



ТЕХНОСФЕРА

Лекция 6 Архитектуры нейронных сетей

Полыковский Даниил
Храбров Кузьма

13 марта 2017 г.

Архитектуры CNN

ImageNet, описание



- ▶ 1000 классов
- ▶ около 1000 изображений в каждом классе
- ▶ около 1 000 000 изображений всего
- ▶ несколько номинаций: таких как распознавание и детектирование/локализация

ImageNet, правила





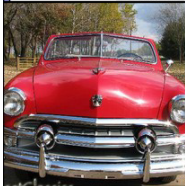



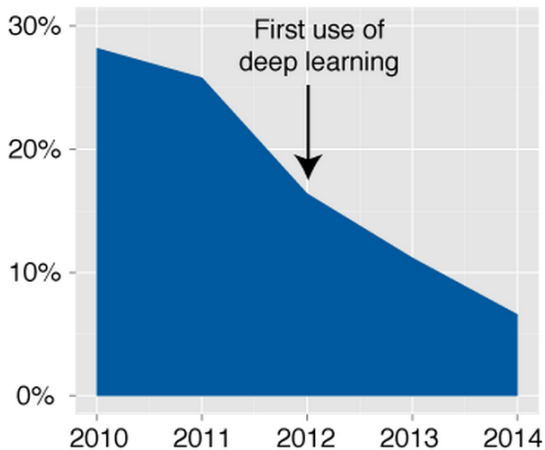
			
mite	container ship	motor scooter	leopard
<div> <div></div> <div>mite</div> <div>black widow</div> <div>cockroach</div> <div>tick</div> <div>starfish</div> </div>	<div> <div></div> <div>container ship</div> <div>lifeboat</div> <div>amphibian</div> <div>fireboat</div> <div>drilling platform</div> </div>	<div> <div></div> <div>motor scooter</div> <div>go-kart</div> <div>moped</div> <div>bumper car</div> <div>golfcart</div> </div>	<div> <div></div> <div>leopard</div> <div>jaguar</div> <div>cheetah</div> <div>snow leopard</div> <div>Egyptian cat</div> </div>
			
grille	mushroom	cherry	Madagascar cat
<div> <div></div> <div>convertible</div> <div>grille</div> <div>pickup</div> <div>beach wagon</div> <div>fire engine</div> </div>	<div> <div></div> <div>agaric</div> <div>mushroom</div> <div>jelly fungus</div> <div>gill fungus</div> <div>dead-man's-fingers</div> </div>	<div> <div></div> <div>dalmatian</div> <div>grape</div> <div>elderberry</div> <div>ffordshire bullterrier</div> <div>currant</div> </div>	<div> <div></div> <div>squirrel monkey</div> <div>spider monkey</div> <div>titi</div> <div>indri</div> <div>howler monkey</div> </div>

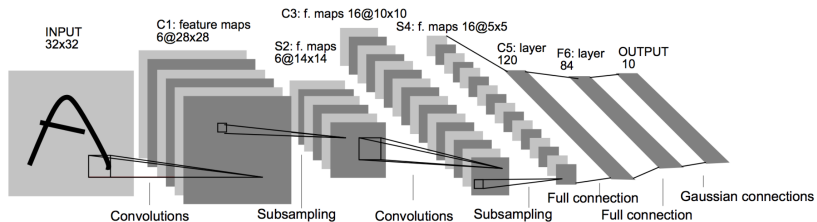
Рис.: Примеры прогнозов

ImageNet, прогресс

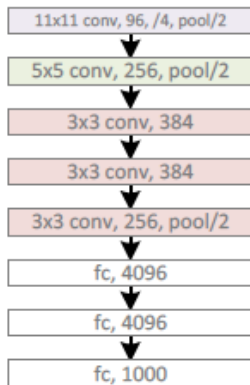
Object classification error rate



Lenet¹



¹<http://yann.lecun.com/exdb/lenet>



- ▶ 5 сверточных и 3 полносвязных слоя
- ▶ 60M параметров, 650к нейронов
- ▶ эффективное распараллеливание на 2 GPU/CUDA
- ▶ свёртки 11x11, 5x5, 3x3
- ▶ ReLU, т.к. не надо вычислять exp

