

C-004

動的部分再構成可能な FPGA を用いた
 複数の入力解像度に対応した超解像 CNN の実装
 Implementation of Super-Resolution CNN for Multiple Input Resolutions
 Using Dynamic Partial Reconfigurable FPGA

宇貞 勇輝[†]
 Yuki Usada

黒木 修隆[†]
 Nobutaka Kuroki

沼 昌宏[†]
 Masahiro Numa

1. はじめに

近年、物体認識や画像・映像の高画質化を、畳み込みニューラルネットワーク(CNN:Convolutional Neural Network)により実現する技術が注目されている。CNN の処理には膨大な計算量が必要であり、その処理を高速化するアクセラレータとして GPGPU (General Purpose Graphics Processing Unit) が一般的に用いられてきたが、消費電力が大きいという問題がある。そこで、GPGPU の約 10 分の 1 程度の消費電力で動作可能な FPGA (Field Programmable Gate Array) 上に CNN 専用の回路を実装することにより、低消費電力化を実現する研究が注目されている。FPGA の限られたリソース上で実装可能な超解像 CNN として、Y. Kim らが提案した HF-SRN (Hardware-Friendly Super-Resolution Network) [1] と、森らが提案したネットワーク [2] に改良を加えた RA-SRN (Resource-Aware Super-Resolution Network) が、少ないパラメータ数での超解像を実現している。しかし、HF-SRN と RA-SRN は複数の入力解像度に対応していない問題がある。

そこで本稿では、実装した回路を動的に変更する FPGA の動的部分再構成機能を用いて、入力解像度に対応した超解像 CNN 回路を再構成することで、複数の超解像 CNN を実装する手法を提案する。

2. 提案手法

2.1 動的部分再構成

FPGA の回路書き換えは、ビットストリームと呼ばれる回路の構成情報を、内部メモリに書き込んで保持した後、それをもとに内部の LUT (Look-Up Table) や FF (Flip Flop) などのロジックセルや内部配線を構成することで行われる。通常、FPGA の回路書き換えを行う前に、FPGA 全体の動作を停止させる必要がある。しかし、動的部分再構成では、FPGA 全体の動作を停止させることなく、回路書き換えを部分的に行うことができる。

図 1 に動的部分再構成の原理を示す。まず、FPGA のロジックセルを、動的部分再構成が行われるリコンフィギュラブル・ロジックの領域と、動的部分再構成が行われないスタティック・ロジック領域の 2 種類に分割する。

そして、パーシャル BIT ファイルと呼ばれる、リコンフィギュラブル・ロジックで実現する構成情報を表すビットストリームを生成し、それを FPGA に書き込む。これにより、リコンフィギュラブル・ロジックのみ回路の書き換えが行われる。その際、スタティック・ロジック部分の動作は停止しない。

また、リコンフィギュラブル・ロジックとして割り当てられたロジックセルは、動的部分再構成を用いて実装する回路で

の利用に限定される。

本稿では、このリコンフィギュラブル・ロジックを用いて超解像 CNN 回路を実装する。

2.2 実装対象とする超解像 CNN

本稿では、動的部分再構成を適用する超解像 CNN として、図 2 に示す RA-SRN を採用する。

森らが提案したネットワーク [2] に対して、超解像品質を維持しながらパラメータ数を削減するため、以下の変更を加えた。

- Residual Block を 2 段の DBBlock (Deep Bottleneck Block) に置換
 - 入力段にある DWConv を 2 段に削減
 - 一部の PWConv に対してグループ化畳み込みを適用
- また、i) の DBBlock については、もとの Residual Block から以下のような変更を施した。
- フィルタサイズ 1×5 の DWConv を 2×2 の DWConv 2 層に置換
 - 入力の DWConv を PWConv に変更
 - 入出力部分の PWConv にグループ化畳み込みを適用
 - ReLU を PWConv に置換

2.3 動的部分再構成の対象とする超解像 CNN 回路

RA-SRN に基づく超解像 CNN 回路に対して、動的部分再構成を適用する。図 3 に、従来手法と提案手法のそれぞれにおける超解像 CNN 回路実装方式を示す。提案手法では、入力映像の解像度に対応する超解像 CNN 回路を適宜再構成

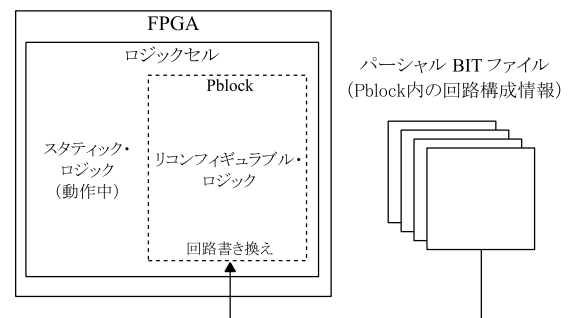


図 1 動的部分再構成の原理

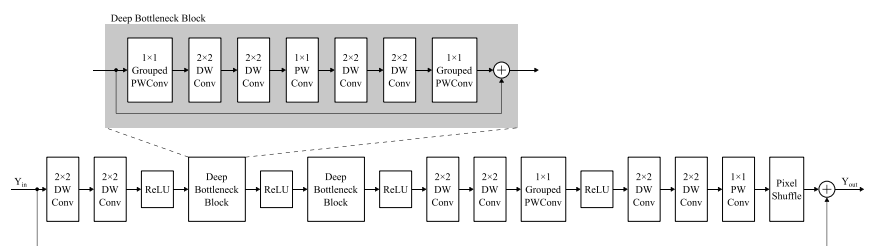


図 2 RA-SRN のネットワーク構造

[†]神戸大学, Kobe University

表 1 従来手法での利用リソース数

リソース	利用数		
	2 倍超解像	3 倍超解像	合計
LUT	58,697	60,961	119,658
LUTRAM	856	856	1,712
FF	31,841	32,603	64,444
DSP	1,977	2,050	4,027
BRAM	389.5	394.5	784

