注的 annotation vocabulary 包含 4800 个 [Knowledge Graph](#) entities (知识图谱实体)。

每一个 entity 至少有 120 个训练视频，平均每个 entity 有 2229 个训练视频。annotation 的定义方式参照 [YouTube Data API](#)。

其中最多的一个 entity 是 [Vehicle](#)，超过了 500K 个训练视频；拥有最少视频的 entity 是 [Somersault](#)，只有 120 个视频。

4800 个 entity 被分为 24 种 **frequent**，最高 frequent 的是 Arts & Entertainment，超过：训练视频。最少 frequent 是 Finance，小于 14,000 个训练视频。

这里，所谓的 [24 top-level verticals](#)，详情可以参考 Youtube-8M 主页中的这个 Explore 页面：[dataset browser](#)，如下图：

关闭



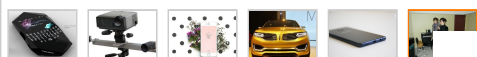




## 建盏大师排名



## 学唱歌



YouTube | 8M

## Vertical

- ✓ All
- Arts & Entertainment
- Autos & Vehicles
- Beauty & Fitness
- Books & Literature
- Business & Industrial
- Computers & Electronics
- Finance
- Food & Drink
- Games
- Health
- Hobbies & Leisure
- Home & Garden
- Internet & Telecom
- Jobs & Education
- Law & Government
- News
- People & Society
- Pets & Animals
- Real Estate
- Reference
- Science
- Shopping
- Sports
- Travel
- Uncategorized

Concert (386872)

Music video (266829)

Football (221721)

Food (188044)

Animal (164711)

Trailer (105288)

Trailer (91808)

Fashion (88723)

Mobile phone (84422)

Minecraft (79834)

Action-adventure game (77649)

Smartphone (77433)

Fishing (68256)

Bollywood (63628)

Cooking (60417)

Google

具体的，如 Arts & Entertainment 这类，其 Entity 如下：



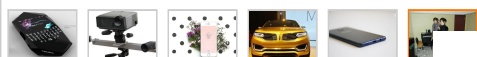
关闭



建盏大师排名



学唱歌



YouTube | 8M

Vertical

Arts & Entertainment

Filter

Entities

- |                          |                      |               |
|--------------------------|----------------------|---------------|
| Concert (386872)         | Animation (290812)   |               |
| Music video (266829)     | Dance (215675)       |               |
| Guitar (105288)          | Disc jockey (100370) |               |
| Trailer (91808)          | Bollywood (63628)    |               |
| Musical ensemble (60355) | Orchestra (60164)    |               |
| Choir (52870)            | Comedy (51919)       | Piano (50206) |
| Cartoon (48432)          | Comics (42490)       |               |
| Album (40472)            | Talent show (33274)  |               |
| Photography (32662)      | Television (26980)   |               |
| Drums (26589)            | Ballet (25565)       |               |
| Cooking show (20406)     | Drawing (20349)      |               |

Google

另外，关于数据集的具体的 labels，其归属的 Vertical，对应的 CSV 文件：[train-labels-histogram](#)：

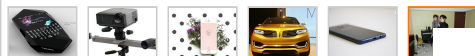




建盏大师排名



学唱歌



	A	B	C	D	E	F
1	NumTrainVideos	KnowledgeGraphID	LabelName	FirstVertical	SecondVertical	ThirdVertical
2		539926 /m/07yv9	Vehicle	/Autos & Vehicles		
3		386872 /m/01jddz	Concert	/Arts & Entertainment		
4		290812 /m/0hcr	Animation	/Arts & Entertainment		
5		266829 /m/0mdxd	Music video	/Arts & Entertainment		
6		252639 /m/01mw1	Video game	/Games		
7		221721 /m/02vx4	Football	/Sports		
8		215675 /m/026bk	Dance	/Arts & Entertainment		
9		188044 /m/02wbm	Food	/Food & Drink		
10		173192 /m/0410tth	Motorsport	/Sports		
11		164711 /m/0jbn	Animal	/Pets & Animals		
12		150413 /m/0k4j	Car	/Autos & Vehicles		
13		105288 /m/0342h	Guitar	/Arts & Entertainment		
14		100370 /m/02dsz	Disc jockey	/Arts & Entertainment		
15		91808 /m/03hdbf	Trailer	/Arts & Entertainment		
16		88723 /m/032t1	Fashion	/Beauty & Fitness		
17		84422 /m/050k8	Mobile phone	/Internet & Telecom		
18		79834 /m/09v6kpg	Minecraft	/Games		
19		77649 /m/02ntfj	Action-adventure game	/Games		
20		77433 /m/0169zh	Smartphone	/Internet & Telecom		
21		68256 /m/094jc	Fishing	/Hobbies & Leisure		
22		63628 /m/01chg	Bollywood	/Arts & Entertainment		
23		60417 /m/01mtb	Cooking	/Food & Drink		
24		60355 /m/05229	Musical ensemble	/Arts & Entertainment		
25		60164 /m/05pd6	Orchestra	/Arts & Entertainment		
26		55405 /m/04_sv	Motorcycle	/Autos & Vehicles		
27		52870 /m/0114jd	Choir	/Arts & Entertainment		
28		52673 /m/0643t	Personal computer	/Computers & Electronics		
29		52188 /m/07_tb	Video game console	/Computers & Electronics		
30		51919 /m/03ykjs9	Comedy	/Arts & Entertainment		
31		50206 /m/05r5c	Piano	/Arts & Entertainment		
32		50121 /m/09qf5	Winter sport	/Sports		
33		49736 /m/018w8	Basketball	/Sports		
34		48777 /m/0138t1	Toy	/Shopping		
35		48432 /m/0215n	Cartoon	/Arts & Entertainment		
36		47879 /m/03bf_rm	Strategy video game	/Games		

