

Sprint15-論文まとめ-

MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

どんなもの？

MobileNetsは、軽量のディープニューラルネットワークを構築するために、depth-wise separable convolutionsを使用する合理化されたモデル。リソースと精度のトレードオフに関する広範な実験を行い、ImageNet分類における他の一般的なモデルと比較して高い性能を示した。

どうやって有効だと検証した？

物体検出、ファイングレイスケール分類、顔属性、大規模ジオローカリゼーションなどの幅広いアプリケーションやユースケースにおいて、モバイルネットの有効性を実証。

技術の手法や肝は？

MobileNetモデルの肝は、通常の畳み込みを depth-wise separable convolutionsと Pointwise Convolution (1x1 Convolution)に分割したことにある。標準的な畳み込みを深さ方向畳み込みと 1x1の畳み込みの2つに分解する手法。

議論はある？

なぜ分解すると早くなるのか。

次に読むべき論文は？

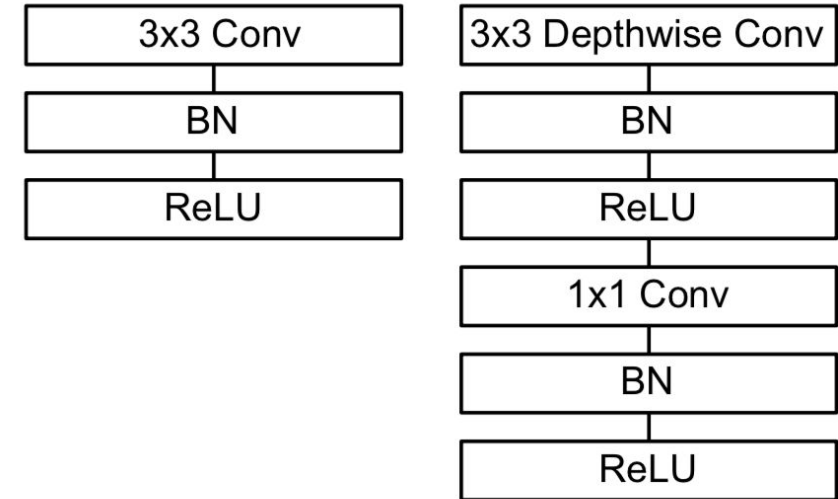


Figure 3. Left: Standard convolutional layer with batchnorm and ReLU. Right: Depthwise Separable convolutions with Depthwise and Pointwise layers followed by batchnorm and ReLU.

先行研究と比べて何がすごい？

MobileNet v1では通常のConvolutionをこのDepthwise Separable Convolutionに変えて、13段重ねることで、約1/8 ~ 1/9に総演算量を削減している。

MobileNetV2: Inverted Residuals and Linear Bottlenecks

どんなもの？

mobile netの改良版。

前回のDepthwise Separable Convolutionにおいて、Pointwise Convolutionの計算量(パラメータ数)が大きいため、これを減らす為に、Depthwise Separable Convolutionに代わってInverted Residualを導入した。

どうやって有効だと検証した？

特徴抽出器としての MobileNetV2とMobileNetV1の物体検出性能 [33]を、シングルショット検出器(SSD)の改良版[34]を用いてCOCOデータセット[2]上で評価・比較した。また、ベースラインとして YOLOv2 [35] とオリジナル SSD (VGG-16 [6] をベースネットワークとする) との比較も行っている。我々はモバイル /リアルタイムモデルに焦点を当てているため、Faster-RCNN [36]やRFCN [37]のような他のアーキテクチャとの性能比較は行っていません。

技術の手法や肝は？

Inverted Residual

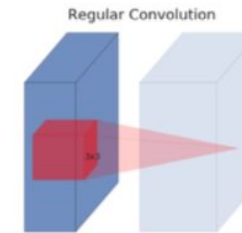
ResNetで使われるResidual を逆にしたもの。

Bottleneck構造と呼ばれていて、3x3Depthwise Conv.を1x1Conv.(Pointwise Conv.)で挟む構造になっており、mobile net1の1x1Conv.(Pointwise Conv.)の計算量を減らしている。

