

兴趣视觉关系检测 ViROI 数据集

【简介】

ViROI 数据集是第一个用于兴趣视觉关系检测（Visual Relation of Interest Detection, VROID）评估的数据集。ViROI 数据集基于 IOID 和 MSCOCO2017 数据集构建，分为训练集和测试集，分别包含 25091 张和 5029 张图片。ViROI 数据集共有 133 类物体，每张图片上的兴趣视觉关系（VROI）用<主语，谓语，宾语>三元组表示，主语和宾语给出物体类别和包围框。

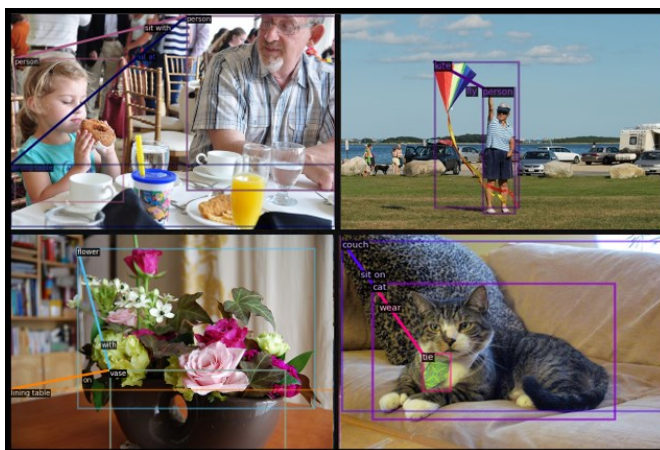


图 1 ViROI 数据集标注示例

【详细说明】

我们基于 IOID 和 MSCOCO 2017 的训练集构建了第一个用于兴趣视觉关系检测（VROID）评估的数据集。IOID 本身是基于 MSCOCO2017 的训练集构建的，共包含 45000 张图片，用于兴趣物体检测。由于兴趣视觉关系存在于兴趣物体之间，我们在 IOID 的基础上继续进行了筛选和标注。在标注过程中，我们利用了 MSCOCO2017 训练集提供的 caption，最终构建了包含 30120 张图片的 ViROI 数据集，其中训练集和测试集分别包含 25091 张和 5029 张图片。我们将图片上的兴趣视觉关系（VROI）用<主语，谓语，宾语>三元组表示，主语和宾语给出物体类别和包围框。ViROI 数据集中包含了 133 个物体类别，与 IOID 和 MSCOCO 2017 中全景分割所提供的类别相同，包括 80 种 thing（如人、球和牛），和 53 种 stuff（如墙、树和山）。ViROI 数据集中包含了 249 个关系谓词类别，包括 77 个动词，17 个介词，5 个空间关系词和 150 个介词短语。

ViROI 数据集保存在 json 格式的文件中：

- train_images_dict.json: 训练集物体标注（格式如下图(a)）

- train_images_triplets_dict.json: 训练集兴趣视觉关系标注（格式如下图(b)）
- test_images_dict.json: 测试集物体标注（格式如下图(a)）
- test_images_triplets_dict.json: 测试集兴趣视觉关系标注（格式如下图(b)）
- class_dict.json: 物体类别字典（格式如下图(c)）
- relation_dict.json: 视觉关系类别字典（格式如下图(d)）

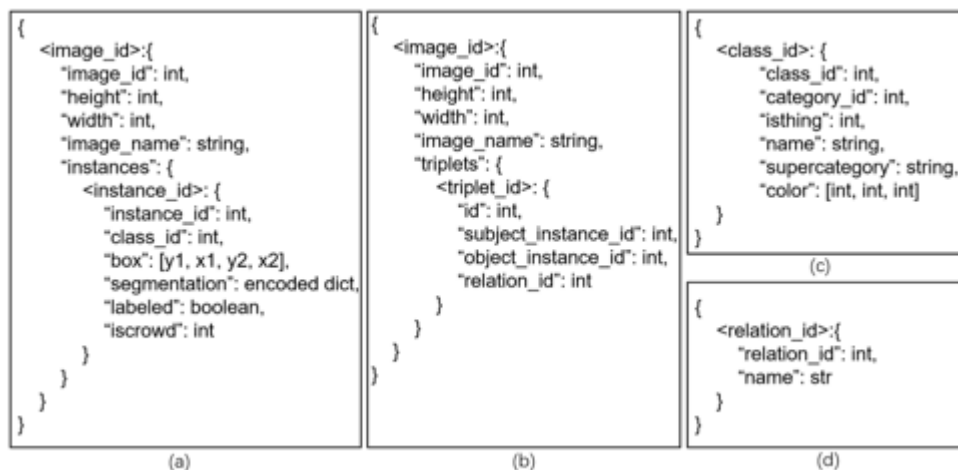


图 2 ViROI 数据集标注格式

【文献引用】

如果您使用 ViROI 数据集，请引用以下文献：

- [1] Fan Yu, Haonan Wang, Tongwei Ren, Jinhui Tang and Gangshan Wu. Visual relation of interest detection. Proceedings of ACM International Conference on Multimedia (MM'20), Seattle, WA, USA, 2020.
- [2] Fan Yu, Haonan Wang, Tongwei Ren, Jinhui Tang, Gangshan Wu, Jingjing Chen, and Zhenzhong Kuang. Reproducibility companion paper: visual relation of interest detection. Proceedings of ACM International Conference on Multimedia Reproducibility Companion Papers (MM'21-Repro), Chengdu, China, 2021.