



# SUPPORT MATRIX FOR TENSORRT

DA-09025-001 \_v4.0.1 | June 2018

## Support Guide



## TABLE OF CONTENTS

Chapter 1. Features For Platforms And Software.....	1
Chapter 2. Layer Features.....	2

# Chapter 1.

## FEATURES FOR PLATFORMS AND SOFTWARE

The following table lists the supported platforms and which version supports a particular feature.

Table 1

Feature	Linux x86-64	Linux aarch64	Android aarch64	QNX aarch64
Supported CUDA versions	8.0, 9.0, <<9.2>>	9.2	9.2	9.2
Supported cuDNN versions	<<7.1>>	7.1	7.1	7.1
TensorRT Python API	Yes	No	No	No
NvUffParser	Yes	Yes	Yes	Yes
NvOnnxParser	Yes	Yes	Yes	Yes



Serialized engines are not portable across platforms or TensorRT versions.

## Chapter 2.

# LAYER FEATURES

The following table lists the TensorRT layers, specifically, which features are supported by a particular layer.

Table 2

Layer	Dimensions of input tensor	Dimensions of output tensor	Does the operation apply to only the innermost 3 dimensions?	Supports broadcast <sup>1</sup>	Supports broadcast across batch <sup>2</sup>
Activation	1-7 dimensions	1-7 dimensions	No	No	No
Concatenation	1-7 dimensions	1-7 dimensions	No	No	No
Constant	0-7 dimensions	0-7 dimensions	No	No	Always
Convolution	3 or more dimensions	3 or more dimensions	Yes	No	No
Deconvolution	3 or more dimensions	3 or more dimensions	Yes	No	No
ElementWise	0-7 dimensions	0-7 dimensions	No	Yes	Yes
FullyConnected	3 or more dimensions	3 or more dimensions	Yes	No	No
Gather	<ul style="list-style-type: none"> <li>Input1: 1-7 dimensions</li> <li>Input2: 1-7 dimensions</li> </ul>	1-7 dimensions	No	No	Yes

Layer	Dimensions of input tensor	Dimensions of output tensor	Does the operation apply to only the innermost 3 dimensions?	Supports broadcast <sup>1</sup>	Supports broadcast across batch <sup>2</sup>
LRN	3 or more dimensions	3 or more dimensions	Yes	No	No
MatrixMultiply	2 or more dimensions	2 or more dimensions	No	Yes	Yes
Padding	3 or more dimensions	3 or more dimensions	Yes	No	No
Plugin	User defined	User defined	User defined	User defined	User defined
Pooling	3 or more dimensions	3 or more dimensions	Yes	Yes	Yes
RaggedSoftMax	<ul style="list-style-type: none"> <li>Input: 2 dimensions</li> <li>Bounds: 2 dimensions</li> </ul>	2 or more dimensions	No	No	Yes
Reduce	1-7 dimensions	0-7 dimensions	No	No	No
RNN	3 dimensions	3 or more dimensions	No	No	No
RNNv2	<ul style="list-style-type: none"> <li>Data/Hidden/Cell: 2 or more dimensions</li> <li>SeqLen: 0 or more dimensions</li> </ul>	Data/Hidden/Cell: 2 or more dimensions	No	No	No
Scale	3 or more dimensions	3 or more dimensions	Yes	No	No
Shuffle	0-7 dimensions	0-7 dimensions	No	No	No
SoftMax	1-7 dimensions	1-7 dimensions	No	No	No

Layer	Dimensions of input tensor	Dimensions of output tensor	Does the operation apply to only the innermost 3 dimensions?	Supports broadcast <sup>1</sup>	Supports broadcast across batch <sup>2</sup>
TopK	1-7 dimensions	<ul style="list-style-type: none"> <li>▶ Output1: 1-7 dimensions</li> <li>▶ Output2: 1-7 dimensions</li> </ul>	Yes	No	Yes
Unary	0-7 dimensions	0-7 dimensions	No	No	No

For more information about each of the TensorRT layers, see [TensorRT Layers](#).

## Notice

THE INFORMATION IN THIS GUIDE AND ALL OTHER INFORMATION CONTAINED IN NVIDIA DOCUMENTATION REFERENCED IN THIS GUIDE IS PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE INFORMATION FOR THE PRODUCT, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the product described in this guide shall be limited in accordance with the NVIDIA terms and conditions of sale for the product.

THE NVIDIA PRODUCT DESCRIBED IN THIS GUIDE IS NOT FAULT TOLERANT AND IS NOT DESIGNED, MANUFACTURED OR INTENDED FOR USE IN CONNECTION WITH THE DESIGN, CONSTRUCTION, MAINTENANCE, AND/OR OPERATION OF ANY SYSTEM WHERE THE USE OR A FAILURE OF SUCH SYSTEM COULD RESULT IN A SITUATION THAT THREATENS THE SAFETY OF HUMAN LIFE OR SEVERE PHYSICAL HARM OR PROPERTY DAMAGE (INCLUDING, FOR EXAMPLE, USE IN CONNECTION WITH ANY NUCLEAR, AVIONICS, LIFE SUPPORT OR OTHER LIFE CRITICAL APPLICATION). NVIDIA EXPRESSLY DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY OF FITNESS FOR SUCH HIGH RISK USES. NVIDIA SHALL NOT BE LIABLE TO CUSTOMER OR ANY THIRD PARTY, IN WHOLE OR IN PART, FOR ANY CLAIMS OR DAMAGES ARISING FROM SUCH HIGH RISK USES.

NVIDIA makes no representation or warranty that the product described in this guide will be suitable for any specified use without further testing or modification. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer's sole responsibility to ensure the product is suitable and fit for the application planned by customer and to do the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer's product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this guide. NVIDIA does not accept any liability related to any default, damage, costs or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this guide, or (ii) customer product designs.

Other than the right for customer to use the information in this guide with the product, no other license, either expressed or implied, is hereby granted by NVIDIA under this guide. Reproduction of information in this guide is permissible only if reproduction is approved by NVIDIA in writing, is reproduced without alteration, and is accompanied by all associated conditions, limitations, and notices.

## Trademarks

NVIDIA, the NVIDIA logo, and cuBLAS, CUDA, cuDNN, cuFFT, cuSPARSE, DIGITS, DGX, DGX-1, Jetson, Kepler, NVIDIA Maxwell, NCCL, NVLink, Pascal, Tegra, TensorRT, and Tesla are trademarks and/or registered trademarks of NVIDIA Corporation in the United States and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

## Copyright

© 2018 NVIDIA Corporation. All rights reserved.