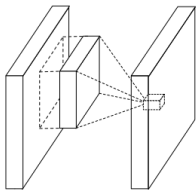
VGG³

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

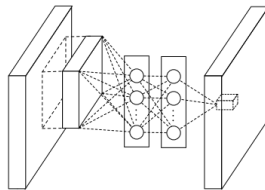
- ▶ VGG-19 (E): 144M параметров
- ▶ Very Deep Convolutional Networks for Large-Scale Image Recognition^a
- ▶ Только свертки 3x3

^aK. Simonyan, A. Zisserman

³<http://arxiv.org/pdf/1409.1556.pdf>

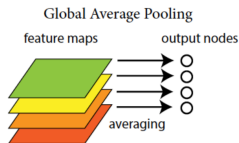
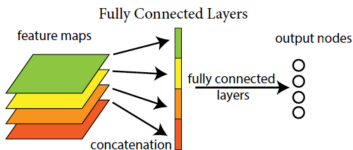


(a) Linear convolution layer



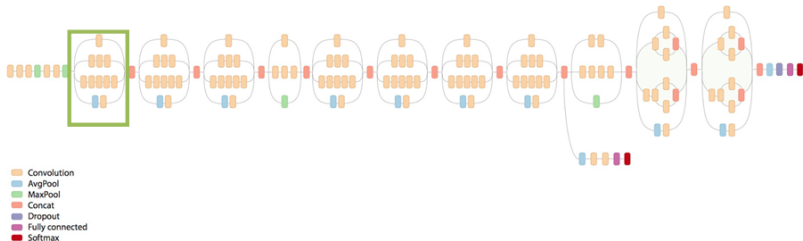
(b) Mlpconv layer

- ▶ “Полносвязные слои” (свертки 1x1) внутри свертки
- ▶ Глобальный пулинг для вытягивания изображения (+ доп. регуляризация)

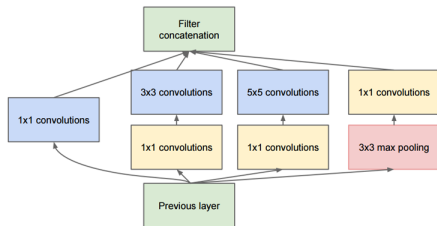


⁴<https://arxiv.org/pdf/1312.4400v3.pdf>

GoogLeNet⁵



Green box shows parallel region of GoogLeNet



Full Inception module

⁵http://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Szegedy_Going_Deeper_With_2015_CVPR_paper.pdf

Revolution of Depth

AlexNet, 8 layers
(ILSVRC 2012)



VGG, 19 layers
(ILSVRC 2014)



ResNet, 152 layers
(ILSVRC 2015)



- ▶ Нагляднее: <http://josephpcohen.com/w/visualizing-cnn-architectures-side-by-side-with-mxnet/>

ResNet⁸

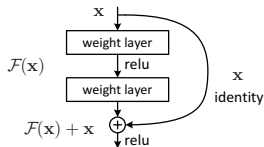


Рис.: Residual block

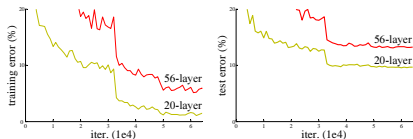


Рис.: Обучение “обычной” сети

- ▶ Результат на 56 слоях хуже. Проблема не в переобучении
- ▶ Решение заведомо существует: 20 слоев, затем $F(x)=x$
- ▶ Выучить $F(x)=x$ тяжело, а $F(x)=0$ просто
- ▶ Residual block решает эту проблему

Существует множество модификаций ResNet: ResNet in ResNet⁶, DenseNet⁷

⁶<https://arxiv.org/pdf/1603.08029v1.pdf>

⁷<https://arxiv.org/pdf/1608.06993v1.pdf>

⁸<https://arxiv.org/pdf/1512.03385v1.pdf>

Case study
National Data Science Bowl

National Data Science Bowl¹⁰

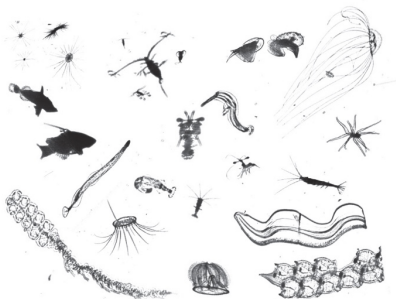


Рис.: Примеры изображений из набора данных: всего 121 несбалансированный класс, 30 000 изображений

Победители⁹ использовали архитектуру VGG-16, а так же ряд специальных трюков.

