

平成28年度 情報工学コース卒業研究報告要旨

高田・本田 研究室	氏 名	岡 本 卓 也
卒業研究題目	SW/HW 協調設計ツールを用いた DNN による手書き文字認識プログラムの実装	
<p>近年, Deep Learning の研究が進み, 文字や画像, 音声の認識を高い精度で行うことが可能になり, 注目を集めている. Deep Learning は多層構造のニューラルネットワーク (DNN) を用いた機械学習の手法である. 認識する対象の学習データをこのネットワークに入力し, 各層で学習が繰り返され, 特徴量が自動で計算される仕組みになっている. Deep Learning は従来技術と比べ大幅に認識の精度を向上させる技術であるが, そのために大量のデータを何度も学習する必要がある. 現在の機械学習では, 学習時は GPU コンピューティングによる計算が一般的となっている. 認識時も GPU を使用することは出来るものの, 機器への組み込みの観点からは, CPU より高速に演算が可能で, GPU より消費電力を抑えることが可能な FPGA を使用することが検討されている.</p> <p>しかし, Deep Learning のような, 発展目覚ましい技術はアルゴリズムの変化も速く, その都度ハードウェア記述言語 (HDL) による構造記述を変化させるのは手間がかかってしまう. 更に, HDL によるハードウェア回路設計はそれ自体を行える設計者が少ないという問題もはらんでいる.</p> <p>そこで注目を浴びているのが, C 言語を用いて FPGA の回路設計を行うことが出来る高位合成である. 従来は HDL による RTL 記述を直接記述していたため専門性も高く, 容易に HDL による構造記述を作り出すことが出来なかったが, 高位合成の台頭により回路設計が容易となった. C 言語から構造記述を生成するため, 直接 HDL で RTL を記述するよりもハードウェアの構成を変化させやすくなる. また, 高位合成を行うことで設計回路の検証や, デバッグの効率を高めることもできる.</p> <p>本研究では, DNN を用いた機械学習の手書き文字認識プログラムを, 高位合成を用いて様々な特徴を持つハードウェアへと変換する. これにより, Deep Learning のような大量のデータを処理するプログラムをいかに効率的にハードウェア化し FPGA に実装できるかを検証する.</p> <p>まず, 元となる Deep Learning のプログラムから, 1 プロセスで6層分の演算を行うプログラム singleloop を作成した. その後, 各層ごとにプロセス化した layer, それらのプロセスを並列化しパイプライン化を施した layer_pipeline, 面積の増加を抑えるよう, 処理時間の短いプロセスをまとめてプロセス数を減らした layer_compact と layer_compact2 を作成し, それぞれの実行時間, 面積などを評価した. また, パイプライン化を施したものについては並列実行が出来ているかを確認するため, プロファイリング行ってプロセスの実行状態を確認した. 結果的に, 各プログラムをハードウェア化することで, 組み込みプロセス上でのオールソフトの実行に比べて, singleloop は 9.17 倍, layer は 3.47 倍の実行速度が得られた. パイプライン化をすることにより, layer_pipeline は layer より 1.45 倍速くなり, layer_compact は layer_pipeline より, 使用面積が 5.85% 小さくなった. layer_compact2 は, layer_compact よりもプロセスの数を減らしたにも関わらず, 面積が増えてしまった.</p> <p>パイプライン化によって高速化は出来たものの, 演算量の違いから各層での処理時間の違いがボトルネックとなり, パイプライン化の恩恵を大きく受けることが出来なかった. 今後の課題としては, メモリアクセス方法の改善による高速化や, layer_compact2 の面積増加の原因究明などが考えられる. 更に, 各変数のビット数を削減し実装した際に, うまく動作しなかったことについての原因を究明していく. また, 手書き文字認識に留まらず, 顔認識プログラムなどにも積極的に取り組んでいきたい.</p>		