

第四次训练结果

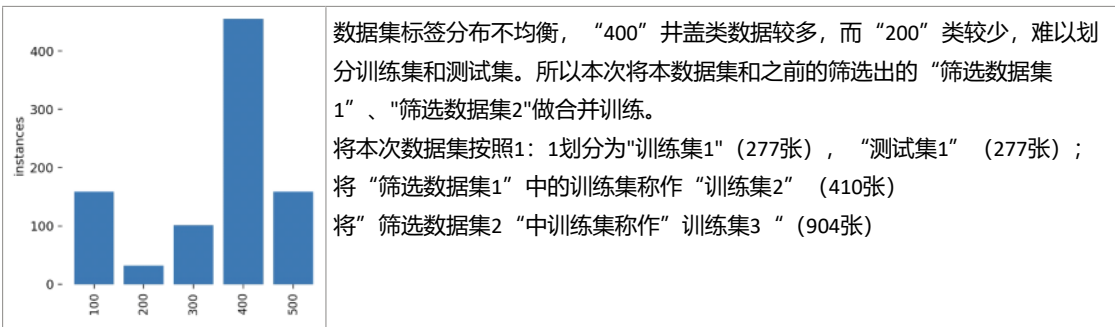
2022年9月21日 10:55

1、数据集

本次接收到训练集来源于JOB-202209-1\JOB-202209，经挑选及整理共得到554张图片（丢弃的数据将在下文中分析）。

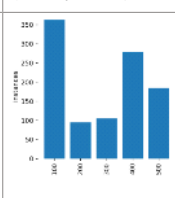
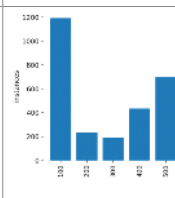
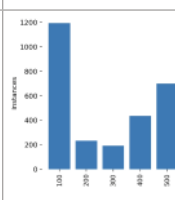
名称	修改日期	类型
JOB-20220901	2022/09/20 15:09	文件夹
JOB-20220902	2022/09/17 16:01	文件夹

