1. **使用方法**

本代码基于resnet50实现垃圾图片分类，在src/main.py的main\_worker函数中加载了ImageNet预训练参数文件，您可以根据自己的需要来选择是否加载预训练参数文件，也可以修改为使用其他模型。

本代码适合在华为云ModelArts平台的notebook和训练作业中运行，如需在您自己的机器上运行，需自行修改代码。

* 1. **notebook中的运行方法**

请查看src/main.py中文件开头的注释。

* 1. **训练作业中的运行方法**

跑训练任务的参数设置如图1所示，注意分析图1中训练作业参数设置和notebook运行参数设置的差异。图1中的参数设定值，如表1所示。“训练作业”只是个名称而已，本质就是运行一套代码，是否要读数据、跑什么代码、是否存结果，都是用户自己决定的，因此用户也可以在训练作业中执行转pb或评价任务。

表1 训练作业参数设定值

|  |  |
| --- | --- |
| 参数名 | 参数值 |
| 数据存储位置（训练数据） | myobsbucket1118/train\_val/，注意OBS路径要选择到train\_val |
| AI引擎 | PyTorch、PyTorch-1.0.0-python3.6 |
| 代码目录（训练代码） | myobsbucket1118/src/ |
| 启动文件 | myobsbucket1118/src/main.py |
| 训练输出位置（模型保存路径） | myobsbucket1118/model\_snapshots/，自己创建一个目录即可 |
| train\_url（设置训练输出位置后会自动生成该参数的值，无需手动设定） | /myobsbucket1118/model\_snapshots/ |
| data\_url（设置数据存储位置后会自动生成该参数的值，无需手动设定） | /myobsbucket1118/train\_val/ |
| num\_classes | 54 |
| deploy\_script\_path | s3://myobsbucket1118/src/deploy\_scripts/ |
| arch | resnet50 |
| workers | 4 |
| epochs | 6 |
| pretrained | True |
| seed | 0 |
| 作业日志路径 | myobsbucket1118/logs/，自己创建一个目录即可 |
| 规格 | CPU: 8核 64GiB GPU : 1\*nvidia-p100 16GiB |

**注意：以上参数中所有路径都要改成自己用的路径。**

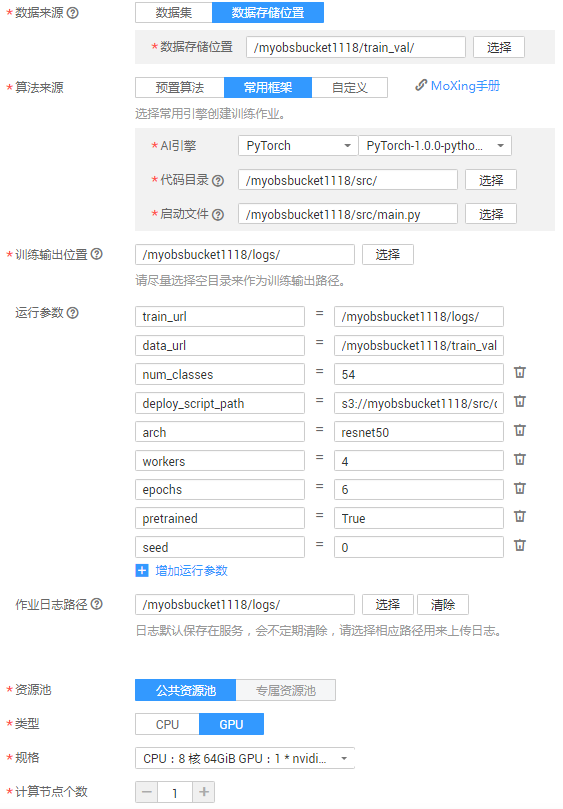


图1 图像分类baseline代码训练作业参数设置

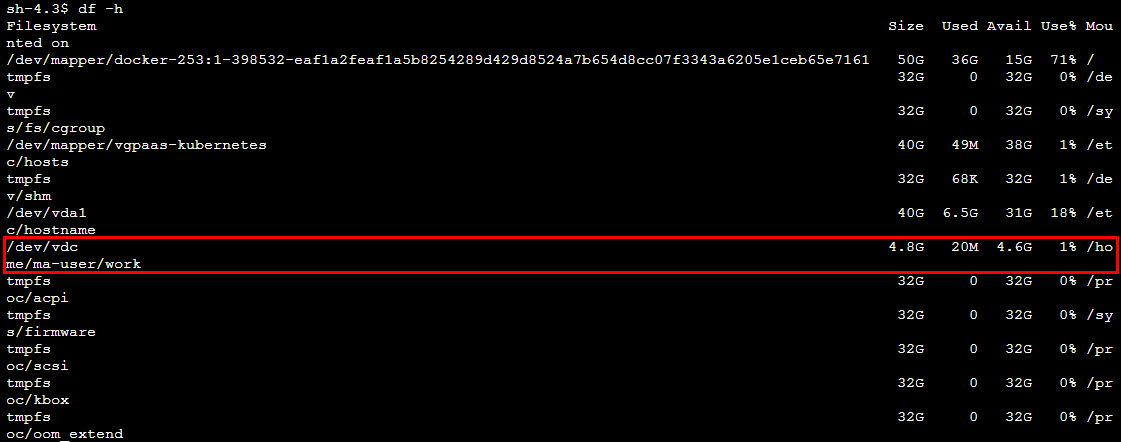
1. **baseline代码参数说明**

请查看src/main.py中参数说明。

1. **调试推理脚本customize\_service.py**

请查看src/inference.py中文件开头的注释。

1. **注意事项**
2. baseline代码的路径参数支持本地路径和OBS路径；
3. 如果某个参数要设置成OBS路径，则必须按照’s3://{桶名}/{绝对路径}’的格式填写，不支持路径中含有‘./’和’../’的相对路径；
4. 创建的notebook默认只有5G的EVS存储空间，很容易用完，请注意及时清理不需要的文件，使用’df –h’命令可查看存储空间的使用情况，如下图所示，空间总大小为4.8G、已用20M、可用4.6G。当然您也可以创建大于5G的EVS notebook，超过5G的部分会收费；



1. 可以分别创建一个CPU和一个GPU的notebook，不必须用到GPU的代码调试用CPU类型的notebook即可，收费更低。完成代码调试后，如需跑训练，推荐将notebook中调试好的代码传输到OBS，然后创建训练作业，从OBS加载训练数据和训练代码进行训练。**使用训练作业跑训练任务有如下几点好处**，这些优点都是notebook不具备的：
2. 训练作业有版本管理功能，会记录某次训练使用的训练数据、训练代码、运行参数、训练输出目录，还有训练日志、运行时长、资源占用情况等信息；
3. 训练作业完成后，可以创建TensorBoard查看训练情况曲线图，有助于分析下一步应该怎样调试模型；
4. 调试模型时，也许一次新的训练就只需要改一个参数，使用训练作业可以在某个版本的训练作业上点击“修改”，然后修改要改的参数即可，其他参数不用动，这样很方便地就创建了一个新的训练作业，而且训练作业有版本溯源功能，能可视化地表示在哪个版本基础上创建了新版本；
5. 训练作业运行完即自动停止，不会继续收费，而notebook必须手动停止或自动停止时间到了之后才停止；
6. 训练作业输出的模型会自动保存到OBS，如需导入模型并部署成RESTAPI服务，可以直接从指定的训练作业中加载模型，省去了一步步选择OBS路径的麻烦。