6、Profiling

本文采用AlexNet网络作为基准，评估Caffe的性能。AlexNet的网络结构如图【来源请求】，其主体由五层卷积层连接三层全连接层构成。为方便开发者测试网络性能，Caffe封装了time函数，对外以argument的形式开放。我们将time函数劫持出来，写到自定义源文件caffe-time.cpp的入口，放到/tools下随Caffe的其他源文件一起编译得到caffe-time.bin。编译完成后，我们采用这个命令作为性能评估基准：

./build/tools/caffe-time.bin --model=./models/bvlc\_alexnet/deploy.prototxt –iterations 50

caffe-time.bin代表含有time函数的可执行文件；model参数赋值为AlexNet的ProtoTXT格式网络描述；iterations为测试的迭代次数。由于第一次迭代开始之前，内存和CPU缓存是冷的，所以需要多次迭代测试取平均值来消除这种误差。“一次迭代”在这个语境下指对整个网络进行一次前向计算和反向传播。该命令等价于

./build/tools/caffe time --model=./models/bvlc\_alexnet/deploy.prototxt –iterations 50

执行命令之后，Caffe的输出范例如下【来源请求，Intel whitepaper】

我们对Armeabi-v7a和Arm64-v8a两种ARM架构进行测试。其中，Armeabi-v7a对应的测试设备为ZTE Nubia Z7；【来源请求，性能参数】Arm64-v8a对应的测试平台为Qualcomm 820开发板。【来源请求，性能参数】两个平台使用【commit号版本】的caffe-android-lib脚本进行交叉编译。考虑到Arm64-v8a的兼容性，使用NDK版本为【NDK r11c】