**【任务】**

复现论文[《LQMFormer: Language-aware Query Mask Transformer for Referring Image Segmentation》](https://openaccess.thecvf.com/content/CVPR2024/papers/Shah_LQMFormer_Language-aware_Query_Mask_Transformer_for_Referring_Image_Segmentation_CVPR_2024_paper.pdf)

**【要求】**

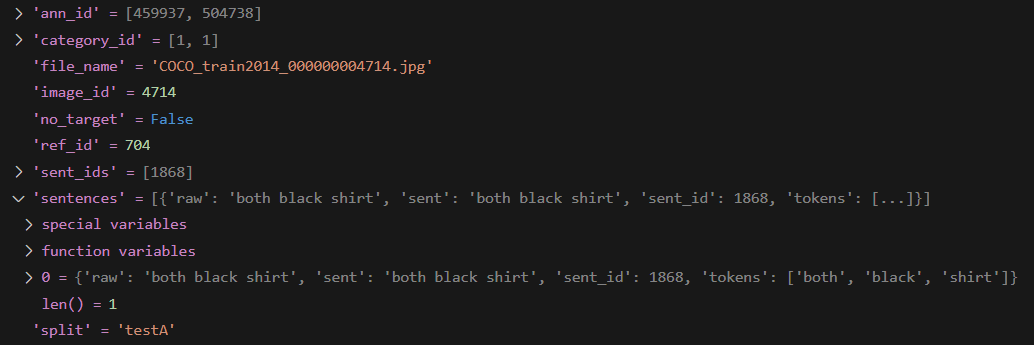
* **实现referring image segmentation任务**
* 根据论文尽量复现代码和模型
* 评价指标：testA和testB上的oIoU，gIoU，和Pr@0.7;Heatmap visualization of cosine similarity between different query mask predictions

**【交作业形式】**

PPT答辩：设计思路、方法、实验结果（参考论文，进行定量和定性分析，鼓励不同的消融实验与对比实验）、方案优劣分析等。

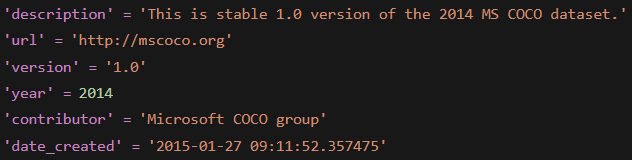
**【数据说明】**

1. 图像来自COCO2014,下载链接为：<https://cocodataset.org/#download>
2. 文本标注和数据划分保存在json文件中，下载链接为<https://drive.google.com/drive/folders/1HSLS5BsUYqwJZ15f92O0iH6d0z9CWUJw?usp=sharing>：
3. 数据以7000：1000：1000：1000划分为“train”，“val”, “testA”和“testB”。用“train”进行训练，“val”进行验证，“testA”和“testB”进行最终性能评估。
4. 每个样本的“ann\_id”,“category\_id”,“file\_name”，“image\_id”,“no\_target”,“ref\_id”,“sent\_ids”,“sentences”,“split”以字典的形式储存在**miami2025.json**中，如下图所示：



1. 数据集的“info”,“image”,“licenses”,“annotations”,“categories”以字典的形式储存在**instances.json**中，细节如下：

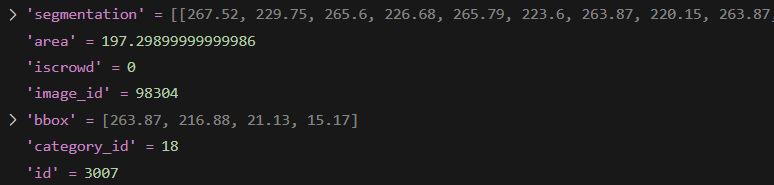
* “info”:数据集整体信息



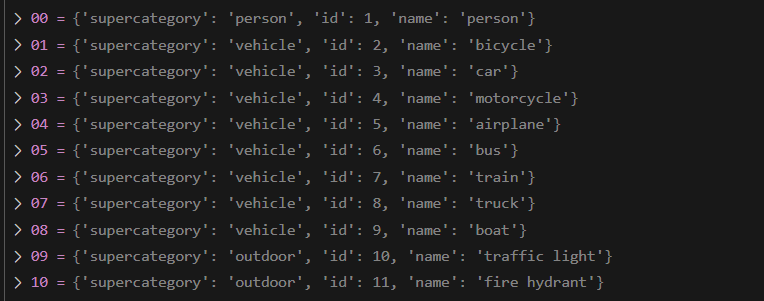
* “image”：保存有每个样本的图像信息，包括文件名，下载链接等



* “licenses”：许可证相关信息
* “annotations”：保存有每个样本的标注信息，包括标注ID和分割掩码



* “categories”：80个类别的类别信息



1. 数据准备说明

* 从[https://cocodataset.org/#download](https://cocodataset.org/#download中下载完整的COCO2014) 中下载完整的COCO2014数据集，或根据instances.json->“image”->“flickr\_url”找到每个样本的下载链接，依次下载图像。
* 根据miami2025.json->“split”将数据分为“train”，“val”, “testA”和“testB”四组。
* 训练时，从miami2025.json中读取“split”为“train”的数据，根据“file\_name”找到对应下载好的图像，根据“sentences”-> “sent”取出对应输入文本(若需分词结果，可从“sentences”-> “tokens”中取出)。根据“ann\_id”找到instances.json->“annotations”->“id”中编号一致的标注，该标注对应的“segmentation”即为分割标注。请注意，**部分“ann\_id”由多个实例组成，最终输出结果需要将多个实例的mask融合成一个。**（中间监督想使用实例级mask的可以自行使用）
* 同上，获取val，testA，testB的图像、文本、和分割标注。