⁹<http://benanne.github.io/2015/03/17/plankton.html>

¹⁰<https://www.kaggle.com/c/datasciencebowl>

Data augmentation, #1

Data augmentation:

- ▶ rotation: random with angle between 0 and 360 degree (uniform)
- ▶ translation: random with shift between -10 and 10 pixels (uniform)
- ▶ rescaling: random with scale factor between $1/1.6$ and 1.6 (log-uniform)
- ▶ flipping: yes or no (bernoulli)
- ▶ shearing: random with angle between -20 and 20 degree (uniform)
- ▶ stretching: random with stretch factor between $1/1.3$ and 1.3 (log-uniform)

Данные нужно аугментировать в реальном времени (иначе — значительное увеличение размера датасета). Не стоит недооценивать test-time аугментацию.

Data augmentation, #2

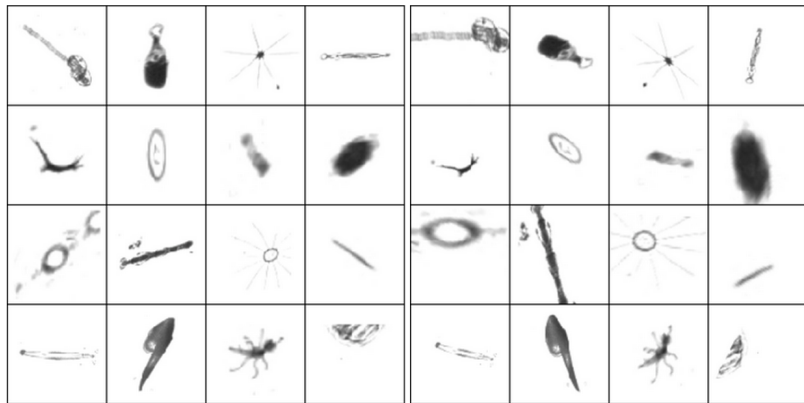


Рис.: Оригиналы и аугментированные образы

Cyclic pooling

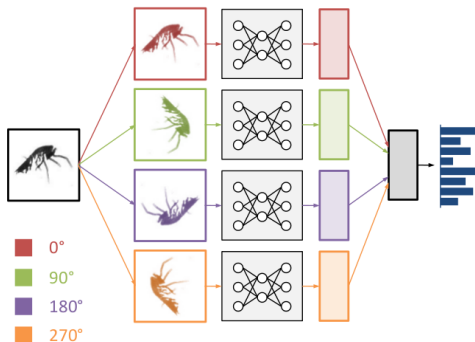


Рис.: Schematic representation of a convnet with cyclic pooling

- ▶ объединение результатов извлечения признаков после прохода по нескольким копиям одной сети параллельно
- ▶ это позволило сократить размер батча в 4 раза (со 128 до 32)
- ▶ root-mean-square pooling оказался эффективнее других

Rolling feature maps

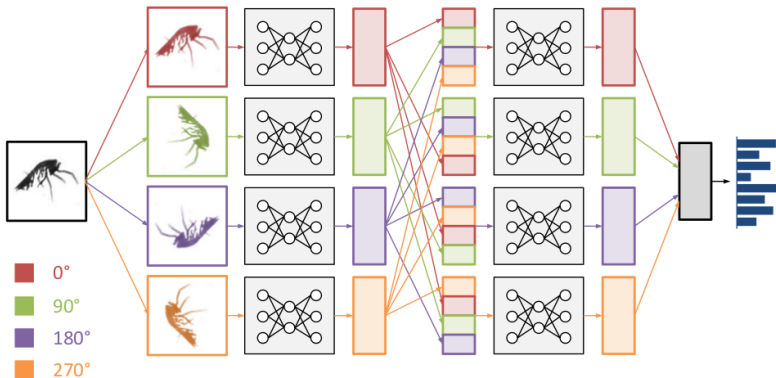


Рис.: Schematic representation of a roll operation inside a convnet with cyclic pooling.

Советы

- ▶ leaky/parameterized ReLU: $f(x) = \max(x, a \cdot x), a \leq 0$
- ▶ добавление к сверточным признакам других признаков перед полносвязным слоем (Hu moments, Zernike moments, atc)
- ▶ использовать предобученные модели
- ▶ self-training с тестовой выборкой
- ▶ аугментировать данные

Вопросы

