Novatek AI2 SDK
Installation and Operation

Table of Content

1   Installation and Operation

1.1   Related Version

This document is based on the following H/W and S/W versions.

1.1.1   For NT9832X/NT9852X/NT9856X

- DLA Engine 2.5
- Nvt AI Tool: V1.1.03 or later
- Model bin format: V4.1
- Nvt AI2 SDK: 02.00.2012070 or later

1.1.2   For NT9833X

- DLA Engine 2.5
- Nvt AI Tool: V00.80.2106250 or later
- Model bin format: V4.1
- Nvt AI2 SDK: 02.02.2107050 or later

1.1.3   For NT98331

- DLA Engine 2.5
- Nvt AI Tool: for 32-bits using V01.03.2203310 or later, for 64-bits using V01.07.2211296 or later
- Model bin format: V4.1
- Nvt AI2 SDK: for 32-bits using 02.05.2204210 or later, for 64-bits using 02.15.2212010 or later

1.1.4   For NT98530

- DLA Engine 2.5   deep learning accelerator engine
- Nvt AI Tool: V01.04.2205170 or later

- Model bin format: V4.1
- Nvt AI2 SDK: 02.07.2205130 or later
- DSP bin: V1.0 or later
    - ☐ Nvt AI Tool: using V01.06.2209070 or later, to generate network with DSP op.
    - ☐ Nvt AI2 SDK: using 02.16.2209020 or later, to running network with DSP op

1.2 Directory of AI2 SDK

1. SDK related directory
   code/hdal/vendor/ai2
   code/hdal/vendor/dsp (only for NT9853X)

2. Sub-directory of AI2
   (Drivers)
   code/hdal/vendor/ai2/drivers
   (Header File)
   code/hdal/vendor/ai2/include
   (Source Code)
   code/hdal/vendor/ai2/source
   (Open Source Code)
   code/hdal/vendor/ai2/source_pub
   code/hdal/vendor/ai2/source_pub/include
   code/hdal/vendor/ai2/source_pub/vendor_ai_cpu
   code/hdal/vendor/ai2/source_pub/vendor_ai_dsp

3. Sub-directory of DSP (only for NT9853X)
   (Drivers)
   code/hdal/vendor/dsp/drivers
   (Header File)
   code/hdal/vendor/dsp/include
   (Source Code)
   code/hdal/vendor/dsp/source

1.3   Compilation Steps

1.3.1   NT9852X/NT9856X/NT98530

1. Download SDK codebase

2. Before compiling, please set the linux environment and switch to EVB project sample project
    $ source build/envsetup.sh
    $ lunch
    (Select :**Linux > cfg_IPCAM1_EVB > arm-ca9-linux-gnueabihf-6.4**)

3. Full compilation (Include AI)
    $ make all

4. The way to compile AI only (if the user has changed the code inside the source_pub directory)
    $ cd code/hdal/vendor/ai2; make clean; make; make install
    $ cd ../../../..
    $ mr
    $ make pack

5. After compiling, you can find F/W bin files in these directories
    - na510xx_linux_sdk/output/packed/FW9852XA.bin   (for 52x)
    - na510xx_linux_sdk/output/packed/FW9856XA.bin   (for 56x)
    - na51102_linux_sdk/output/packed/FW9853XA.bin   (for 53x)

6. Compilation steps of running network with DSP op (only for NT9853X)
    a.   Open file code/hdal/vendor/ai2/source_pub2/Makefile at about line 49:
       Change this setting yes or no; default is yes
       nn_dli_dsp = yes
       (If not require these DSP op, user could turn off it to reduce library code size.)
    b.   Execute above step 3 or 4 to build code

1.3.2 NT9832X/NT9833X/NT98331

1. Download SDK codebase

2. Before compiling, please set the linux environment and switch to EVB project sample project
   $ source build/envsetup.sh
   $ lunch
   (for 32-bits, Select : Linux > cfg_98332_8GX2_NAND_DVR_SYS
      > arm-ca53-linux-gnueabihf-8.4)
   (for 64-bits, Select: Linux > cfg_98332_8GX2_NAND_DVR_SYS_A64 > aarch64-ca53-linux-gnueabihf-8.4.01)

