Lecture 26

Agenda

Convolutional Neural Networks (CNN)

- CNN Architectures
  - LeNet
  - AlexNet
  - VGG
  - GoogLeNet
  - ResNet
- Pixel Importance
  - Up-Convolution
  - Shapley Values

LeNet 1998

PROC. OF THE IEEE, NOVEMBER 1998

Gradient-Based Learning Applied to Document Recognition

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner

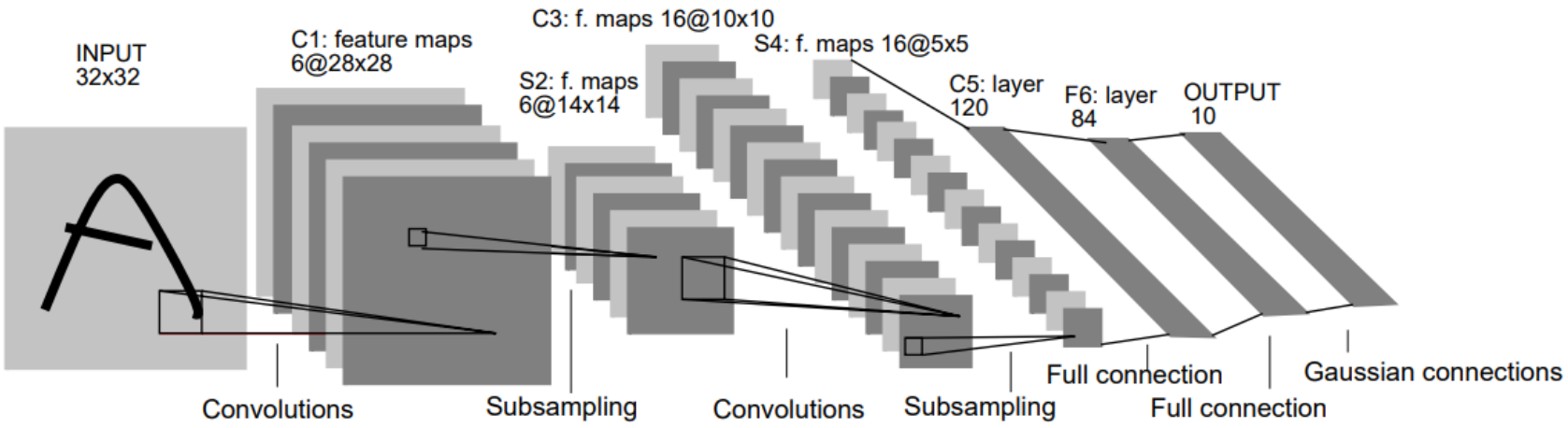


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf

R·I·T

Computer Science @ RIT

LeNet 1998

- Trained on MNIST
- End-to-end
- Augmentation was used
- About 60,000 parameters
- Test error 0.8

