

# MobileNet

Пигасин Д. А. 18-АС

# История

MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

Дата публикации: 17 апреля 2017

Исследователи из Google:

Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang,  
Tobias Weyand, Marco Andreetto, Hartwig Adam

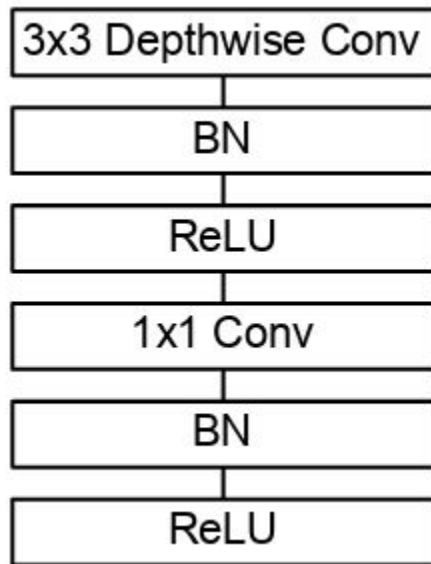
# Цель

При помощи mobilenet задачи вроде идентификации объектов на изображениях можно выполнять на маломощных устройствах.

Например, на мобильных телефонах.

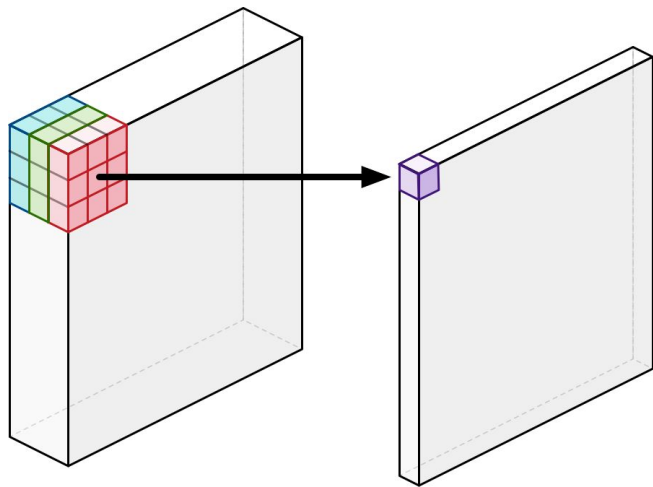
Использование данной сети позволяет уменьшить количество обращений к серверу, что обеспечивает возможность автономной работы, более быстрый отклик, меньшее потребление памяти и аккумулятора, а также повышает конфиденциальность пользователей.

# Структура

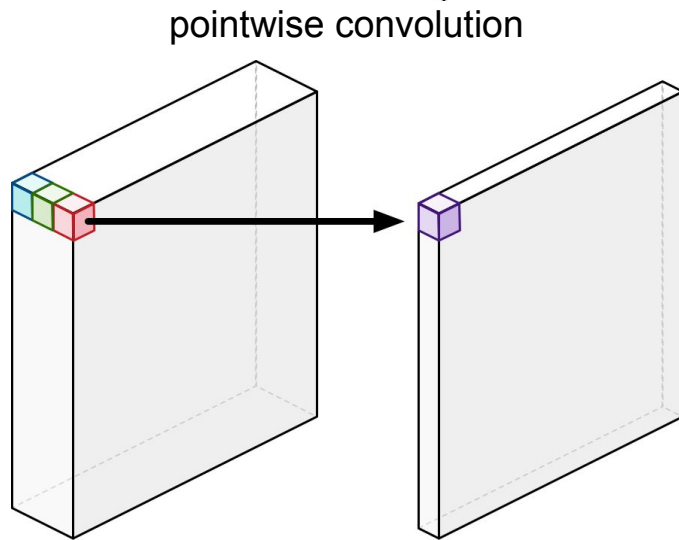
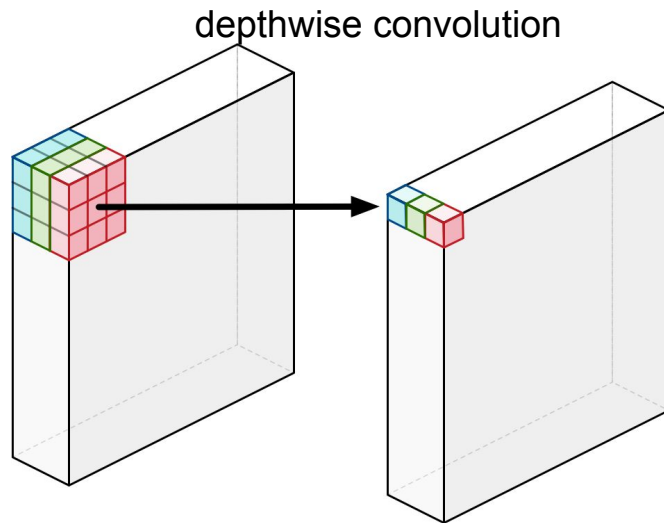


Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
5×	Conv dw / s1 $3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
	Conv / s1 $1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool $7 \times 7$	$7 \times 7 \times 1024$
FC / s1	$1024 \times 1000$	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

# Как это работает?



Стандартная свертка



## Пример

Пусть мы сворачиваем изображение с 16 каналами свёрточным слоем с 32 фильтрами. Суммарно этот слой будет иметь  $16 * 32 * 3 * 3 = 4608$  весов, так как у нас будет  $16 * 32$  свёрток  $3 \times 3$ .

Сколько же весов будет в аналогичном depthwise separable convolution блоке?

Во-первых, у нас будет  $32 * 3 * 3 = 288$  весов у depthwise convolution.

Во-вторых,  $16 * 32 * 1 * 1 = 512$  весов у pointwise convolution .

В сумме получим 800 весов, что намного меньше, чем у обычного свёрточного слоя.

# Особенности

- Использование **Depthwise separable convolutions**
- Отсутствие pooling-слоев. Для снижения пространственной размерности используется свертка с параметром stride, равным 2.
- Гиперпараметры  $a$  (множитель ширины) и  $p$  (множитель глубины или множитель разрешения).

# Гиперпараметры

- **Width multiplier (множитель ширины  $a$ )** : это уменьшает количество каналов. Если множитель ширины равен 1, сеть начинается с 32 каналов и заканчивается 1024.
- **Resolution multiplier (множитель разрешения  $p$ )**: это уменьшает размеры входного изображения. Размер входного сигнала по умолчанию составляет  $224 \times 224$  пикселя.

Эти настройки можно использовать, чтобы сделать сеть меньше — и, следовательно, быстрее — но за счет точности предсказания.



# Сравнение

Архитектура сети	Количество параметров	Top-1 accuracy	Top-5 accuracy
Xception	22.91M	0.790	0.945
VGG16	138.35M	0.715	0.901
MobileNetV1 (alpha=1, rho=1)	4.20M	0.709	0.899
MobileNetV1 (alpha=0.75, rho=0.85)	2.59M	0.672	0.873
MobileNetV1 (alpha=0.25, rho=0.57)	0.47M	0.415	0.663
MobileNetV2 (alpha=1.4, rho=1)	6.06M	<b>0.750</b>	<b>0.925</b>
MobileNetV2 (alpha=1, rho=1)	3.47M	0.718	0.910
MobileNetV2 (alpha=0.35, rho=0.43)	1.66M	0.455	0.704

# Источники

- <https://arxiv.org/abs/1704.04861>
- <https://machinethink.net/blog/googles-mobile-net-architecture-on-iphone/>
- <https://habr.com/ru/post/352804/>
- <https://habr.com/ru/post/347564/>