# 知识增强的视频语义理解 技术评测文档

目录

[知识增强的视频语义理解 技术评测文档 1](#_Toc68632128)

[1、背景 1](#_Toc68632129)

[2、任务描述 1](#_Toc68632130)

[3、数据集描述 3](#_Toc68632131)

[3.1、视频语义标签数据 3](#_Toc68632132)

[3.2、 知识数据 4](#_Toc68632133)

[4、评测指标 5](#_Toc68632134)

[4.1、评测指标说明 5](#_Toc68632135)

[4.2、计算方法说明 5](#_Toc68632136)

[5、任务提交方式 6](#_Toc68632137)

[5、奖励设置 6](#_Toc68632138)

[7、时间安排 6](#_Toc68632139)

[8、组织者信息 7](#_Toc68632140)

[附录 7](#_Toc68632141)

## 1、背景

在移动互联网、大数据的时代背景下，互联网上的视频数据呈现爆发式增长，作为日益丰富的信息承载媒介，视频的深度语义理解是诸多视频智能应用的基础，具有重要的研究意义和实际应用价值。传统基于感知的视频内容分析缺乏语义化理解能力，而充分利用知识图谱的语义化知识并结合多模态学习和知识推理技术，有望实现更深入的视频语义理解。

知识增强的视频语义理解技术评测，期望融合知识、NLP、视觉、语音等相关技术和多模态信息，为视频生成刻画主旨信息的语义标签，从而实现视频的语义理解。

## 2、任务描述

知识增强的视频语义理解任务，以百度好看/全民小视频资源为对象，在感知内容分析的基础上，融合知识、语言、视觉、语音等多模信息，结合知识计算与推理，为视频生成相应的语义标签。

具体来讲，首先实现对视觉基础内容的分析，然后利用知识进行计算与推理，对百度好看/全民小视频从分类标签、语义标签（包括：实体/概念/事件/实体属性等维度）层面进行理解，并为其生成这几个层面相应的语义标签结果。

此次任务的输入输出定义如下：

#### 输入：

#### 视频数据

#### 原始视频数据（平均时长<=5分钟）

#### 基础感知解析结果（OCR、语音识别、人脸识别结果等）

#### 输出：

#### 视频语义标签，包括以下几类标签：

#### 分类标签：二层体系的封闭集（详见数据集描述）

#### 语义标签：实体/概念/事件/实体属性等标签（详见数据集描述）

同时，我们也提供了任务所需的统一的知识数据。

#### 知识数据

#### 筛选后的知识库（筛选策略，详见数据集描述）

示例结果数据如下：

#### 示例1：

#### 输入：



#### 输出：

#### 分类标签：生活家居(一级)、装修(二级)

#### 语义标签：阳台、阳台设计、阳台装修

#### 示例2：

#### 输入：



#### 输出：

#### 分类标签：舞蹈（一级）、少儿舞蹈（二级）

#### 语义标签：小跳蛙、小敏老师、跳舞、舞蹈、舞蹈教学

## 3、数据集描述

### 3.1、视频语义标签数据

视频语义标签数据集由训练集、验证集和测试集组成，整体标注数据大约8万条左右，按6:1:1比例分配（最终发布的数据规模以及比例可能存在变动，以最终发布为准），三份数据均通过百度众包标注生成，详细标注质量将会在数据发布时一并给出。

