EfficientNet:

Rethinking Model Scaling for Convolutional Neural Networks

사공진

목차

Abstract - 두 마리 토끼를 모두 잡은 EfficientNet

1.Neural Architecture Search를 사용해 baseline 네트워크를 만들고, scaling을 수행

2.이때, depth, width, resolution을 동시에 scaling

3.기존의 ConvNet보다 8.4배 작고, 6.1배 빠르다.

Introduction

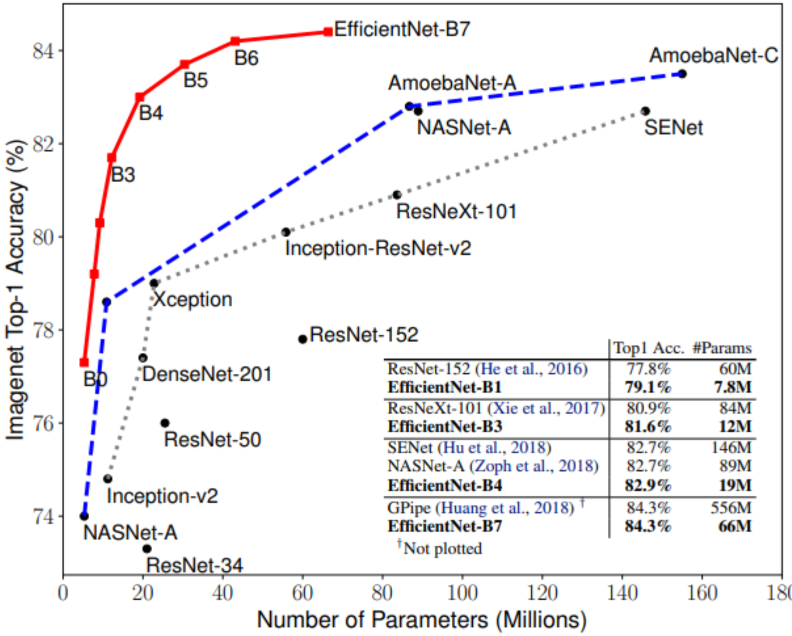


Figure 1. **Model Size vs. ImageNet Accuracy.** All numbers are for single-crop, single-model. Our EfficientNets significantly outperform other ConvNets. In particular, EfficientNet-B7 achieves new state-of-the-art 84.3% top-1 accuracy but being 8.4x smaller and 6.1x faster than GPipe. EfficientNet-B1 is 7.6x smaller and 5.7x faster than ResNet-152. Details are in Table 2 and 4.

1.기존의 ConvNet은 세 가지 중 나머지를 고정하고, 한 가지를 늘리는 방식. 주로, depth.

