講演番号：124

FPGA によるバイナリニューラルネットワークの高速化

D－2　　Acceleration of Binarized Neural Network by FPGA

島田　源†　黒木　啓之†
Gen SHIMADA†　Takashi KUROKI†
† 東京都立産業技術高等専門学校
† Tokyo Metropolitan College of Industrial Technology

1. はじめに

近年，機械学習の分野においてニューラルネットワークが多く用いられている．ニューラルネットワークを扱うには，膨大な積和演算を行う必要がある．この場合，多くは GPU を用いて計算を行うが，積和演算を行う新たなデバイスとして FPGA が注目されている．FPGA は高速・小型であり，様々な組み込み機器への実装が期待される．

Courbariaux ら[1]は，重みと活性化値を二値化したニューラルネットワークである BNN(Binarized Neural Network)の学習手法を報告している．また Umuroglu ら[2]は，BNN を FPGA に効率的に実装できるフレームワークを報告している．しかしながら高速化にはあまり言及されていない．そこで本研究では，FPGA を用いた BNN の推論における更なる高速化を目的とする．

2. FPGA による推論手法

ニューラルネットワークを扱う際には，計算精度は必要にならないことが多い．そこで，FPGA でニューラルネットワークを扱う場合には，重みと活性化値を 1 と -1 に二値化(1bit 化)し，計算を簡略化する手法を採用することで，FPGA のリソースを節約し，比較的大きいニューラルネットワークの実装を可能にしている．二値化した際の計算は XNOR を用いることによって実現できる．図 1 のように積和演算を，前層からの入力と重みで XNOR を取り，立っているビットの数をカウントすることに置き換えている．その結果が閾値を超えていれば次の層に出力を行う．

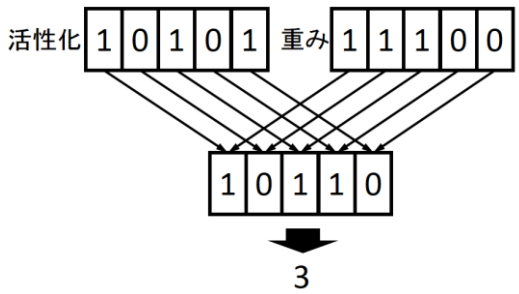

図 1 XNOR による積和演算

これを FPGA に実装するには，MVTU(Matrix Vector Threshold Unit) と呼ばれる回路を用い，ニューラルネットワークの層 1 層に対し MVTU1 つで計算を行う．また，MVTU 中にある PE(Processing Element) でユニット 1 つごとの積和演算を行う．この PE の数や，PE への入力数を適切に設定することで，効率化を図ることができる．

3. FPGA ボード

今回使用する FPGA ボードは Xilinx の PYNQ-Z1 で，PYNQ プロジェクトのハードウェアプラットフォームである．PYNQ とは，Python を用いて設計を行うことができる Xilinx のオープンソースプロジェクトである．

4. 推論時間の比較

Cifar10 学習済みモデルを用い，ARM CPU と FPGA の推論時間を比較した．表 1 にその結果を示す．ARM CPU を用いた場合より FPGA を用いたほうが 895 倍速い．この結果より，FPGA を用いることでニューラルネットの大幅な高速化が期待できることを確認できた．

また，学習させるデータセットをきゅうりの画像に変更し，ARM CPU と FPGA の推論時間を比較した．同じく表 1 にその結果を示す．ARM CPU を用いた場合より FPGA を用いたほうが 892 倍速い．この結果より，任意のデータにおいて高速化が期待できることを確認できた．

表 1 推論時間の比較

計算装置	推論対象	
	Cifar10 [μs]	きゅうり [μs]
ARM CPU	1451154	1443087
FPGA	1622	1618

5. まとめ

FPGA を用いてバイナリニューラルネットワークの推論を行い，FPGA の有用性を確認できた．今後は更なる高速化のため，FPGA への実装に適したネットワーク構造の検討を行う．

参考文献

[1] Matthieu Courbariaux et al. Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1. eprint ArXiv: 1602.02830, 2016.
[2] Yaman Umuroglu et al. FINN: A Framework for Fast, Scalable Binarized Neural Network Inference. In Proc. ACM/SIGDA ISFPGA, 2017.