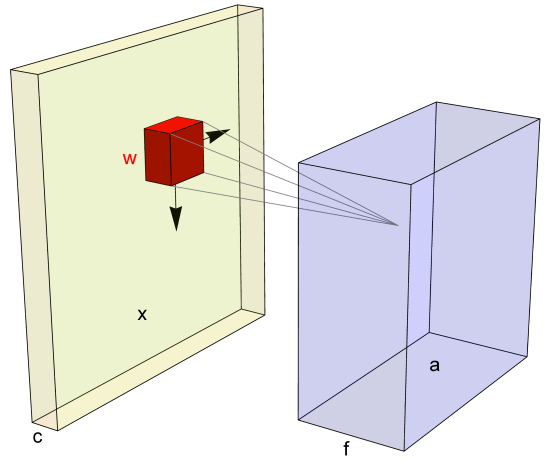
Convolutional Neural Networks (CNNs) Überblick

MoSim II Vorlesung

Christian B. Mendl

31. Mai 2018

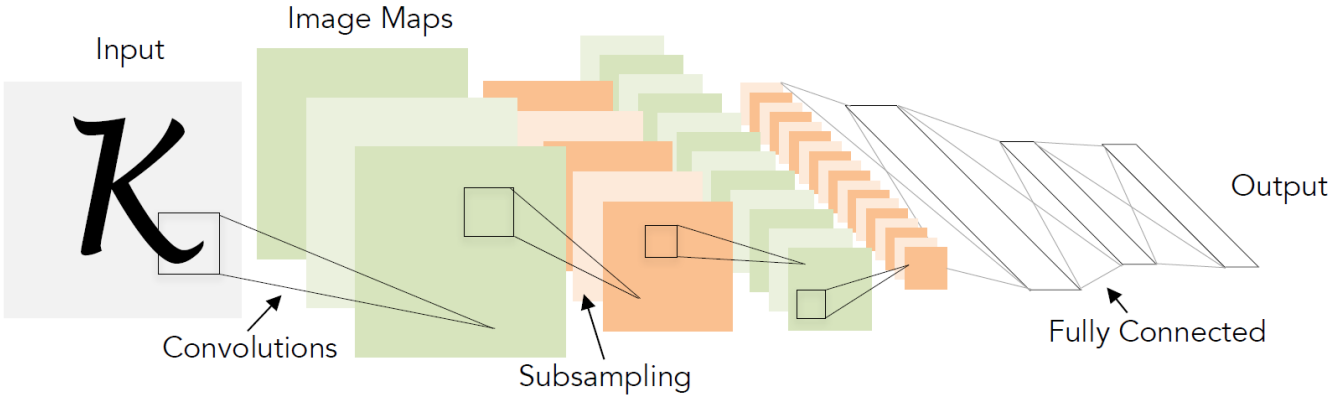


CNN Entwicklung

1998

LeCun et al.

"LeNet-5"



# of transistors

10^6

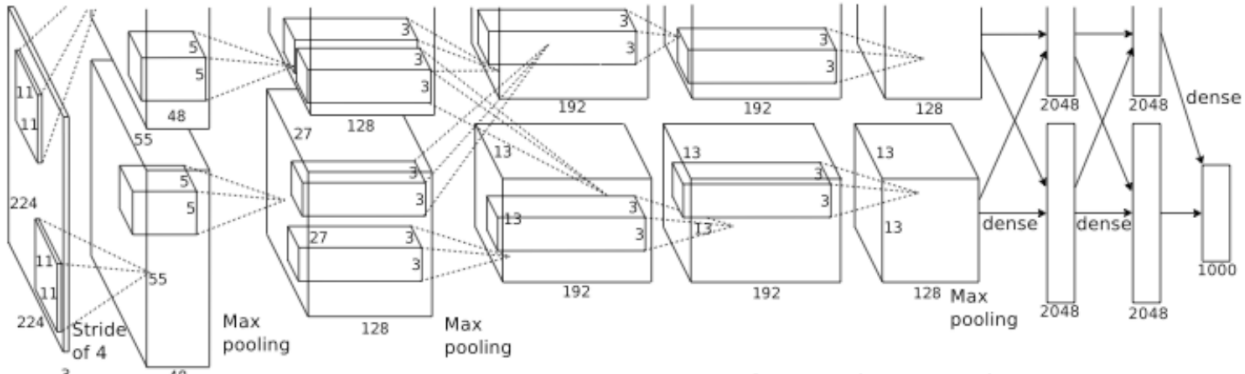
# of pixels used in training

10^7 NIST

2012

Krizhevsky et al.

"AlexNet"



# of transistors

10^9

GPUs

# of pixels used in training

10^{14} IMAGENET

Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

Source: http://cs231n.stanford.edu/syllabus.html

ImageNet Online-Datenbank

IMAGENET

http://image-net.org

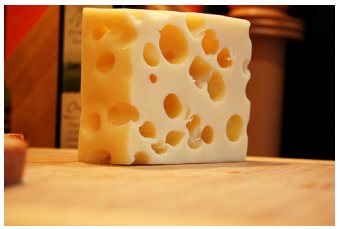
14 Millionen Bilder
22 Tausend Kategorien (WordNet "synonym sets")