3. Full compilation (Include AI)
   $ make all

4. The way to compile AI only (if the user has changed the code inside the source_pub directory)
   $ cd code/hdal/vendor/ai2; make clean; make; make install
   $ cd ../../../..
   $ mr
   $ make pack

5. After compiling, you can find F/W bin files in these directories
   - na510xx_linux_sdk/output/packed/FW9832XA.bin   (for 32x)
   - na510xx_linux_sdk/output/packed/FW9833XA.bin   (for 33x)
   - na510xx_linux_sdk/output/packed/FW98331A.bin   (for 331)

1.4 Directory of Samples & Test Data

1. The AI-related samples
   code/hdal/samples/ai_*

2. The AI-related samples include the following files:
   (network processing with 1 input & 1 classify output)

code/hdal/samples/ai_net

(network processing with 1 input & 1 detection output)

code/hdal/samples/ai_dnet

(2 network processing)

code/hdal/samples/ai_2net

(network processing with multiple input & multiple output)

code/hdal/samples/ai_net_with_mblob

(network processing with dynamic multiple scale)

code/hdal/samples/ai_net_with_mscale

(network processing with dynamic multiple batch)

code/hdal/samples/ai_net_with_mbatch

(single operation)

code/hdal/samples/ai_op

(network processing with single operation)

code/hdal/samples/ai_net_with_op

(hdal liveview with 1 network processing)

code/hdal/samples/ai_video_liveview_with_net

(hdal liveview with 2 network processing)

code/hdal/samples/ai_video_liveview_with_2net

(network processing with DSP op) (only for NT9853X)

code/hdal/samples/ai_net_dsp

3. The models and data for test

   code/hdal/samples/ai_pattern

4. The models and data for test include:

   (The nvt_model bin file of net)

   code/hdal/samples/ai_pattern/para/nvt_model.bin

   code/hdal/samples/ai_pattern/para/nvt_model_*.bin

   (The yuv or rgb image files of input)

   code/hdal/samples/ai_pattern/jpg/*.rgb

   code/hdal/samples/ai_pattern/jpg/YUV420_SP_W512H376.bin

   (The referenced label file for output of classification net)

   code/hdal/samples/ai_pattern/accuracy/label.txt

   (The DSP firmware bin file) (only for NT9853X and running network with DSP op)

   code/hdal/samples/ai_pattern/DSP/DSP1.bin

1.5   Preparation of SD Card / USB Removable Drive

1.5.1   NT9852X/NT9856X/NT98530

1. Format SD card (/mnd/sd)

2. Copy the loader bin to SD card. (LD9852x.bin is for 52x, LD9856XT.bin is for 56x, LD9853xT.bin is for 53x)

3. Copy F/W bin to SD card. (FW9852x.bin is for 52x, FW9856x is for 56x, FW9853x is for 53x)

4. Copy the test files in the ai_pattern directory to the SD card..
   code/hdal/samples/ai_pattern

5. Make DSP directory in the SD card, and copy DSP1.bin to the directory (only for NT9853X and running network with DSP op).

6. Make sure these files under the /mnd/sd root directory (take 52x as an example)
   LD9852x.bin
   FW9852x.bin
   para/nvt_model.bin
   jpg/YUV420_SP_W512H376.bin
   accuracy/label.txt
   DSP/DSP1.bin (only for NT9853X and running network with DSP op)

