个人信息

- 董启凡/男/1990.08.17
- 手机： 18616265727
- Email： dongqifan1990@gmail.com

教育经历

2009.09-2013.06	本科	山东大学信息科学与工程学院	工学学士
2013.09-2016.06	研究生	山东大学信息科学与工程学院	工学硕士(推免)

工作经历

2016.06-至今　　酷芯微电子有限公司　　(机器视觉算法与优化工程师)

项目经历

1. CNN算法的嵌入式移植与优化 （2016.12 - 2018.04）

该项目是公司的重点项目，基本上一人完成，实现了将浮点caffe网络移植到DSP上的工作,支持AlexNet,GoogleNet,ResNet,MobileNet,Yolo,
MobileNet-SSD,MTCNN等主流网络.为下一代芯片奠定了运行CNN的软件基础,在芯片上实测macc利用率50%-60%左右（整个网络平均），主要包括两大部分:

- *1.1 CDNN*算法的*PC*端预处理
  Generator将PC端将预先训练好的caffemodel转换为dsp端读取的binary model，主要包括以下几点:
  - 权重系数的重排
    不同kernel的运算并行，在进行卷积计算之前，将n个kernel的系数重排，可以同时获取n的kernel的权重
  - 权重系数的量化
    基于文章Ristretto的dynamic fixed point方法，实现了generator端的权重系数的量化。包括16bit与8bit权重
  - caffe框架的修改
    基于caffelib来解析权重文件,完成量化与重排序工作，生成用于dsp端计算的网络模型.
- *1.2 CNN*算法的*DSP*端实现
  - CNN算法各个Layer的功能实现
    - ConvolutionLayer(包括group,depth wise以及不同的kernel与优化方法)
    - BatchnormLayer(与convolution layer,bias,scale,activation合并计算)
    - RegionLayer(YOLO),其中使用了expf的快速计算方法,在此基础上完成定点化、向量化

- Layer计算的速度优化
    - 针对128KB DTCM的片上内存的速度优化(**最终相对于512KB DTCM仅有低于5%的额外耗时**)
        - 新的Tile划分方式（channel方向划分）
        - 新的DMA搬运方式，充分复用小的DTCM,减少CPU负载，提升DMA带宽利用率，减少了queue memory的尺寸
    - 针对卷积kernel计算的优化
        - 通过合理设置pattern，使得权重的重排序不需要补0，将macc效率从75%提升到100%
        - 最大化利用每次load的数据，减少for循环次数
        - 合理设置stride，充分利用16个bank读写的并行
        - 11x11,7x7,5x5,3x3,1x1等各种情况(stride,512kb DTCM,128kb DTCM)的卷积优化
        - 8bit权重计算优化
        - 溢出问题的合理解决
        - 反汇编了CDNN2.0(ceva提供的收费版本)的卷积优化,弄清楚其优化策略并代码实现
    - 卷积优化最优的评价标准制定
        - 查看汇编代码看VPU,LSU单元是否全部被并行利用并软件流水
        - 计算macc利用率看是否达到预期
        - 计算数据搬运时间与数据处理时间，以及如何在mem限制的情况下最优划分tile策略
        - 搭建基于veloce的硬件仿真环境，通过dump波形查看性能瓶颈。并针对性修改
- CEVA软硬件bug定位并解决
    - CEVA编译器的bug的定位
      编译器在两条vector指令之间少插了nop，导致在寄存器的值尚未被写回的时候就被读取。通过dump波形，查看-o4优化选项的汇编代码并人肉定位到源代码位置并解决
    - DMA queue的约束
      现象是DMA搬运卡住，经过debug，发现必须queue_desc与queue_base没有做好读写先后的保护，导致desc的记数增加的同时desc没来得及写到mem。dma搬运不可预知。

2. 扫地机项目的支持 （2017.12 - 2018.03）

负责扫地机项目的识别算法的训练（前期）与嵌入式移植：

- 算法训练：
    - 数据增强：反转、噪声、random crop等
    - 修改网络结构，保证精度的前提下减少flops
- tiny-yolo网络算法的嵌入式移植
    - 将最新训练好的权重，最快部署到芯片，用于demo
    - 自动化浮点与定点评估工具

3. 下一代芯片的demo （2018.03 - 2018.04）

在软件SDK尚未ready的情况下，完成了tiny-yolo算法在裸版上的demo，能够从摄像头采集图像并最终将识别的结果显示在VGA接口的显示器上.该部分主要工作有：

- Baremetal CEVA多核的程序编写以及调试
- 芯片多核worst case的程序编写以及多核功耗测量
- ARM与CEVA多核通信
- Demo程序的编写以及调试

4. Miscellaneous (2016.07 - 2018.04)

- 协助DLA硬件组熟悉CNN算法并协同制定spec
- CEVA-XM4 dsp的培训文档、实验例程与移植guideline整理
- 协助ISP组移植畸变矫正、双目立体视觉匹配算法到CEVA
- Clang++编译器的研究，直接编译C/C++ codes并与ARM性能比较
- SLAM Eigen库的研究与优化方法探索，协助移植到CEVA
- 基于MobileNet-SSD，MTCNN等多个项目（物体检测、人脸检测）的支持

技能清单

- 语言水平：CET6
- 编程语言：C/C++,Python,Shell,Halide,Verilog
- 效率工具：GNU Make,CMake,QMake,Markdown,Git
- 开发环境：Visual Studio,Eclipse,PyCharm,Qt Creator,ModelSim,Cl,Gcc,Clang++,LLVM,Gdb
- 计算机视觉：OpenCV,Dlib,Caffe,Tensorflow,Pytorch
- 深度学习网络: AlexNet,GoogleNet,ResNet,SqueezeNet,MobileNet,ShuffleNet,Faster-RCNN,YOLO,SSD,MTCNN,FCN,Pruning,Quantization,Compression,IR
- 自我评价: Strong **Debugging,Learning,Engineering** Capability

荣誉与证书

- 2017.12　　上海市酷芯微电子公司优秀新人奖
- 2016.05　　山东大学2016届优秀毕业生
- 2015.12　　山东大学2015年度优秀研究生,山东大学2015年度光华奖学金
- 2012.10　　山东大学优秀学生奖学金
- 2011.10　　山东大学优秀学生奖学金、山东大学潍柴动力奖学金

文章作品

- Fall Alarm and Inactivity Detection System Design and Implementation on Raspberry Pi International Conference on Advanced Communication Technology, ***Advanced Communication Technology (ICACT)***, *2015 17th International Conference on*
- Sliced integral histogram: an efficient histogram computing algorithm and its FPGA implementation, ***Multimedia Tools and Applications***,*June 2017, Volume 76, Issue 12, pp 14327–14344*