1. **使用方法**

本代码基于resnet50实现垃圾图片分类，在src/models/resnet50.py的ResNet50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，也可以在src/models/resnet50.py的model\_fn中修改为其他模型。

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/run.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\_data/，注意OBS路径要选择到train\_data |
| AI引擎 | TensorFlow、TF-1.13.1-python3.6 |
| 代码目录（训练代码） | myobsbucket1118/src/ |
| 启动文件 | myobsbucket1118/src/run.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\_data/ |
| num\_classes | 54 |
| deploy\_script\_path | s3://myobsbucket1118/src/deploy\_scripts/ |
| max\_epochs | 6 |
| 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**

当前baseline代码可以执行训练、转pb和评价任务，如表2所示是支持的参数设置。

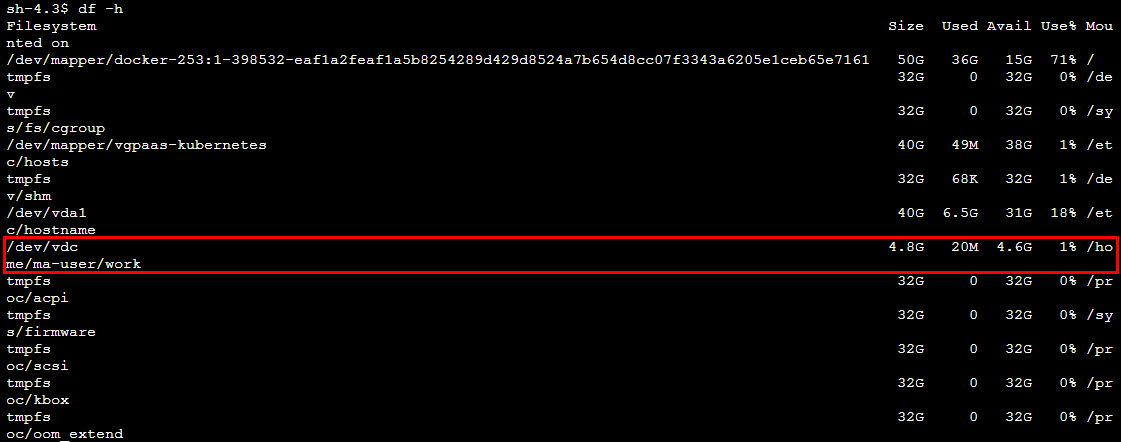
表2 baseline代码参数说明

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名 | 是否必选 | 默认值 | 参数用途 |
| mode | 是 | train | 运行模式，只能设为’train’、’save\_pb’和’eval’三个值，分别代表本次任务是训练、保存pb模型还是评价模型 |
| local\_data\_root | 否 | '/cache/' | 本地数据根目录。ModelArts上主要涉及两类数据存储路径，一类是OBS路径，另一类是docker容器内路径。创建一个notebook或训练作业，都是以docker的形式运行，其内部存储空间（下文统称为本地）是与OBS存储空间分开的，可以形象地理解为两个独立的硬盘。执行训练任务时，可以将训练数据从OBS一次性全部拷贝到本地，再从本地加载，以免每一张图都从OBS加载引起网络开销。其他使用频次较高的文件也可以先从OBS拷贝到本地再使用。该参数的默认值是’/cache/’，在训练作业中，该目录是多用户共享的4T存储空间，在notebook中该目录并不大，一般只有10~20G。/cache目录下的数据会在notebook关闭后或训练作业结束后全部被清空，如需永久存储本地数据，请使用存储配置为EVS的notebook，并设置该参数为/home/ma-user/work下的某个路径。/home/ma-user/work目录默认的存储空间为5G，如需更大的存储空间，请在创建notebook时调大磁盘规格，**一旦创建之后规格不可改变** |
| data\_url | 是 | 空字符串 | 训练数据路径，在notebook中执行训练任务时，必须手动指定该参数，在创建训练作业时，选择“数据来源”时就已经指定了该参数，不需要再手动指定 |
| restore\_model\_path | 否 | 空字符串 | 恢复训练模型文件路径，如果设置了该参数，将会加载该模型，在此基础上继续训练。与预训练模型的区别是，恢复训练模型是自己训练输出的历史模型 |
| train\_url | 是 | 空字符串 | 保存训练结果的目录路径，在notebook中执行训练任务时，必须手动指定该参数，在创建训练作业时，选择“训练输出位置”就已经指定了该参数，不需要再手动指定。如果该目录路径不存在，则会按照设置的参数值自动创建目录 |
| keep\_weights\_file\_num | 否 | 20 | 保存模型文件的最大数量，默认值为20，只保留最新的20个模型，如果设为-1，则没有限制。notebook EVS的存储容量有限，可以根据需要设置合理的keep\_weights\_file\_num值，以免导致训练过程因保存模型失败而被中断。该参数只对本地路径的模型文件数做限制，对OBS路径上存储的模型文件数不做限制。 |
