

Aggregated Residual Transformations for Deep Neural Networks

Анна Воронцова

ФКН НИУ ВШЭ

Москва, 2018

Модульность: последовательность базовых блоков одинаковой топологии →
→ ограничение на кол-во возможных значений гиперпараметров
→ меньше риск переобучения
→ масштабируемость

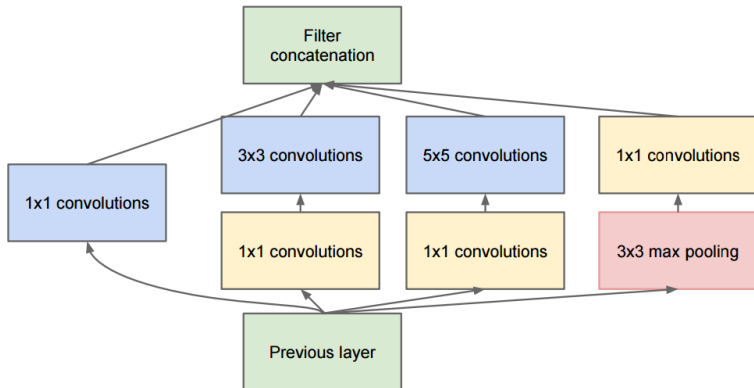
K. Simonyan, A. Zisserman. Very Deep Convolutional Networks for Large-scale Image Recognition.

Принцип split-transform-merge:

- ▶ разбить (split) вход на несколько низкоразмерных представлений (1×1 свертка)
- ▶ преобразовать (transform) множеством фильтров (3×3 , 5×5 , etc.)
- ▶ конкатенировать пути (merge).

Ch. Szegedy et al. Going deeper with convolutions.

Inception



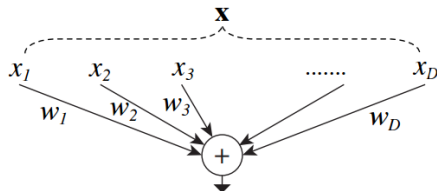
Проблема: оптимизация гиперпараметров (число фильтров и их размеры)

Простые нейроны

Inner product как split-transform-aggregate:

- ▶ разбить вход x на x_i
- ▶ преобразовать и масштабировать: $w_i x_i$

- ▶ агрегировать:
$$F(x) = \sum_{i=1}^D w_i x_i.$$



Network-in-Neuron

$$F(x) = \sum_{i=1}^D w_i x_i.$$

Можем использовать более сложную функцию вместо $w_i x_i$ – например, нейросеть:

$$F(x) = \sum_{i=1}^C T_i(x)$$

Split-transform-aggregate:

$$F(x) = \sum_{i=1}^C T_i(x).$$

C – **кардинальность** – кол-во преобразований.

VGG + Inception + ResNet

VGG → модульность, масштабируемость

Inception → split-transform-merge, ветвление

ResNet → residual connections

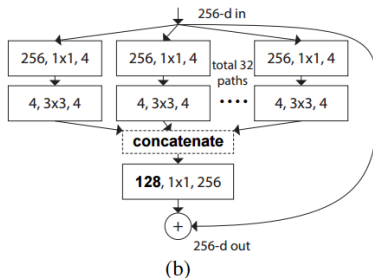
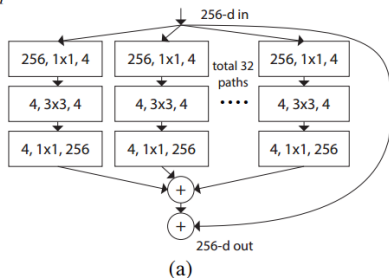
Aggregated Transformations:

$$y = \mathbf{x} + \sum_{i=1}^C T_i(\mathbf{x}),$$

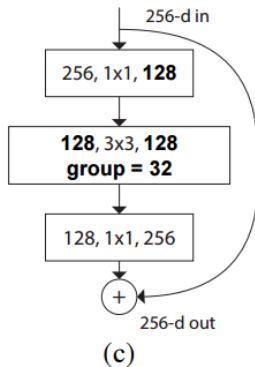
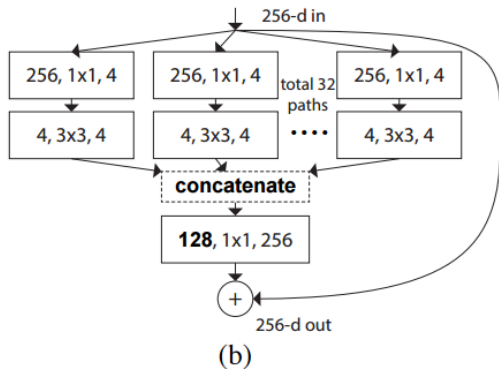
T_i имеют **одинаковую топологию**.

