Vitis AI Runtime v1.3

Introduction

Vitis AI Run time enables applications to use the unified high-level runtime API for both cloud and edge. Therefore, making cloud-to-edge deployments seamless and efficient. The Vitis AI Runtime API features are:

- Asynchronous submission of jobs to the accelerator
- Asynchronous collection of jobs from the accelerator
- C++ and Python implementations
- Support for multi-threading and multi-process execution

For edge users, click Quick Start For Edge to get started quickly.

For cloud users, click Quick Start For Alveo to get started quickly.

Vitis AI Runtime directory structure introduction

```
VART
├── README.md
├── adas_detection
│   ├── build.sh
│   └── src
├── common
│   ├── common.cpp
│   └── common.h
├── inception_v1_mt_py
│   ├── inception_v1.py
│   └── words.txt
├── pose_detection
│   ├── build.sh
│   └── src
├── resnet50
│   ├── build.sh
│   ├── src
│   └── words.txt
├── resnet50_mt_py
│   ├── resnet50.py
│   └── words.txt
├── segmentation
│   ├── build.sh
│   └── src
├── squeezenet_pytorch
│   ├── build.sh
│   ├── src
│   └── words.txt
└── video_analysis
        ├── build.sh
        └── src
```

Quick Start For Edge

Setting Up the Host

Follow Setting Up the Host to set up the host for edge.

Setting Up the Target

Follow Setting Up the Target to set up the target.

Running Vitis AI Examples

Follow Running Vitis AI Examples to run Vitis AI examples.

Quick Start For Alveo

Setting Up the Host

1. Click Setup Alveo Accelerator Card with HBM for DPUCAHX8H/L to set up the Alveo Card.

2. Download the xclbin files from here. Untar it, choose the Alveo card and install it. Take `U50` as an example.

```
cd /workspace
wget https://www.xilinx.com/bin/public/openDownload?filename=alveo_xclbin-1.3.1.t
tar -xzvf alveo_xclbin-1.3.1.tar.gz
cd alveo_xclbin-1.3.1/U50/6E300M
sudo cp dpu.xclbin hbm_address_assignment.txt /usr/lib
```

This step is also described in DPUCAHX8H/L Overlay Usage.

Running Vitis AI Examples

Suppose you have downloaded `Vitis-AI`, entered `Vitis-AI` directory, and then started Docker. Thus, `VART` is located in the path of `/workspace/demo/VART/` in the docker system.

`/workspace/demo/VART/` is the path for the following example.

If you encounter any path errors in running examples, check to see if you follow the steps above.

1. Download the vitis_ai_runtime_r1.3.0_image_video.tar.gz package and unzip it.

```
cd /workspace/demo
wget https://www.xilinx.com/bin/public/openDownload?filename=vitis_ai_runtime
tar -xzvf vitis_ai_runtime_r*1.3*_image_video.tar.gz -C VART
```

2. Download the model. For each model, there will be a yaml file which is used for describe all the details about the model. In the yaml, you will find the model's download links for different platforms. Please choose the corresponding model and

download it. Click Xilinx AI Model Zoo to view all the models.

- Take `resnet50` of U50 as an example.

```
cd /workspace
wget https://www.xilinx.com/bin/public/openDownload?filename=resnet50-u50-r
```

- Install the model package.
  If the `/usr/share/vitis_ai_library/models` folder does not exist, create it first.

```
sudo mkdir /usr/share/vitis_ai_library/models
```

Then install the model package.

```
tar -xzvf resnet50-u50-r1.3.1.tar.gz
sudo cp resnet50 /usr/share/vitis_ai_library/models -r
```

3. Compile the sample, take `resnet50` as an example.

```
cd /workspace/demo/VART/resnet50
bash -x build.sh
```

4. Run the example, take `U50` platform as an example.

```
./resnet50 /usr/share/vitis_ai_library/models/resnet50/resnet50.xmodel
```

Note that different alveo cards correspond to different model files, which cannot be used alternately.

Launching Commands for VART Samples on U50/U50lv/U280

No.	Example Name	Command
1	resnet50	./resnet50 /usr/share/vitis_ai_library/models/resnet50/r
2	resnet50_mt_py	/usr/bin/python3 resnet50.py 1 /usr/share/vitis_ai_libra
3	inception_v1_mt_py	/usr/bin/python3 inception_v1.py 1 /usr/share/vitis_ai_library/models/inception_v1_tf/incep
4	pose_detection	./pose_detection video/pose.webm /usr/share/vitis_ai_l /usr/share/vitis_ai_library/models/ssd_pedestrian_prune
5	video_analysis	./video_analysis video/structure.webm /usr/share/vitis_ai_library/models/ssd_traffic_pruned_0_

No.	Example Name	Command
6	adas_detection	./adas_detection video/adas.webm /usr/share/vitis_ai_library/models/yolov3_adas_pruned_
7	segmentation	./segmentation video/traffic.webm /usr/share/vitis_ai_li
8	squeezenet_pytorch	./squeezenet_pytorch /usr/share/vitis_ai_library/models