retinaNet违建测试

-预训练模型、anchors、FPN层等对比

# 预训练模型不同

## R-101.pkl

### 实验1: (576,1024)、p2-p7、O4、S4、R6、trainval

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**参数设置：**

**网络结构**：configs/illbuld/ retinanet\_R-101-FPN\_P7\_2\_R6.yaml

**预训练模型**为R-101.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576，1024，2个GPU，batchsize总共为4。

下面是训练和测试的结果：（retinanet\_R-101-FPN\_P7\_2\_R6.yaml）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | benchmark (349) | |
| AP | AP50 | AP | AP50 | AP | AP50 |
| 1万次 | 1.78🡪0.05 | 0.7799 | 0.9498 | 0.8037 | 0.9790 | 0.0760 | 0.1852 |
| 3万次 | 1.78🡪0.004 | 0.9434 | 0.9984 | 0.9469 | 0.9997 | 0.0759 | 0.1790 |
| 4万次 | 1.78🡪0.004 | 0.9644 | 0.9985 | 0.9625 | 0.9998 | 0.0758 | 0.1779 |
| 6万次 | 1.78🡪0.004 | 0.9675 | 0.9985 | 0.9634 | 0.9997 | 0.0751 | 0.1783 |

耗时： 249ms

**备注：**AP是IOU0.5到IOU0.95的均值，AP50是IOU0.5。

20号球基的测试集：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1万次 | 3万次 | 4万次 | 6万次 |
| AP | 0.0333 | 0.0341 | 0.0318 | 0.0310 |
| AP50 | 0.1076 | 0.1067 | 0.1053 | 0.1036 |

## model\_final.pkl

### 实验2: (768,896)、p2-p7、O4、S4、R6、trainval

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**参数设置：**

**网络结构**：configs/illbuild/ retinanet\_R-101-FPN\_P7\_2\_R6\_ft.yaml

**预训练模型**为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, 0.5, 0.2),6个ratios；

**图片大小**：配置为768，896，2个GPU，batchsize总共为4。

**备注：**虽然model\_final.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\_R-101-FPN\_P7\_2\_R6\_ft.yaml）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | benchmark (349) | |
| AP | AP50 | AP | AP50 | AP | AP50 |
| 1万次 | 1.78🡪0.05 | 0.8272 | 0.9949 | 0.8452 | 0.9973 | 0.1376 | 0.2710 |
| 3万次 | 1.78🡪0.004 | 0.9518 | 0.9984 | 0.9552 | 0.9987 | 0.1393 | 0.2783 |
| 6万次 | 1.78🡪0.004 | 0.9664 | 0.9975 | 0.9694 | 0.9988 | 0.1360 | 0.2707 |
| 10万次 | 1.78🡪0.004 | 0.9664 | 0.9975 | 0.9691 | 0.9988 | 0.1357 | 0.2712 |

耗时： 228ms

20号球基的测试集：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1万次 | 3万次 | 6万次 | 10万次 |
| AP | 0.0654 | 0.0557 | 0.0546, | 0.0546 |
| AP50 | 0.1597 | 0.1495 | 0.1471 | 0.1472 |

### 实验3: (900,1408)、p2-p7、O4、S4、R6、trainval

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**参数设置：**

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

**预训练模型**为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, 0.5, 0.2),6个ratios；

**图片大小**：配置为900，1408，2个GPU，batchsize总共为2。（896\*1408）

**备注：**虽然model\_final.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\_R-101-FPN\_p7\_2\_r6\_L\_ft.yaml）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | 测试集（310） | |
| AP | AP50 | AP | AP50 | AP | AP50 |
| 1万次 | 1.78🡪0.05 | 0.7771 | 0.9967 | 0.7328 | 0.9524 | 0.0991 | 0.2073 |
| 3万次 | 1.78🡪0.004 | 0.9247 | 0.9997 | 0.8815 | 0.9998 | 0.0964 | 0.2047 |
| 4万次 | 1.78🡪0.004 | 0.9408 | 0.9997 | 0.9287 | 0.9999 | 0.0942 | 0.2000 |
| 4.5万次 | 1.78🡪0.004 | 0.9432 | 0.9997 | 0.9326 | 0.9999 | 0.0945 | 0.2001 |

测试集（310）是20号球机的图片

耗时：340ms

### 实验9: (900,1408)、 p2-p7、O4、S4、R6、train

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

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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 (756) | | 验证集 (188) | | benchmark (310)  20号球机 | |
| AP | AP50 | AP | AP50 | AP | AP50 |
| 1万次 |  | 0.8233 | 0.9982 | 0.4844 | 0.7885 | 0.0988 | 0.2044 |
| 2万次 |  | 0.9211 | 0.9997 | 0.4817 | 0.7818 | 0.0922 | 0.1995 |
| 3万次 |  | 0.9532 | 0.9997 | 0.4743 | 0.7827 | 0.0901 | 0.1913 |
| 5万次 |  | 0.9741 | 0.9998 | 0.4782 | 0.7858 | 0.0860 | 0.1888 |
| 6万次 |  | 0.9187 | 0.9998 | 0.4851 | 0.7933 | **0.1002** | **0.2165** |
|  |  |  |  |  |  |  |  |