1.5.2   NT9832X/NT9833X/NT98331

1. Format USB removable drive (USB disk)

2. Copy the test files in the ai_pattern directory to the USB removable drive (USB disk)
   code/hdal/samples/ai_pattern

---

3. Make sure that there are these files under the /mnd/sd root directory (take 32x as an example)

para/nvt_model.bin

jpg/YUV420_SP_W512H376.bin

accuracy/label.txt

1.6   Boot and Insert Linux AI module

1.6.1   NT9852X/NT9856X

1. Insert the SD card into the EVB and press the power or reset button to power on

2. After booting, please unmount the following old ko files of AI1 (if any)

    $ rmmod netflowsample

3. Insmod ko files of AI2

    $ insmod /lib/modules/4.19.91/hdal/kdrv_ai/kdrv_ai.ko

    $ insmod /lib/modules/4.19.91/hdal/kflow_ai/kflow_ai.ko

1.6.2   NT9832X

1. Insert the USB disk into EVB and press the power or reset button to power on

2. After booting, please unmount the following old ko files of AI1 (if any)

    $ rmmod netflowsample

3. Insmod ko files of AI2

    $ insmod /lib/modules/4.9.118/hdal/kdrv_ai/kdrv_ai.ko

    $ insmod /lib/modules/4.9.118/hdal/kflow_ai/kflow_ai.ko

1.6.3   NT9833X/NT98331

1. Insert the USB disk into EVB and press the power or reset button to power on

2. Insmod ko files of AI2
   $ insmod /lib/modules/4.19.148/hdal/kdrv_ai/kdrv_ai.ko
   $ insmod /lib/modules/4.19.148/hdal/kflow_ai/kflow_ai.ko

1.6.4   NT9853X

1. Insert the SD card into the EVB and press the power or reset button to power on

2. insert ko files of AI2
   $ insmod /lib/modules/4.19.148/hdal/kdrv_ai/kdrv_ai.ko
   $ insmod /lib/modules/4.19.148/hdal/kflow_ai/kflow_ai.ko

3. insert ko files of DSP (only for running network with DSP op)
   $ insmod /lib/modules/4.19.148/hdal/kdrv_rpc/kdrv_rpc.ko
   $ insmod /lib/modules/4.19.148/hdal/kflow_rpc/nvt_ipc.ko
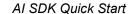   *$ insmod /lib/modules/4.19.148/hdal/dsp/nvt_dsp.ko*

1.7   Run the Basic Network Sample

1. Insert SD card into EVB and press power (or reset) botton to boot.

2. Execute the following steps:
   $ cd /mnt/sd
   $ ai_net 11 0 0

   Execute the following steps instead when running network with DSP op on NT9853x:
   $ cd /mnt/sd
   $ vendor_dsp_msg 0 &   (optional step: enable DSP log)
   $ ai_net_dsp 11 0 0

3. Then, the tips will be printed as follows:

usage:

enter q: exit

enter r: run

4. Please press *r* to start running the network, and press *q* to exit.

5. After running the network, the UART will output the following Top-5 results of classification:

Classification Results:

```
1. no=65, label=sea snake, score=0.284158

2. no=58, label=water snake, score=0.224791

3. no=67, label=diamondback, diamondback rattlesnake, Crotalus adamanteus, score=0.093708

4. no=54, label=hognose snake, puff adder, sand viper, score=0.068559

5. no=60, label=night snake, Hypsiglena torquata, score=0.045666
```

6. Parameter description for this basic (ai_net) sample

Parameter	Detail
0, 1, 10, 11	Job optimize mode
0, -1*	Job sync mode
-1, 0, 1, 2, 3	Buffer optimize option

*: -1 is only support with job optimize 0

2   Revision History

Revision	Date	Author	Changes
0.0.1	2020/10/03	Jeah Yen	1st version
0.0.2	2020/12/15	Jeah Yen	Modify SDK verson.
0.0.3	2021/1/19	Jeah Yen	Support NT9856X.
0.0.4	2021/1/26	Adam SY Su	Modify the description of insmod AI2 ko.
0.0.5	2021/6/30	K L Chu	Support NT9833X
0.0.6	2021/7/14	Boyan Huang	Modify the description of insmod AI2 ko.
0.0.7	2022/01/26	K L Chu	Support NT98331
0.0.8	2022/03/30	Jeah Yen	Update NT9853X
0.0.9	2022/05/30	Jeah Yen	Support NT9853X
0.1.0	2022/09/07	B B Jiang Omega Yu	Support DSP of NT9853X
0.1.1	2022/12/02	Grace TY Lee	Support A64 of NT98331