| num\_classes | 是 | 0 | 分类任务需要区分的类别数，需要根据具体任务来设定，必选参数 |
| input\_size | 否 | 224 | 模型的输入图像大小，一般输入图像为正方形，因此只需设置一个整数值。本代码使用resnet50模型，因此输入图像大小默认为224 |
| batch\_size | 否 | 128 | 每一步训练使用的图像样本数，一般设为2的n次方值，如16、32、64、128、256等，由于计算机硬件的因素，设为非2的幂值可能会降低训练速度。 |
| learning\_rate | 否 | 1e-4 | 学习率 |
| max\_epochs | 否 | 6 | 整个训练集的训练迭代次数，默认值为6，值越大，训练时间越长，一些收敛较快的训练任务不必设置太大的max\_epochs值 |
| deploy\_script\_path | 否 | 空字符串 | 模型部署脚本目录路径，如果设置了该参数，模型部署脚本config.json和customize\_service.py将会拷贝到pb模型所在的目录。模型部署的含义是，可以在ModelArts管理控制台的模型管理页面中，导入一个pb模型就可以将其发布成一个RESTAPI服务。但是，要成功部署模型需要遵循一定的规范，详情请查看[模型包规范介绍](https://support.huaweicloud.com/engineers-modelarts/modelarts_23_0091.html)。如果mode=’train’且设置了该参数，则在完成模型训练后，当前最新的模型将会被保存为pb文件，并将模型部署脚本拷贝到pb文件所在目录。  如果mode=’save\_pb’，则该参数必选 |
| freeze\_weights\_file\_path | 否 | 空字符串 | 要转成pb的参数文件路径，如果mode=’save\_pb’，则该参数必选，且只有在mode=’save\_pb’的情况下生效。save\_pb模式将会生成一个含有pb文件的model目录，并保存到freeze\_weights\_file\_path所在的同级目录 |
| eval\_weights\_path | 否 | 空字符串 | 需评价的参数文件的路径。如果该参数是文件路径，则只评价该文件，如果目录路径，则评价该目录中的所有参数文件。由test\_data\_url参数指定评价数据集，评价结果会被保存为与参数文件相同前缀名的txt文件，且与参数文件在同级目录。如果mode=’eval’，则eval\_weights\_path和eval\_pb\_path参数必须二选一 |
| eval\_pb\_path | 否 | 空字符串 | 需评价的pb文件路径。由test\_data\_url参数指定评价数据集，评价结果会被保存为与pb文件相同前缀名的txt文件，且与pb文件在同级目录。如果mode=’eval’，则eval\_weights\_path和eval\_pb\_path参数必须二选一 |
| test\_data\_url | 否 | 空字符串 | 测试数据路径。如果mode=’train’且设置了该参数，在完成模型训练后，将会评估当前模型在测试数据上的准确率。如果mode=’eval’，则该参数必选 |
| data\_local | 否 | 空字符串 | 训练数据本地缓存路径，如果data\_url为OBS路径，则会将训练数据拷贝到本地，再从本地加载、训练 |
| train\_local | 否 | 空字符串 | 训练结果本地缓存路径，如果train\_url为OBS路径，则会将训练结果保存到本地，再同步到train\_url |
| test\_data\_local | 否 | 空字符串 | 测试数据本地缓存路径，如果test\_data\_url为OBS路径，则会将测试数据拷贝到本地，再从本地加载、测试 |
| tmp | 否 | 空字符串 | 临时数据本地缓存路径，其他需要与OBS进行数据交互的数据，利用tmp目录进行中转缓存 |

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. 以上所有可设置路径的参数都支持本地路径和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。