

如何在 PYNQ 和 ZYNQ 上用 FPGA 加速神经网络

- 1、 在 PYNQ 上顶层用 Python 调用，但是准备工作是框架（如：tensorflow）和 IP（如：CNN），IP 部分要自己在 Vivado HLS（基于 OpenCL）里面用 C/C++ 或图形化界面进行设计，最后生成一个 .bit 和一个 .tcl 文件，再通过 Jupyter 把这俩文件上传至 PYNQ 的 memory。也就是把新的框架和 IP 自己手动放入 memory，然后就可以在顶层用 Python 完成所有调用，实现 ARM+FPGA。
- 2、 目前有两个例程，都是 ARM+FPGA：a、Spark+PYNQ+LR 进行 MNIST 模型训练；b、Theano+CNN 进行 MNIST 模型预测。且都对比了有 FPGA 加速和仅用 ARM 时的速度。
- 3、 ZYNQ 加速和 PYNQ 加速大体基本一致，就是在顶层用 C/C++ 调用已在 ZYNQ 上配置好的新框架和 IP

例程地址：a、[GitHub - AcceleratedCloud/SPynq: Spark on Pynq](#)

b、[GitHub - awai54st/PYNQ-Classification: Python on Zynq FPGA for Convolutional Neural Networks](#)