**耗时：**

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 10000 | 0.4844 | 0.7885 | 0.5279 | 0.3461 | 0.4297 | 0.5235 |
| 20000 | 0.4817 | 0.7818 | 0.5334 | 0.3264 | 0.4163 | 0.5570 |
| 30000 | 0.4743 | 0.7827 | 0.5151 | 0.2667 | 0.4118 | 0.5480 |
| 50000 | 0.4782 | 0.7858 | 0.5189 | 0.2696 | 0.4151 | 0.5543 |
| 60000 | 0.4851 | 0.7933 | 0.5178 | 0.3131 | 0.4285 | 0.5478 |

## model\_final\_X.pkl

### 实验4:(900,1280)、p2-p7、O4、S4、R6、trainval

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**参数设置：**

**网络结构**：configs/illbuild/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4。（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）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | 测试集（310） | |
| AP | AP50 | AP | AP50 | AP | AP50 |
| 1万次 | 2.05🡪0.05 | 0.8109 | 0.9982 | 0.8195 | 0.9995 | 0.1012 | 0.2230 |
| 3万次 | 2.05🡪0.004 | 0.9404 | 0.9997 | 0.9480 | 1.0000 | 0.0906 | 0.2141 |
| 5万次 | 2.05🡪0.004 | 0.9654 | 0.9997 | 0.9599 | 1.0000 | 0.0897 | 0.2175 |

耗时： 240ms

### 实验11: (900,1280)、p2-p7、O4、S4、R6、train

**存放位置（214）：**/data/yushan/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：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：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | 测试集（310） | |
| AP | AP50 | AP | AP50 | AP | AP50 |
| 1万次 | 1.81🡪0.02 | 0.5916 | 0.9440 | 0.4267 | 0.7415 | 0.0902 | 0.2144 |
| 2万次 | 1.81🡪0.01 | 0.6040 | 0.9484 | 0.4281 | 0.7460 | 0.0886 | 0.1986 |
| 3万次 | 1.81🡪0.003 | 0.6024 | 0.9464 | 0.4216 | 0.7407 | 0.0802 | 0.1877 |
| 4.5万次 | 1.81🡪0.003 | 0.6061 | 0.9439 | 0.4236 | 0.7413 | 0.0807 | 0.1853 |

## 实验对比

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **预训练模型** | **其它** | **图片大小** | **训练集 (756)** | | **验证集 (188)** | | **benchmark (310)** | | **耗时**  **（ms）** | **备注** |
| **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **R-101.pkl** | **p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**、11类** | （576,1024） | 0.943 | 0.998 | 0.947 | 0.999 | 0.0341 | 0.1067 | 249 | [实验1](#_实验1:_(576,1024)、p2-p7、O4、S4、R6)  3万次  ys  trainval |
| **model\_final.pkl** | **p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**、11类** | （768,896） | 0.952 | 0.998 | 0.955 | 0.999 | 0.0557 | 0.1495 | 228 | [实验2](#_实验2:_(768,896)、p2-p7、O4、S4、R6)  3万次  ys  trainval |
| **p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**、11类** | （900,1408） | 0.943 | 0.9997 | 0.9326 | 0.9999 | 0.0964 | 0.2047 | 340 | [实验3](#_实验3:_(900,1408)、p2-p7、O4、S4、R6)  ys  trainval |
| **p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**、11类** | （900,1408） | 0.919 | 0.9998 | 0.4851 | 0.7933 | 0.1002 | 0.2165 | 340 | [实验9](#_实验9:_(900,1408)、_p2-p7、O4、S4、R6)  6万次  zj  train |
| **model\_final\_X.pkl** | **p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**、11类** | （900,1280） | 0.965 | 0.9997 | 0.9599 | 1.0000 | 0.0897 | 0.2175 | 240 | [实验4](#_实验4:(900,1280)、p2-p7、O4、S4、R6)  5万次  ys  trainval |
| **p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**、11类** | （900,1280） | 0.592 | 0.9440 | 0.4267 | 0.7415 | 0.0902 | 0.2144 | 240 | [实验11](#_实验11:_(900,1280)、p2-p7、O4、S4、R6)  ys  train |

### 结论：

预训练模型model\_final.pkl效果最好。数据集小对model\_final\_X.pkl的影响比较大，训练集AP一直上不去。

**备注**：

1）configs/12\_2017\_baselines/ retinanet\_R-101-FPN\_1x.yaml中的TRAIN: DATASETS:

('coco\_2007\_trainval', 'coco\_2007\_val')训练时，该参数提供的所有数据集都参与训练，可以只写一个数据集。

2）实验1、2、3、4：训练集、验证集都参与训练；

实验9、11：只训练集参与训练。

3）作者默认的OCTAVE: 3、SCALE: 4、ASPECT\_RATIOS: (1.0, 2.0, 0.5)、p3~p7。

4) 图片大小（1080,1920）、（900,1408），一个gpu的batchsize=1。

5) benchmark（310）是20号球机的数据（310张、1920\*1080），benchmark（349）是crop

过的数据（含有网络图片、图片长宽比基本保持1:1）。

# ASPECT\_RATIOS不同

## ASPECT\_RATIOS: (1.0, 2.0, 0.5)

### 实验5: model\_final.pkl、p3-p7、（756,1280）、O3、S4、trainval– lg

**存放位置（213）：**/data/ligang/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**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