ImageNet Competition 2010–2017

`https://image-net.org/challenges/LSVRC/`

- About 1,000,000 images
- 1000 classes

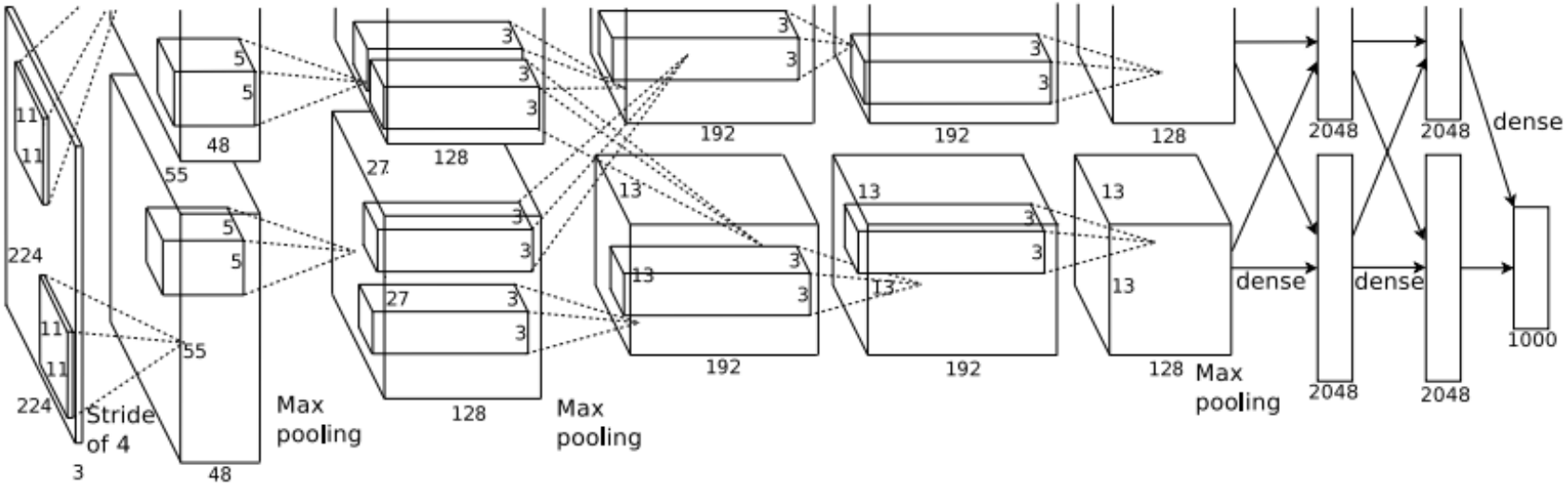
AlexNet 2012

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto

Ilya Sutskever
University of Toronto

Geoffrey E. Hinton
University of Toronto



https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

AlexNet 2012

- ReLu, augmentations, dropout
- Gradient descent with momentum
- GPU (5 days)
- About 1,000,000,000 parameters
- Test error 17%