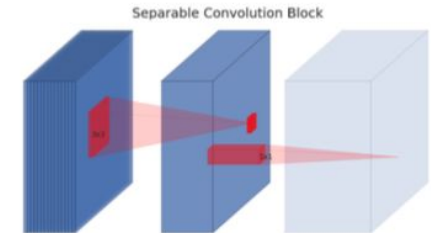
議論はある？

特になし。分類、物体検出などで比較しており、十分な比較ができていると思われる。

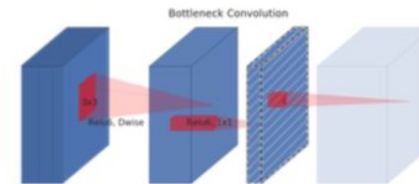
(a) Regular



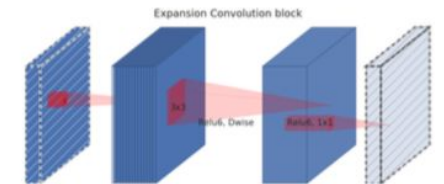
(b) Separable



(c) Separable with linear bottleneck



(d) Bottleneck with expansion layer



先行研究と比べて何がすごい？

SSDLite(SSDに今回の mobile netを組み合わせたモデル)は、Yolov2よりも計算量が20倍、パラメータが10倍削減されます。

次に読むべき論文は？

MobileNetV3

Searching for MobileNetV3

どんなもの？

この論文では、次世代の高精度で効率的なニューラル・ネットワーク・モデルを提供するために、MobileNetV3 LargeモデルとSmallモデルを開発し、オンデバイス・コンピュータ・ビジョンを実現するためのアプローチについて説明します。この新しいネットワークは、最先端の技術を前進させ、効率的なモデルを構築するために自動化された検索と新しいアーキテクチャの利点を組み合わせる方法を実証しています。

どうやって有効だと検証した？

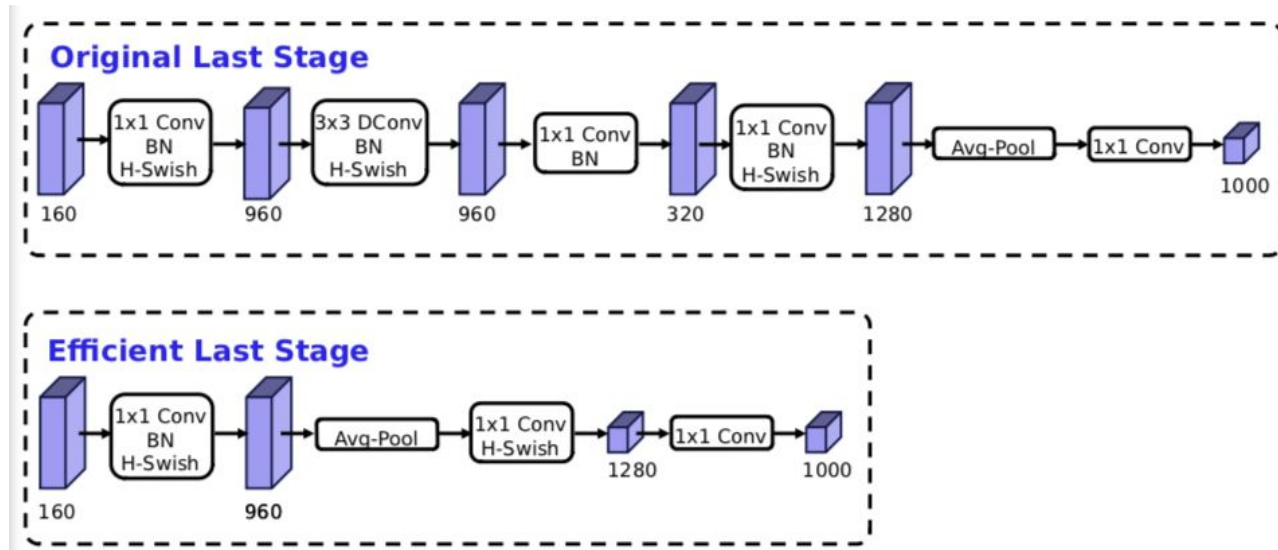
新しいMobileNetV3モデルの有効性を実証するための実験結果を示す。
分類、検出、セグメンテーションを実施した。