3，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为3，

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0, 0.5),3个ratios；

**图片大小**：配置为756，1280，2个GPU，batchsize总共为4。

下面是训练和测试的结果：（retinanet\_R-101-FPN\_2x\_illbuild.yaml）

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | | benchmark (349) | |
| AP | AP50 | AP | AP50 | Time | AP | AP50 |
| 1万次 |  | 0.8274 | 0.9739 | 0.8347 | 0.9742 |  | 0.1893 | 0.3781 |
| 2万次 |  | 0.8891 | 0.9684 | 0.8898 | 0.9652 |  | 0.1820 | 0.3654 |
| 3万次 |  | 0.9219 | 0.9709 | 0.9204 | 0.9701 |  | 0.1810 | 0.3648 |
| 4万次 |  | 0.9255 | 0.9602 | 0.9264 | 0.9548 | 236ms | 0.1776 | 0.3584 |
| 5万次 |  | 0.9370 | 0.9679 | 0.9418 | 0.9639 | 236ms | 0.1804 | 0.3611 |
| 6万次 |  | 0.9398 | 0.9686 | 0.9347 | 0.9654 | 233ms | 0.1784 | 0.3593 |
| 7万次 |  | 0.9382 | 0.97 | 0.9423 | 0.9660 | 241ms | 0.1818 | 0.3666 |
| 8万次 |  | 0.9453 | 0.9719 | 0.9486 | 0.9699 | 229ms | 0.1791 | 0.3647 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 10000 | 0.8347 | 0.9742 | 0.9510 | 0.1360 | 0.8505 | 0.8668 |
| 20000 | 0.8898 | 0.9652 | 0.9551 | 0.126 | 0.9163 | 0.9106 |
| 30000 | 0.9204 | 0.9701 | 0.965 | 0.1389 | 0.9497 | 0.9303 |
| 40000 | 0.9264 | 0.9548 | 0.9521 | 0.1111 | 0.9558 | 0.9326 |
| 50000 | 0.9418 | 0.9639 | 0.9593 | 0.1205 | 0.9668 | 0.9508 |
| 60000 | 0.9347 | 0.9654 | 0.9616 | 0.0916 | 0.9685 | 0.9403 |
| 70000 | 0.9423 | 0.9660 | 0.9641 | 0.0916 | 0.9685 | 0.9484 |
| 80000 | 0.9486 | 0.9699 | 0.9679 | 0.091 | 0.9724 | 0.9566 |

## ASPECT\_RATIOS: (1.0, 2.0, 3.0, 5.0, 0.2, 0.5)

### 实验6: model\_final.pkl、p3-p7、（756,1344）、O3、S4、train

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

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**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

3，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 3，

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

**图片大小**：配置为756，1344，2个GPU，batchsize总共为4。

下面是训练和测试的结果：（retinanet\_R-101-FPN\_2x\_illbuild.yaml）

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | |
| AP | AP50 | AP | AP50 | Time |
| 1万次 |  | 0.8584 | 0.9927 | 0.4752 | 0.7940 | 230 |
| 2万次 |  | 0.9352 | 0.9934 | 0.4638 | 0.7852 | 228 |
| 3万次 |  | 0.9521 | 0.9943 | 0.4568 | 0.7798 | 233 |
| 4万次 |  | 0.9623 | 0.9943 | 0.4543 | 0.7808 | 229 |
| 5万次 |  | 0.9696 | 0.9942 | 0.4572 | 0.7822 | 213 |
| 6万次 |  | 0.9719 | 0.9942 | 0.4581 | 0.7817 | 210 |
| 7万次 |  | 0.9743 | 0.9944 | 0.4557 | 0.7809 | 203 |
| 8万次 |  | 0.9743 | 0.9944 | 0.4574 | 0.7809 | 204 |
| 9万次 |  | 0.9746 | 0.9944 | 0.455 | 0.7808 | 202 |

**备注**：该实验只有训练集参与训练。

## 实验对比

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ASPECT\_RATIOS** | **其它** | **图片大小** | **训练集 (756)** | | **验证集 (188)** | | **benchmark (349)** | | **耗时**  **（ms）** | **备注** |
| **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **R3**（1.0、2.0、0.5）  默认 | model\_final.pkl、p3-p7、O3、S4、11类 | **（756,1280）** | 0.938 | 0.97 | 0.942 | 0.966 | 0.182 | 0.367 | 236 | [实验5](#_实验5:_model_final.pkl、p3-p7、（756,128)  7万次  lg  trainval |
| **R6**（1.0、2.0、3.0、5.0、0.2、0.5） | model\_final.pkl、p3-p7、O3、S4、、11类 | **（756,1344）** | 0.858 | 0.993 | 0.4752 | 0.794 | / | / | 230 | [实验6](#_实验6:_model_final.pkl、p3-p7、（756,134)  1万次  zj  train |

### 结论：

由于不小心将实验6的训练模型删除了，所以没法补充benchmark的测试结果，也就不好给出对比结果。

下面表格是之前的违建实验，可以看出6个ratios的效果比3个ratios的好，AP能提高不到1个点。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验 | 实验细节  coco的模型：P3~P7, 3个ratios | Loss | 训练集  (3406) | 验证集  (850) | 测试集  (349) |
| 1 | 预训练模型：coco训练的模型  P3~P7, 3个ratios | 1.41🡪0.02 | 0.9827 | 0.9878 | 0.7707 |
| 2 | 预训练模型：coco训练的模型  P3~P7, 6个ratios | 1.72🡪0.01 | 0.9942 | 0.9963 | 0.7716 |
| 6 | 预训练模型：R-101.pkl  P2~P7, 3个ratios | 1.63🡪0.01 | 0.9775 | 0.9855 | 0.7424 |
| 7 | 预训练模型：R-101.pkl  P2~P7, 6个ratios | 1.69🡪0.04 | 0.9926 | 0.9968 | 0.7504 |