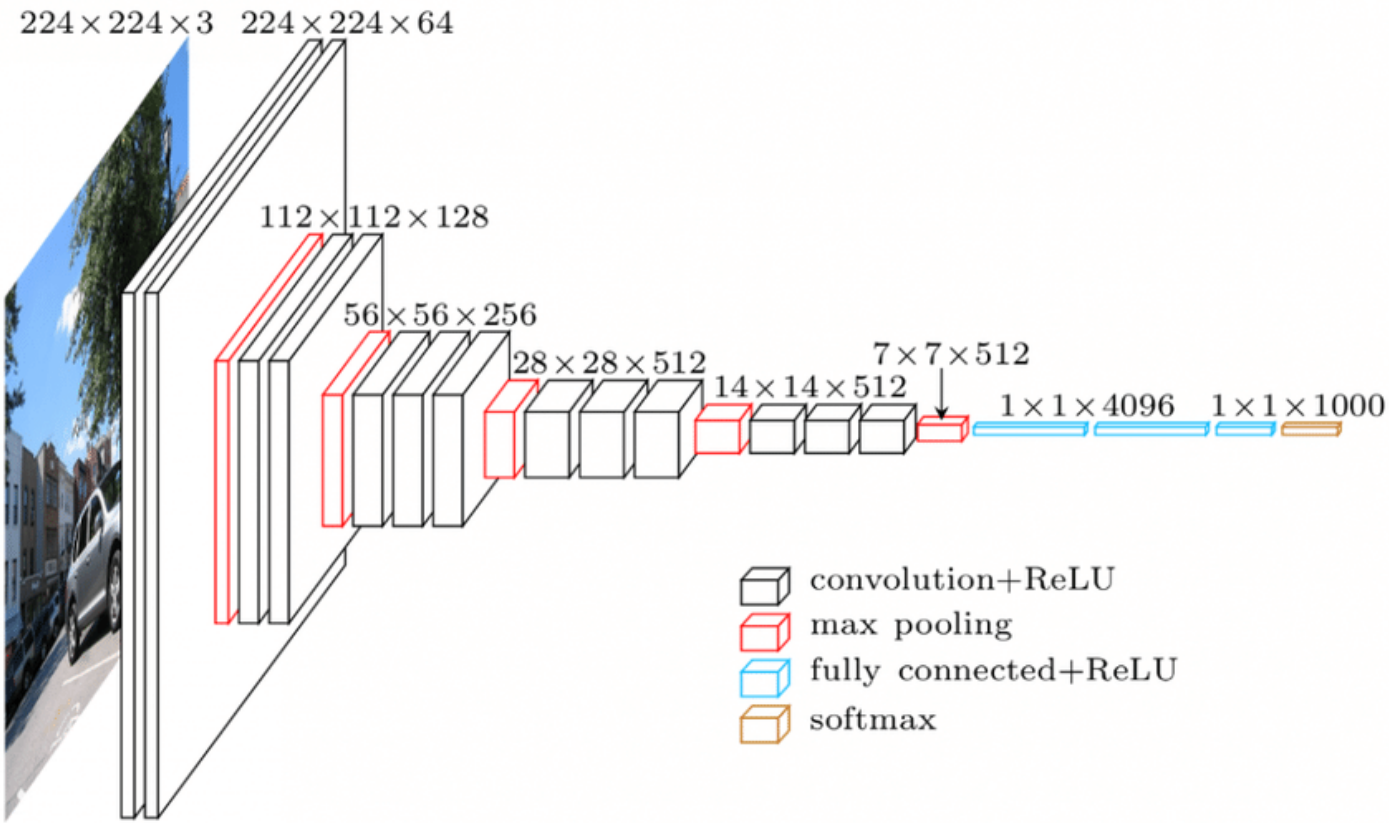
VGG 2014

Published as a conference paper at ICLR 2015

VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION

Karen Simonyan* **& Andrew Zisserman+**
Visual Geometry Group, Department of Engineering Science, University of Oxford

https://arxiv.org/pdf/1409.1556.pdf



VGG 2014

- Small kernels
- Gradient descent with momentum (inertia)
- Dropout
- About 147,000,000,000 parameters
- Specific initialization
- Test error 8%

GoogLeNet (Inception) 2014

Going deeper with convolutions

Christian Szegedy
Google Inc.

Wei Liu
University of North Carolina, Chapel Hill

Yangqing Jia
Google Inc.

Pierre Sermanet
Google Inc.

Scott Reed
University of Michigan

Dragomir Anguelov
Google Inc.