视频语义标签数据集中的视频数据主要来自于百度好看/全民小视频物料库,并从中筛选部分涉及知识类的视频作为目标数据。标注数据集中每条数据的格式为：

{

 "@id": "1",

 "@fromurl": "http://www.baidu.com/video.html",

 "title": "好几万元一平米的空间，你就只用来只放拖把吗？",

 "objUrl": "http://vd4.bdstatic.com/mda-iffsemnuj5r9sac.mp4",

 "perception": [

 "ocr":["ocr结果（json格式）"], *// 字幕识别结果*

 "asr":["asr结果（json格式）"], *// 语音识别结果*

 "face\_detect":["人脸检测结果的json"], *// 人脸识别结果*

 ],

 "resourceType": ["video"],

 "category": [ *// 视频分类标签*

 {

 "@value": "生活家居",

 "@meta": {"type": "level1"}

 },

 {

 "@value": "装修",

 "@meta": {"type": "level2"}

 }

 ],

 "tag": [ *// 视频语义标签*

 {

 "@value": "阳台",

 },

 {

 "@value": "阳台设计",

 },

 {

 "@value": "阳台装修",

 },

 ],

}

 语义标签tag字段，包括描述这个视频的实体、概念、事件及实体属性等不同类型的标签，详细标准见附录《**不同类型语义标签解释**》。

### 知识数据

#### 3.2.1、知识库

 该任务知识库来自百度百科知识库，知识库中的每个实体都包含一个subject\_id(知识库id)，一个subject名称，实体的别名，对应的概念类型，以及与此实体相关的一系列二元组< predicate，object>（<属性，属性值>）信息。知识库中每行代表知识库的一条记录（一个实体信息），每条记录为json数据格式。

示例如下所示：

{

 "subject\_id":"10001",

 "subject":"碟中谍5",

 "alias":[

 "碟中谍5:神秘国度",

 "碟中谍:失控国度",

 "职业特工队5",

 "不可能的任务5"

 ],

 "type":[

 "Movie"

 ],

 "data":[

 {

 "predicate":"外文名",

 "object":" Mission: Impossible - Rogue Nation,

"

 },

 {

 "predicate":"发行公司",

 "object":"派拉蒙影业公司"

 },

 {

 "predicate":"类型",

 "object":"动作, 惊悚, 冒险"

 },

 {

 "predicate":"导演",

 "object":"克里斯托夫·迈考利"

 },

 {

 "predicate":"主演",

 "object": [

 "汤姆·克鲁斯",

 "杰瑞米·雷纳",

 "西蒙·佩吉",

 "丽贝卡·弗格森",

 "文·瑞姆斯"

 }

 ]

}

#### 3.2.2、视频分类体系数据

**视频分类体系人工精心设计的二层结构，总计近300个有效类别。视频分类体系数据**详见附录《视频分类部分体系》。

## 4、评测指标

### 4.1、评测指标说明

本任务采用传统的F1-Measure作为评测指标，我们通过将输出结果与人工标准集合进行比较来计算F1 分值。

对于给定的视频，评测集中的标签结果为：$TG=\{t\_{1},t\_{2},…,t\_{n}\}$，$n$为评测集的标签数目,参赛系统预测出的标签结果为：$TP =\{t'\_{1},t'\_{2},…,t'\_{m}\}$，$m$为预测标签数目。$N$为评测集的视频数目。计算公式如下：

* + $P=\frac{\sum\_{i\in N}^{}|TG(i) ∩ TP(i)|}{\sum\_{i\in N}^{}|TP(j)|}$
	+ $R=\frac{\sum\_{i\in N}^{}|TG(i) ∩ TP(i)|}{\sum\_{i\in N}^{}|TG(i)|}$
	+ $F1=\frac{2\*P\*R}{P+R}$

其中, P（Precision）为准确率、R（Recall）为召回率。

### 4.2、计算方法说明

本任务会输出两类标签结果（分类标签和语义标签），首先会针对每次标签单独计算F1（即分类标签F1-a、实体/概念/事件/实体属性标签F1-b），然后通过加权求和计算最终 F1-final。计算公式如下：

$$F1\_{final}=W\_{1}\*F1\_{a}+ W\_{2}\*F1\_{b}$$

其中，权重$W\_{1} W\_{2} $会根据标签分布情况、标签预测的难易程度、关注标签技术类型等因素选择（权重数据可能存在变动，以最终发布为准）

## 5、任务提交方式

本次任务将采取排行榜的方式，验证集发布后，允许参赛队伍每天至多向平台提交三次结果，文件命名为“result.txt”，格式与任务描述中的示例输出相同，排名每十分钟更新一次。参赛队伍可在测试集发布之前随时上传验证集的计算结果，管理系统会及时更新各队伍的最新排名情况，验证集榜不计入比赛最终成绩，我们会准备部分百度周边礼品奖励周冠军和周最佳进步团队。

测试集为最终评估比赛成绩的有效评测集，待测试集发布后，允许参赛队伍每天至多提交一次，总共至多提交五次评测结果。

**所有获奖团队需额外提交相关材料以供资格审查，要求如下**：

1. 测试集结果文件，用result.txt命名（UTF-8格式），格式如下：
2. 相关代码及说明
3. 方法描述文档（非评测论文，评测论文撰写要求见CCKS 2021官网）