ResNeXt

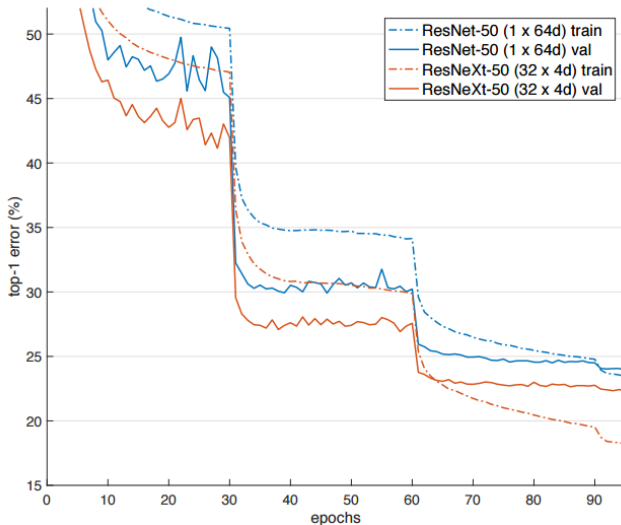
equivalent



ResNeXt



Эксперименты: ImageNet-1K



Кардинальность vs ширина

	setting	top-1 error (%)
ResNet-50	$1 \times 64d$	23.9
ResNeXt-50	$2 \times 40d$	23.0
ResNeXt-50	$4 \times 24d$	22.6
ResNeXt-50	$8 \times 14d$	22.3
ResNeXt-50	$32 \times 4d$	22.2
ResNet-101	$1 \times 64d$	22.0
ResNeXt-101	$2 \times 40d$	21.7
ResNeXt-101	$4 \times 24d$	21.4
ResNeXt-101	$8 \times 14d$	21.3
ResNeXt-101	$32 \times 4d$	21.2

setting: $C \times$ ширина

Сложность (FLOPs)

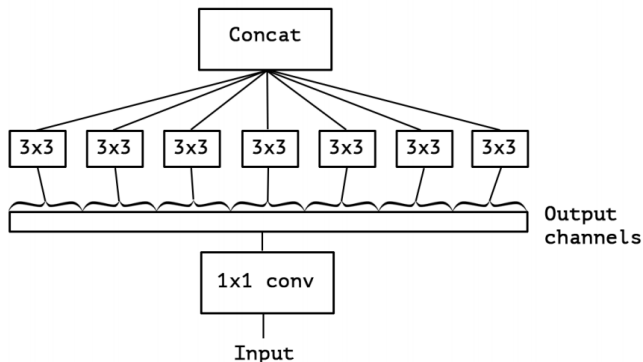
	setting	top-1 err (%)	top-5 err (%)
<i>1 × complexity references:</i>			
ResNet-101	$1 \times 64d$	22.0	6.0
ResNeXt-101	$32 \times 4d$	21.2	5.6
<i>2 × complexity models follow:</i>			
ResNet- 200 [15]	$1 \times 64d$	21.7	5.8
ResNet-101, wider	$1 \times \mathbf{100d}$	21.3	5.7
ResNeXt-101	$\mathbf{2} \times 64d$	20.7	5.5
ResNeXt-101	$\mathbf{64} \times 4d$	20.4	5.3

Residual connections

	setting	w/ residual	w/o residual
ResNet-50	$1 \times 64d$	23.9	31.2
ResNeXt-50	$32 \times 4d$	22.2	26.1

Родственные архитектуры: Xception

Хсептион, октябрь 2016 (ResNext – ноябрь 2016)



F. Chollet. Xception: Deep Learning with Depthwise Separable Convolutions.

Заключение

ResNeXt:

- ▶ н/с для классификации изображений
- ▶ простая: однородная, с небольшим кол-вом гиперпараметров
- ▶ модульная: базовые блоки, преобразования одинаковой топологии
- ▶ кардинальность – кол-во преобразований как еще одна размерность н/с
- ▶ при фикс. теор. сложности лучше увеличить кардинальность, чем глубину или ширину (эмпирически, ImageNet-1K)

Список литературы

- ▶ S. Xie et al. Aggregated Residual Transformations for Deep Neural Networks. [1611.05431]
- ▶ K. Simonyan, A. Zisserman. Very Deep Convolutional Networks for Large-scale Image Recognition. [1409.1556]
- ▶ Ch. Szegedy et al. Going deeper with convolutions. [1409.4842]
- ▶ F. Chollet. Xception: Deep Learning with Depthwise Separable Convolutions. [1610.02357]