

畳み込みの多段階分解による Semantic Segmentation の高速化

Fast semantic segmentation with multistage separable convolution

野田 紘司 久徳 遙矢 川西 康友 出口 大輔 井手 一郎 村瀬 洋
K. Noda H. Kyutoku Y. Kawanishi D. Deguchi I. Ide H. Murase

名古屋大学
Nagoya University

1 はじめに

近年、自動運転や高度運転支援を目的とした高速かつ高精度な車両周囲環境認識技術の重要性が高まっている。その中でも Semantic Segmentation は画像の画素単位でラベル推定を行なうシーン理解の要素技術として注目を集めている。しかし、Semantic Segmentation は畳み込み演算を含む深層学習の枠組みを用いた手法であり、高精度な Semantic Segmentation 手法の計算コストは高く、自動運転などの実時間処理を要する応用を考えた場合、実用可能なレベルの実行速度は得られていない。従って、精度を維持したまま、より高速に計算する手法が求められている。

Semantic Segmentation の処理は、畳み込み演算の計算時間が大部分を占める。そこで本発表では、畳み込み演算の計算時間削減による Semantic Segmentation の高速化手法を提案する。

2 フィルタの垂直・水平方向への分解

畳み込み演算の計算時間は、フィルタのパラメータ数と相関があることが知られている。そこで、パラメータ数を減少させることで積和演算の回数を削減し、計算時間の削減を目指す。

高精度な Semantic Segmentation 手法として知られている DeepLab v3+[1] の Xception モジュール (図 1) では、通常の畳み込み演算を、空間方向の畳み込み (Depthwise convolution) とチャンネル方向の畳み込み (Pointwise convolution) からなる Separable convolution (図 1 中 sep conv) にすることでパラメータ数を削減している。一方、図 2 のように通常の畳み込み層を垂直・水平方向へ分解することにより、パラメータ数を削減する手法が提案されている [2]。一般に、垂直・水平方向の分解により、斜め方向の特徴を捉えることが困難になり、精度が低下すると言われている。ここで道路環境を考えると、街灯や電柱、標識などのように垂直・水平方向の構造物が多く存在することに気づく。従って、それらの構造物は垂直・水平方向の特徴が認識精度に大きく寄与し、斜め方向の特徴の影響は小さいと考えられる。そこで本手法では、Xception モジュールの Depthwise convolution を垂直・水平方向に分解することで畳み込み演算のさらなる高速化を試みる。具体的には、図 1 中の赤枠の部分分割する。この変更により、パラメータ数は分割した

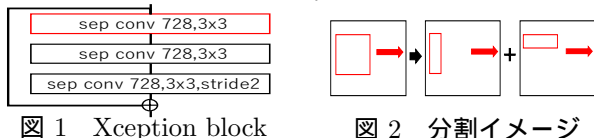


図 1 Xception block

図 2 分割イメージ

表 1 mIoU, パラメータ数及び学習・推定時間の変化

	mIoU	パラメータ数	学習 [sec]	推定 [sec/枚]
比較	0.772	37,682,112	4,831	18.56
提案	0.743	37,647,168	4,663	17.86

層のチャンネル数の 3 倍ずつ減少する。つまり、1 層ごとに $728 \times 3 = 2,184$ ずつ、合計 $2,184 \times 17 = 37,128$ のパラメータを削減する。

3 実験及び考察

提案手法の有効性を確認するために、ラベル推定精度及び実行時間の評価実験を行った。比較手法を従来の Xception を利用する DeepLab v3+ とし、データセットには Cityscapes[3] を用いた。フルスクラッチの学習には時間を要するため、あらかじめ学習されたモデルを利用し、提案手法による変更を伴う層のみを再学習した。実行時間の測定には Intel Core i7-3770K CPU @ 3.50GHz を使用し、100 反復の学習時間、画像 500 枚の平均推定時間を評価した。結果を表 1 に示す。提案手法により、mIoU の低下を 0.03 以下に抑えつつ、パラメータ数と実行時間を削減可能なことを確認した。本手法によるセグメンテーション結果例を図 3 に示す。

4 むすび

本発表では、Xception における畳み込み層の垂直・水平方向への分解による Semantic Segmentation の高速化手法を提案した。実験により、精度の低下を極力抑えつつ、処理の高速化が可能なことを確認した。今後の課題として、畳み込み層の更なる演算量の削減、時間方向の類似度を利用した手法の検討などが挙げられる。謝辞 本研究の一部は科学研究費補助金による。

参考文献

- [1] L. Chen et al., "Encoder-decoder with atrous separable convolution for semantic image segmentation," ECCV, 2018
- [2] A. Paszke et al., "ENet: A deep neural network architecture for real-time semantic segmentation," arXiv, 2016
- [3] M. Cordts et al., "The Cityscapes dataset for semantic urban scene understanding," CVPR, 2016



図 3 提案手法によるセグメンテーション結果例