Ex)From ResNet-18 to ResNet-200

2.저자는 3가지를 골고루 scaling up하면 좋지 않을까 하고 생각

3.How to?
Scaling each of them with constant ratio

Introduction

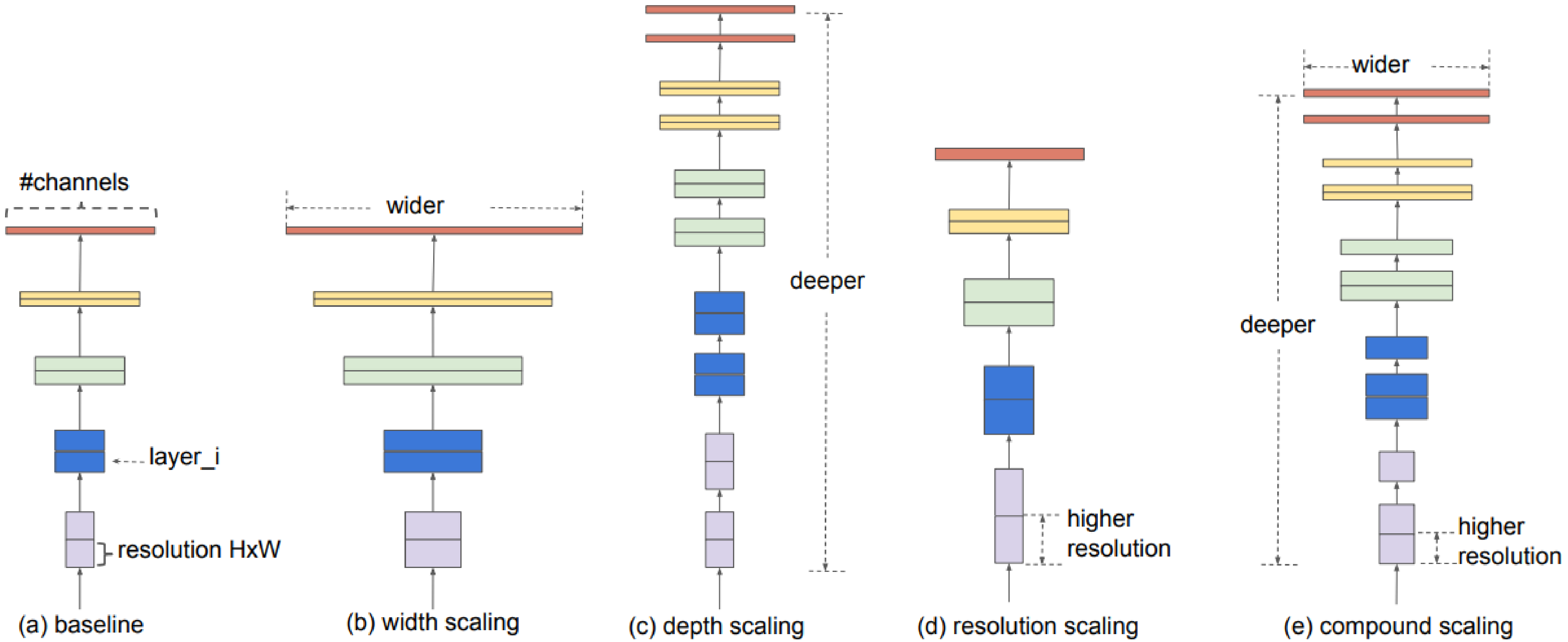


Figure 2. **Model Scaling.** (a) is a baseline network example; (b)-(d) are conventional scaling that only increases one dimension of network width, depth, or resolution. (e) is our proposed compound scaling method that uniformly scales all three dimensions with a fixed ratio.

Related Work

1.Accuracy

성능이 좋아짐에 따라 무거워질 수 밖에 없음. 이제는 Efficiency를 고려해야 할 때!

GoogleNet: 74.8% accuracy with 6.8m parameters

SENet: 82.7% accuracy with 145m parameters

Gpipe: 84.3% accuracy with 557m parameters

2.Efficiency

주로 Model Compression을 통해 효율성을 높임. But, 큰 모델에 적용하기엔 불확실

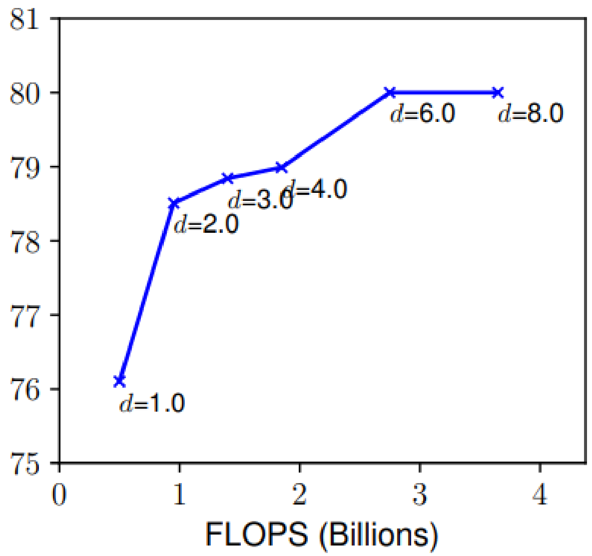
Ex)SqueezeNets, MobileNets, ShuffleNets and MNasNet which is basis of EfficientNet

Compound Model Scaling

1.Scaling dimensions - depth

Depth가 깊어짐에 따라 성능이 비례해서 올라가나, 일정 depth부터 별다른 개선 없음

Ex) ResNet-101과 ResNet-1000의 경우, accuracy에서 별반 차이가 없음.



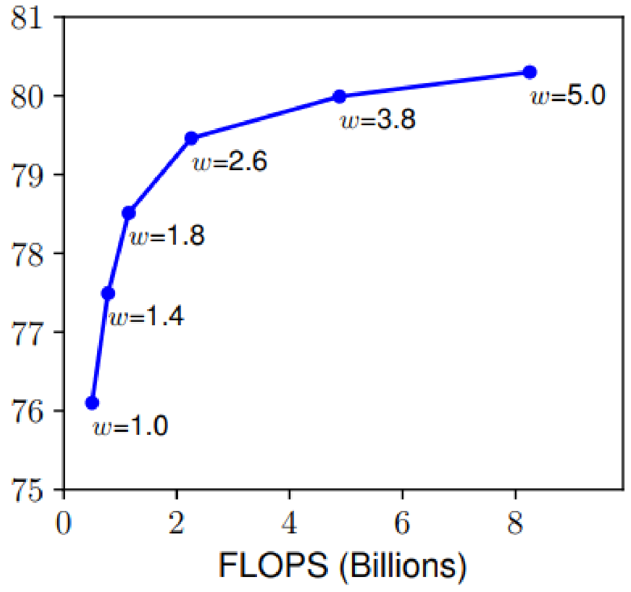
Compound Model Scaling

2.Scaling dimensions – width

Width가 넓어짐에 따라 fine_grained feature를 얻기 쉽고, train하기 쉬움.