beading plane

brown root rot fungus

common roundworm

scalded milk



Swiss cheese

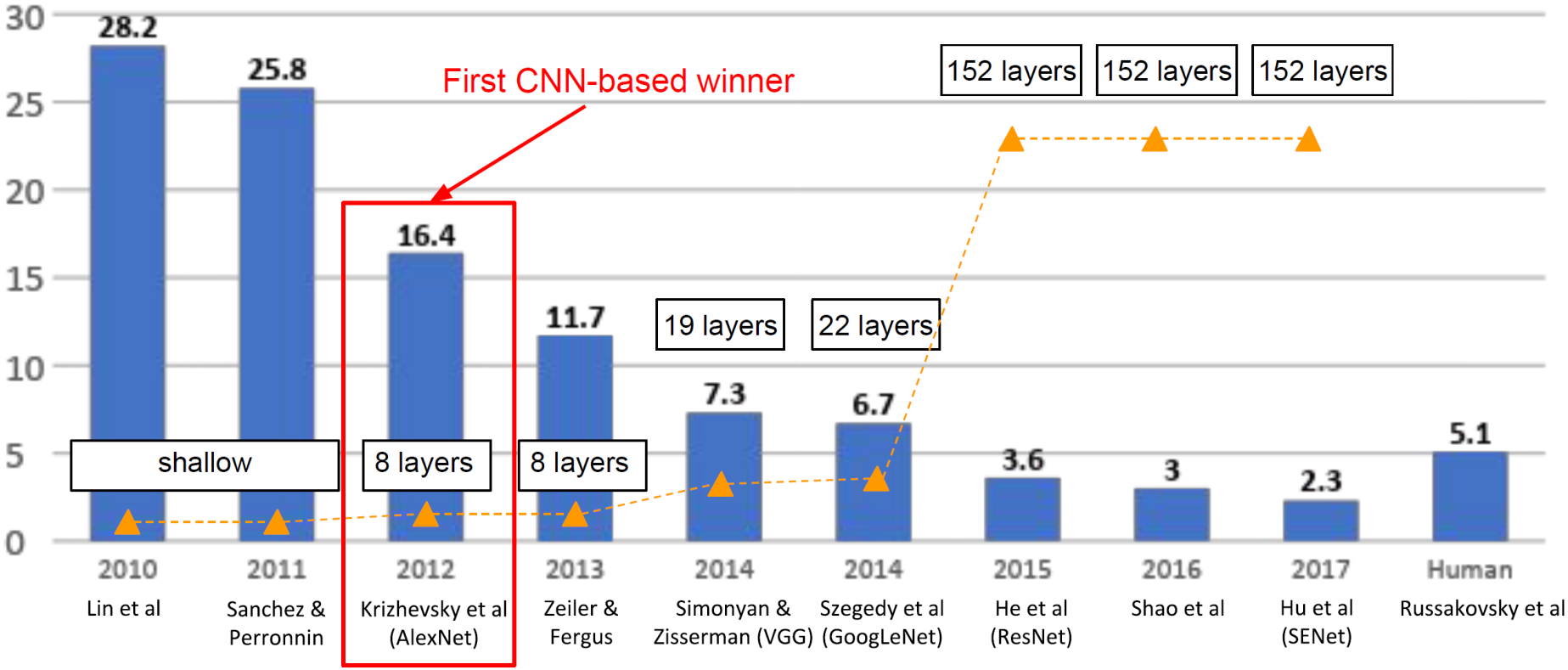
Human labeling (Amazon's Mechanical Turk crowd-sourcing)

J. Deng, W. Dong et al., ImageNet: A large-scale hierarchical image database. CVPR (2009)

ImageNet – Large Scale Visual Recognition Challenge

Jährlicher Wettbewerb

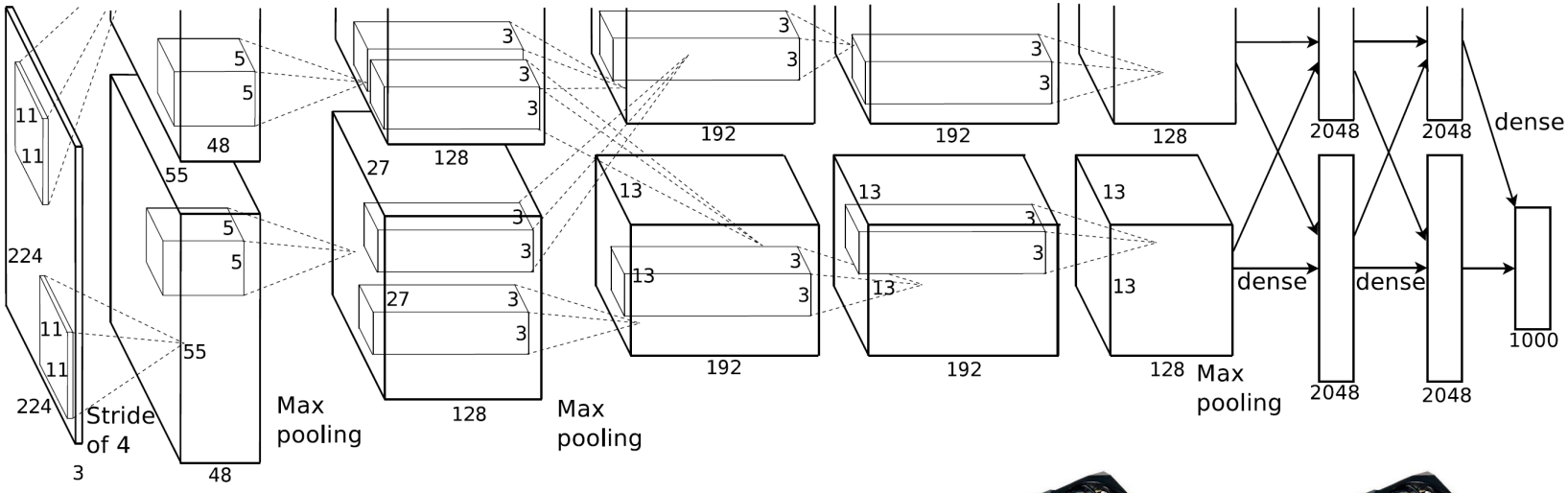
1000 Kategorien
1.4 Millionen Bilder



Source: http://cs231n.stanford.edu/syllabus.html

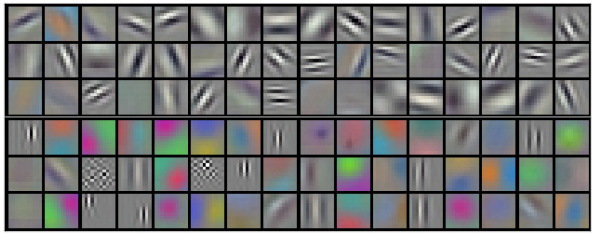
O. Russakovsky, J. Deng et al., IJCV (2015)

AlexNet (2012) Netzwerk-Architektur



Conv1: 48+48 Filter (11×11)