**备注**：

1）configs/12\_2017\_baselines/ retinanet\_R-101-FPN\_1x.yaml中的TRAIN: DATASETS:

('coco\_2007\_trainval', 'coco\_2007\_val')训练时，该参数提供的所有数据集都参与训练，可以只写一个数据集。

2）实验5：训练集、验证集都参与训练；

实验6：只训练集参与训练。

3) benchmark（310）是20号球机的数据（310张、1920\*1080），benchmark（349）是crop过的数据（含有网络图片、图片长宽比基本保持1:1）。

# 输入图片大小不同

## (1080,1920)

### 实验8:model\_final.pkl、p3-p7、O3、S4、R3、trainval – lg

**存放位置（213）：**/data/ligang/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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**：G

**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

3，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 3，

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0, 0.5),3个ratios；

**图片大小**：配置为1080，1920，2个GPU，batchsize总共为2。

下面是训练和测试的结果：（retinanet\_R-101-FPN\_2x\_illbuild.yaml）

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | | benchmark (349) | |
| AP | AP50 | AP | AP50 | Time | AP | AP50 |
| 1万次 | 0.04406 |  |  | 0.7826 | 0.9841 | 343 | 0.1822 | 0.3568 |
| 2万次 | 0.00865 |  |  | 0.8768 | 0.9831 | 343 | 0.1689 | 0.3428 |
| 3万次 | 0.00426 |  |  | 0.9171 | 0.9805 | 352 | 0.1691 | 0.3310 |
| 4万次 | 0.00300 |  |  | 0.9339 | 0.9797 | 345 | 0.1707 | 0.3403 |
| 5万次 | 0.00283 |  |  | 0.9440 | 0.9794 | 350 | 0.1680 | 0.3291 |
| 6万次 | 0.00188 |  |  | 0.9469 | 0.9813 | 365 | 0.1675 | 0.3316 |
| 7万次 | 0.00120 |  |  | 0.9509 | 0.9786 | 349 | 0.1665 | 0.3346 |
| 8万次 | 0.00185 |  |  | 0.9510 | 0.9756 | 347 | 0.1679 | 0.3319 |

验证集

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 10000 | 0.7826 | 0.9841 | 0.9332 | 0.4210 | 0.8011 | 0.8084 |
| 20000 | 0.8768 | 0.9831 | 0.9770 | 0.5762 | 0.8890 | 0.8874 |
| 30000 | 0.9171 | 0.9805 | 0.9787 | 0.6551 | 0.9269 | 0.9227 |
| 40000 | 0.9339 | 0.9797 | 0.9621 | 0.7755 | 0.9452 | 0.9340 |
| 50000 | 0.9440 | 0.9794 | 0.9652 | 0.7879 | 0.9508 | 0.9490 |
| 60000 | 0.9469 | 0.9813 | 0.9804 | 0.8299 | 0.9607 | 0.9417 |
| 70000 | 0.9509 | 0.9786 | 0.9778 | 0.8224 | 0.9644 | 0.9453 |
| 80000 | 0.9510 | 0.9756 | 0.9622 | 0.8567 | 0.9599 | 0.9521 |

### 实验10: model\_final.pkl、p3-p7、O3、S4、R3、train – lg

**存放位置（213）：**/data/ligang/Detectron

**容器：**detectron\_gpu

**镜像：**/home/tar/detectron\_gpu.tar

**训练图片：**756训练集，验证集188， (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

3，COARSEST\_STRIDE为128； RETINANET中 SCALES\_PER\_OCTAVE为 3

ANCHOR\_SCALE为 4，ASPECT\_RATIOS为 (1.0, 2.0,,0.5),3个ratio;

**图片大小**：配置为1080，1920，2个GPU，batchsize总共为2。

下面是训练和测试的结果：（retinanet\_R-101-FPN\_2x\_illbuild.yaml）

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (756) | | 验证集 (188) | | | benchmark (310) | |
| AP | AP50 | AP | AP50 | Time | AP | AP50 |
| 1万次 | 0.04406 | 0.8362 | 0.9836 | 0.4874 | 0.7926 | 360 | 0.094 | 0.1855 |
| 2万次 | 0.00865 | 0.9069 | 0.9853 | 0.4866 | 0.7941 |  | 0.0902 | 0.1924 |
| 3万次 | 0.00426 |  |  | 0.4837 | 0.7870 |  | 0.0867 | 0.1837 |
| 4万次 | 0.00300 |  |  | 0.4818 | 0.7883 |  |  |  |
| 5万次 | 0.00283 | 0.9599 | 0.9832 | 0.4817 | 0.7868 |  | 0.0808 | 0.1754 |
| 6万次 | 0.00188 |  |  | 0.4746 | 0.7865 |  |  |  |
| 7万次 | 0.00120 |  |  | 0.4747 | 0.7859 |  |  |  |
| 8万次 | 0.00185 | 0.9652 | 0.9823 | 0.4773 | 0.7866 |  | 0.0793 | 0.1740 |