数据集标签分布如下：



2、训练结果

这里展示多次训练后最佳的几个训练结果。

序号	训练集	测试集	训练集标签分布	结果																																																		
1	"训练集1" (277张) "训练集2" (410张)	"测试集1" (277张)		<table><tr><th>Class</th><th>Images</th><th>Labels</th><th>P</th><th>R</th><th>mAP@.5</th><th>mAP@.5:.95:</th></tr><tr><td>all</td><td>277</td><td>472</td><td>0.607</td><td>0.456</td><td>0.468</td><td>0.279</td></tr><tr><td>100</td><td>277</td><td>86</td><td>0.595</td><td>0.307</td><td>0.353</td><td>0.147</td></tr><tr><td>200</td><td>277</td><td>15</td><td>0.362</td><td>0.267</td><td>0.155</td><td>0.1</td></tr><tr><td>300</td><td>277</td><td>51</td><td>0.66</td><td>0.431</td><td>0.489</td><td>0.344</td></tr><tr><td>400</td><td>277</td><td>236</td><td>0.789</td><td>0.808</td><td>0.843</td><td>0.458</td></tr><tr><td>500</td><td>277</td><td>84</td><td>0.63</td><td>0.464</td><td>0.499</td><td>0.346</td></tr></table>	Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:	all	277	472	0.607	0.456	0.468	0.279	100	277	86	0.595	0.307	0.353	0.147	200	277	15	0.362	0.267	0.155	0.1	300	277	51	0.66	0.431	0.489	0.344	400	277	236	0.789	0.808	0.843	0.458	500	277	84	0.63	0.464	0.499	0.346	
Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:																																																
all	277	472	0.607	0.456	0.468	0.279																																																
100	277	86	0.595	0.307	0.353	0.147																																																
200	277	15	0.362	0.267	0.155	0.1																																																
300	277	51	0.66	0.431	0.489	0.344																																																
400	277	236	0.789	0.808	0.843	0.458																																																
500	277	84	0.63	0.464	0.499	0.346																																																
2	"训练集1" (277张) "训练集2" (410张) "训练集3" (904张)	"测试集1" (277张)		<table><tr><th>Class</th><th>Images</th><th>Labels</th><th>P</th><th>R</th><th>mAP@.5</th><th>mAP@.5:.95:</th></tr><tr><td>all</td><td>277</td><td>472</td><td>0.688</td><td>0.48</td><td>0.5</td><td>0.303</td></tr><tr><td>100</td><td>277</td><td>86</td><td>0.542</td><td>0.314</td><td>0.354</td><td>0.166</td></tr><tr><td>200</td><td>277</td><td>15</td><td>0.47</td><td>0.333</td><td>0.189</td><td>0.106</td></tr><tr><td>300</td><td>277</td><td>51</td><td>0.874</td><td>0.412</td><td>0.509</td><td>0.375</td></tr><tr><td>400</td><td>277</td><td>236</td><td>0.87</td><td>0.767</td><td>0.849</td><td>0.444</td></tr><tr><td>500</td><td>277</td><td>84</td><td>0.682</td><td>0.571</td><td>0.599</td><td>0.425</td></tr></table>	Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:	all	277	472	0.688	0.48	0.5	0.303	100	277	86	0.542	0.314	0.354	0.166	200	277	15	0.47	0.333	0.189	0.106	300	277	51	0.874	0.412	0.509	0.375	400	277	236	0.87	0.767	0.849	0.444	500	277	84	0.682	0.571	0.599	0.425	
Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:																																																
all	277	472	0.688	0.48	0.5	0.303																																																
100	277	86	0.542	0.314	0.354	0.166																																																
200	277	15	0.47	0.333	0.189	0.106																																																
300	277	51	0.874	0.412	0.509	0.375																																																
400	277	236	0.87	0.767	0.849	0.444																																																
500	277	84	0.682	0.571	0.599	0.425																																																
3	"训练集1" (277张) "训练集2" (410张) "训练集3" (904张)	"测试集1" (277张)		<table><tr><th>Class</th><th>Images</th><th>Labels</th><th>P</th><th>R</th><th>mAP@.5</th><th>mAP@.5:.95:</th></tr><tr><td>all</td><td>277</td><td>472</td><td>0.724</td><td>0.426</td><td>0.486</td><td>0.294</td></tr><tr><td>100</td><td>277</td><td>86</td><td>0.709</td><td>0.267</td><td>0.376</td><td>0.183</td></tr><tr><td>200</td><td>277</td><td>15</td><td>0.486</td><td>0.2</td><td>0.156</td><td>0.0915</td></tr><tr><td>300</td><td>277</td><td>51</td><td>0.864</td><td>0.392</td><td>0.501</td><td>0.368</td></tr><tr><td>400</td><td>277</td><td>236</td><td>0.858</td><td>0.763</td><td>0.842</td><td>0.443</td></tr><tr><td>500</td><td>277</td><td>84</td><td>0.703</td><td>0.507</td><td>0.556</td><td>0.386</td></tr></table>	Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:	all	277	472	0.724	0.426	0.486	0.294	100	277	86	0.709	0.267	0.376	0.183	200	277	15	0.486	0.2	0.156	0.0915	300	277	51	0.864	0.392	0.501	0.368	400	277	236	0.858	0.763	0.842	0.443	500	277	84	0.703	0.507	0.556	0.386	与序号2 训练集、 测试集相 同
Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:																																																
all	277	472	0.724	0.426	0.486	0.294																																																
100	277	86	0.709	0.267	0.376	0.183																																																
200	277	15	0.486	0.2	0.156	0.0915																																																
300	277	51	0.864	0.392	0.501	0.368																																																
400	277	236	0.858	0.763	0.842	0.443																																																
500	277	84	0.703	0.507	0.556	0.386																																																

根据上述结果，在测试集上准确率最高为第三次的训练结果，达72.4%，除 “200” 类均达到70%以上的准确率，但召回率只有 “400” 类别达到了76%。这是因为测试集中都是新采集的图片，且400类的数据占比大。相比之下，序号2准确率稍低，但召回率较序号3高。

3、本次数据集存在的问题及建议

①命名问题

在JOB-20220901和JOB-20220902文件夹中，图片均是按照同一个前缀和起始编号命名，这样两个文件夹中的图片都是从foo-0000.jpg开始编号，存在大量相同命名

的图片，建议给视频分割时**更改不同的前缀或者起始编号**，以免给后续数据集汇总、划分带来不便。

②数据集挑选问题

本次接收到的数据集中，井盖类占大多数，而其他类别较少，可以在分割时**提高采样帧数**，即一秒内分割出更多的图片，**模糊的图片不要进行标注**，某类别较多的可以少标注一些，例如数据集中井盖类数据较多，为平衡标签，可以少标注一些，而其余类别尽量多标注一些。

下面本次数据集中是一些**不适合**作为数据集的图片实例，标注时应该尽量避免标注有类似问题的图：

\\0919第四次训练\delete\foo-00211.jpg [3 / 66]