$\rightsquigarrow (227 - 11)/4 + 1 = 55$



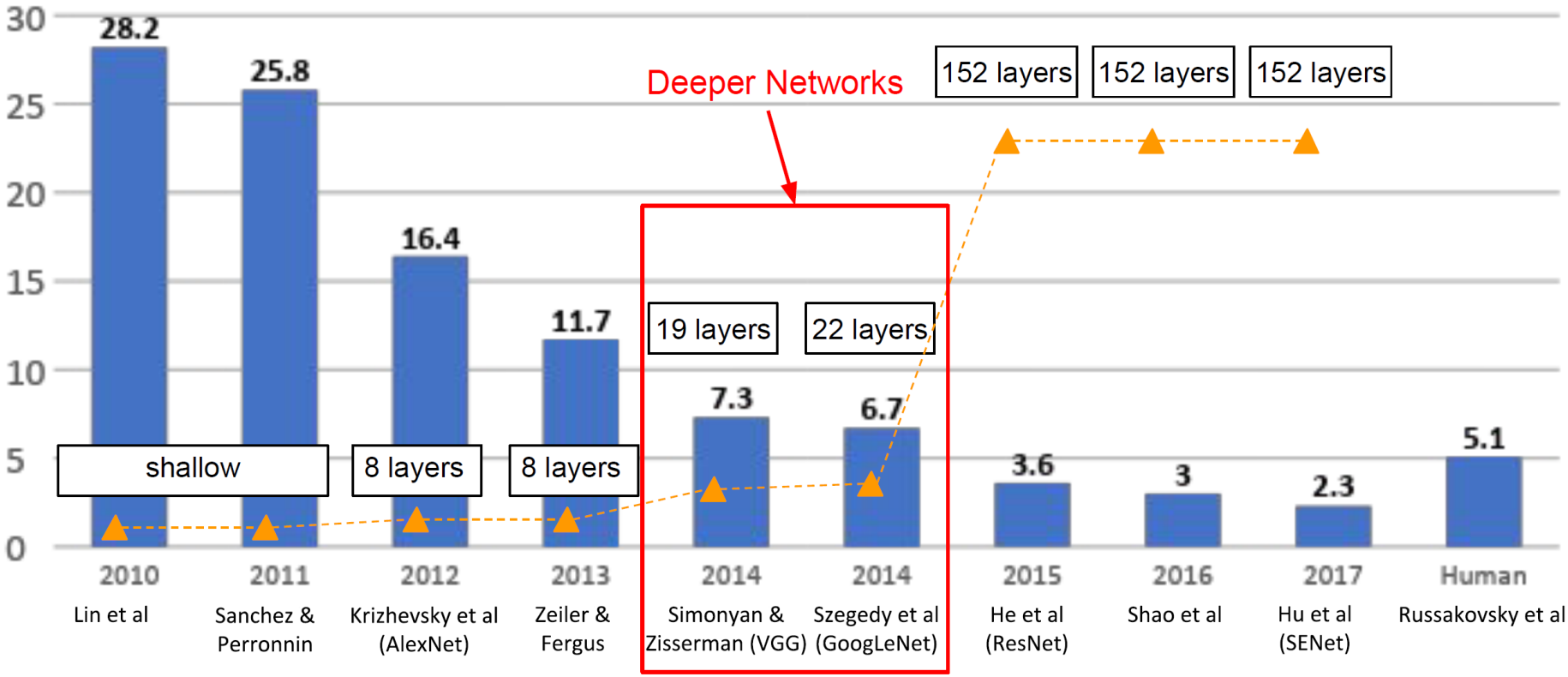
Zwei Nvidia GTX 580 3GB GPUs

Top-5 Fehlerrate: 16.4%

Erstmalige Verwendung der ReLU (Rectified Linear Unit) Aktivierungsfunktion

A. Krizhevsky, I. Sutskever, G. E. Hinton, NIPS (2012)

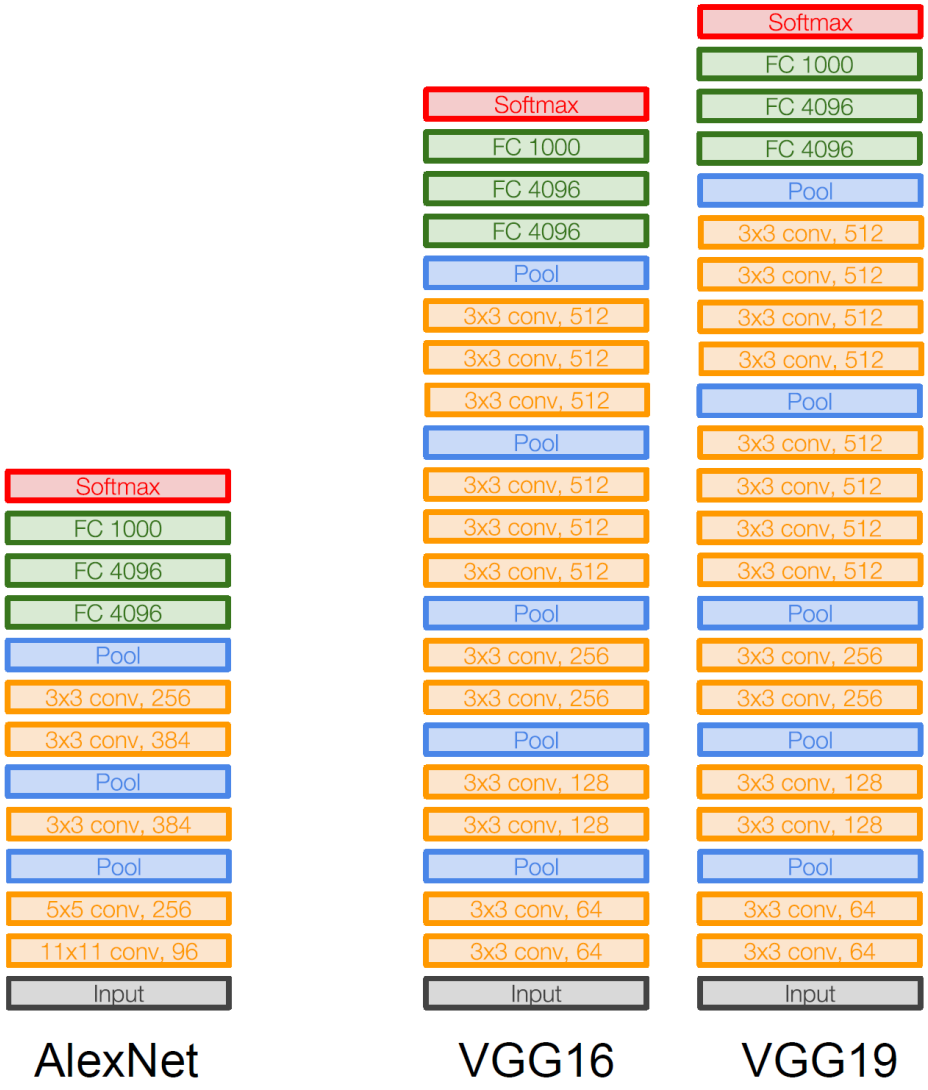
ImageNet – Large Scale Visual Recognition Challenge



