

# RK3568 正点原子开发板简介

林宏杰

2025-4-2

# I 概览

- 如何选择和烧写镜像
- 开发环境准备
- 安装NPU驱动和 `librknn`
- 模型量化, 转换
- 板上推理

## ■ 概览

开发板的Spec如下

- CPU : RK3568, 4x Cortex-A55 @ 2.0GHz (还行)
- GPU : Mali-G52 (比较拉. OpenCL支持不太完善)
- NPU: 1TOPS (int8) (还行)

总结: 不想折腾Cross Compile的话, 直接板上编译也不会太慢。

# ■ 选择镜像

## 大方向

- 如果你会编译 `Linux` 内核，直接用 `Buildroot` 测试或者从 `Linux SDK` 开始。
- 否则，直接用完整发行版。

## 发行版已知问题

- Ubuntu: 没有无线网卡驱动。(少内核模块)
- Debian: 驱动齐。 `rknpu` 驱动版本略低，未测试。

这个板子也支持鸿蒙，但是没玩过。

## ■ 烧写镜像

文档有详细介绍。参考官方资料。

实验室的板子有板载闪存，不自带SD卡也可以。

### 插哪个口？

- `OTG` : 后面是SoC的USB口，一般来说应该用这个。
- `UART` : 后面是 `CH340` 的串口。debug用。

## ■ 开发环境准备

改好账户密码。镜像一般自带 `sshd` , 连上网就可以愉快的用 `code-server` 了。

### 怎么联网?

- github有前辈写的三方drcom认证客户端。可以自行编译 (获得软路由!).
- 宿舍 / 教学区无线连接可以走旁路认证。 `curl` 任意 `http` 网页, 弄到劫持的重定向链接即可。
- 拿自己的电脑桥接。没必要。

## ■ 开发环境准备

板子镜像自带rknpu的内核模块 (驱动). 你可以拷个新的librknn什么的。

搞AI大概率用Python. 个人建议装Conda, 然后在里面用pip.

### 交叉编译?

官方资料包里有x86 -> aarch64的交叉编译器, 也可以上网下个新点的。

交叉编译的坑比较大。如果你需要编译Python的二进制库, 不如找找别的现成轮子。

## ■ 模型量化和转换

<https://github.com/airockchip/rknn-toolkit2>

人工智能模型需要转换成 `rknn` 专有格式才能在板上运行。以上repo包括了

- 板上Runtime ( `librknn` , etc.)
- 模型转换工具
- 示例工程

`rknn-toolkit-lite` 是对C库的Python封装，提供板上的Python 推理 API.



## ■ 模型量化和转换

`rknn-toolkit` 是运行在桌面上的模型处理工具链，执行量化和转换两种工程。

开始之前，你需要将你的模型转换为 `ONNX` 格式，顺便检查一下模型有没有NPU不支持的算子。

转换的配置参见手册。

**注意** Torch导出到ONNX的功能似乎为社区提供，需要较老的Python版本 (~3.8)

## 板上推理

安装了 `librknn` 就可以用 C / C++ API了。

要用 Python 调 API, 选择合适的 `rknn-toolkit-lite` 版本, 安装repo里边的prebuilt wheel.