retinanet违建测试-类别对比测试

# 详细实验

## 实验1: model\_final.pkl、p2-p7、（900,1408）、O4、S4、R6 – zj

**存放位置（100）：**/home/zhangjing/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**1004训练集，验证集250， (8:2，1920\*1080h)；测试集349（扣过的图片）

**类别**：1+10[autotruck、crane、forklift、mixerTruck、person、colorPlate、pit、bricksPile、mound、car]

**备注**：

person包括戴帽子的工人、人；

car包括油罐车、小汽车、客车；

bricksPile包括砖堆、柱状堆料、袋装堆料；

forklift包括铲车、挖掘机。

**GPU内存**：6.9G + 6.9G

**参数设置：**

**网络结构**：configs/12\_2017\_baselines/ retinanet\_R-101-FPN\_1x.yaml

**预训练模型：**model\_final.pkl

FPN中 MULTILEVEL\_RPN为 True，并且RPN\_MAX\_LEVEL为 7，RPN\_MIN\_LEVEL为

2，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 4

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0, 3.0,5.0,0.2,0.5),6个ratio;

**图片大小**：配置为900，1408，2个GPU，batchsize总共为2。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 5000 |  |  |  | 0.4821 | 0.8403 |
| 6000 |  |  |  | 0.4992 | 0.8446 |
| 8000 |  |  |  | 0.5095 | 0.8529 |
| 1万 |  | 0.7624 | 0.9951 | 0.5089 | 0.8521 |
| 2万 |  | 0.8642 | 0.9985 | 0.5106 | 0.8438 |
| 3万 |  | 0.9109 | 0.9986 | 0.5089 | 0.8342 |
| 6万 |  | 0.9625 | 0.9988 | 0.5037 | 0.8313 |
| 10万 |  | 0.9766 | 0.9988 | 0.5052 | 0.8291 |
| 15万 |  | 0.9764 | 0.9988 | 0.5057 | 0.8292 |

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 5000 | 0.4821 | 0.8403 | 0.4782 | 0.3159 | 0.4113 | 0.5439 |
| 6000 | 0.4992 | 0.8446 | 0.5163 | 0.3048 | 0.4304 | 0.5461 |
| 8000 | 0.5095 | 0.8529 | 0.5285 | 0.3017 | 0.4331 | 0.5766 |
| 1万 | 0.5089 | 0.8521 | 0.5224 | 0.2752 | 0.4245 | 0.5741 |
| 2万 | 0.5106 | 0.8438 | 0.5409 | 0.3543 | 0.4246 | 0.5869 |
| 3万 | 0.5089 | 0.8342 | 0.5408 | 0.3079 | 0.4183 | 0.5885 |
| 6万 | 0.5037 | 0.8313 | 0.5466 | 0.3515 | 0.4171 | 0.5826 |
| 10万 | 0.5052 | 0.8291 | 0.5513 | 0.3093 | 0.4188 | 0.5849 |
| 15万 | 0.5057 | 0.8292 | 0.5507, | 0.3093 | 0.4184 | 0.5870 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **索引** | **类别** | **训练集**  （1004图、**4980bbox**） | **验证集**  （250图、1201bbox） | **1万迭代、AP** | |
| **tain** | **val** |
| **1** | autotruck | 628 | 161 | 79.8 | 58.9 |
| **2** | crane | 11 | 0 | 75.0 | nan |
| **3** | forklift | 1051 | 239 | 76.5 | 59.7 |
| **4** | mixerTruck | 40 | 10 | 78.3 | 43.3 |
| **5** | person | 354 | 87 | 67.8 | 31.2 |
| **6** | colorPlate | 1124 | 288 | 74.2 | 54.6 |
| **7** | pit | 153 | 39 | 78.3 | 70.5 |
| **8** | bricksPile | 516 | 103 | 78.0 | 40.4 |
| **9** | mound | 657 | 160 | 78.1 | 60.9 |
| **10** | car | 446 | 114 | 76.4 | 38.5 |

## 实验2: model\_final.pkl、p2-p7、（900,1408）、O4、S4、R6 – lg

**存放位置（213）：**/data/ligang/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**1004训练集，验证集250， (8:2，1920\*1080h)；测试集349（扣过的图片）

**类别**：1+12：[autotruck、crane、forklift、mixerTruck、person、colorPlate、pit、bricksPile、

mound、car、Pile、digger]

**备注**：person包括戴帽子的工人、人；

car包括油罐车、小汽车、客车；

Pile包括柱状堆料、袋装堆料。

**GPU内存**：6.9G + 6.9G

**参数设置：**

**网络结构**：configs/12\_2017\_baselines/ retinanet\_R-101-FPN\_2x\_illbuild\_P2\_P7\_S4\_anchors6.yaml

**预训练模型：**model\_final.pkl

FPN中 MULTILEVEL\_RPN为 True，并且RPN\_MAX\_LEVEL为 7，RPN\_MIN\_LEVEL为

2，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 4

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0, 3.0,5.0,0.2,0.5),6个ratio;

**图片大小**：配置为900，1408，2个GPU，batchsize总共为2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 1万次 | 0.065 | 0.6987 | 0.9091 | 0.4709 | 0.7832 |
| 2万次 | 0.018 | 0.7968 | 0.9154 | 0.4604 | 0.7691 |
| 3万次 | 0.00906 | 0.8285 | 0.9155 | 0.4642 | 0.7659 |
| 4万次 | 0.00504 | 0.8558 | 0.9156 | 0.4622 | 0.7717 |
| 5万次 | 0.00441 | 0.8780 | 0.9156 | 0.4610 | 0.77 |
| 6万次 | 0.00316 | 0.8854 | 0.9156 | 0.4565 | 0.7594 |
| 7万次 | 0.00312 | 0.8879 | 0.9157 | 0.4541 | 0.7572 |
| 8万次 | 0.00185 | 0.8953 | 0.9157 | 0.4590 | 0.7615 |
| 10万次 | 0.00249 | 0.8974 | 0.9157 | 0.4550 | 0.7596 |
| 14万次 | 0.00119 | 0.9004 | 0.9157 | 0.4579 | 0.7679 |