验证集

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 10000 | 0.4874 | 0.7926 | 0.5106 | 0.2272 | 0.4515 | 0.5316 |
| 20000 | 0.4866 | 0.7941 | 0.5335 | 0.2171 | 0.4467 | 0.5388 |
| 30000 | 0.4837 | 0.7870 | 0.5123 | 0.1953 | 0.4427 | 0.5338 |
| 40000 | 0.4818 | 0.7883 | 0.5045 | 0.2203 | 0.4421 | 0.5291 |
| 50000 | 0.4817 | 0.7868 | 0.5053 | 0.2144 | 0.4407 | 0.5324 |
| 60000 | 0.4746 | 0.7865 | 0.5023 | 0.2140 | 0.4351 | 0.5278 |
| 70000 | 0.4747 | 0.7859 | 0.4905 | 0.2265 | 0.4363 | 0.5256 |
| 80000 | 0.4773 | 0.7866 | 0.5015 | 0.2140 | 0.4383 | 0.5241 |

## 实验对比

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **图片大小** | **其它** | **训练集 (756)** | | **验证集 (188)** | | **benchmark (310)** | | **备注** |
|  | **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **（768,896）** | model\_final.pkl、p2-p7、O4、S4、R6（1.0、2.0、3.0、5.0、0.2、0.5）、11类 | 0.952 | 0.998 | 0.955 | 0.999 | 0.0557 | 0.1495 | [实验2](#_实验2:_model_final.pkl、p2-p7、（768,896)  3万次  ys  trainval |
| **（900,1408）** | model\_final.pkl、p2-p7、O4、S4、R6（1.0、2.0、3.0、5.0、0.2、0.5）、11类 | 0.943 | 0.9997 | 0.9326 | 0.9999 | 0.0964 | 0.2047 | [实验3](#_实验3:_model_final.pkl、p2-p7、（900,140)  ys  trainval |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **图片大小** | **其它** | **训练集 (756)** | | **验证集 (188)** | | **benchmark (349)** | | **备注** |
|  | **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **（756,1280）** | model\_final.pkl、p3-p7、O3、S4、R3（1.0、2.0、0.5）、11类 | 0.938 | 0.97 | 0.942 | 0.966 | 0.182 | 0.367 | [实验5](#_实验5:_model_final.pkl、p3-p7、（756,128)  7万次  lg  trainval |
| **（1080,1920）** | model\_final.pkl、p3-p7、O3、S4、R3（1.0、2.0、0.5）、11类 | 0.796 | 0.9854 | 0.7826 | 0.9841 | 0.1822 | 0.3568 | [实验8](#_实验8:_model_final.pkl、p3-p7、（1080,19)  1万次  lg  trainval |

### 结论：

从上面表格可以看出随着输入图片的增大，AP有所提高。

**备注**：

1）configs/12\_2017\_baselines/ retinanet\_R-101-FPN\_1x.yaml中的TRAIN: DATASETS:

('coco\_2007\_trainval', 'coco\_2007\_val')训练时，该参数提供的所有数据集都参与训练，可以只写一个数据集。

2）实验2、3、5、8：训练集、验证集都参与训练；

3）作者默认的OCTAVE: 3、SCALE: 4、ASPECT\_RATIOS: (1.0, 2.0, 0.5)、p3~p7。

4) 图片大小（1080,1920）、（900,1408），一个gpu 的batchsize=1；

图片大小（768,896）、（756,1280），一个gpu 的batchsize=2。

5) benchmark（310）是20号球机的数据（310张、1920\*1080），benchmark（349）是crop

过的数据（含有网络图片、图片长宽比基本保持1:1）。

# 分类回归不同：有无P2层

## 结论

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **实验** | **训练集 (756)** | | **验证集 (188)** | | **benchmark (310)** | | **备注** |
| **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **model\_final.pkl、p3-p7、O3、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**（756,1344）、11类** | 0.858 | 0.993 | 0.4752 | 0.794 | / | / | [实验6](#_实验6:_model_final.pkl、p3-p7、（756,134)  1万次  zj  train |
| **model\_final.pkl、p2-p7、O4、S4、R6**（1.0、2.0、3.0、5.0、0.2、0.5）**（900,1408）、11类** | 0.919 | 0.9998 | 0.4851 | 0.7933 | 0.1002 | 0.2165 | [实验9](#_实验9:_model_final.pkl、p2-p7、（900,140)  6万次  zj  train |
| **model\_final.pkl、p3-p7、O3、S4、R3**（1.0、2.0、0.5）**（1080,1920）、11类** | 0.907 | 0.9853 | 0.4866 | 0.7941 | 0.0902 | 0.1924 | [实验10](#_实验10:_model_final.pkl、p3-p7、（1080,1)  lg  train |

**备注**：

1）configs/12\_2017\_baselines/ retinanet\_R-101-FPN\_1x.yaml中的TRAIN: DATASETS:

('coco\_2007\_trainval', 'coco\_2007\_val')训练时，该参数提供的所有数据集都参与训练，可以只写一个数据集。

2）实验6、9、10：只训练集参与训练，所以可以比较验证集指标。（实验6不小心将训练

模型删除了，所以没法补充benchmark的测试）

3) benchmark（310）是20号球机的数据（310张、1920\*1080），benchmark（349）是crop过的数据（含有网络图片、图片长宽比基本保持1:1）。

