

# CENG5030 Lab 06

## TensorRT

### 1 Sample Codes

#### Sample Codes

- Example codes:
  - `./Lab06-code/ceng5030.py`
  - `./Lab06-code/lab_utils.py`
  - `./Lab06-code/run-exp.sh`
- Installation script: `./Lab06-code/install-test.sh`
- pip requirements: `./Lab06-code/requirements.txt`
- Example image: `./Lab06-code/dog.jpg`

### 2 Assignment

#### Q1 Run VGG13 via TensorRT:

- Model: `https://pytorch.org/hub/pytorch_vision_vgg/`  
`torch.hub.load('pytorch/vision:v0.6.0', 'vgg13', pretrained=True)`
- Outputs:
  - inference time
  - classification label

#### Q2 Report layer names via `get_layer`:

- In Q1, you have called `builder.create_network` to create a network.
- Based on the returned network object, please print the layer names while generating the engine.
- Some useful functions or variables:
  - `tensorrt.Builder.create_network`
  - `tensorrt.INetworkDefinition.get_layer`
  - `tensorrt.ILayer.name`
- You can find them in `https://docs.nvidia.com/deeplearning/tensorrt/api/python_api/index.html`

**Submissions :**

- Source code of your implementations.
- Shell script files to run your code.
- A short .txt log file, in which you need to record your results.
- Put the above materials in a .zip file, and submit the zip file via blackboard.