But, 넓어질수록 high level feature를 얻기 어려움

Depth를 늘릴때와 마찬가지로 일정 width부터 별다른 개선 없음



Compound Model Scaling

3.Scaling dimensions – resolution

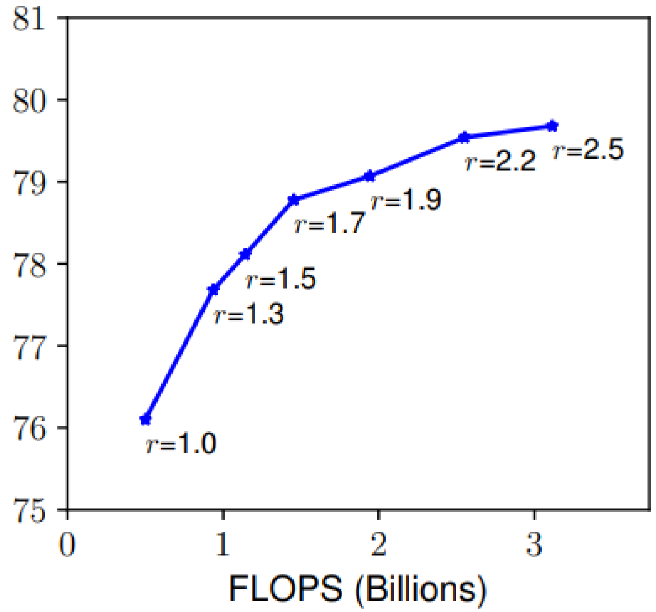
위 사례와 유사

Early ConvNets – 224 x 224

Inception v3 – 299 x 299

ResNet – 331 x 331

Gpipe – 480 x 480

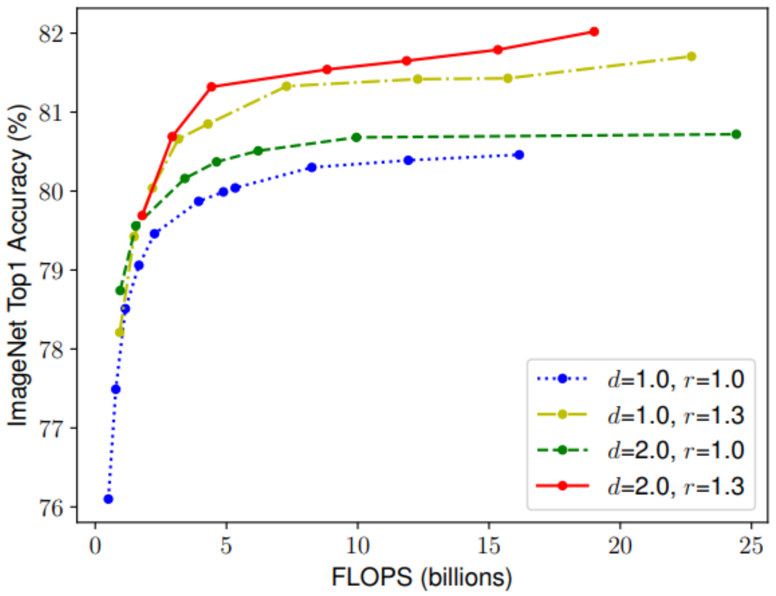


- Observation: 세 가지 경우 모두 모델이 커짐에 따라, 성능이 좋아지지만 점점 증가량이 감소한다.

Compound Model Scaling

Bigger Input image, More layers, More channels

- Observation: 밸런스를 맞추는 게 중요하다



Compound Model Scaling

- Compund Scaling Method

$$
\begin{aligned}
\text{depth: } d &= \alpha^{\phi} \\
\text{width: } w &= \beta^{\phi} \\
\text{resolution: } r &= \gamma^{\phi} \\
\text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\
\alpha \geq 1, \beta &\geq 1, \gamma \geq 1
\end{aligned}
$$

a, b, c – small grid search 통해 결정

Pi is a user_specified coefficient that controls how many more resources are available for model scaling