4）图片大小（1080,1920）、（900,1408），一个gpu 的batchsize=1；

图片大小（768,896）、（756,1280），一个gpu 的batchsize=2。

本次违建实验没有做有无P2层分类回归的实验对比，下面表格是之前的违建实验。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验 | 实验细节 (768,896)  coco的模型：P3~P7, 3个ratios | Loss | 训练集  (3406) | 验证集  (850) | 测试集  (349) |
| 3 | 预训练模型：coco训练的模型  P3~P7, 6个ratios，O4,S4 | 1.72🡪0.01 | 0.9942 | 0.9963 | 0.7716 |
| 2 | 预训练模型：coco训练的模型  P2~P7, 6个ratios, O4,S4  不加载FPN预测层的所有卷积层 | 1.69🡪0.02 | 0.9911 | 0.9969 | 0.7587 |
| 5 | 预训练模型：coco训练的模型  P2~P7, 6个ratios, O3,S4  加载FPN预测层的所有卷积层 | 1.61🡪0.03 | 0.9884 | 0.9967 | 0.7679 |
| 6 | 预训练模型：R-101.pkl  P3~P7, 3个ratios, O3,S4 | 1.61🡪0.02 | 0.9778 | 0.9843 | 0.7204 |
| 7 | 预训练模型：R-101.pkl  P2~P7, 3个ratios, O3,S4 | 1.63🡪0.01 | 0.9775 | 0.9855 | **0.7424** |

实验5中P2层没有预训练权重，FPN的分类回归层（P2~P7）的权重是加载了预训练模型的分类回归层(P3~P7)的权重。

实验2和实验5对比，coco预训练的模型所有能加载的权重都加载（实验5）要比不加载FPN的分类回归层权重（实验3）的指标高一些。

实验3和实验5对比，可以看出，如果P2层和后面的分类回归层不能提供好的预训练模型的权重，反而不如使用和作者预训练模型一样网络结构的效果。

实验6和实验7对比，可以看出，预训练模型为R-101.pkl，有P2层效果会好些。

# SCALES\_PER\_OCTAVE不同

## 结论

本次违建实验没有做SCALES\_PER\_OCTAVE不同的实验对比，下面表格是之前的违建实验，可以看出SCALES\_PER\_OCTAVE=4会比SCALES\_PER\_OCTAVE=3效果好些。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验 | 实验细节 (768,896)  coco的模型：P3~P7, 3个ratios | Loss | 训练集  (3406) | 验证集  (850) | 测试集  (349) |
| 3 | 预训练模型：coco训练的模型  P3~P7, 6个ratios，O4,S4 | 1.72🡪0.01 | 0.9942 | 0.9963 | 0.7716 |
| 4 | 预训练模型：coco训练的模型  P3~P7, 6个ratios, O3,S4 | 1.41🡪0.02 | 0.9827 | 0.9878 | 0.7707 |

# ANCHOR\_SCALE不同

## 实验16: S3、model\_final.pkl、p2-p7、（900,1408）、O4、R6

**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 3，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 | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 8000 | 0.7248 | 0.9859 | 0.5032 | 0.8433 |
| 1万 | 0.7563 | 0.9930 | 0.5009 | 0.8438 |
| 2万 | 0.8317 | 0.9979 | 0.5006 | 0.8304 |
| 3万 | 0.8981 | 0.9985 | 0.5087 | 0.8404 |
| 4万 | 0.9387 | 0.9988 | 0.5092 | 0.8418 |
| 5万 | 0.9536 | 0.9988 | 0.4998 | 0.8424 |

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 8000 | 0.5032 | 0.8433 | 0.5319 | 0.2873 | 0.4284 | 0.5567 |
| 1万 | 0.5009 | 0.8438 | 0.5178 | 0.3121 | 0.4206 | 0.5666 |
| 2万 | 0.5006 | 0.8304 | 0.5149 | 0.2221 | 0.4237 | 0.5756 |
| 3万 | 0.5087 | 0.8404 | 0.5233 | 0.2780 | 0.4204 | 0.5810 |
| 4万 | 0.5092 | 0.8418 | 0.5402 | 0.2709 | 0.4271 | 0.5760 |
| 5万 | 0.4998 | 0.8424 | 0.5152 | 0.2687 | 0.4169 | 0.5553 |

## 实验12: S4、model\_final.pkl、p2-p7、（900,1408）、O4、R6

**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 | 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17: S6、model\_final.pkl、p2-p7、（900,1408）、O4、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+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\_2x\_illbuild\_P2\_P7\_S4\_anchors6\_ANCHOR\_SCALE6.yam

**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 6，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 | 0.08056 | 0.7377 | 0.9913 | 0.4943 | 0.8535 |
| 2万次 | 0.02634 | 0.8388 | 0.9945 | 0.5083 | 0.8401 |
| 3万次 | 0.01451 | 0.8880 | 0.9942 | 0.5064 | 0.8379 |
| 4万次 | 0.00635 | 0.9206 | 0.9963 | 0.5072 | 0.8378 |
| 5万次 | 0.00401 | 0.9532 | 0.9965 | 0.5029 | 0.8383 |
| 6万次 | 0.00428 | 0.9572 | 0.9974 | 0.5001 | 0.8487 |

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 20000 | **0.5083** | **0.8401** | **0.5494** | **0.2130** | **0.4266** | **0.571** |

## 实验18: S2、model\_final.pkl、p2-p7、（900,1408）、O4、R6

**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)；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/illbuild/ retinanet\_R-101-FPN\_p7\_2\_r6\_L\_ft.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 2，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。

**训练模型的存放位置**：retinanet\_R-101-FPN\_p7\_2\_r6\_L\_ft\_a2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 5千次 | 0.11967 |  |  | 0.4658 | 0.8188 |
| 1万次 | 0.07246 |  |  | 0.4707 | 0.8155 |
| 2万次 | 0.02094 | 0.8449 | 0.9971 | 0.4835 | 0.8211 |
| 3万次 | 0.00955 | 0.8707 | 0.9975 | 0.4882 | 0.8304 |
| 3.5万次 | 0.00756 | 0.8877 | 0.9977 | 0.4912 | 0.8264 |
| 4万次 | 0.00712 | 0.8907 | 0.9977 | 0.4873 | 0.8216 |

