

VOLTA

Repository with several multimodal encoders: LXMERT, ViLBERT, VisualBERT, etc.

How to install it in LISA cluster by SurfSara

1. Connect to LISA gpu
 - a. `$ ssh <yourname>@login-gpu.lisa.surfsara.nl`
2. Clone VOLTA git repository with submodules
 - a. `$ git clone --recurse-submodules https://github.com/e-bug/volta.git`
3. Load needed modules
 - a. `$ module load 2020`
 - b. `$ module load Anaconda3/2020.02`
4. Create conda virtual environment with python3.6 [I named it 'voltaenv' but can be any]
 - a. `$ conda create -n voltaenv python=3.6`
5. Activate virtual env (note that conda activate doesn't work)
 - a. `$ source activate .conda/envs/voltaenv` OR `$ source activate voltaenv`
6. Install VOLTA requirements
 - a. `$ cd volta`
 - b. `$ pip install -r requirements.txt`
7. Check pytorch and torchvision versions match required ones (if not, \$ conda install them)
 - a. `$ python`
 - b. `$ import torch`
 - c. `$ torch.__version__`
8. Install required condatoolkit
 - a. `$ conda install cudatoolkit=10.1 -c pytorch`
9. Load required modules too have CUDA 10.1.243 (same as pytorch) and GCC 8.3.0
 - a. `$ module load 2019`
 - b. `$ module load CUDA/10.1.243-GCC-8.3.0 [TBC]`
10. Install APEX
 - a. `$ cd apex/`
 - b. Try: `$ pip install -v --disable-pip-version-check --no-cache-dir --global-option="--cpp_ext" --global-option="--cuda_ext" ./` [NOTE: works with **CUDA/10.1.243-GCC-8.3.0**]
11. Setup the refer submodule for referring expression comprehension + install it
 - a. `$ cd tools/refer; make`
 - b. `$ python setup.py develop` [didn't work for me; see workaround below]
 - c. Workaround:
 - i. `$ vim setup.py`
 - ii. comment first line: `# from distutils.core import setup`
 - iii. add the line: `from setuptools import setup`
 - iv. save and quit
 - d. `$ python setup.py develop` [worked for me]

Case study: retrieval with Flick30k

1. This task (TASK 8 in config) is the only one which allows evaluating models pretrained on Conceptual Captions (CC) but without any finetuning with the option `--zero_shot`; see `~/volta/eval_retrieval.py`
2. We need to download Flick30k data
 - a. We create a folder in project space: `$ mkdir /project/dmg_data/data-volta`
 - b. `$ cd data-volta`
 - c. `$ mkdir flickr30k`
 - d. `$ cd flickr30k`
 - e. `$ wget data`
 - f. Run scripts to obtain captions as described at <https://github.com/e-bug/volta/tree/main/data/flickr30k>
 - g. Organize data following the structure described here (bottom page): <https://github.com/e-bug/volta/tree/main/data/flickr30k>
NOTE1: *.tsv files are not given
NOTE2: imdb contains two *.mdb files: data and lock
3. We create folder `sandro_examples` in `~/volta/`
4. We modify the file `~/volta/config_tasks` with our correct paths
 - a. `$ vim ctrl_test_tasks.yml`
5. We download pretrained models on CC; we will start with ViLBERT, which was the one originally tested in zero-shot setting (see: <https://arxiv.org/pdf/1908.02265.pdf>)

Pooled_output: the final representation (~[CLS] in BERT), not at each layer

To deactivate virtual environment:
`$ conda deactivate`