ISSN 2186-5647

P-7

MobileNet を用いた超解像技術の提案

日大生産工　○大村 健斗　　日大生産工　山内　ゆかり

1. まえがき

現在、様々な表示機器において、大画面化・高精細化が必要とされている。従来の補完法ではボケやジャギーなど画像の劣化が生じる問題がある。そこで、元の画像に含まれない高周波成分を推定することで解像度を向上させる超解像と呼ばれる技術が生まれた。そのような超解像において畳み込みニューラルネットワーク(CNN)を利用し、高精度に細部の情報を推定・復元する研究が行われてきた。浦添らは、より少ないパラメータ数で構成されるCNNアーキテクチャ[1]を提案し、従来のMulti-Channel CNN (MCH)[2]と比べ、精度を示す同等のPSNRを維持しながら処理時間を1/4に削減したと報告されている。しかし、依然としてニューラルネットワークをベースとする超解像は線形補間方式や事例参照型と比べると計算コストが大きく、処理時間に対するPSNRは優れているとはいえない。そこで本研究ではMobileNet[3]を導入することでPSNRを維持しつつ処理時間の短縮を図り、CNN型超解像の性能向上を目指すアーキテクチャを提案する。

2. 従来研究

まず初めに、今回の構造の基本となるMCHを図1に紹介する。K倍の拡大をするにあたりK^2個の画素を出力する。例えば2倍拡大する場合、一つのCNNの2^2個の出力チャンネルから生成される。この構造の問題点として、PSNRの値に対して過剰な処理を行っていることがわかった。具体的には高周波成分を低層CNNで生成してしまっていた。

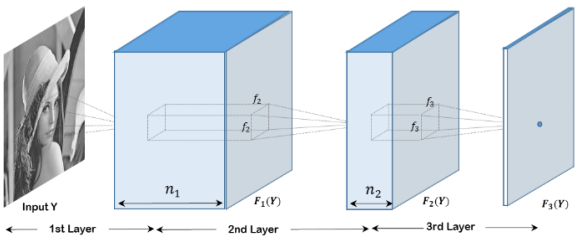

図 1　MCH の構造

上の問題点を解決するため、低周波成分は低層で、高周波成分は深層と伝搬経路を分ける構造にすることで、画質性能に対する相対的なパラメータの削減を目指したMulti-Path CNNが発案された。その中で最も良い結果を示した3Path-CNNの構造を図2に示す。これは低周波成分を伝搬するPath Low、高周波成分を伝搬するPath High、入力画像Yを伝搬するPath Originの3つの経路で構成されている。また畳み込みの計算は以下のように行う。WはCNNのフィルタ係数、Bはバイアス、Actは活性化関数を表しPReLUを用いる。

$$F_1(Y) = \mathrm{Act}(W_1 * Y + B_1) \quad (1)$$
$$F_2(Y) = \mathrm{Act}(W_2 * F_1(Y) + B_2) \quad (2)$$

二層目では特徴マップ$F_1(Y)$、$F_2(Y)$に加え入力画像Yを特徴軸方向に合成させ$F_{concat}(Y)$とする。

$$F_{concat}(Y) = F_2(Y) + F_1(Y) + Y \quad (3)$$
$$F_3(Y) = W_3 * F_{concat}(Y) + B_3 \quad (4)$$

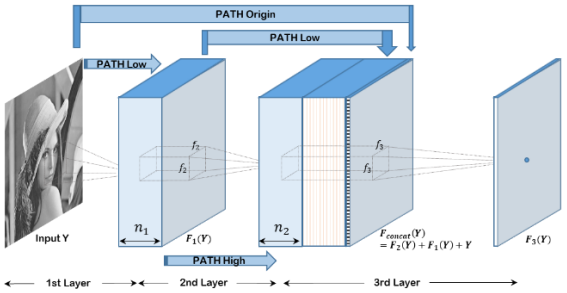

図 2　3Path-CNN の構造

この構造は周波数帯域を伝搬経路で分けるためフィルタW_1とB_1の事前学習が必要となる。そこで3Path-CNNの前に2Path-CNN、2Path-CNNの前に1Path-CNNを学習させ、値を継承させてから実験を行う。

次に、今回提案手法として組み込むMobileNetについて説明する。通常の畳み込み処理をDepthwise Conv.とPointwise Conv.の二段階に分けることでパラメータ数を削減し、処理時間の短縮を行う。この二つを合わせてDepthwise Separable Conv.と呼ぶ。一般的な畳み込みとの構成の違いを図3に示す。

