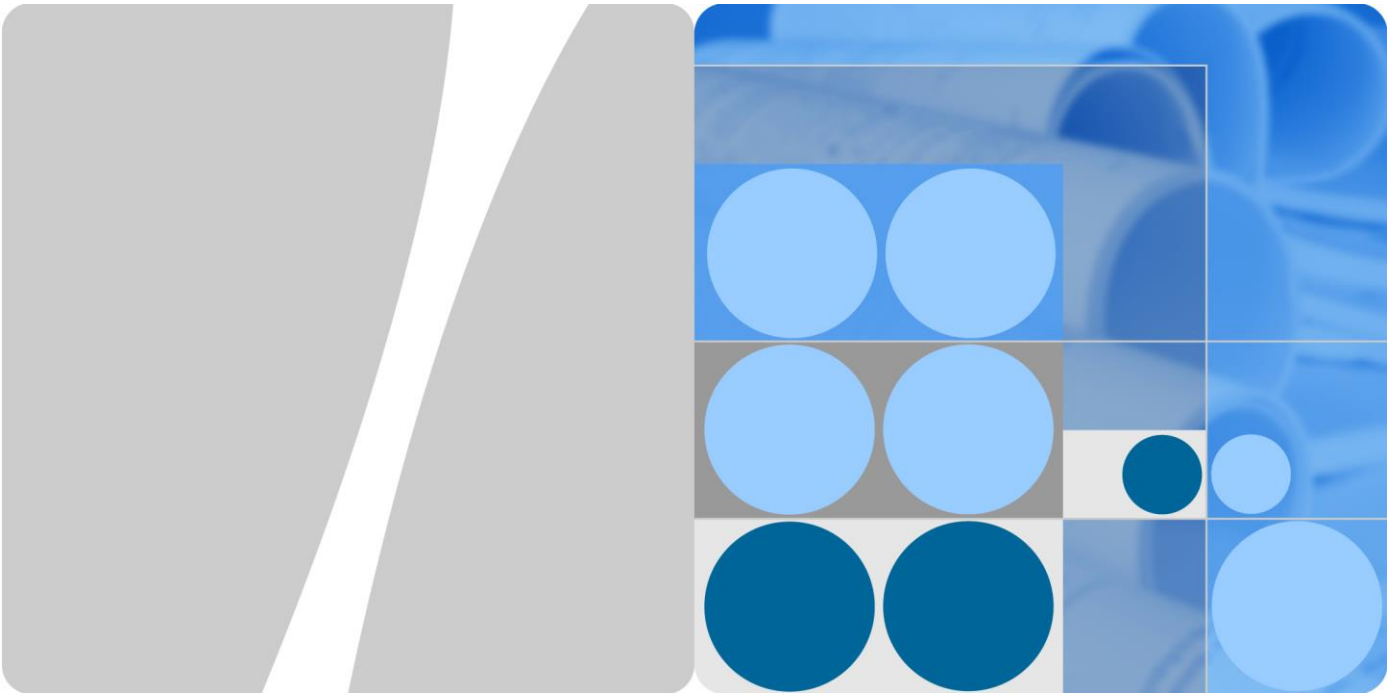

HiAI DDK V320

Release Notes

Issue 03

Date 2020-02-28

HUAWEI TECHNOLOGIES CO., LTD.

Huawei HiAI Application

Send an application email to developer@huawei.com.

Email subject: HUAWEI HiAI + Company name + Product name

Email body: Cooperation company + Contact person + Phone number + Email address

We will reply to you within 5 working days.

Official website: https://developer.huawei.com/consumer/en/

About This Document

Change History

Date	Version	Change Description
2020-07-22	04	Added the description of HiAI DDK V320.030.
2020-02-28	03	Added the description of the system debug tool and general-purpose Arm processor.
2019-12-31	02	Added the description of HiAI DDK V320.
2019-09-04	01	Added the description of HiAI DDK V310.

Contents

1 Version Mapping

1.1 Product Version Information

Product Name	Version	Release Date
HiAI DDK V320	100.320.030.010	2020-07

1.2 Document List

File Name	Path
Huawei HiAI DDK V320 IR Model Building Instructions	document\
Huawei HiAI DDK V320 Quick Start	document\
Huawei HiAI DDK V320 Model Inference and Integration Instructions	document\
Huawei HiAI DDK V320 Operator Specifications	document\
Huawei HiAI DDK V320 OMG Tool Instructions	tools\tools_omg\
Huawei HiAI DDK V320 Lightweight Tool Instructions	tools\tools_dopt\
Huawei HiAI DDK V320 Release Notes	document\
Huawei HiAI DDK V320 Acronyms and Abbreviations	document\
Huawei HiAI DDK V320 FAQs	document\

File Name	Path
Huawei HiAI DDK V320 System Debug Tool Instructions	tools\tool_sysdbg\

2 Feature Changes

2.1 100.320.030.010 Feature Changes

No.	Change Description	Impact
1	Added the **hiai::op::** namespace for IR operator definition.	The operator definition in the old **ge::op** namespace is moved to **./graph/compatible/**. For details, see the *Huawei HiAI DDK V320 IR Model Building Instructions*
2	Removed the Windows version OMG.	The Windows version OMG is no longer available.
3	Implemented zero-copy between NPUs and other compute units.	Two APIs are added. For details, see sections 3.4.1 and 3.5.1.22 in the *Huawei HiAI DDK V320 Model Inference Integration Guide*.

