CUDA详解

0.CUDA说明

CUDA的全称是Compute Unified Device Architecture，即计算统一设备架构。它是由NVIDIA英伟达公司开发的一个并行计算平台和编程模型，允许软件开发者和研究人员利用NVIDIA的GPU（图形处理单元）进行高性能计算。

CUDA的设计目的是为了简化并加速科学计算、数据分析、机器学习等领域的复杂计算任务。通过CUDA，开发者可以编写程序来执行这些任务，而这些程序可以在NVIDIA的GPU上高效运行。相比于传统的CPU，GPU拥有更多的核心，能够同时处理多个线程，非常适合用于数据并行的工作负载。

CUDA提供了C、C++、Fortran等语言的扩展，让开发者可以通过这些语言直接访问GPU的底层资源。此外，还有多种高级编程语言和框架（如Python中的PyCUDA、深度学习库PyTorch和TensorFlow等）支持CUDA，使得开发者能够在更高层次上利用GPU的能力。

CUDA的核心组成部分包括：

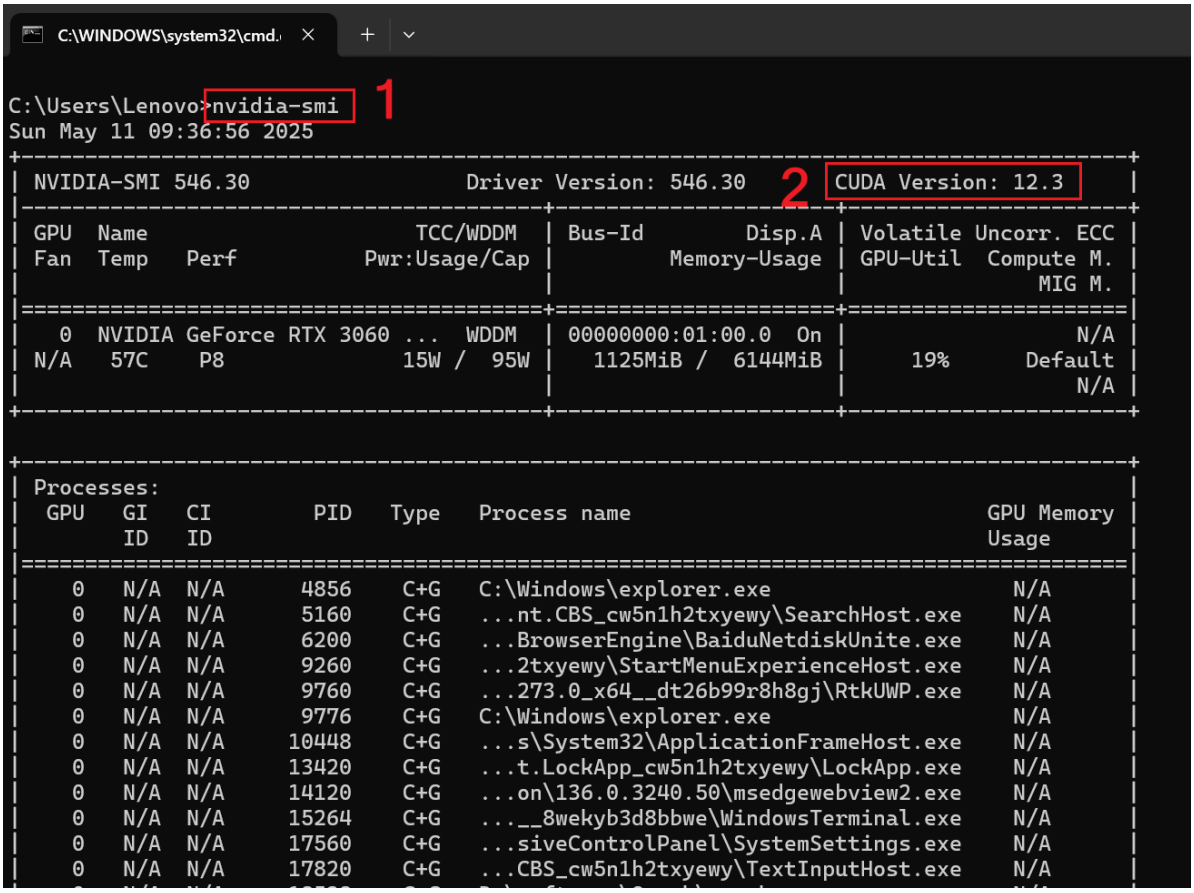
硬件层面：包含了一系列支持CUDA技术的NVIDIA GPU，它们具有大量的CUDA核心，用于执行并行计算任务。