出力側のチャンネル数をM、フィルタサイズをKとする通常の畳み込みにMobileNetを導入

Super-Resolution with MobileNet

Kento OMURA and Yukari YAMAUCHI

すると、演算量を$\frac{M+K^2}{MK^2}$倍に削減することができる。例えばM = 16、K = 5の畳み込みがあるときにMobileNetでは約1/10の演算量になる。

Depthwise Conv.では1チャンネルに1つのフィルタが対応しており、格チャンネルで畳み込みを行う。このとき、入力と出力のチャンネル数は変化しない。Pointwise Conv.では通常と同じ畳み込み処理を1×1のフィルタで行う。

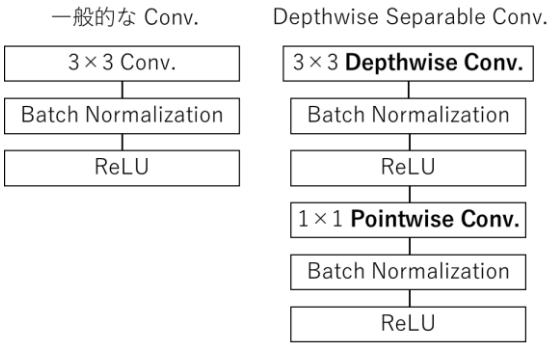

図 3　畳み込みの構成の違い

3.　提案手法

解決したい問題点は画質性能に対する処理時間の改善であり、従来手法の一般的な畳み込み処理にMobileNetを導入することで解決できると考える。これは事前学習を含めて解決したい問題点であるため、1Path-CNNや2Path-CNNも対象としている。3Path-CNN内でMoblieNetを導入する箇所を図4に示す。PathOriginは畳み込み処理を行っていないため手を加えず、従来通りの経路とする。

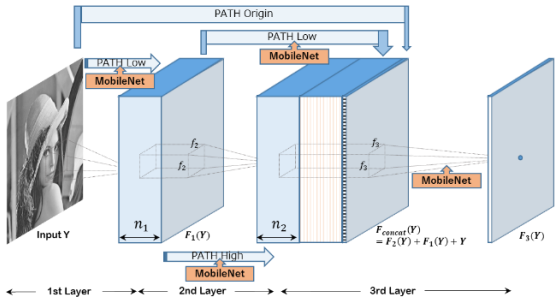

図 4　3Path-CNN の MobileNet 化

ここで複数の伝搬経路を持つ CNN 構造の表記を示す。例えば$n_1 = 8$、$n_2 = 24$のときの 2 倍拡大 3Path-CNN は(8-24+8+1-4)と表記することとする。

4.　実験および検討

従来手法と今回提案したMobileNetを導入したそれぞれの3Path-CNNで比較する。比較項目はPSNRと処理時間とする。2倍拡大の超解像を行うため、学習用画像をBicubic法で1/2に縮小した低解像度画像を入力信号として学習を行い、学習後にPSNRで評価する。学習率は最終層以外を10^{-4}、最終層を10^{-5}とし、フィルタサイズは中間層では5、出力層では3として行う。特徴マップ数(8-24+8+1-4)を用意する。

MobileNetを導入することで、図3に示した通り計算過程は増えるものの、計算量自体は削減されるが、[3]から精度は落ちてしまうと考えられる。

5.　まとめ

本研究ではＣＮＮを用いた超解像技術の処理時間の削減を目的に取り組んだ。一般的な畳み込み処理にMobileNetを導入することで従来手法より処理時間が短縮されることが予想されるが、高性能な線形補間法や事例参照型とも比較する必要があると考える。

今回の実験では三層のCNNを用いるため、[3]のMobileNetのモデル構造より十層ほど少ない構造になっている。そのためDepthwise Separable Conv.に変換すると精度自体が大きく損なわれる可能性が考えられるので、実装時には処理時間に対する精度のトレードオフも調べる必要がある。

参考文献

[1]浦添和哉, 黒木修隆, 加藤裕, 大谷真也, 廣瀬哲也, 沼昌宏, 複数の伝搬経路を持つ畳み込みニューラルネットワークによる超解像,電気学会論文誌C,2020年,140巻,6号,p.636-650
[2]Shinya OHATANI, Yu KATO, Nobutaka KUROKI, Tetsuya HIROSE, and Masahiro NUMA, Multi-Channnel Neural Networks for Image Super-Resolution, IEICE Transactions on Fundeamentals of Electronics, Communications and Computer Sciences, E100-A No.2, pp572-580, (2017)
[3]Andrew G. Howard, Menglong Zhu, Bo Chan, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hatwig Adam, Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications, CoRR, abs/1704.04861, 2017