通过Modelarts训练作业

调用模型移植服务文档

# 文档说明

文档编排格式说明：

文档中黄底字体需要根据实际情况修改。

文档中灰色背景表示执行命令或代码段内容。

文档中红色字体表示醒目提示，需重点关注。

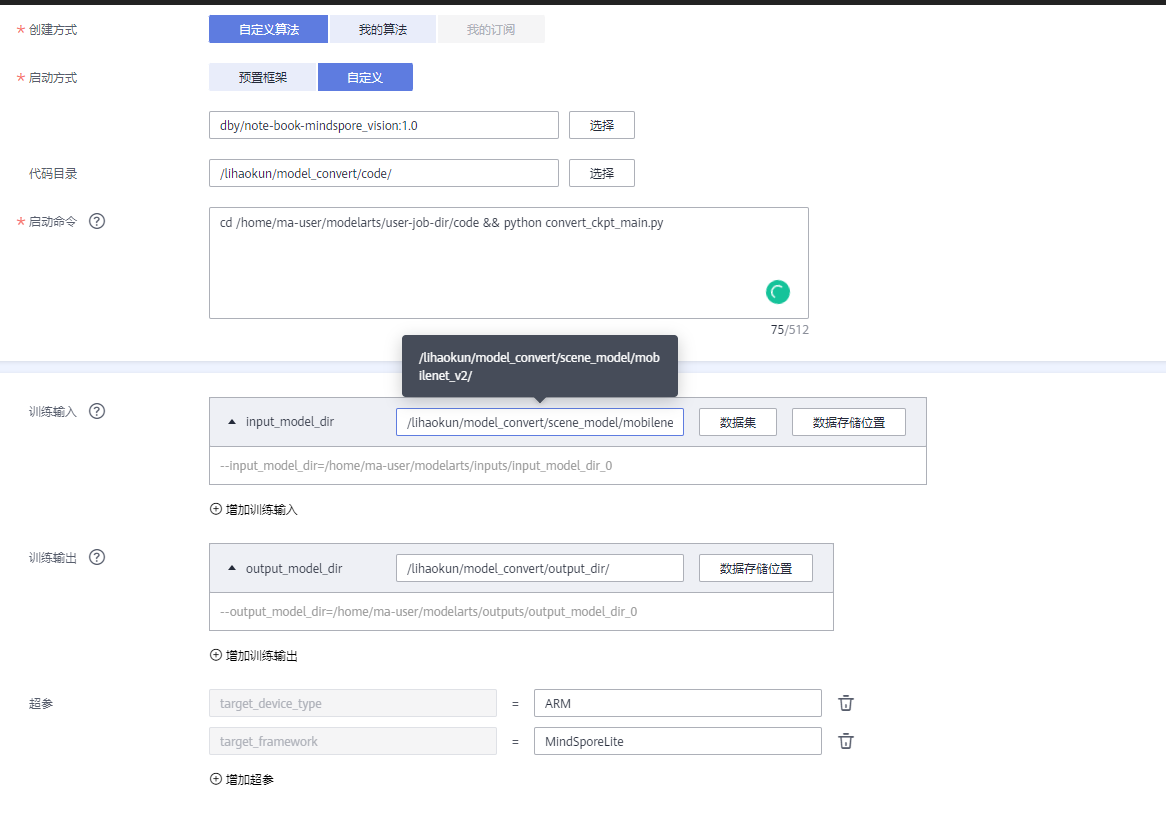
# 接口设计

## 模型移植服务接口分组

### 模型移植接口

|  |  |  |  |
| --- | --- | --- | --- |
| **接口描述** | 调用应用，把MindSpore模型移植到CPU端、GPU端、ARM端、FPGA、DSP、鹏城自研芯片，返回移植后的模型文件的存储路径。 | | |
| **HTTP方法** | POST | | |
| **URI** | xxx/model-convert/v1 | | |
| **参数名称** | **必须** | **参数类型** | **参数描述** |
| input\_model\_dir | 是 | string | 待移植的MindSpore模型，例如MindSpore IR文件（xxx.mindir格式）存储的文件夹 |
| target\_device\_type | 是 | string | 移植的目标设备类型（6种）：Inter-CPU，Nvidia-GPU，ARM，FPGA，DSP，鹏城自研芯片 |