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 3.5万次 | 0.4912 | 0.8264 | 0.5137 | 0.2678 | 0.4110 | 0.5429 |

## 实验对比

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ANCHOR\_SCALE** | **其它** | **训练集 (1004)** | | **验证集 (250)** | | **benchmark (310)** | | **耗时**  **（ms）** | **备注** |
| **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **S2** | model\_final.pkl、p2-p7、O4、R6、（900,1408）、11类 | 0.8877 | 0.9977 | 0.4912 | 0.8264 | / | / |  | [实验18](#_实验18:_S2、model_final.pkl、p2-p7、（900)  4万次  zj  train |
| **S3** | model\_final.pkl、p2-p7、O4、R6、（900,1408）、11类 | 0.9387 | 0.9988 | 0.5092 | 0.8418 | / | / | 230 | [实验16](#_实验16:_S3、model_final.pkl、p2-p7、（900)  4万次  zj  train |
| **S4** 默认 | model\_final.pkl、p2-p7、O4、R6、（900,1408）、11类 | 0.8642 | 0.9985 | 0.5106 | 0.8438 | / | / |  | [实验12](#_实验12:_S4、model_final.pkl、p2-p7、（900)  2万次  zj  train |
| **S6** | model\_final.pkl、p2-p7、O4、R6、（900,1408）、11类 | 0.8388 | 0.9945 | 0.5083 | 0.8401 | / | / |  | [实验17](#_实验17:_S6、model_final.pkl、p2-p7、（900)  6万次  lg  train |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | **ANCHOR\_SCALE** | AP | AP50 | AP75 | APs | APm | APl |
| 3.5万次 | S2 | 0.4912 | 0.8264 | 0.5137 | 0.2678 | 0.4110 | 0.5429 |
| 4万 | S3 | 0.5092 | 0.8418 | 0.5402 | 0.2709 | 0.4271 | 0.5760 |
| 2万 | **S4** | 0.5106 | 0.8438 | 0.5409 | 0.3543 | 0.4246 | 0.5869 |
| 2万 | S6 | 0.5083 | 0.8401 | 0.5494 | 0.2130 | 0.4266 | 0.571 |

### 结论：

从上面表格可以S4的效果最佳。S3,S6的效果和S4相差不多，S2的效果较差。

**备注：**

1）作者默认的OCTAVE: 3、SCALE: 4、ASPECT\_RATIOS: (1.0, 2.0, 0.5)、p3~p7。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SCALES\_PER\_OCTAVE | ANCHOR\_SCALE | 每一层对应的 1:1的anchors的宽高对应特征图上的像素个数分别为：(anchor面积，向上取整) | | | |
| 4 | 2 | 2 | 2\*1.19=3 | 2\*1.414=3 | 2\*1.68=4 |
| 4 | 3 | 3 | 3\*1.19=4 | 3\*1.414=5 | 3\*1.68=6 |
| 4 | 4 | 4 | 4\*1.19=5 | 4\*1.414=6 | 4\*1.68=7 |
| 4 | 6 | 6 | 6\*1.19=8 | 6\*1.414=9 | 6\*1.68=11 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SCALES\_PER\_OCTAVE | ANCHOR\_SCALE | 每一层对应的 1:1的anchors的宽高对应特征图上的像素个数分别为：(anchor面积，向上取整) | | |
| 3 | 2 | 2 | 2\*1.26=3 | 2\*1.59=4 |
| 3 | 3 | 3 | 3\*1.26=4 | 3\*1.59=5 |
| 3 | 4 | 4 | 4\*1.26=6 | 4\*1.59=7 |
| 3 | 6 | 6 | 6\*1.26=8 | 6\*1.59=10 |

# NUM\_CONVS不同

## 实验22: NUM\_CONVS3、model\_final.pkl、p2-p7、（900,1408）、S4、O4、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; NUM\_CONVS为3

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 迭代次数 | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 1万 | 0.7616 | 0.9901 | 0.4992 | 0.8443 |
| 2万 | 0.8647 | 0.9986 | 0.5083 | 0.8392 |
| 3万 | 0.8936 | 0.9986 | 0.5026 | 0.8349 |
| 4万 | 0.9278 | 0.9988 | 0.5039 | 0.8337 |
| 5万 | 0.9548 | 0.9988 | 0.5008 | 0.8308 |
| 10万 | 0.9747 | 0.9988 | 0.4972 | 0.8318 |

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 1万 | 0.4992 | 0.8443 | 0.5244 | 0.2150 | 0.4220 | 0.5554 |
| 2万 | 0.5083 | 0.8392 | 0.5512 | 0.2531 | 0.4277 | 0.5735 |
| 3万 | 0.5026 | 0.8349 | 0.5371 | 0.2446 | 0.4194 | 0.5693 |
| 4万 | 0.5039 | 0.8337 | 0.5302 | 0.2557 | 0.4183 | 0.5745 |
| 5万 | 0.5008 | 0.8308 | 0.5132 | 0.2268 | 0.4177 | 0.5761 |
| 10万 | 0.4972 | 0.8318 | 0.5222 | 0.2751 | 0.4165 | 0.5705 |

