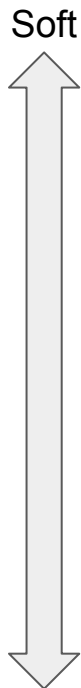


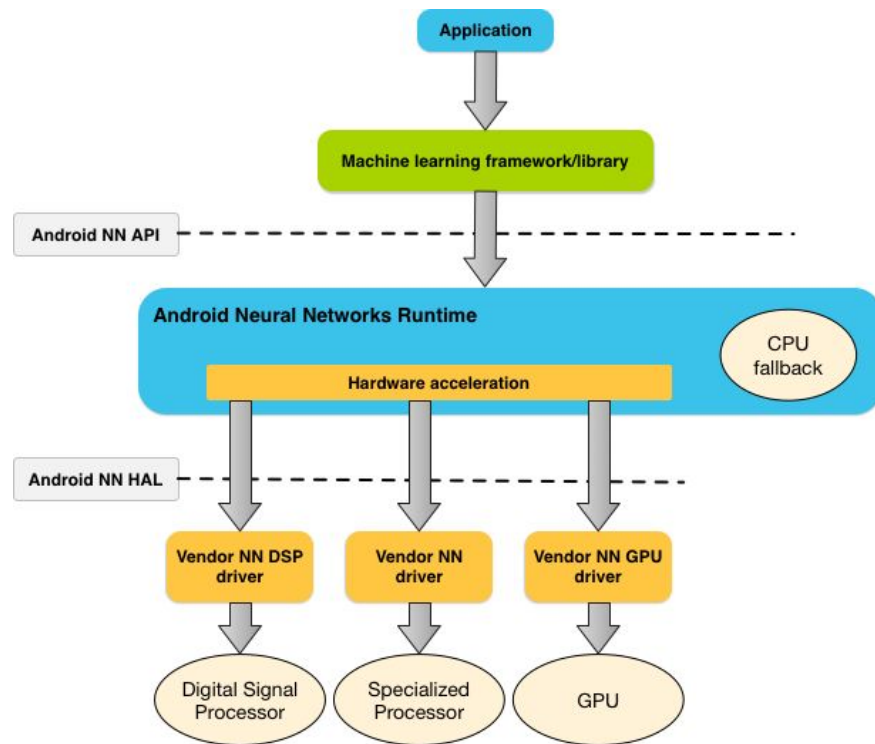
Entry Level Training for AI hardware Design

Chixiao Chen

AI Hardware Perspectives



1. **Algorithm level**
Run algorithms on HW constraints
2. **Embedded programming level**
Interface with your host devices and AI cores
3. **Hardware driver level**
According to the intermediate layer
Write lib for different functions
4. **RTL / HDL level** for deep learning
5. **Chip / Board level**, make it small!



Algorithm Level

-- Understanding the basic ideas of deep learning

Reading Material

The first CNN paper: Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, november 1998

Hands-on Labs:

(1) Run MNIST on Tensorflow / Caffe <https://www.tensorflow.org/tutorials/layers>

Extra Materials: Stanford Courses on DNN (Youtube) <https://youtu.be/vT1JzLTH4G4>

Embedded Program / Hardware SDK Level

-- Try to explore run a state-of-art network on you mobile phone

Reading Material

MobileNet: <https://arxiv.org/abs/1704.04861>

Hands-on Lab -- Build a mobileNet on you phone

(google/baidu some example online)

Tensorflow iOS: https://www.tensorflow.org/mobile/ios_build

Tensorflow Android: https://www.tensorflow.org/mobile/android_build

RTL / HDL level

To understand what's the difference between AI cores and conventional GPU/DSP

Reading Materials: (search and read the two papers via ieeexplore.org)

/////////
Vivienne Sze

Designing Hardware for Machine Learning

The important role played by circuit designers

Embedded Deep Neural Network Processing

*Algorithmic and processor techniques
bring deep learning to IoT and edge devices*

Answer the Question: Why do we need to design DNN specific circuits?
(compared with CPU and GPU)

RTL / Chip / Board Level

To investigate the what a real hardware communicate with your devices

Reading & Lab

Go through a reference PCIe design

mini PCIe FPGA: <https://github.com/RHSResearchLLC/PicoEVB>

And Investigate the linux kernel development in software:

PCIe DMA driver: <https://github.com/RHSResearchLLC/XilinxAR65444>

Assignment

-- Pick 3 out of 4 tracks and complete all the reading and labs.

(Embedded Program / Hardware SDK Level is mandatory!!!)

-- Prepare slides that demonstrates your achievements

(20-30 pages, not too much)

-- Due date: March 4th