技術の手法や肝は？

bottle_neck構造にSqueeze-and-Excitationと言う手法を追加することで表現を増やしている。
InputデータからGAPで各チャンネルごとの代表値をとって(Squeeze)、それをinputとして全結合層(上図に置ける奥の経路)に入れて各チャンネルの重みを計算したのち最終的には元々のInputデータとその重みを掛け算する
GAPとは各チャンネルごとに平均値をとり、それを最終のSoftmaxする。
この手法により、最終的な計算量を減らしている。

議論はある？

特になし。



先行研究と比べて何がすごい？

mobile net2に比べ、3は精度の向上とデータセットによっては高速化も実現できている。

次に読むべき論文は？

EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

どんなもの？

モデルの精度を上げるために、そのモデルの大きさをスケールアップするというのは常套手段である、多くのリソースが利用可能な場合には精度を向上させるためにスケールアップされる。ConvNetsのスケールアップ方法は多種多様でありあまり理解が進んでいない。本論文では、モデルのスケールアップを系統的に研究し、ネットワークの深さ、幅、およびリソースのバランスを慎重にとることで、より良い性能が得られることを明らかにした。

どうやって有効だと検証した？

MobileNets と ResNets にCompound Scaling(複合スケールアップ)を用いて調整し、実験した。

技術の手法や肝は？

Compound Scaling(複合スケールアップ)

ネットワーク幅(w)、奥行き(d)、解像度(r)を個別に増やすだけでは効果が出ないため、それぞれをバランスよく調整するための手法。

議論はある？

他のモデルでも実験したい。逆に複合スケールアップ法が有効でないモデルはあるのか。

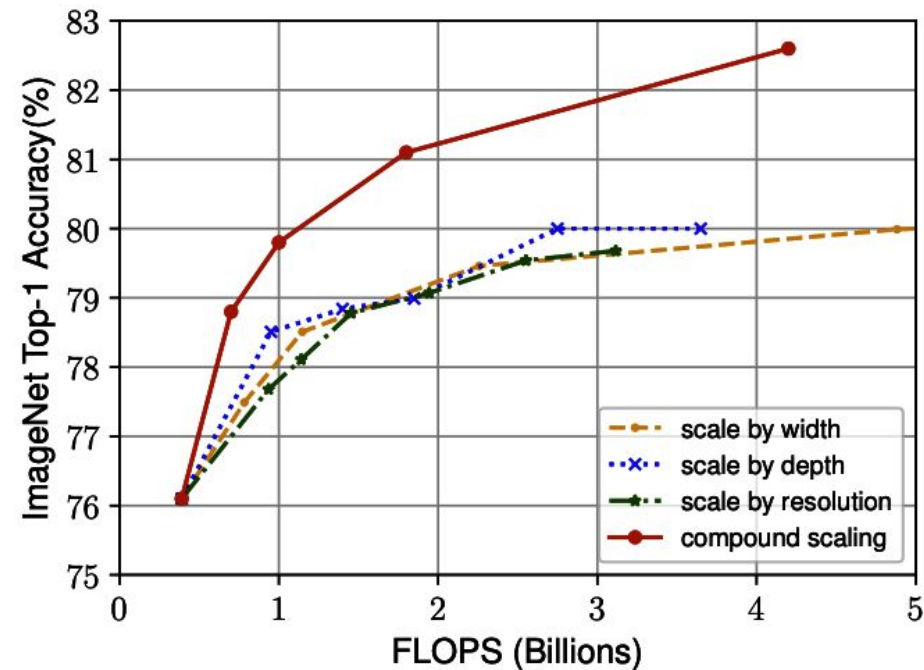


Figure 8. Scaling Up EfficientNet-B0 with Different Methods.