Source: http://cs231n.stanford.edu/syllabus.html

O. Russakovsky, J. Deng et al., IJCV (2015)

VGGNet (2014) Architektur



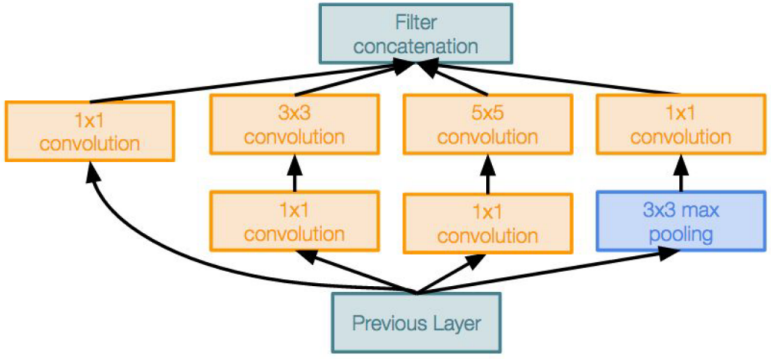
Source: http://cs231n.stanford.edu/syllabus.html

Hintereinanderschaltung von drei 3×3 Conv-Layers $\rightsquigarrow$ effektives rezeptives Feld: 7×7

Vorteil: weniger Parameter

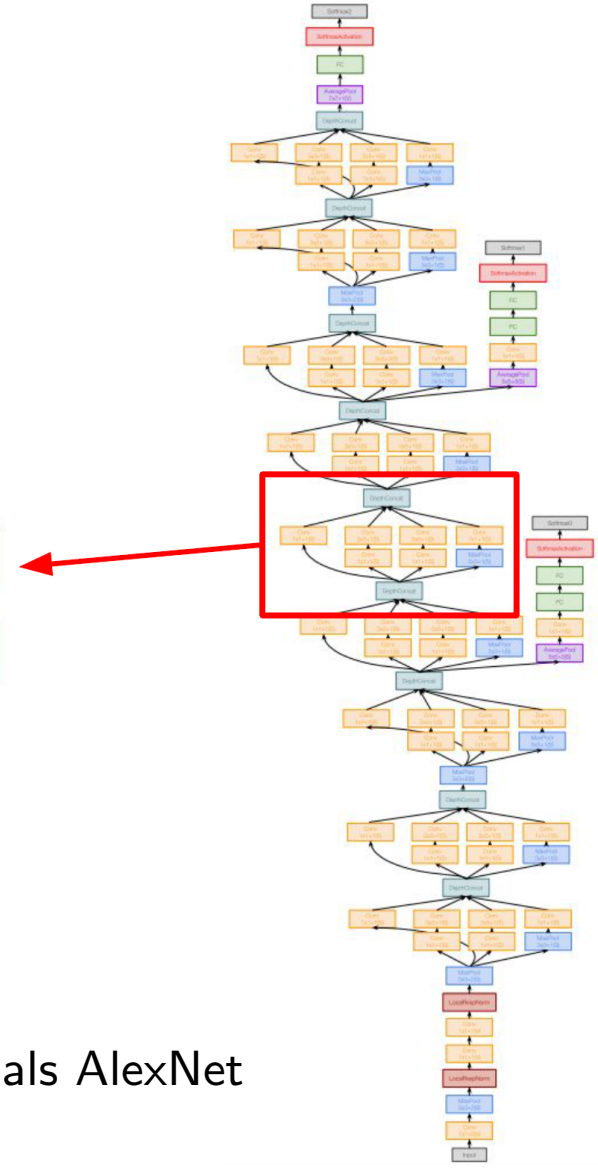
Simonyan und Zisserman, arXiv:1409.1556 (2014)

