Coral

Get started with the USB Accelerator

The Coral USB Accelerator is a USB device that provides an Edge TPU as a coprocessor for your computer. It accelerates inferencing for your machine learning models when attached to a Linux host computer.

This page is your guide to get started. All you need to do is download our Edge TPU runtime and Python library to the host computer where you'll connect the USB Accelerator. Then you can begin running TensorFlow Lite models.

If you want to learn more about the USB Accelerator, see the datasheet.

Requirements

- Any Linux computer with a USB port
  - Debian 6.0 or higher, or any derivative thereof (such as Ubuntu 10.0+)
  - System architecture of either x86_64 or ARM64 with ARMv8 instruction set

And yes, this means Raspberry Pi is supported. However, it must be Raspberry Pi 2/3 Model B/B+ running Raspbian (or another Debian derivative).

Also note that you should use a USB 3.0 port for the best inference speeds (unfortunately, Raspberry Pi has USB 2.0 ports only).

Setup for Linux or Raspberry Pi

To get started, perform the following steps on your Linux machine or Raspberry Pi that will connect to the Accelerator.

1. Install the Edge TPU runtime and Python library:

```
wget http://storage.googleapis.com/cloud-iot-edge-pretrained-models/edgetpu_api.tar.gz

tar xzf edgetpu_api.tar.gz

cd python-tflite-source

bash ./install.sh
```

> Caution: During installation, you'll be asked, "Would you like to enable the maximum operating frequency?" Enabling this option improves the the inferencing speed but it also causes the USB Accelerator to become very hot to the touch during operation and might cause burn injuries. If you're not sure you need the increased performance, type `N` and press Enter to use the default operating frequency.

> Help! If you see the message `./install.sh: line 116: python3.5: command not found`, then the install failed because you don't have Python 3.5 installed. So type `python3 --version` and press Enter. If it prints `Python 3.6` or higher, then open the `install.sh` script and edit the very last line to use `python3` instead of `python3.5`. However, if your Python version is lower than 3.5, you need to install Python 3.5.

2. Plug in the Accelerator using the provided USB 3.0 cable. (If you already plugged it in, remove it and replug it so the just-installed udev rule can take effect.)

Run a model on the Edge TPU

Now that your USB Accelerator is setup, you can start running TensorFlow Lite models on the Edge TPU.

For example, here's a demo that performs image classification with the parrot image in figure 1:

```
# From the python-tflite-source directory
cd edgetpu/

python3 demo/classify_image.py \
--model test_data/mobilenet_v2_1.0_224_inat_bird_quant_edgetpu.tflite \
--label test_data/inat_bird_labels.txt \
--image test_data/parrot.jpg
```

Figure 1. parrot.jpg

The results should look like this:

```
---------------------------
Ara macao (Scarlet Macaw)
Score :  0.613281
---------------------------
Platycercus elegans (Crimson Rosella)
Score :  0.152344
```

Congrats! You've just performed an inference on the Edge TPU.

Help! If you see the following message:

```
ERROR: Failed to retrieve TPU context
```

```
ERROR: Node number 0 (edgetpu-custom-op) failed to prepare
```

Then either your USB Accelerator is not plugged in or you cannot access it because your user account is not in the `plugdev` system group. Ask your system admin to add your account to the `plugdev` group, and then restart your computer for it to take effect.

This demo uses a Python API we created that makes it easy to perform an image classification or object detection inference on the Edge TPU. To learn more about the API, see the Edge TPU API overview & demos.

For details about how to create compatible TensorFlow Lite models, read TensorFlow Models on the Edge TPU.