

图像标注指南

版本号: 20180302

1 标注系统介绍

1.1 登录标注网站

网址为 [REDACTED] 登录界面如图 1 所示。
用户使用自己的账号和密码登录。



图 1 标注网站登录界面

1.2 按要求标注

用户登录之后，网页会跳转到图 2 界面。请用户根据如下要求按顺序进行标注。



图 2 登录后的标注界面，请给出能够描述图像内容的一个句子与多个标签

1.2.1 句子标注

1.2.1.1 选择推荐句子

系统会推荐 5 个句子，如果有较为符合图像表述的句子，可以点选该句子，点选后的句子将会出现在下方的输入框中，然后可以基于输入

框中的句子进行修改；如果推荐的句子都表述不当，则可以不选。注意推荐句子大多是通过机器翻译产生的，可能存在细微的语法错误，如图2中的“双女”应该“女双”，标注时请注意观察并作出修改。

1.2.1.2 输入描述句子

如果在上一步点选了推荐句子，则在已有的句子上进行修改；如果没有点选，则直接用一个句子描述图像内容，要求“能包含图片所呈现的主要对象（subject 可以是人或物）在什么场景（where）做什么事（action）”且“客观”，不少于 20 字，否则不能提交，输入框下方的句子字数部分会对输入框中输入的句子做实时字数统计。更详细的要求参见附录。请在保证句子通顺的前提下，使句子尽可能详尽。

1.2.2 标签标注

1.2.2.1 选择推荐标签

系统会推荐 5 个标签，如果有符合图像内容的标签，可以点选该标签，点选后的标签将会出现在下方的输入框中。注意如果推荐的标签不符合图像内容，或者不够准确，如“动物”并没有表述清楚是猫、狗还是大象，则不进行选择。

1.2.2.2 输入其他标签

在选择标签后，可以接着直接在输入框中输入标签，即输入一个中文词并按回车键。标签既用于描述图像中的主要物体、场景或事件，也用于在图像中出现但不适合放入句子的内容，如图2 中的“观众”，如果放入句子会影响句子的通顺程度，但是也属于图像的内容，可以加入到标签当中。要求至少输入 3 个标签，标签要尽可能详尽，使用最细粒度的标注词汇，如图 2 的场景，可以表述为“网球场”就不要标注“运动场”，且标签不能被输入的句子完全包含。

1.2.3 提交标注结果

完成句子标注和标签标注后，可以点击提交按钮。若有不符合规范的操作，会在右下角弹出相应的提示，如图 3 所示，只有改正错误之后才能顺利提交。



图 3 提交后弹出错误提示的示例

用户可以通过左上角的【已标注图片数目】了解到自己已经标注的图像数目；

通过右上角的【浏览】观察自己的标注历史，如图 4 所示，点击标注历史中相应的图像可以对标注不当的图像进行修改重新提交，在标注历史页面点击右上角的【back】可以回到正常的标注界面。



图 4 标注历史页面，可以点击图像修改已经提交的标注句子和标签

2 标注示例

以下为一些供用户参考的标注示例。

2.1 正确示例



 <p>COCO_val2014_000000515176</p> <p>玻璃杯 洗衣机 房间 鞋</p> <p>房间里白色的洗衣机旁边有一台银色的饮水机</p>	 <p>COCO_train2014_000000313517</p> <p>长颈鹿 沙滩 湖水 树木</p> <p>一只长颈鹿走在沙滩上，前面的湖水中的绿地里生长着许多树木</p>
 <p>COCO_train2014_000000543263</p> <p>女孩 长椅 公园 草坪 游人</p> <p>公园里一个穿粉色短袖上衣戴着太阳镜的女子坐在一条蝴蝶形的长椅上</p>	 <p>COCO_train2014_000000005823</p> <p>雪地 滑雪 雪橇 女孩 护目镜</p> <p>一个穿着红色格子上衣、蓝色裤子的女孩戴着护目镜在雪地里滑雪</p>

2.2 错误示例



COCO_val2014_000000025228

日落 海洋 女人 小船

一个女人坐在海里的小船上

✗ 句子不足 20 个字



COCO_train2014_000000041135

网球 女人

一个女人在一场网球比赛中挥舞球拍击球

✗ 标签不足 3 个



COCO_train2014_000000559356

刀 菜板 西红柿 蔬菜

桌上的菜板上放着西红柿和一些切碎的蔬菜，旁边有一盘胡萝卜丝和一把刀

✗ 标签被句子完全包含



COCO_val2014_000000070460

火车 涂鸦 蓝天 白云

雪山

蓝天白云下一列有涂鸦的火车停
止站台边，远处可以看到雪山

✗ 语句不通顺



3 联系我们

在标注过程中如遇到问题，或者有意见建议需要反馈，请联系 王晓旭（wangxiaoxuruc@163.com）。

附录：句子标注细则

- 1 只能用一句话描述图像内容，且不少于 **20** 个字。
 - 1.1 尽量采用书面语，避免口语化的表述，不能使用语气词（“啊”，“哦”）。
 - 1.2 避免错别字。
 - 1.3 保证语句通顺，语法正确，如采用正确的量词（“一辆车”，“一列火车”）等。
 - 1.4 必须是句子，不能写一个词组，比如“……的……”（如“一辆银色的汽车”）这样的形式。
- 2 句子中只能有客观事实。
 - 2.1 不要写将要可能发生的事情。不要写推测。不要有声音方面的描写。
 - 2.2 不要写任何不确定的东西，确信写下的句子正确的描述了图像。
 - 2.3 不要表达对图像的看法或感受。
 - 2.4 尽量不要用“这里”、“这是”等指示代词。
- 3 不要使用专有名词进行描述，包括国家名、地名、建筑名、人名等，如“大本钟”、“奥巴马”、“伦敦”等。
- 4 只需要描述图像的主体，无须描述图像中的每一个细节。