

**임베디드 AI 구현을 위한
Neural Accelerator 적용 및 딥러닝 네트워크
최적화 분석 과제**

지능통신시스템 연구실

개발 환경

[HW]

- Single Board Computer (CPU) : Raspberry Pi 4
- Neural Accelerator : Intel Neural Compute Stick 2 (NCS 2)

[SW – SBC/Neural Accelerator]

- OS : Raspbian (Debian Linux-based OS)
- Neural Accelerator 연동 플랫폼 : OpenVINO
- 가상환경 플랫폼 : Anaconda (사용 가능시 권장됨)
- 사용 언어 : Python 3

[SW – Host PC]

- 딥러닝 플랫폼 : PyTorch (버전 : 1.12 / Stable)
- 딥러닝 모델 변환 플랫폼 : ONNX
- 가상환경 플랫폼 : Anaconda
- 사용 언어 : Python 3 (버전 : 3.8 이상)

과제 1 : Raspberry Pi – NCS 2 연동 환경 구성

- ① Raspberry Pi 4에 Raspbian OS를 설치하여 Linux 개발 환경 구성
- ② NCS 2를 Raspberry Pi 4에 연결한 후 디바이스 인식
- ③ NCS 2를 이용한 딥러닝 예제 실행

[참고자료]

- RaspberryPI + Movidius NCS | OpenVINO™ toolkit | Ep. 37 | Intel Software
: https://www.youtube.com/watch?v=PNmH_ugW6Zw
- Get Started with Intel Neural Compute Stick 2 – Linux 또는 Raspbian 참고
: <https://www.intel.com/content/www/us/en/developer/articles/guide/get-started-with-neural-compute-stick.html>
- Raspberry Pi 4 and Intel Neural Compute Stick 2 Setup
: <https://www.intel.com/content/www/us/en/developer/articles/guide/raspberry-pi-4-and-intel-neural-compute-stick-2-setup.html>
- Arm7 Single Board Computers and the Intel Neural Compute Stick 2 (Intel NCS 2)
: <https://www.intel.com/content/www/us/en/developer/articles/technical/arm-sbc-and-ncs2.html>

과제 2 : NCS 2를 이용한 딥러닝 모델 로딩

① Raspberry Pi 4 – NCS 2 개발 환경 구성 후

아래 예시 자료를 참고하여 사전에 학습된 모델을 NCS 2에 로딩

② NCS 2에 사전 학습된 모델을 로딩하기 위한 과정 파악

[요구사항]

- NCS 2에 딥러닝 모델을 로딩하기 위한 필요사항 파악

(어떤 라이브러리를 가져와야하는지? / 모델을 로딩하기 위한 함수는 어떻게 사용하는지? / 딥러닝 모델은 어떤 종류여야 하는지?)

- NCS 2에 딥러닝 모델을 로딩하기 위한 과정 파악

(핵심 필요 함수 또는 API를 어떤 순서대로 사용해야하는가?)

[예시 자료]

- Raspberry Pi and Movidius NCS Face Recognition

: <https://pyimagesearch.com/2020/01/06/raspberry-pi-and-movidius-ncs-face-recognition/>

과제 3 : NCS 2를 이용한 YOLO 모델 실행

- ① Raspberry Pi 4 – NCS 2 개발 환경 구성 후
아래 예시 자료를 참고하여 사전에 YOLO 모델을 NCS 2에 로딩
- ② NCS 2에 사전 학습된 모델을 로딩하기 위한 과정 파악
- ③ NCS 2에 모델이 감당할 수 있는 딥러닝 네트워크 규모 파악

[요구사항]

- NCS 2에 YOLO 모델을 로딩하기 위한 필요사항 파악
(어떤 라이브러리를 가져와야하는지? / 모델을 로딩하기 위한 함수는 어떻게 사용하는지? / 딥러닝 모델은 어떤 종류여야 하는지?)
- NCS 2에 다양한 버전의 YOLO를 로딩하여 최대 감당할 수 있는 네트워크 규모 파악
(NCS 2와 같은 Neural Accelerator는 어느 정도 규모의 딥러닝 모델을 감당할 수 있는가?)

[예시 자료]

- YOLO and Tiny-YOLO object detection on the Raspberry Pi and Movidius NCS
: <https://pyimagesearch.com/2020/01/27/yolo-and-tiny-yolo-object-detection-on-the-raspberry-pi-and-movidius-ncs/>

과제 4 : NCS 2를 이용한 PyTorch 모델 실행

- ① Host PC에서 PyTorch Hub를 통해 Yolo5 모델을 준비함
- ② ONNX를 이용하여 PyTorch Yolo5 모델을 NCS 2 로딩에 적합하게 변환함
- ② NCS 2에 변환된 PyTorch 모델을 로딩하여 실행함

[요구사항]

- NCS 2에 PyTorch 모델을 로딩하기 위한 필요사항 파악

(PyTorch 모델을 로딩하기 위한 사전 작업 및 ONNX 변환 방법 파악?)

[참고 자료]

- How Can a PyTorch Model to Be Used with OpenVINO?
: <https://www.intel.com/content/www/us/en/support/articles/000089611/software/development-software.html>
- Converting a PyTorch Model - Typical Steps to Convert PyTorch Model
: https://docs.openvino.ai/latest/openvino_docs_MO_DG_prepare_model_convert_model_Convert_Model_From_PyTorch.html
- PyTorch – Yolo5
: https://pytorch.org/hub/ultralytics_yolov5/