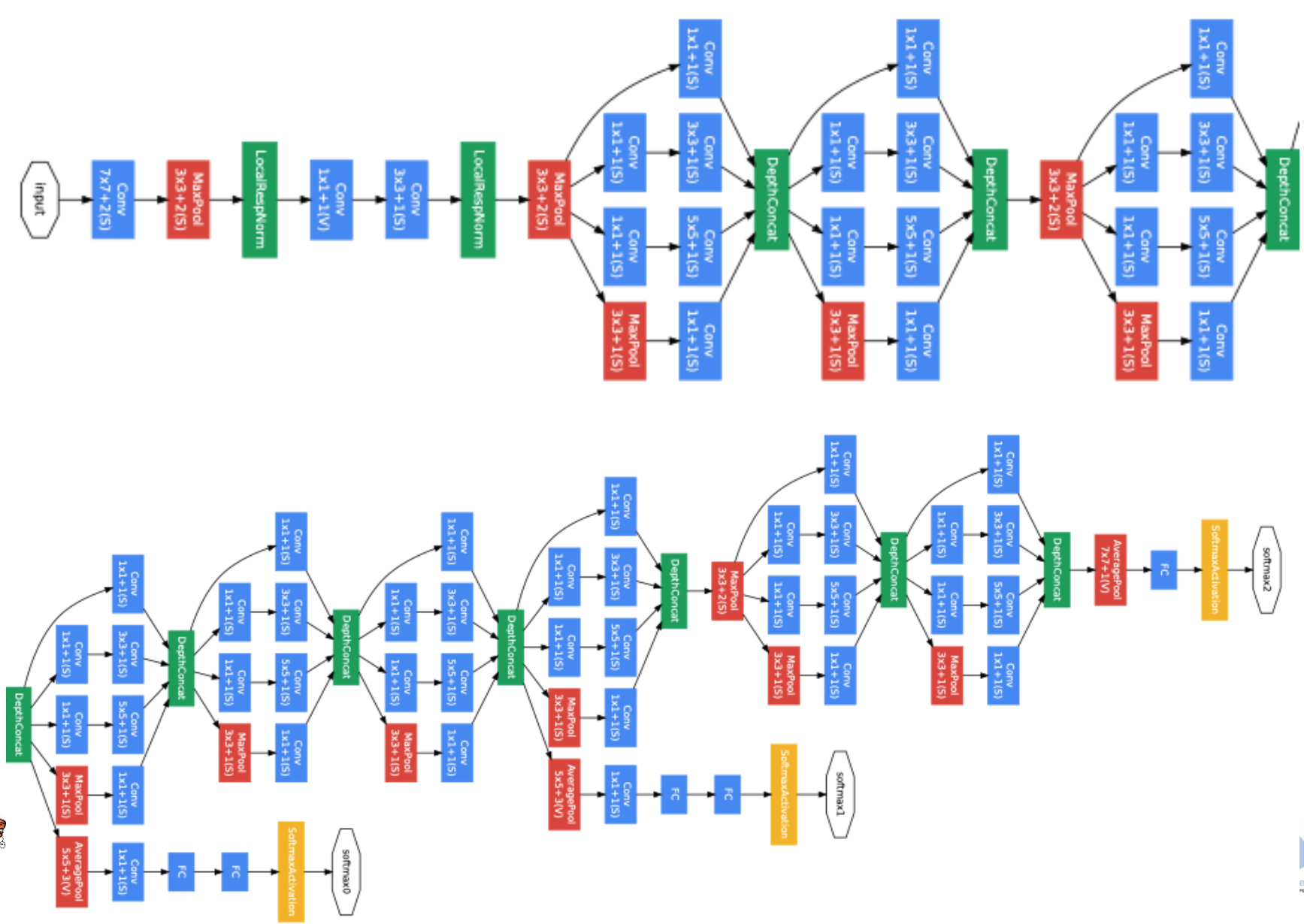
Dumitru Erhan
Google Inc.

Vincent Vanhoucke
Google Inc.

Andrew Rabinovich
Google Inc.

https://arxiv.org/abs/1409.4842

GoogLeNet (Inception) 2014



GoogLeNet 2014

- 1 x 1 convolutions
- More than one outputs
- Gradient descent with inertia
- About 25,000,000,000 parameters
- Test error 6.67%