| target\_framework | 是 | String | 移植的目标框架：OpenVINO、TensorRT、MindSporeLite、TensorFlowLite |
| output\_model\_dir | 是 | string | 移植后的模型文件存储的文件夹路径 |
| other\_config | 是 | string | 其他与具体算法相关的私有参数，使用Base64编码为string，具体步骤如下：   1. 用户选择targetDeviceType后，页面展示对应的otherConfig模板，有默认参数和注释。 2. 用户可在页面对otherConfig模板文件（如.yaml文件）进行编辑，修改默认参数。 3. otherConfig文件经过Base64编码为string后，作为“otherConfig”参数发送给服务器，在服务器解码后存储到特定位置供算法访问。 |

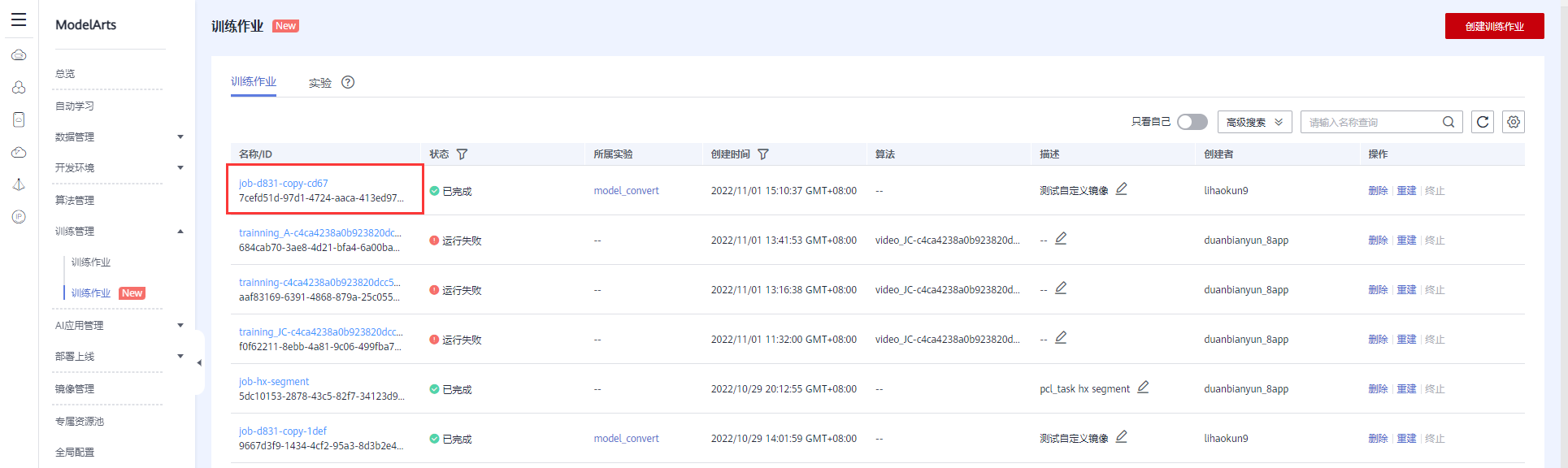
Modelarts训练作业配置示例：



代码、资源包在Modelarts OBS上的存储路径：

obs://lihaokun/model\_convert

一个成功的Modelarts练作业（名称：job-d831-copy-cd67 ID：7cefd51d-97d1-4724-aaca-413ed97de8ac），可在Modelarts的训练作业界面点击ID进入详情页面查看，如下图所示：



训练作业详情页面如下：