将以上三个文件在任务提交截止日期前发送至邮箱ccks2021\_baidukg@126.com。邮件的标题为：“CCKS2021-参赛队名称”，例如“CCKS2021-TeamKG队”。代码及其文档需打包成一个文件（tar，zip，gzip，rar等均可），作为邮件附件传送，要求提交所有的程序代码及相关的配置说明，**程序应当可以运行且所得结果与result.txt相符**。如果方法使用了额外资源，要求说明并提供资源文件或地址。

## 5、奖励设置

提供3,5000元奖金给获奖团队，奖金设置如下：

第一名：20,000元

第二名：10,000元

第三名：5,000元

创新奖：5,000元

## 7、时间安排

评测任务发布：4月15日

报名时间：4月15日—7月15日

训练及验证数据发布：暂定4月30日（以最终发布时间为准）

测试数据发布：7月15日

提交测试结果：7月20日

评测论文提交：8月5日

CCKS会议日期(评测报告及颁奖)：8月18日—21日

## 8、组织者信息

知识增强的视频语义理解 组委会邮箱：ccks2021\_baidukg@126.com

汪琦：wangqi31@baidu.com

冯知凡：fengzhifan@baidu.com

柴春光：chaichunguang@baidu.com

杨虎：yanghu03@baidu.com

## 附录

**视频分类部分体系**，部分示例如下（完整体系将随正式数据发布）：

|  |  |
| --- | --- |
| 一级分类 | 二级分类 |
| 影视 | 影视花絮 |
| 影视片段 |
| 影视评论|影视解说 |
| 影视剧情介绍 |
| 影视资讯 |
| 综艺节目 | 脱口秀 |
| 真人秀 |
| 文物鉴宝 |
| 明星娱乐 |
| 音乐综艺 |
| 访谈节目 |
| 法制节目 |
| 舞蹈综艺 |
| 动漫 | 动漫花絮 |
| 动漫片段 |
| 动漫资讯 |
| 动漫剧情介绍 |
| 动漫周边|动漫衍生 |
| 二次元文化 |
| 纪录片 | 美食纪录片 |
| 历史纪录片 |
| 法制纪录片 |
| 科技纪录片 |
| 自然纪录片 |
| 军事纪录片 |
| 纪录片制作 |
| 摩托车 | 摩托车资讯 |
| 摩托车配件 |
| 摩托车展示 |
| 汽车 | 汽车资讯 |
| 汽车评测|汽车选购 |
| 汽车配件 |
| 汽车维修 |
| 汽车展示 |
| 汽车租赁 |
| 汽车改装 |
| 用车技巧 |
| 汽车安全 |
| 汽车养护 |
| 超跑|跑车 |
| 越野 |
| 轿车 |
| 驾考 |
| 教育|教育培训 | 学前教育 |
| 义务教育|基础教育 |
| 大学教育 |
| 成人教育 |
| 职业教育 |
| 兴趣爱好学习 |
| 考试|资格考试 |
| 在线教育 |
| 留学 |
| 考研 |
| 教育资讯 |
| 职场教育 |
| … | … |

 **不同类型语义标签解释：**

|  |  |
| --- | --- |
| **实体** | 包括实例与概念，**1、实例**是客观存在的具体唯一的物体，不可再分。如：一个人『刘德华』、一个组织机构『北京大学』、一部电影『美丽心灵』、一本书『少年维特之烦恼』；**2、概念**一般分为以下几种：2.1、类别概念（一般是指一类实体的总称，如苹果、水果、植物、钱、明星，类别型概念是可以实例化为客观存在的实例对象）2.2、抽象概念（如红色、跳跃、情绪等，简单来说就是多个物质与物质的联系）2.3、其它（一些不可再分的类别也是概念，如“菜名”、“产品名”等） |
| **主题** | **又称为topic**，主题是指描述一个文档内容主旨的名词或名词集，一篇文章通常包含多种主题。如：人工智能、无人车、爱情片、明星、性感舞蹈等。 |
| **事件** | 在特定的时空下(时间、空间)，由一个或多个角色(事件主体)参与的，围绕某个主题开展的一系列活动。这里的事件一般是反映事件元，而非具体的事件实例，如：接吻、拥抱、唱歌、跳舞、英语教学、维修汽车、飙车等 |
| **实体属性** | 一般是描述一个实体或概念自身具备的某种特性，包括：实体的属性、实体关系、等。如：刘德华的老婆、红烧肉的做法、无线充电的原理、明星街拍等 |