软件层面：CUDA Toolkit，这是一个包含了编译器、调试工具和性能分析工具的集合，帮助开发者编写和优化CUDA应用程序。

> **注意1:**电脑如果是Mac平台对CUDA的支持是有限的，因为CUDA是由NVIDIA开发的并行计算平台和编程模型，它主要用于NVIDIA的GPU上。由于苹果公司的Mac计算机通常不使用NVIDIA的显卡，而是采用集成的Intel GPU或者自家的Apple Silicon芯片（如M1、M2等），因此标准的macOS系统并不直接支持CUDA。
>
> 不过，如果你确实拥有一个带有NVIDIA GPU的Mac，比如通过外接的eGPU盒子连接的NVIDIA显卡，并且你的macOS版本支持eGPU的话，那么在一定程度上你可以在Mac上使用CUDA。需要注意的是，macOS High Sierra (10.13) 及之后的版本才开始正式支持eGPU设备，但即使如此，NVIDIA官方已经停止了对macOS的新版Web Driver和CUDA驱动程序的支持，这意味着最新的CUDA版本可能无法在Mac上运行。
>
> 对于使用Apple Silicon芯片的Mac用户来说，苹果推荐使用Metal API来代替CUDA进行GPU加速计算。PyTorch等深度学习框架也已经开始支持使用Apple的Metal Performance Shaders (MPS)作为后端，这为在这些机器上进行GPU加速提供了替代方案。
>
> 总结一下，除非通过外部GPU (eGPU) 并且有合适的旧版驱动支持，否则普通的Mac电脑不支持CUDA。
>
> **注意2**:电脑如果是Windows平台下的Nvidia GPU的用户，可以选择配置Pytorch的CUDA版本
>
> 斌哥在此友情提醒: 以下均为windows版本配置步骤

1.检查电脑是否支持GPU

> win + r 输入命令 nvidia-smi
>
> 斌哥这里显示的是cuda最高版本12.3,你和我的不一定一样哦!!!
>
> 如果觉得自己的驱动低,斌哥告诉你可以去官网升级: https://www.nvidia.cn/drivers/lookup/



2.CUDA Toolkit 下载并安装

> 斌哥友情提醒官网链接: https://developer.nvidia.com/cuda-toolkit-archive

下载

CUDA Toolkit 12.8.1 (March 2025), Versioned Online Documentation
CUDA Toolkit 12.8.0 (January 2025), Versioned Online Documentation
CUDA Toolkit 12.6.3 (November 2024), Versioned Online Documentation
CUDA Toolkit 12.6.2 (October 2024), Versioned Online Documentation
CUDA Toolkit 12.6.1 (August 2024), Versioned Online Documentation
CUDA Toolkit 12.6.0 (August 2024), Versioned Online Documentation
CUDA Toolkit 12.5.1 (July 2024), Versioned Online Documentation
CUDA Toolkit 12.5.0 (May 2024), Versioned Online Documentation
CUDA Toolkit 12.4.1 (April 2024), Versioned Online Documentation
CUDA Toolkit 12.4.0 (March 2024), Versioned Online Documentation
CUDA Toolkit 12.3.2 (January 2024), Versioned Online Documentation
CUDA Toolkit 12.3.1 (November 2023), Versioned Online Documentation
CUDA Toolkit 12.3.0 (October 2023), Versioned Online Documentation
CUDA Toolkit 12.2.2 (August 2023), Versioned Online Documentation
CUDA Toolkit 12.2.1 (July 2023), Versioned Online Documentation
CUDA Toolkit 12.2.0 (June 2023), Versioned Online Documentation
CUDA Toolkit 12.1.1 (April 2023), Versioned Online Documentation
CUDA Toolkit 12.1.0 (February 2023), Versioned Online Documentation
CUDA Toolkit 12.0.1 (January 2023), Versioned Online Documentation
CUDA Toolkit 12.0.0 (December 2022), Versioned Online Documentation
CUDA Toolkit 11.8.0 (October 2022), Versioned Online Documentation
CUDA Toolkit 11.7.1 (August 2022), Versioned Online Documentation
CUDA Toolkit 11.7.0 (May 2022), Versioned Online Documentation
CUDA Toolkit 11.6.2 (March 2022), Versioned Online Documentation
CUDA Toolkit 11.6.1 (February 2022), Versioned Online Documentation
CUDA Toolkit 11.6.0 (January 2022), Versioned Online Documentation
CUDA Toolkit 11.5.2 (February 2022), Versioned Online Documentation
CUDA Toolkit 11.5.1 (November 2021), Versioned Online Documentation
CUDA Toolkit 11.5.0 (October 2021), Versioned Online Documentation
CUDA Toolkit 11.4.4 (February 2022), Versioned Online Documentation
CUDA Toolkit 11.4.3 (November 2021), Versioned Online Documentation
CUDA Toolkit 11.4.2 (September 2021), Versioned Online Documentation
CUDA Toolkit 11.4.1 (August 2021), Versioned Online Documentation