## 实验3: model\_final\_X.pkl、p2-p7、（900,1280）、O4、S4、R6 – ys

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**1004训练集，验证集250， (8:2，1920\*1080h)；

**类别**：1+10[autotruck、crane、forklift、mixerTruck、person、colorPlate、pit、bricksPile、

mound、car]

**备注**：person包括戴帽子的工人、人；

car包括油罐车、小汽车、客车；

bricksPile包括砖堆、柱状堆料、袋装堆料；

forklift包括铲车、挖掘机。

**GPU内存**：G

**参数设置：**

**网络结构**：configs/illbuld/ retinanet\_X-101-64x4d-FPN\_p7\_2\_r6\_L\_ft.yaml

**预训练模型：**model\_final\_X.pkl

FPN中 MULTILEVEL\_RPN为 True，并且RPN\_MAX\_LEVEL为 7，RPN\_MIN\_LEVEL为

2，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 4，

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0, 3.0, 5.0, 0.5, 0.2),6个ratios；

**图片大小**：配置为900，1280，2个GPU，batchsize总共为2。（768\*1280）

**备注：**虽然model\_final\_X.pkl保存的权重是从P3开始的，这里修改代码，让其强

行加载到P2层。

下面是训练和测试的结果：（retinanet\_X-101-64x4d-FPN\_p7\_2\_r6\_L\_ft.yaml）

前面的实验，训练集和验证集都参与了训练。下面是只有训练集参与训练的测试结果：模型存储在retinanet\_X-101-64x4d-FPN\_p7\_2\_r6\_L\_ft\_class10文件夹中：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 5千次 | 1.81🡪0.1 | 0.5312 | 0.8982 | 0.4472 | 0.8126 |
| 1万次 | 1.81🡪0.04 | 0.5813 | 0.9472 | 0.4401 | 0.8008 |
| 2万次 | 1.81🡪0.01 | 0.5864 | 0.9496 | 0.4395 | 0.8266 |
| 3万次 | 1.81🡪0.005 | 0.6014 | 0.9466 | 0.4415 | 0.7952 |
| 4万次 | 1.81🡪0.005 | 0.5971 | 0.9446 | 0.4417 | 0.8142 |
| 5万次 | 1.81🡪0.004 | 0.5966 | 0.9439 | 0.4406 | 0.8133 |
| 10万次 | 1.81🡪0.003 | 0.5956 | 0.9441 | 0.4398 | 0.8129 |

## 实验4: model\_final\_X.pkl、p2-p7、（900,1280）、O4、S4、R6 – ys

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**1004训练集，验证集250， (8:2，1920\*1080h)；

**类别**：1+12[autotruck、crane、forklift、mixerTruck、person、colorPlate、pit、bricksPile、

mound、car]

**备注**：person包括戴帽子的工人、人；

car包括油罐车、小汽车、客车；

Pile包括柱状堆料、袋装堆料。

**GPU内存**：G

**参数设置：**

**网络结构**：configs/illbuld/ retinanet\_X-101-64x4d-FPN\_p7\_2\_r6\_L\_ft\_cls12.yaml

**预训练模型：**model\_final\_X.pkl

FPN中 MULTILEVEL\_RPN为 True，并且RPN\_MAX\_LEVEL为 7，RPN\_MIN\_LEVEL为

2，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 4，

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0, 3.0, 5.0, 0.5, 0.2),6个ratios；

**图片大小**：配置为900，1280，2个GPU，batchsize总共为2。（768\*1280）

**备注：**虽然model\_final\_X.pkl保存的权重是从P3开始的，这里修改代码，让其强

行加载到P2层。

下面是训练和测试的结果：（retinanet\_X-101-64x4d-FPN\_p7\_2\_r6\_L\_ft.yaml）

前面的实验，训练集和验证集都参与了训练。下面是只有训练集参与训练的测试结果：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 5千次 | 1.73🡪0.1 | 0.4862 | 0.8164 | 0.4162 | 0.7430 |
| 1万次 | 1.73🡪0.04 | 0.5437 | 0.8706 | 0.4238 | 0.7675 |
| 2万次 | 1.73🡪0.01 | 0.5472 | 0.8689 | 0.4101 | 0.7399 |
| 3万次 | 1.73🡪0.005 | 0.5469 | 0.8635 | 0.4042 | 0.7399 |
| 5万次 | 1.73🡪0.004 | 0.5445 | 0.8617 | 0.4030 | 0.7382 |
| 12万次 | 1.73🡪0.003 | 0.5439 | 0.8615 | 0.4011 | 0.7374 |

# 汇总

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验名称 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 实验1（11类，resnet101） | 0.065 | 0.8642 | 0.9985 | 0.5106 | 0.8438 |
| 实验2（13类，resnet101） | 0.005 | 0.6987 | 0.9091 | 0.4709 | 0.7832 |
| 实验3（11类，resnet101\_neXt） | 0.01 | 0.5971 | 0.9446 | 0.4417 | 0.8142 |
| 实验4（13类, resnet101\_neXt） | 0.04 | 0.5437 | 0.8706 | 0.4238 | 0.7675 |

# 结论

通过上述4个实验表明，11类的实验结果高于10类的实验结果，由于数据少，所以暂时resnet101网络结构比resnet101\_neXt网络结构的结果高。