GoogLeNet (2014): Inception Module



Inception module

Netzwerk innerhalb Netzwerk

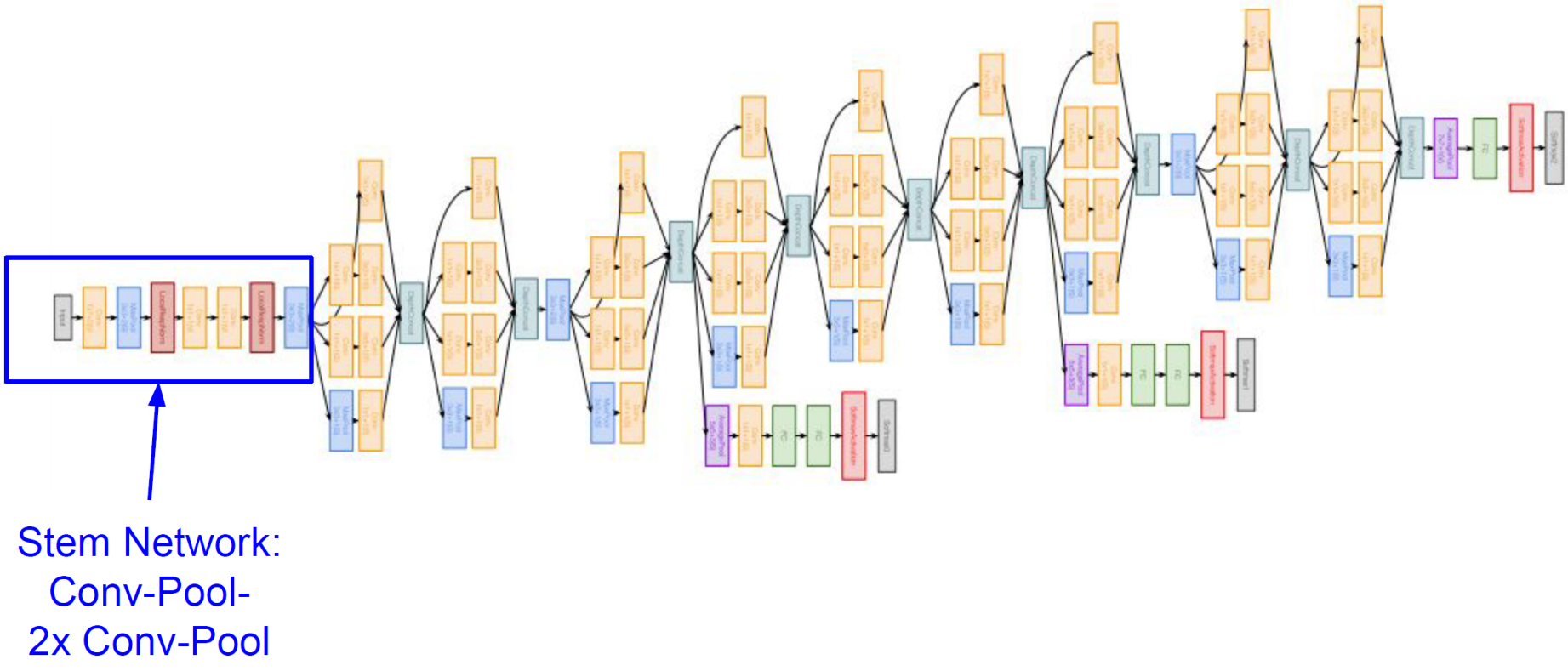


"Nur" 5 Millionen Parameter, 12× weniger als AlexNet

Szegedy et al., CVPR (2014)

GoogLeNet (2014): Gesamtarchitektur

22 Schichten insgesamt

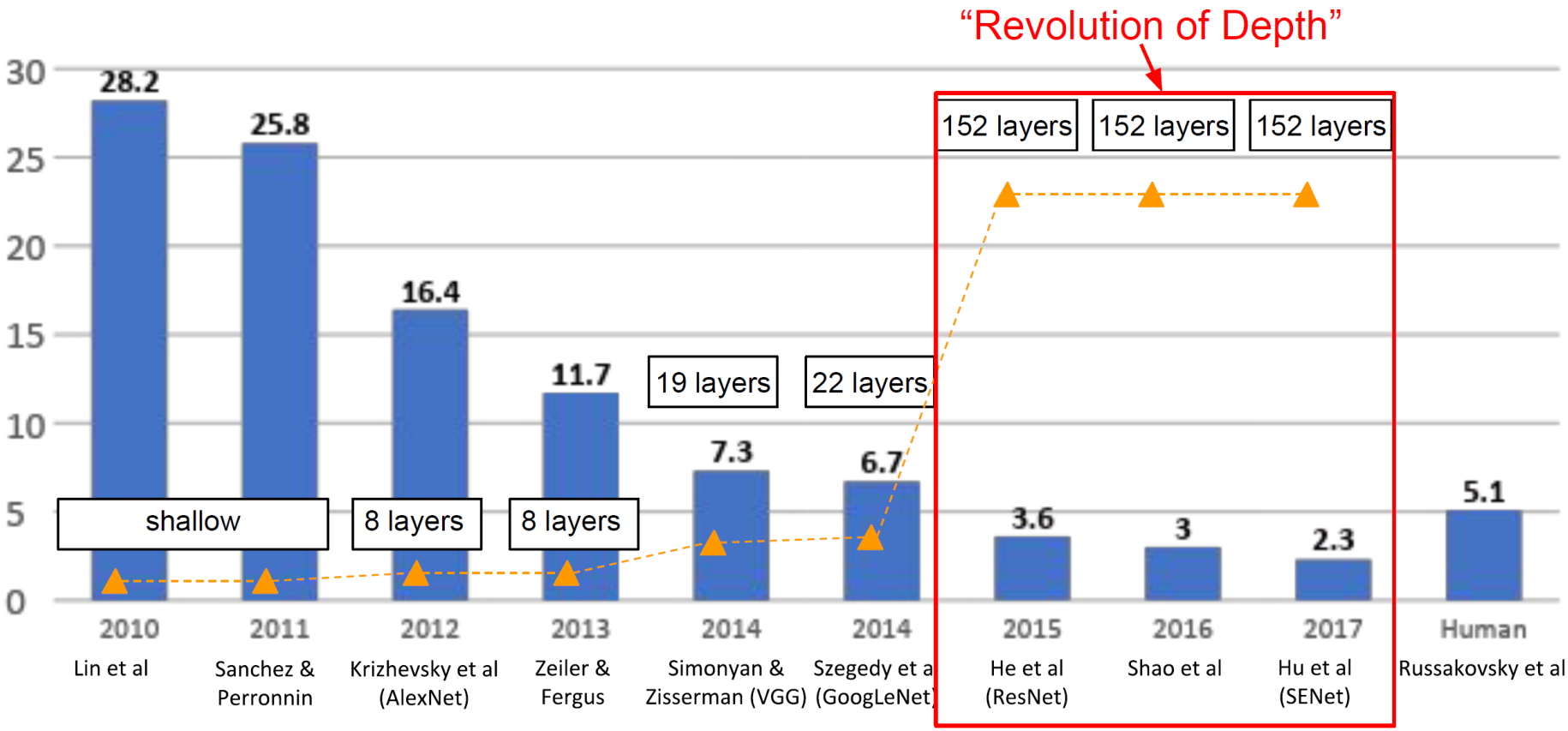


Stem Network:
Conv-Pool-
2x Conv-Pool

Source: http://cs231n.stanford.edu/syllabus.html

Szegedy et al., CVPR (2014)