推荐两个版本

CUDA Toolkit 12.1 Downloads

Select Target Platform

Click on the green buttons that describe your target platform. Only supported platforms will be shown. By downloading and using the software, you agree to fully comply with the terms and conditions of the CUDA EULA.

Operating System	Linux Windows
Architecture	x86_64
Version	10 11 Server 2019 Server 2022
Installer Type	exe (local) exe (network)

Download Installer for Windows 11 x86_64

The base installer is available for download below.

> Base Installer — Download (3.1 GB)

安装



3.下载cuDNN并配置

> 斌哥友情提醒官网链接: https://developer.nvidia.com/rdp/cudnn-archive

下载

cuDNN Archive

NVIDIA cuDNN is a GPU-accelerated library of primitives for deep neural networks.

- Download cuDNN v8.9.7 (December 5th, 2023), for CUDA 12.x
- Download cuDNN v8.9.7 (December 5th, 2023), for CUDA 11.x

根据之前安装的CUDA,
选择对应cuDNN版本

- Download cuDNN v8.9.6 (November 1st, 2023), for CUDA 12.x
- Download cuDNN v8.9.6 (November 1st, 2023), for CUDA 11.x
- Download cuDNN v8.9.5 (October 27th, 2023), for CUDA 12.x
- Download cuDNN v8.9.5 (October 27th, 2023), for CUDA 11.x
- Download cuDNN v8.9.4 (August 8th, 2023), for CUDA 12.x
- Download cuDNN v8.9.4 (August 8th, 2023), for CUDA 11.x
- Download cuDNN v8.9.3 (July 11th, 2023), for CUDA 12.x
- Download cuDNN v8.9.3 (July 11th, 2023), for CUDA 11.x
- Download cuDNN v8.9.2 (June 1st, 2023), for CUDA 12.x

Download cuDNN v8.9.7 (December 5th, 2023), for CUDA 12.x

Local Installers for Windows and Linux, Ubuntu(x86_64, armsbsa)

Local Installer for Windows (Zip)

Local Installer for Linux x86_64 (Tar)

Local Installer for Linux PPC (Tar)

Local Installer for Linux SBSA (Tar)

Local Installer for Debian 11 (Deb)

Local Installer for Ubuntu20.04 x86_64 (Deb)

Local Installer for Ubuntu22.04 x86_64 (Deb)

Local Installer for Ubuntu20.04 aarch64sbsa (Deb)

Local Installer for Ubuntu22.04 aarch64sbsa (Deb)

Local Installer for Ubuntu20.04 cross-sbsa (Deb)

Local Installer for Ubuntu22.04 cross-sbsa (Deb)