## 实验12: NUM\_CONVS4、model\_final.pkl、p2-p7、（900,1408）、S4、O4、R6

**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; NUM\_CONVS为4

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 迭代次数 | 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23: NUM\_CONVS5、model\_final.pkl、p2-p7、（900,1408）、S4、O4、R6-lg

**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; NUM\_CONVS为5

**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 | Time |
| 1万次 | 0.0806 | 0.7458 | 0.9846 | 0.5124 | 0.85 |  |
| 2万次 | 0.0234 | 0.8619 | 0.9965 | 0.5073 | 0.8291 |  |
| 3万次 | 0.0123 | 0.9047 | 0.9986 | 0.5046 | 0.8279 |  |
| 4万次 | 0.0084 | 0.9161 | 0.9986 | 0.5152 | 0.8472 |  |
| 5万次 | 0.0036 | 0.9514 | 0.9987 | 0.5095 | 0.8401 |  |
| 6万次 | 0.0033 | 0.9635 | 0.9988 | 0.5087 | 0.8373 |  |
| 7万次 | 0.0028 | 0.9655 | 0.9988 | 0.5076 | 0.8335 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **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 | 90.7 | 57.2 |
| **2** | crane | 11 | 0 | 95.4 | Nan |
| **3** | forklift | 1051 | 239 | 90.9 | 60.6 |
| **4** | mixerTruck | 40 | 10 | 91.3 | 46.3 |
| **5** | person | 354 | 87 | 91.1 | 31.6 |
| **6** | colorPlate | 1124 | 288 | 90.5 | 54.2 |
| **7** | pit | 153 | 39 | 91.4 | 71.5 |
| **8** | bricksPile | 516 | 103 | 91.6 | 42.8 |
| **9** | mound | 657 | 160 | 89.9 | 63.2 |
| **10** | car | 446 | 114 | 93.4 | 36.3 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 40000 | 0.5152 | 0.8472 | 0.5481 | 0.2393 | 0.4308 | 0.5913 |

## 实验24: NUM\_CONVS6、model\_final.pkl、p2-p7、（900,1408）、S4、O4、R6-ys

**存放位置（214）：**/home/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)；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/illbuild/retinanet\_R-101-FPN\_p7\_2\_r6\_L\_ft.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; NUM\_CONVS为6

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

**模型保存位置：**retinanet\_R-101-FPN\_p7\_2\_r6\_L\_ft\_num6

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 迭代次数 | Loss | 训练集 (1004) | | 验证集 (250) | |
| AP | AP50 | AP | AP50 |
| 1万次 | 0.09 | 0.7400, | 0.9912 | 0.4955 | 0.8373 |
| 2万次 | 0.04 | 0.8591 | 0.9984 | 0.5085 | 0.8394 |
| 3万次 | 0.008 | 0.8947 | 0.9984 | 0.5084 | 0.8424 |
| 4万次 | 0.008 | 0.9266 | 0.9987 | 0.5044 | 0.8382 |
| 5万次 | 0.006 | 0.9340 | 0.9987 | 0.5039 | 0.8390 |
| 8万次 | 0.005 | 0.9352 | 0.9987 | 0.5030 | 0.8374 |
| 12万次 | 0.004 | 0.9356 | 0.9988 | 0.5027 | 0.8375 |

**验证集**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | AP | AP50 | AP75 | APs | APm | APl |
| 3万 | 0.5084 | 0.8424 | 0.5422 | 0.2317 | 0.4255 | 0.5761 |

## 实验对比

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **NUM\_CONVS** | **其它** | **训练集 (1004)** | | **验证集 (250)** | | **benchmark (310)** | | **耗时**  **（ms）** | **备注** |
| **AP** | **AP50** | **AP** | **AP50** | **AP** | **AP50** |
| **3** | model\_final.pkl、p2-p7、S4、O4、R6、（900,1408）、11类 | 0.8647 | 0.9986 | 0.5083 | 0.8392 | / | / |  | [实验22](#_实验22:_NUM_CONVS3、model_final.pkl、p2)  2万次  zj  train |
| **4(默认)** | model\_final.pkl、p2-p7、S4、O4、R6、（900,1408）、11类 | 0.8642 | 0.9985 | 0.5106 | 0.8438 | / | / | 230 | [实验12](#_实验12:_S4、model_final.pkl、p2-p7、（900)  ？万次  zj  train |
| **5** | model\_final.pkl、p2-p7、S4、O4、R6、（900,1408）、11类 | 0.9161 | 0.9986 | 0.5152 | 0.8472 | / | / |  | [实验23](#_实验23:_NUM_CONVS5、model_final.pkl、p2)  4万次  lg  train |
| **6** | model\_final.pkl、p2-p7、S4、O4、R6、（900,1408）、11类 | 0.8947 | 0.9984 | 0.5084 | 0.8424 | / | / |  | [实验24](#_实验24:_NUM_CONVS6、model_final.pkl、p2)  ？万次  ys  train |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 迭代次数 | **NUM\_CONVS** | AP | AP50 | AP75 | APs | APm | APl |
| 2万 | 3 | 0.5083 | 0.8392 | 0.5512 | 0.2531 | 0.4277 | 0.5735 |
| 2万 | **4（默认）** | 0.5106 | 0.8438 | 0.5409 | 0.3543 | 0.4246 | 0.5869 |
| 4万 | 5 | 0.5152 | 0.8472 | 0.5481 | 0.2393 | 0.4308 | 0.5913 |
| 3万 | 6 | 0.5084 | 0.8424 | 0.5422 | 0.2317 | 0.4255 | 0.5761 |

### 结论：