ImageNet – Large Scale Visual Recognition Challenge

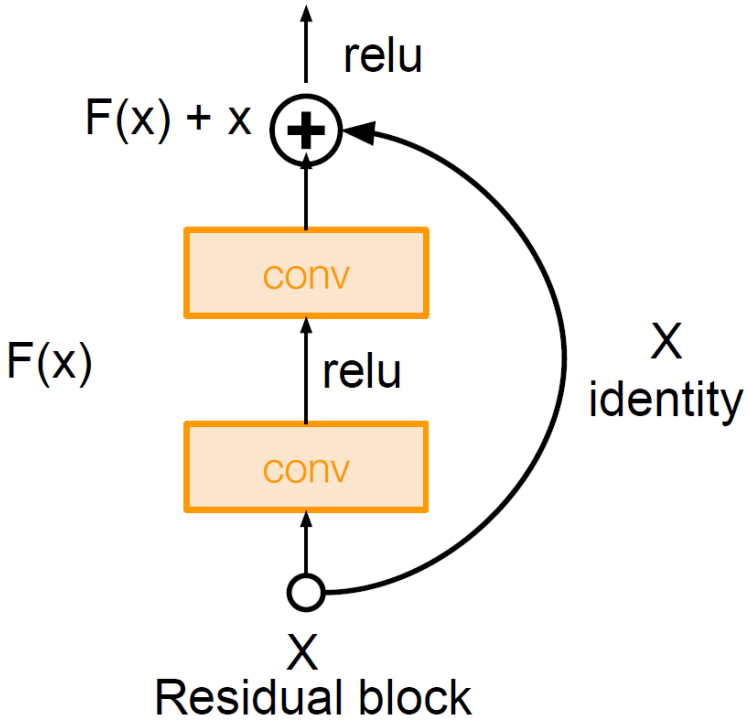


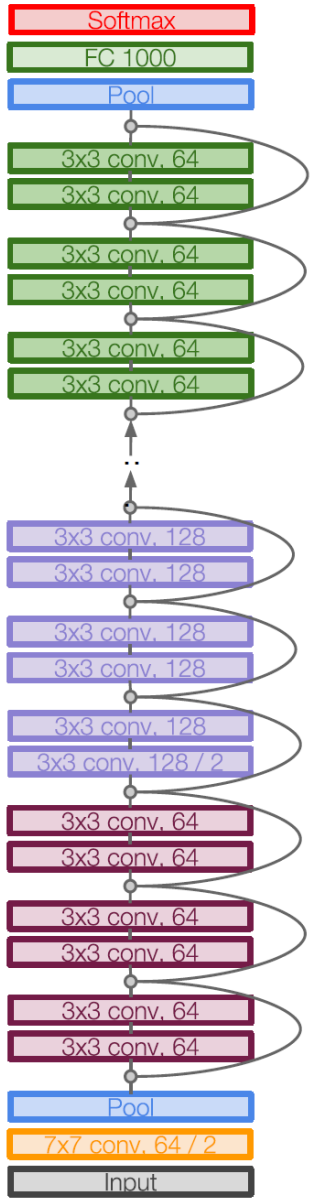
Source: http://cs231n.stanford.edu/syllabus.html

O. Russakovsky, J. Deng et al., IJCV (2015)

ResNet (2015)

Vermeidet “Instabile Gradienten”-Problem bei tiefen Netzen (siehe A26)