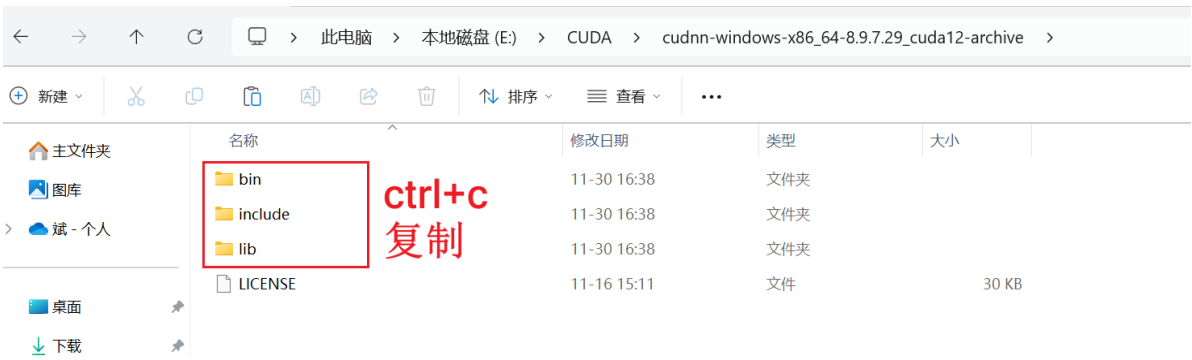




斌哥友情提醒然后输入自定义信息并确认即可

配置

解压下载的zip包并复制



粘贴到第2步安装的cuda对应目录下

> C:\Program Files\NVIDIA GPU Computing Toolkit\CUDA\v12.1



4.pytorch的cuda版本安装

离线安装

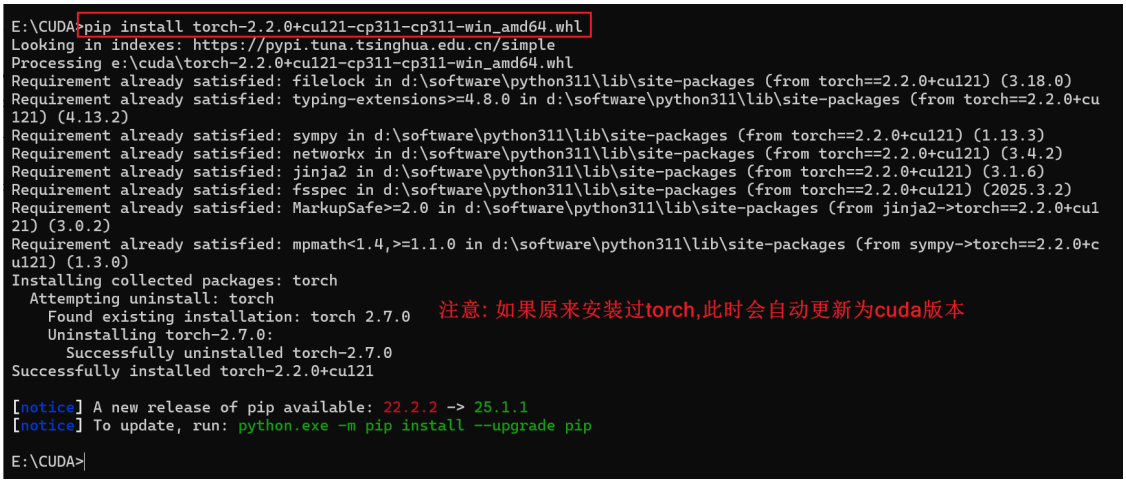
先下载

> 斌哥友情提醒下载网站: https://download.pytorch.org/whl/torch_stable.html
>
> 组成是: cuda版本+torch版本+python版本+平台
>
> ![image-20250511110408256](CUDA斌子教程.assets/image-20250511110408256.png



再安装

> 进入离线下载的文件目录中,直接执行命令即可
>
> ```
> pip install torch-2.2.0+cu121-cp311-cp311-win_amd64.whl
> ```



在线安装

> 官方网址: https://pytorch.org/get-started/locally/
>
> ```
> pip install torch torchvision torchaudio --index-url
> https://download.pytorch.org/whl/cu121
> ```

5.校验cuda是否成功

> 执行以下代码:
>
> ```
> import torch.cuda
>
> # 默认没有设置过cuda时,CUDA available结果是False,如果设置后是True
> print("CUDA是否可用:", torch.cuda.is_available())
> print("CUDA版本:", torch.version.cuda)
> print("PyTorch版本:", torch.__version__)
> print("电脑上GPU数量:", torch.cuda.device_count())
> if torch.cuda.is_available():
>     print("当前使用的GPU版本:",
> torch.cuda.get_device_name(torch.cuda.current_device()))
> ```
>
> 这段代码会输出一系列信息，帮助你确认PyTorch是否正确识别到了CUDA。具体来说：
>
> - `torch.cuda.is_available()` 应该返回 `True`，这表明CUDA是可用的。
> - `torch.version.cuda` 会显示CUDA的版本号。
> - `torch.__version__` 显示了当前安装的PyTorch版本。
> - `torch.cuda.device_count()` 返回系统中可用的GPU数量。
> - 如果有可用的GPU，`torch.cuda.get_device_name(torch.cuda.current_device())` 会显示当前使用的GPU名称。
>
> 斌哥友情提醒如果你运行上述代码时遇到了错误或者`torch.cuda.is_available()`返回`False`，那么可能存在以下几种情况：
>
> 1. CUDA驱动或工具包未正确安装。

> 2. 安装的PyTorch版本与CUDA版本不匹配。
> 3. ...
>
> 确保你的系统已经正确安装了适合你GPU型号和计算能力的CUDA版本，并且PyTorch的版本是支持该CUDA版本的。如果一切正常，你应该能够看到关于CUDA和GPU的相关信息被打印出来。

###