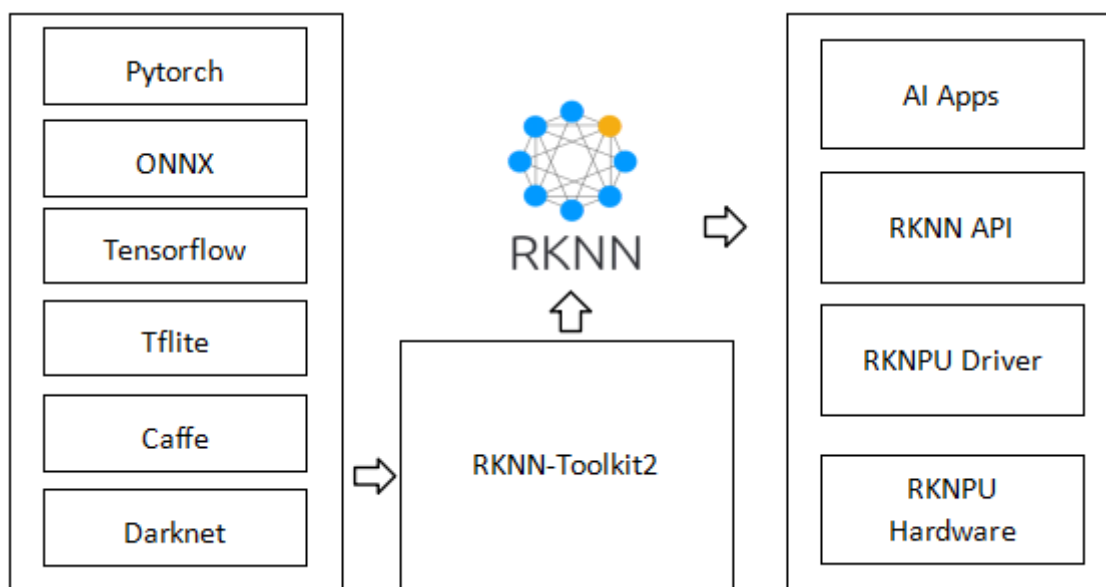


# Description

---

RKNN software stack can help users to quickly deploy AI models to Rockchip chips. The overall framework is as follows:



In order to use RKNPU, users need to first run the RKNN-Toolkit2 tool on the computer, convert the trained model into an RKNN format model, and then inference on the development board using the RKNN C API or Python API.

- RKNN-Toolkit2 is a software development kit for users to perform model conversion, inference and performance evaluation on PC and Rockchip NPU platforms.
- RKNN-Toolkit-Lite2 provides Python programming interfaces for Rockchip NPU platform to help users deploy RKNN models and accelerate the implementation of AI applications.
- RKNN Runtime provides C/C++ programming interfaces for Rockchip NPU platform to help users deploy RKNN models and accelerate the implementation of AI applications.
- RKNPU kernel driver is responsible for interacting with NPU hardware. It has been open source and can be found in the Rockchip kernel code.

# Support Platform

---

- RK3588 Series
- RK3576 Series
- RK3566/RK3568 Series
- RK3562 Series
- RV1103/RV1106
- RV1103B/RV1106B

- RK2118

Note:

**For RK1808/RV1109/RV1126/RK3399Pro, please refer to :**

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

<https://github.com/airockchip/rknpu>

[https://github.com/airockchip/RK3399Pro\\_npu](https://github.com/airockchip/RK3399Pro_npu)

## Download

---

- You can also download all packages, docker image, examples, docs and platform-tools from [RKNPU2\\_SDK](#), fetch code: rknn
- You can get more examples from [rknn mode zoo](#)

## Notes

---

- RKNN-Toolkit2 is not compatible with [RKNN-Toolkit](#)
- The supported Python versions are:
  - Python 3.6
  - Python 3.7
  - Python 3.8
  - Python 3.9
  - Python 3.10
  - Python 3.11
  - Python 3.12
- Latest version:v2.3.0

## RKNN LLM

---

If you want to deploy LLM (Large Language Model), we have introduced a new SDK called RKNN-LLM. For details, please refer to:

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

# CHANGELOG

---

## 1. v2.3.0

---

- RKNN-Toolkit2 support ARM64 architecture
- RKNN-Toolkit-Lite2 support installation via pip
- Add support for W4A16 symmetric quantization (RK3576)
- Operator optimization, such as LayerNorm, LSTM, Transpose, MatMul, etc.

for older version, please refer [CHANGELOG](#)

## Feedback and Community Support

---

- [Redmine](#) (Feedback recommended, Please consult our sales or FAE for the redmine account)
- QQ Group Chat: 1025468710 (full, please join group 3)
- QQ Group Chat2: 547021958 (full, please join group 3)
- QQ Group Chat3: 469385426

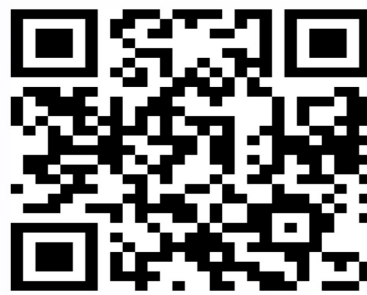
### RKNN交流群#1

群号: 1025468710



### RKNN交流群#2

群号: 547021958



### RKNN交流群#3

群号: 469385426