关闭

## 1.8 Avg.Labels/Video

每个视频平均有 1.8 个 labels , 这些 ground truth labels 来自于 content、metadata、contextual、user signals 对每个视频进行 a

每个视频的 label 个数从 1 到 39 , 平均每个视频有 1.8 个 label labels 个数在 2 ~ 3 个。

app开发报价单

OA办公系统

短信验证码接口

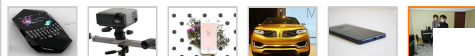
电脑初学入门



建盏大师排名



学唱歌



# Postscripts

上面只是我对这个数据集的一点翻译式的记录描述，使用时具体的详情，请参见 Google 对这份 YouTube-8M 的 [Technical Report: YouTube-8M: A Large-Scale Video Classification Benchmark](#)

arXiv.org > cs > arXiv:1609.08675

Computer Science > Computer Vision and Pattern Recognition

## YouTube-8M: A Large-Scale Video Classification Benchmark

Sami Abu-El-Haija, Nisarg Kothari, Joonseok Lee, Paul Natsev, George Toderici, Balakrishnan Varadarajan, Sudheendra Vijaya

(Submitted on 27 Sep 2016)

Many recent advancements in Computer Vision are attributed to large datasets. Open-source software packages for Machine Learning and inexpensive commodity hardware have reduced the barrier of entry for exploring novel approaches at scale. It is possible to train models over million within a few days. Although large-scale datasets exist for image understanding, such as ImageNet, there are no comparable size video classification datasets.

In this paper, we introduce YouTube-8M, the largest multi-label video classification dataset, composed of ~8 million videos (500K hours) annotated with a vocabulary of 4800 visual entities. To get the videos and their labels, we used a YouTube video annotation system, which includes their main topics. While the labels are machine-generated, they have high-precision and are derived from a variety of human-based annotations including metadata and query click signals. We filtered the video labels (Knowledge Graph entities) using both automated and manual curation strategies, including asking human raters if the labels are visually recognizable. Then, we decoded each video at one-frame-per-second, Deep CNN pre-trained on ImageNet to extract the hidden representation immediately prior to the final layer, and make both the features and video-level labels available for download.

We trained various (modest) classification models on the dataset, evaluated them using popular metrics. Despite the size of the dataset, some of our models train to convergence in less than a day on a single GPU. We provide code for training a TensorFlow model and for computing metrics.

顶  
3

踩  
0

关闭



app开发报价单

OA办公系统

短信验证码接口

电脑初学入门

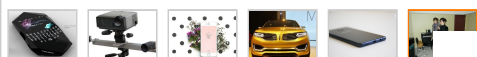




## 建盏大师排名



## 学唱歌



- 上一篇 [Open Image 数据集简介](#)
- 下一篇 [LeetCode: 字符串转换到整数](#)

### 相关文章推荐

- youtube8m视频识别竞赛
- MySQL在微信支付下的高可用运营--莫晓东
- Hive实战之Youtube数据集
- 容器技术在58同城的实践--姚远
- YouTube-8M 数据集介绍与使用总结
- SDCC 2017之容器技术实战线上峰会
- 【数据库】FaceDataset常用的人脸数据库
- SDCC 2017之数据库技术实战线上峰会
- 肤色检测&人脸检测数据集等链接大集合（持续更...
- 腾讯云容器服务架构实现介绍--董晓杰
- 13.3 time--时间操作和转换功能
- 微博热点事件背后的数据库运维心得--张
- SparkSQL实战之Youtube数据集
- 使用Python找出历史上5星好评数最多的5
- 人脸识别数据集之MS-Celeb-1M
- MacroSoft数据集MS-Celeb-1M



人脸识别算法



短信验证码接



显卡回收



吉他学习

### 查看评论

暂无评论

关闭



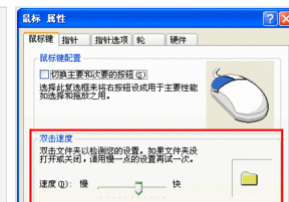
app开发报价单



OA办公系统



短信验证码接口



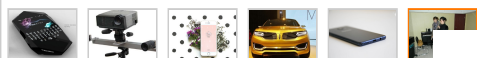
电脑初学入门



## 建盏大师排名



## 学唱歌



您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

\* 以上用户言论只代表其个人观点,不代表CSDN网站的观点或立场

方式 | 版权声明 | 法律顾问 | 问题报告 | 合作伙伴 | 论坛反馈

webmaster@csdn.net 400-660-0108 | 北京创新乐知信息技术有限公司 版权所有 | 江苏知之为计算机有限公司 | 江苏乐:

SDN.NET, All Rights Reserved



关闭

