Simple Framework for Contrastive Learning

An illustration of SimCLR (from our blog here).

Enviroment setup

The code can be run on multiple GPUs or TPUs with different distribution strategies. See the TensorFlow distributed training guide for an overview of `tf.distribute`.

The code is compatible with TensorFlow 2.4+. See requirements.txt for all prerequisites, and you can also install them using the following command. `pip install -r ./official/requirements.txt`

Pretraining

To pretrain the model on Imagenet, try the following command:

```
python3 -m official.vision.beta.projects.simclr.train \
  --mode=train_and_eval \
  --experiment=simclr_pretraining \
  --model_dir={MODEL_DIR} \
  --config_file={CONFIG_FILE}
```

An example of the config file can be found here

Semi-supervised learning and fine-tuning the whole network

You can access 1% and 10% ImageNet subsets used for semi-supervised learning via tensorflow datasets. You can also find image IDs of these subsets in `imagenet_subsets/`.

To fine-tune the whole network, refer to the following command:

```
python3 -m official.vision.beta.projects.simclr.train \
  --mode=train_and_eval \
  --experiment=simclr_finetuning \
  --model_dir={MODEL_DIR} \
  --config_file={CONFIG_FILE}
```

An example of the config file can be found here.

Cite

SimCLR paper:

```
@article{chen2020simple,
  title={A Simple Framework for Contrastive Learning of Visual Representations},
  author={Chen, Ting and Kornblith, Simon and Norouzi, Mohammad and Hinton, Geoffrey},
  journal={arXiv preprint arXiv:2002.05709},
  year={2020}
}
```

SimCLRv2 paper:

```
@article{chen2020big,
  title={Big Self-Supervised Models are Strong Semi-Supervised Learners},
  author={Chen, Ting and Kornblith, Simon and Swersky, Kevin and Norouzi, Mohammad and
Hinton, Geoffrey},
  journal={arXiv preprint arXiv:2006.10029},
  year={2020}
}
```