Rice IP Applied to Neural Networks

(An Overview)

Introduction

Rice Digital Signal Processing (DSP) IP has significant neural network utility. In particular, the Company has developed unique frequency-domain processing capability. This includes;

- Novel “Numeric Building Blocks”
- Hierarchical structures composed of such Building Blocks

The relevance of this IP to Convolutional Neural Networks (CNNs) is summarized below.

Processing IP for CNNs

CNNs expend the bulk of their computations and power on the convolution process itself. This poses extreme computational requirements for either programmable or specialized processors. Modern CNNs often restrict convolution filter (kernel) sizes and weight resolution to compensate for this basic problem. This can reduce flexibility and complicate “training”.

A concept addressing these issues is convolution in the “frequency domain”. Mathematically, this provides a more efficient approach to performing convolution. However, the concept typically requires complex “DFT” or “FFT” structures. Addressing this issue, the Rice Processing IP provides structures of unique efficiency for frequency domain convolution.

As discussed elsewhere on the site, this includes novel “hierarchical” Processing Architectures. Examples of such Architectures are powerful Fast Fourier Transforms (FFTs), where;

- The FFT is constructed using smaller “DFT-like” sections
- The DFTs are based on specialized Numeric Building Blocks
- Said Building Blocks use only addition operations (no multiplies)

The proprietary Building Blocks minimize complexity, and afford extremely high performance. They enable basic DFT computations (of 4 to 20 bit operands) with the equivalent of two to three fixed-point adder circuits, (while not using publicly known techniques such as “Cordic arithmetic” or “residue number systems”).

This reduces complexity by orders-of-magnitude compare to conventional DSP processors, and enables;

- Extended dimensions and weight precision of convolution filters
- Acceleration of the convolution process
- Substantial reduction of size/power consumption of the CNN processor

The Company refers to these critical IP Building Blocks above as “Trigonometric Multipliers” (TMs). These simplistic structures (consisting of only a few hundred gates) can generate a trigonometric value, multiply said value by a random scalar input, and accumulate the product in 2 clock cycles. This allows DFT and FFT implementations of unparalleled efficiency. Performance can be comparable to very large scale FPGA cores, or conventional DSP processors. At the same time, the IP Building Block can be rendered with extremely compact ASIC or custom circuitry.

Summary

Use of frequency domain processing for CNN convolution layer(s) is well documented. Also, open literature has described the utility of "spectral pooling" (i.e. a pooling layer implemented in the frequency domain). Accordingly, the Company's IP has relevance to multiple steps (layers) in advanced CNN processing.

Rice Processing IP can complement or enhance future CNN architectures. Potential advantages of such a synergistic approach include:

- Accelerated processing
- Reduction of size, weight and power consumption
- Enhanced functionality (improvements in training and recognition performance)

The IP can therefore be of significant utility in the development of advanced Neural Network processors.

Rice Electronics

ricetronics@gmail.com

Filename: Neural Net Note 9-2023