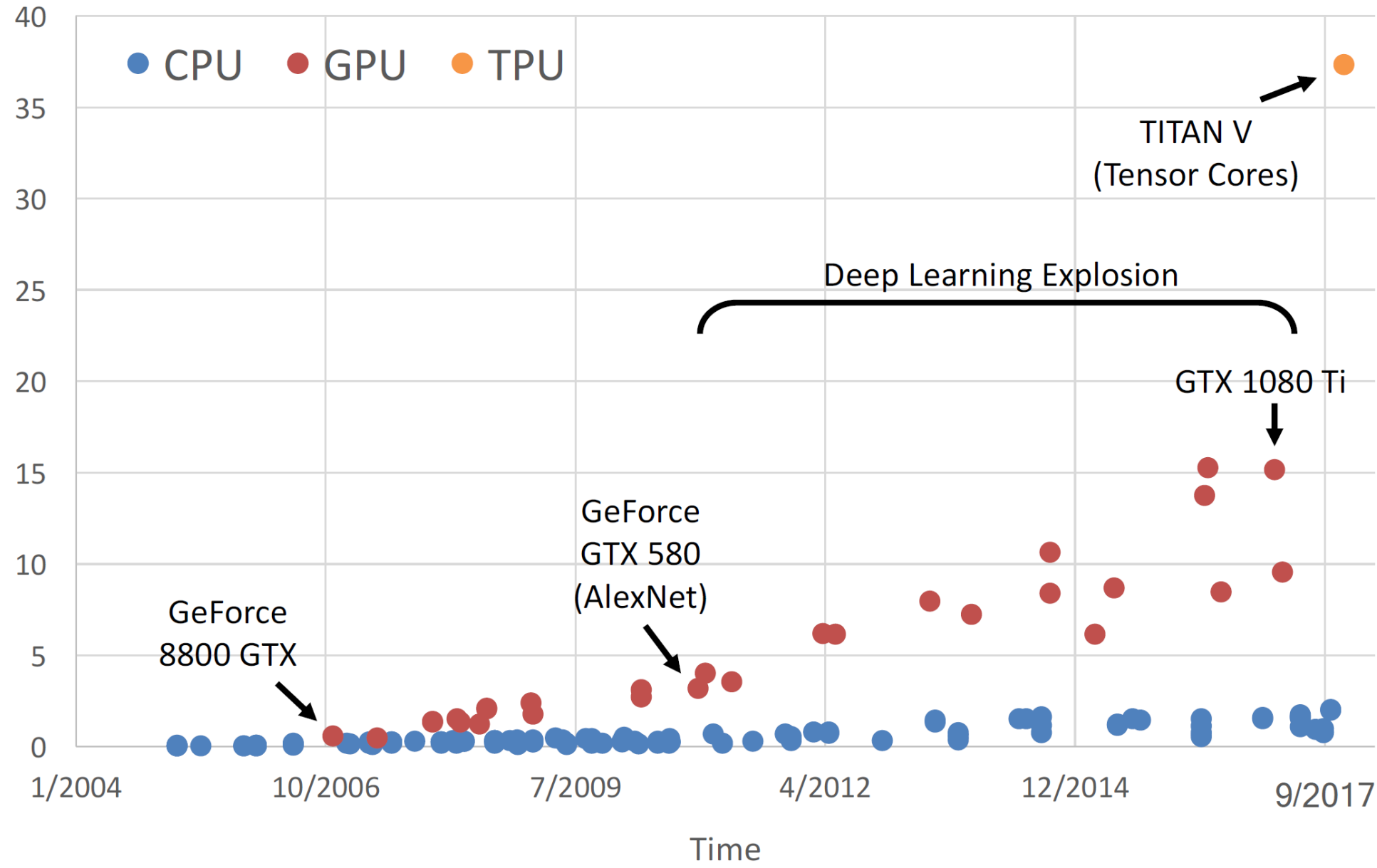




Source: http://cs231n.stanford.edu/syllabus.html

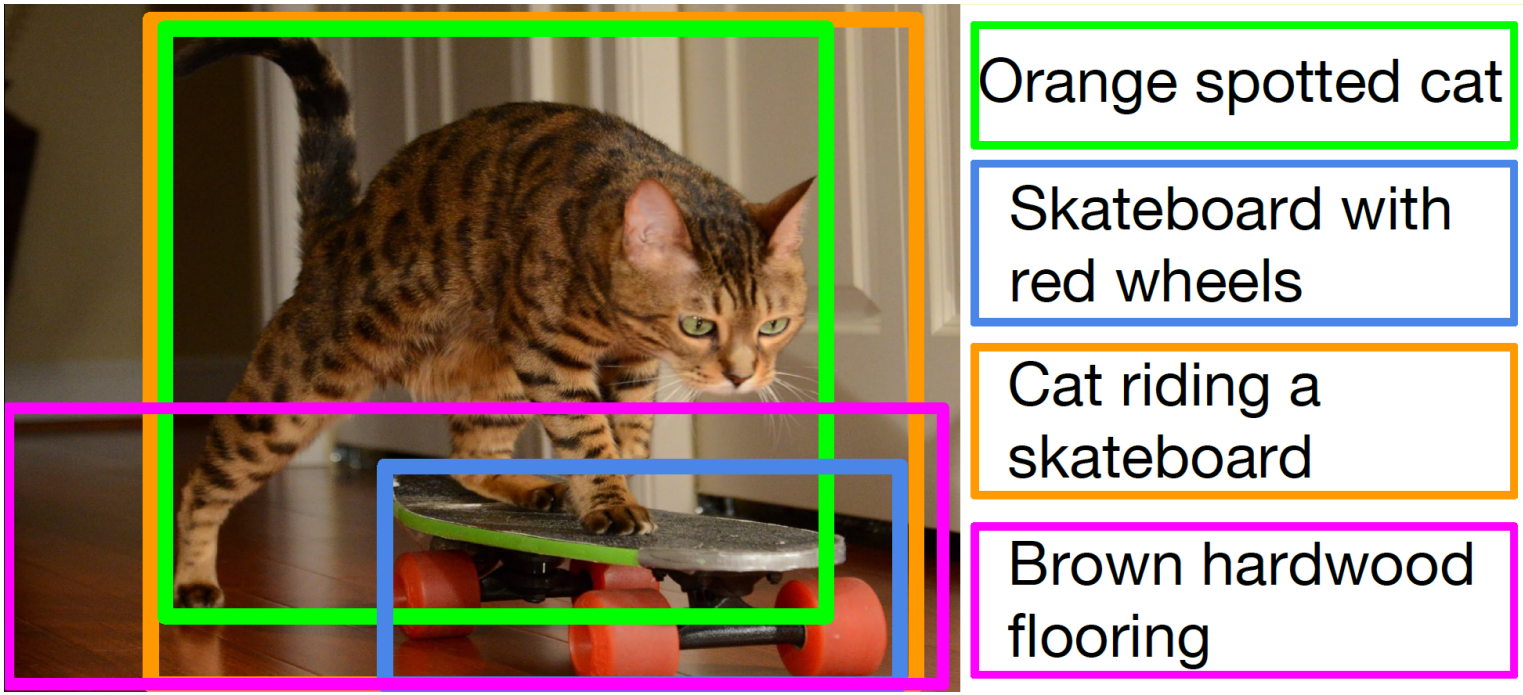
K. He et al., arXiv:1512.03385 (2015)

GigaFLOPS per Dollar



Source: http://cs231n.stanford.edu/syllabus.html

Anwendungen: Bilderkennung und -klassifizierung

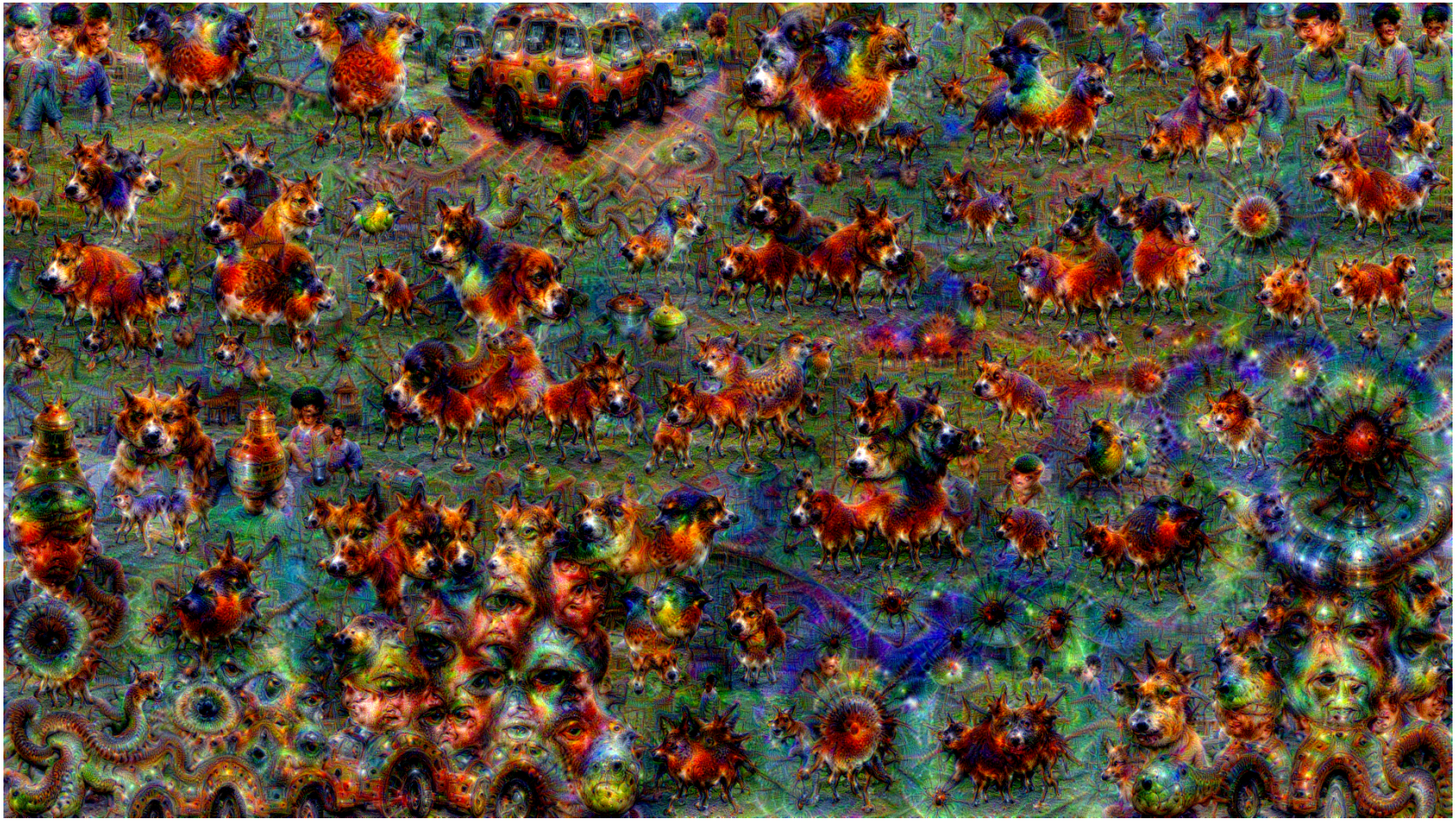


Source: A. Karpathy PhD thesis (2016)