2.2 100.320.020.010 Feature Changes

No.	Change Description	Impact
1	The network architecture search (NAS) function is added.	The NAS tool is added to the lightweight toolkit. For details, see the *Huawei HiAI DDK V320 Lightweight Tool Instructions*.
2	The support of IR model generation is added to the Linux version OMG.	The support of IR model generation is added to the Linux version OMG. For details, see the *Huawei HiAI DDK V320 OMG Tool Instructions*.

No.	Change Description	Impact
3	Inference using offline models on the general-purpose Arm processor is supported,	Models with restricted operators can run on the CPU. For details about the scope of operators that can run on the CPU, see the *Huawei HiAI DDK V320 Operator Specifications*. For details about conversion using OMG, see the *Huawei HiAI DDK V320 OMG Tool Instructions*.
4	17 CPU operators are newly added.	17 CPU operators are newly supported. For details, see the *Huawei HiAI DDK V320 Operator Specifications*.
5	The Profiling tool is added.	The Profiling tool is added to collect and analyze performance statistics on model inference. For details, see the *Huawei HiAI DDK V320 System Debug Tool Instructions*.

2.3 100.320.010.010 Feature Changes

No.	Change Description	Impact
1	80 more NPU operators are supported.	80 more NPU operators are supported. For details, see the *Huawei HiAI DDK V320 Operator Specifications.*
2	The AIPP feature is added.	For details about external APIs of AIPP, see the *Huawei HiAI DDK V320 Model Inference and Integration Instructions*. For details about the model conversion with AIPP, see the *Huawei HiAI DDK V320 OMG Tool Instructions.*
3	The lightweight tool function is added.	The optimization support policy is newly added. For details, see the *Huawei HiAI DDK V320 Lightweight Tool Instructions.*
4	The OMG tool supporting the Windows OS is added.	The OMG offline conversion tool running the Windows OS is added. For details, see the *Huawei HiAI DDK V320 OMG Tool Instructions.*

2.4 100.310.010.010 Feature Changes

No.	Change Description	Impact
1	45 more NPU operators are supported.	45 more NPU operators are supported. For details, see the *Huawei HiAI DDK V310 Operator Specifications.*
2	The AIPP feature is added.	The scenario of initializing the tensor based on the N, H, W, and image format is added. For details, see section 3.4.1 in the *Huawei HiAI DDK V310 Model Inference Integration Instructions.* For details about model conversion with AIPP, see section 2.4 in the *Huawei HiAI DDK V310 OMG Tool Instructions.*
3	FP16 and INT8 module parameters are added.	The configuration of model input and output formats is supported. For details, see chapter 3 in the *Huawei HiAI DDK V310 OMG Tool Instructions.*
4	The parameter for specifying the path of the low-bit quantization configuration file is added.	Low-bit quantization for high-precision data is supported to save network storage space, reduce transfer latency, and improve calculation efficiency. For details, see section 5.2.11 in the *Huawei HiAI DDK V310 OMG Tool Instructions.* The lightweight tool is required for performing low-bit quantization.
5	Instructions for the lightweight tool are provided.	The parameter for specifying the path of the low-bit quantization configuration file is configured by referring to the instructions. For details, see the *Huawei HiAI DDK V310 Lightweight Tool Instructions.*

2.5 100.300.010.010 Feature Changes

No.	Change Description	Impact
1	178 NPU operators are newly added.	178 NPU operators are newly supported. For details, see the *Huawei HiAI DDK V300 Operator Specifications*.
2	The offline conversion tool OMG is provided.	The offline conversion tool OMG is provided. For details, see the *Huawei HiAI DDK V300*

No.	Change Description	Impact
		OMG Tool Instructions.
3	The IR interface is newly supported.	The integration of operators of third-party frameworks is supported. The IR interface can be used to generate model files supported by V300. For details, see the *Huawei HiAI DDK V300 IR Model Building Instructions*.

3 Resolved Issues

3.1 Resolved Issues in 100.320.020.010

Issue Description	Workaround	Progress
Using a HiAI DDK demo with a multi-output model, after the model is successfully loaded, inference fails and an error indicating that the model does not exist is reported.	None	The issue has been resolved.
The Da Vinci offline model converted from a TensorFlow model containing the Depthwise operator by using the OMG is oversized.	Do not use the Depthwise operator that does not meet optimization requirements.	The issue has been resolved.
The AI framework requires 32-bit libhiai_ir.so. In the earlier version, only 64-bit libhiai_ir.so is supported.	None	The issue has been resolved.
When the HiAI DDK inference demo is loaded on the application, repeated loading and memory allocation times have to be reduced.	None	The issue has been resolved.
The PriorBox operator fails in the Da Vinci offline model conversion by using the OMG.	None	The issue has been resolved.

4 Known Issues

4.1 100.320.030.010 Known Issues

None

