小谈RV1126 AI模型部署流程：

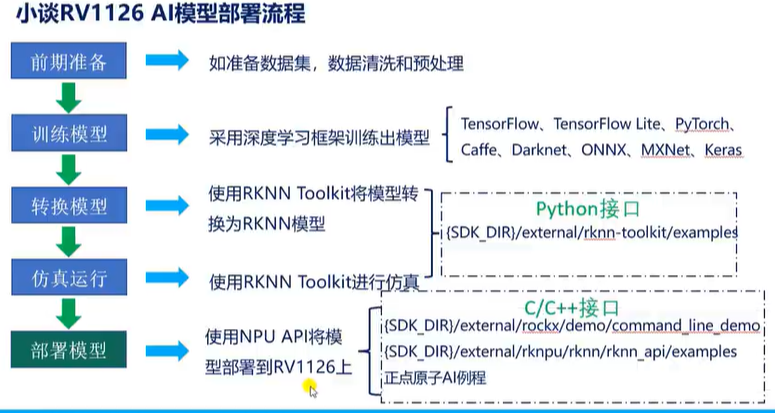
根据AI应用需求进行前期的准备工作，如：准备数据集，数据清洗和预处理。

AI网络模型：采用深度学习框架训练出模型：深度学习框架包括：TensorFlow、TensorFlow Lite、PyTorch、Caffe、Darknet、ONNX、MXNet、Keras。假如使用TensorFlow训练出一个模型后，在PC端去运行该模型，假设在PC端的测试集的损失值地较低，正确率比较高，说明该模型的泛化比较好。现在就需要将模型进行转换操作

也就是将处于TensorFlow框架下的这个模型，使用RKNN Toolkit把它转化为RKNN框架模型。因为RK的NPU不能直接运行其它深度学习下的模型，他只能运行RKNN下的框架模型。所以需要将基于TensorFlow框架下训练出来的模型 转化为 RKNN框架下的模型。这样RKNPU才能运行该模型。转化工具就是RKNN Toolkit.

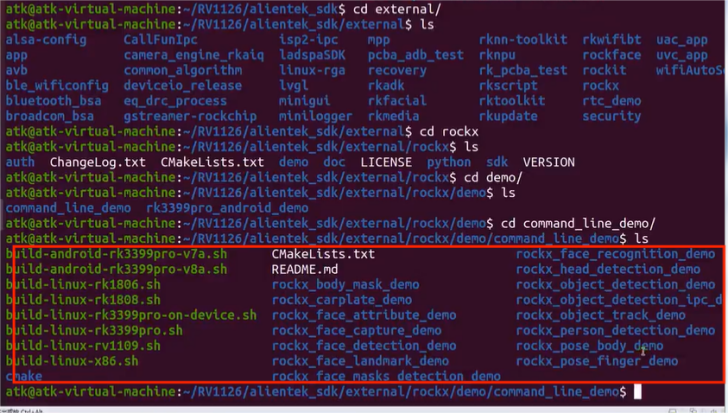
把RKNN模型部署到RV1126上。在部署之前，我们可以先进行仿真测试。也就是使用RKNN Toolkit工具进行模拟仿真，也就是在PC端仿真该模型，或将该模型分发到RKNPU中进行推理仿真。我们可以查看仿真的效果。

仿真通过后，就可以进行模型的部署：使用NPU API将模型部署到RV1126上。这个阶段就需要调用到RKNN的C API （c编写的API接口），所以在RV1126上主要跑的就是C API的代码。如果RV1126跑的是Debain或ubuntu这种文件系统，那么就可以在RV1126上使用rknn-toolkit这种python接口的例程。

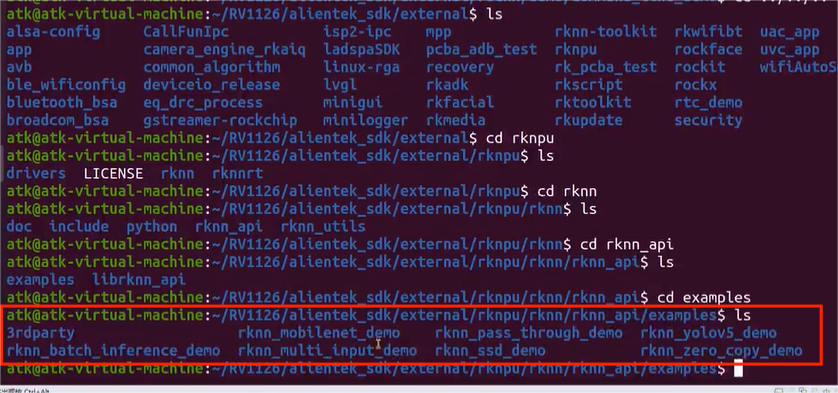


但，由于目前官方给Rv1126适配的buildroot文件系统，所以我们使用RKNN的C API来编写对应的AI程序。

关于RKNN的C API代码，我们可以参考SDK的demo例程，路径详细看图片：



也可以查看rknn的API的例程，路径见图：



以及正点原子发布的AI例程。