Anwendungen: Artistic Style Transfer

Source: https://deepart.io

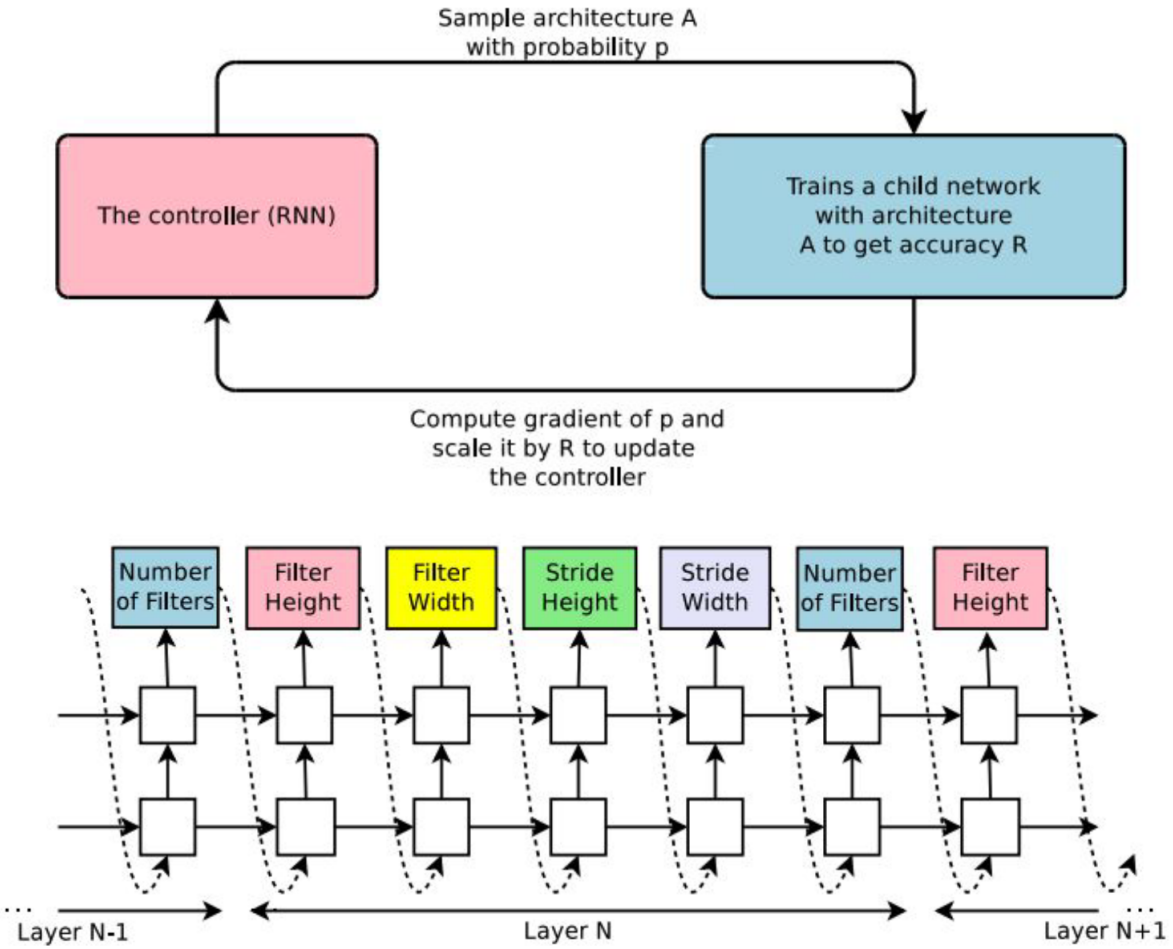
Anwendungen: Deep Dream



Source: http://cs231n.stanford.edu/syllabus.html

Ausblick: Meta-Learning

"Trainiere das Trainieren"
Kontrollnetzwerk lernt eine gute Netzwerkarchitektur zu erstellen



B. Zoph, Q. V. Le, arXiv:1611.01578 (2016)

Literatur

J. Deng u. a. “ImageNet: A large-scale hierarchical image database”. In: *Computer Vision and Pattern Recognition (CVPR)* (2009). DOI: 10.1109/CVPR.2009.5206848.

K. He u. a. “Deep residual learning for image recognition”. In: *arXiv:1512.03385* (2015). URL: https://arxiv.org/abs/1512.03385.

A. Karpathy. “Connecting images and natural language”. Diss. Stanford, 2016.

A. Krizhevsky, I. Sutskever und G. E. Hinton. “ImageNet classification with deep convolutional neural networks”. In: *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems*. 2012, S. 1097–1105. URL: https://dl.acm.org/citation.cfm?id=2999134.2999257.

O. Russakovsky u. a. “ImageNet Large Scale Visual Recognition Challenge”. In: *International Journal of Computer Vision (IJCV)* 115 (2015), S. 211–252. DOI: 10.1007/s11263-015-0816-y.

K. Simonyan und A. Zisserman. “Very deep convolutional networks for large-scale image recognition”. In: *arXiv:1409.1556* (2014). URL: https://arxiv.org/abs/1409.1556.

C. Szegedy u. a. “Going deeper with convolutions”. In: *Computer Vision and Pattern Recognition (CVPR)*. 2015. URL: https://arxiv.org/abs/1409.4842.

B. Zoph und Q. V. Le. “Neural Architecture Search with Reinforcement Learning”. In: *arXiv:1611.01578* (2016). URL: https://arxiv.org/abs/1611.01578.