表 2 提案手法適用後の利用リソース数

リソース	利用数		
	2 倍超解像	3 倍超解像	リコンフィギュラブル・ロジック
LUT	60,676	63,571	233,760
LUTRAM	1,131	1,131	123,840
FF	34,677	35,413	467,520
DSP	1,979	2,050	2,232
BRAM	389.5	394.5	768

表 3 従来手法での利用リソース数 (DSP 削減後)

リソース	利用数		
	2 倍超解像	3 倍超解像	合計
LUT	83,929	89,999	173,928
LUTRAM	856	856	1,712
FF	42,956	45,186	88,142
DSP	1,401	1,399	2,800
BRAM	389.5	394.5	784

表 4 提案手法適用後の利用リソース数 (DSP 削減後)

リソース	利用数		
	2 倍超解像	3 倍超解像	リコンフィギュラブル・ロジック
LUT	86,124	92,355	155,520
LUTRAM	1,131	1,131	82,560
FF	45,671	47,966	311,040
DSP	1,403	1,399	1,440
BRAM	389.5	394.5	528

して実装することで、複数入力解像度に対応可能とする。

3. 評価実験と考察

3.1 実験内容

本稿では、動的部分再構成を用いた超解像 CNN 回路の複数入力解像度への対応手法を提案した。複数の超解像 CNN 回路を FPGA 上に複数実装する手法を従来手法として、利用リソース数を評価する 2 種類の実験を行った。

- i) 2 種類の超解像倍率をもつ、DSP 削減前の超解像 CNN 回路を用いた従来手法と提案手法との比較実験
- ii) DSP 削減後の超解像 CNN 回路を用いた、従来手法と提案手法との比較実験

DSP 削減については、乗算器を DSP ではなく LUT を用いて実装することにより、DSP 利用数を削減した。

3.2 実験結果と考察

DSP 削減前の超解像 CNN 回路に対して、表 1 に従来手法を適用した場合のマッピング結果を、表 2 に提案手法を適用した場合のマッピング結果を示す。従来手法では、2 種類の超解像 CNN 回路を実装するために必要なリソース数は、それぞれの超解像 CNN 回路で利用するリソース数の合計となる。一方で提案手法では、リソース利用数が多い 3 倍超解像 CNN

回路で利用するリソースのみである。そのため、提案手法では従来手法と比較して利用リソース数を約 30~50% 削減していることを確認した。

DSP 削減を行った超解像 CNN 回路に対して、表 3 に従来手法を適用した場合のマッピング結果を、表 4 に提案手法を適用した場合のマッピング結果を示す。両手法において LUT や FF の利用数は約 1.5 倍に増加するが、DSP 利用数を約 600 削減可能となった。その結果、提案手法について、DSP 削減前の超解像 CNN 回路と比較してリコンフィギュラブル・ロジックを 33% 削減可能であることを確認した。

4. まとめ

本稿では、FPGA に実装した回路を動的に書き換える動的な部分再構成機能を活用することで、入力映像の解像度に応じて超解像 CNN 回路を再構成する手法を提案した。

提案手法に基づく超解像 CNN 回路に関して実験評価を行った結果、従来手法と比べて約 30~50% 少ないリソース数で複数入力解像度への対応が可能であることを確認した。また、DSP の一部を LUT に置換して、DSP 利用数を削減することでリコンフィギュラブル・ロジックを約 33% 削減でき、周辺回路設計に必要なリソースを確保できることを確認した。

今後の課題として、回路全体のアーキテクチャ設計と、入力映像の種類に適した超解像 CNN 回路を再構成することによる超解像品質向上手法の検討が挙げられる。

参考文献

- [1] Y. Kim, J.-S. Choi and M. Kim, "A real-time convolutional neural network for super-resolution on FPGA with applications to 4K UHD 60 fps video services," IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 8, pp. 2521-2534, 2018.
- [2] 森 一紀, "FPGA 実装に向けた超解像 CNN のパラメータ削減", FIT2020, C-006, 2020 年 9 月.

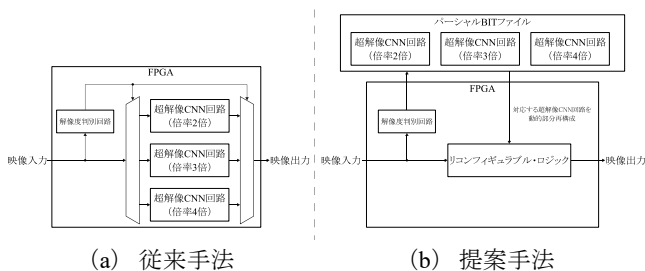


図 3 従来手法と提案手法での超解像 CNN 回路実装方式