先行研究と比べて何がすごい？

モデルを変えずに、ネットワーク幅(w)、奥行き(d)、解像度(r)のスケールアップだけで、計算量を増やさずに、精度を上げることができる。

次に読むべき論文は？

参考文献で読んだのをここにお書き

Searching Beyond MobileNetV3

どんなもの？

今日まで、すべてのモバイルメソッドは、主に CPUのレイテンシに焦点を当てています。GPUは、後者は、しかし、そのために、実際にははるかに好ましいです。ターゲットとするハードウェアを念頭に置いて、我々は最初のモバイル

Mobile GPU-Aware (MoGA)のニューラルアーキテクチャは、以下のような側面で進めています

:モバイルCPUからモバイルGPUへの検索傾向のシフト

:従来の多目的最適化を重み付きフィットネス戦略

どうやって有効だと検証した？

mobile net3やMnas_netなどの最新のモデルと比較した。

技術の手法や肝は？

MoGA

GPUでの処理に特化した探索pipeline

高速な評価器として訓練されたスーパーネット、

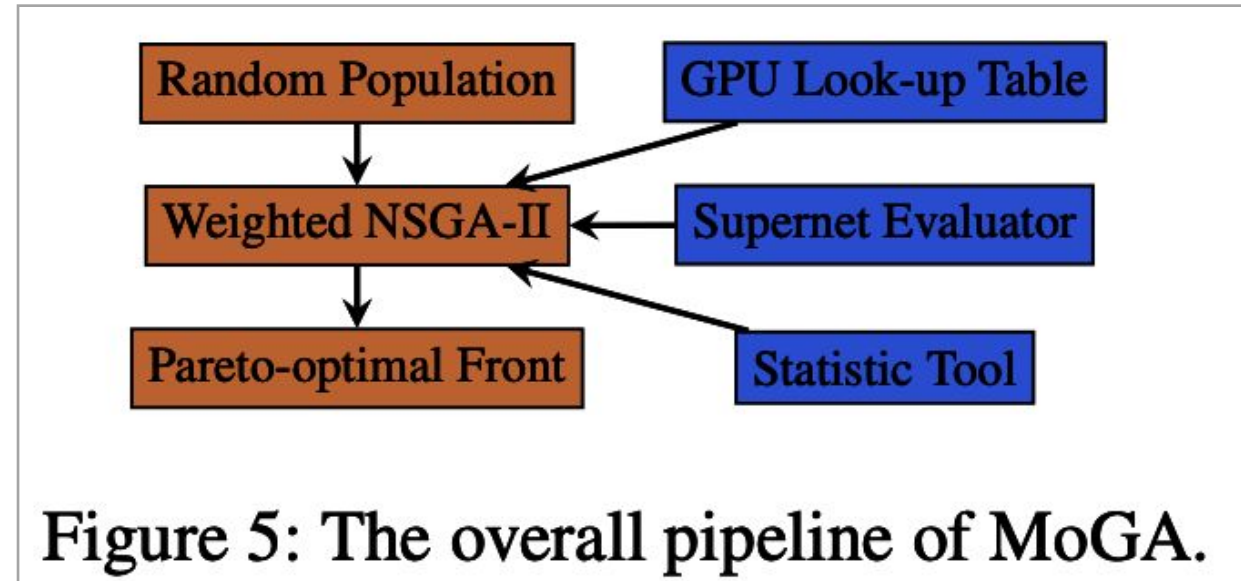
GPUレイテンシルックアップテーブル、およびパラメータの数を計算するための統計ツールがあります。

初期のランダム母集団は、有意な速度を向上させることができます。

パイプラインは120世代を母集団サイズ70で進化させ、これら8400のモデルを評価するのに約1.5 GPU日しかかかりません。

議論はある？

特になし



先行研究と比べて何がすごい？

MnasNetより200倍少ないGPU日数で、以下の条件でMobileNetV3を上回るモデルが得られます。つまり、MoGA-A は ImageNet 上で 75.9%のトップ 1 精度を達成し、MoGA-B は 75.5%を達成しています。

次に読むべき論文は？

Squeeze-and-Excitation Networks