



Tina NPU 开发部署说明

版本号: 1.0
发布日期: 2021.07.21

版本历史

版本号	日期	制/修订人	内容描述
1.0	2021.07.21	PDC	NPU 开发部署说明文档

目 录

1 前言	1
1.1 读者对象	1
1.2 约定	1
1.2.1 符号约定	1
2 正文	2
2.1 NPU 开发简介	2
2.2 开发流程	2
2.3 模型训练	2
2.4 模型转换	3
2.5 程序开发	3
2.6 acuity Toolkit	3
3 结束	4

1 前言

1.1 读者对象

本文档（本指南）主要适用于以下人员：

- 技术支持工程师
- 软件开发工程师
- AI 应用案客户

1.2 约定

1.2.1 符号约定

本文中可能出现的符号如下：



警告

警告

 **技巧**

1. 技巧
2. 小常识

 **说明**

说明

2 正文

2.1 NPU 开发简介

- 支持 int8/uint8/int16 量化精度，运算性能可达 1TOPS.
- 相较于 GPU 作为 AI 运算单元的大型芯片方案，功耗不到 GPU 所需要的 1%.
- 可直接导入 Caffe, TensorFlow, Onnx, TFLite, Keras, Darknet, pyTorch 等模型格式.
- 提供 AI 开发工具：支持模型快速转换、支持开发板端侧转换 API、支持 TensorFlow, TF Lite, Caffe, ONNX, Darknet, pyTorch 等模型.
- 提供 AI 应用开发接口：提供 NPU 跨平台 API.

2.2 开发流程

NPU 开发完整的流程如下图所示：

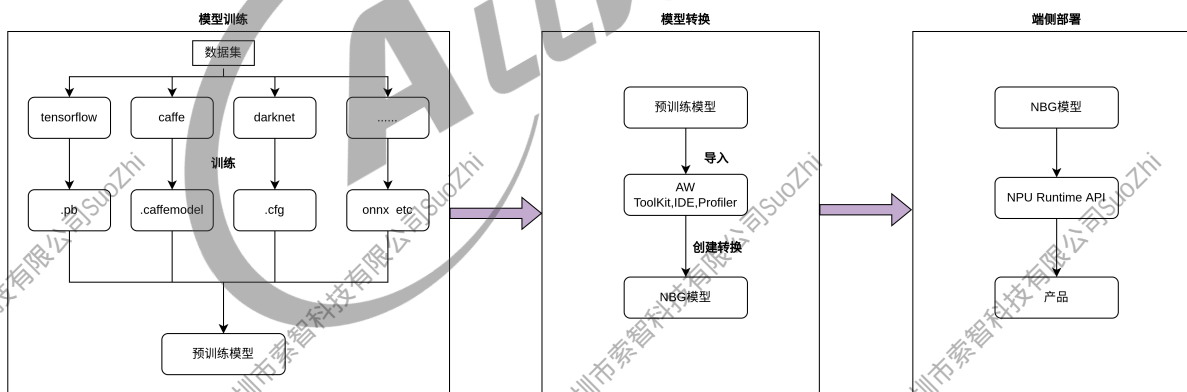


图 2-1: npu_1.png

2.3 模型训练

在模型训练阶段，用户根据需求和实际情况选择合适的框架（如 Caffe、TensorFlow 等）进行训练得到符合需求的模型。也可直接使用已经训练好的模型，对于基于已有的算法模型部署来讲，可以不用经过模型训练阶段。

2.4 模型转换

此阶段为通过 Acuity Toolkit 把模型训练中得到的模型转换为 NPU 可用的模型 NBG 文件。

2.5 程序开发

最后阶段为基于 VIPLite API 开发程序实现业务逻辑。

2.6 acuity Toolkit

Allwinner 提供 acuity toolkit 开发套件进行模型转换、推理运行和性能评估。

用户通过提供的 python 接口可以便捷地完成以下功能：

- 1) 模型转换：支持 Caffe, TensorFlow Lite, Tensorflow, ONNXDarknet NBG 模型导入导出，后续能够在硬件平台上加载使用。
- 2) 模型推理：能够在 PC 上模拟运行模型并获取推理结果，也可以在指定硬件平台上运行模型并获取推理结果。
- 3) 性能评估：能够在 PC 上模拟运行并获取模型总耗时及每一层的耗时信息，也可以通过联机调试的方式在指定硬件平台上运行模型，并获取模型在硬件上运行时的总时间和每一层的耗时信息。

此文档主要介绍模型转换和基于 NPU 程序开发，不涉及模型训练的内容。

3 结束



著作权声明

版权所有 © 2022 珠海全志科技股份有限公司。保留一切权利。

本文档及内容受著作权法保护，其著作权由珠海全志科技股份有限公司（“全志”）拥有并保留一切权利。

本文档是全志的原创作品和版权财产，未经全志书面许可，任何单位和个人不得擅自摘抄、复制、修改、发表或传播本文档内容的部分或全部，且不得以任何形式传播。

商标声明

、 **全志科技** （不完全列举）均为珠海全志科技股份有限公司的商标或者注册商标。在本文档描述的产品中出现的其它商标，产品名称，和服务名称，均由其各自所有人拥有。

免责声明

您购买的产品、服务或特性应受您与珠海全志科技股份有限公司（“全志”）之间签署的商业合同和条款的约束。本文档中描述的全部或部分产品、服务或特性可能不在您所购买或使用的范围内。使用前请认真阅读合同条款和相关说明，并严格遵循本文档的使用说明。您将自行承担任何不当使用行为（包括但不限于如超压，超频，超温使用）造成的不利后果，全志概不负责。

本文档作为使用指导仅供参考。由于产品版本升级或其他原因，本文档内容有可能修改，如有变更，恕不另行通知。全志尽全力在本文档中提供准确的信息，但并不确保内容完全没有错误，因使用本文档而发生损害（包括但不限于间接的、偶然的、特殊的损失）或发生侵犯第三方权利事件，全志概不负责。本文档中的所有陈述、信息和建议并不构成任何明示或暗示的保证或承诺。

本文档未以明示或暗示或其他方式授予全志的任何专利或知识产权。在您实施方案或使用产品的过程中，可能需要获得第三方的权利许可。请您自行向第三方权利人获取相关的许可。全志不承担也不代为支付任何关于获取第三方许可的许可费或版税（专利税）。全志不对您所使用的第三方许可技术做出任何保证、赔偿或承担其他义务。