ResNet 2015

Deep Residual Learning for Image Recognition

Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun
Microsoft Research

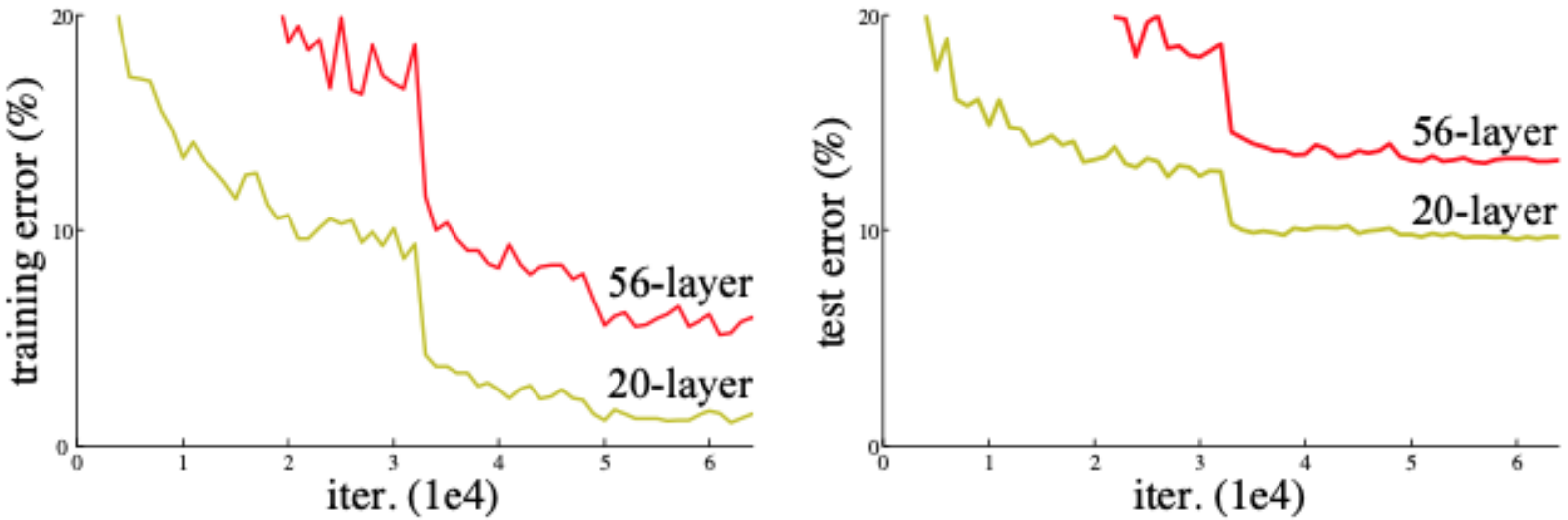


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer "plain" networks. The deeper network has higher training error, and thus test error. Similar phenomena on ImageNet is presented in Fig. 4.

https://arxiv.org/pdf/1512.03385.pdf

ResNet 2015

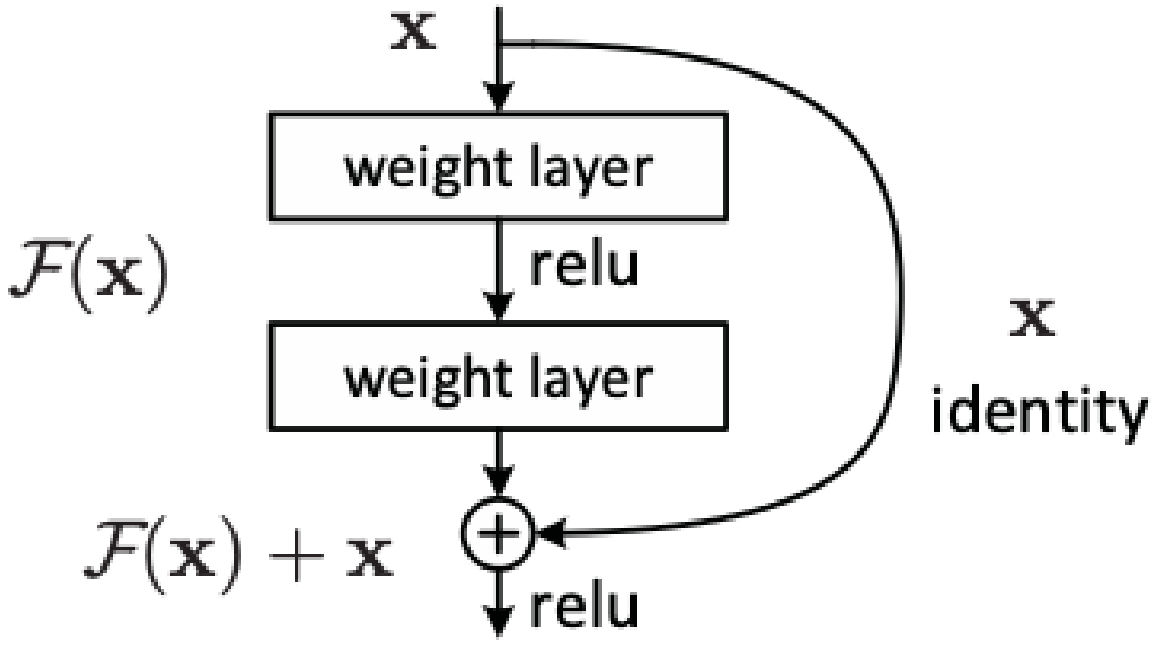
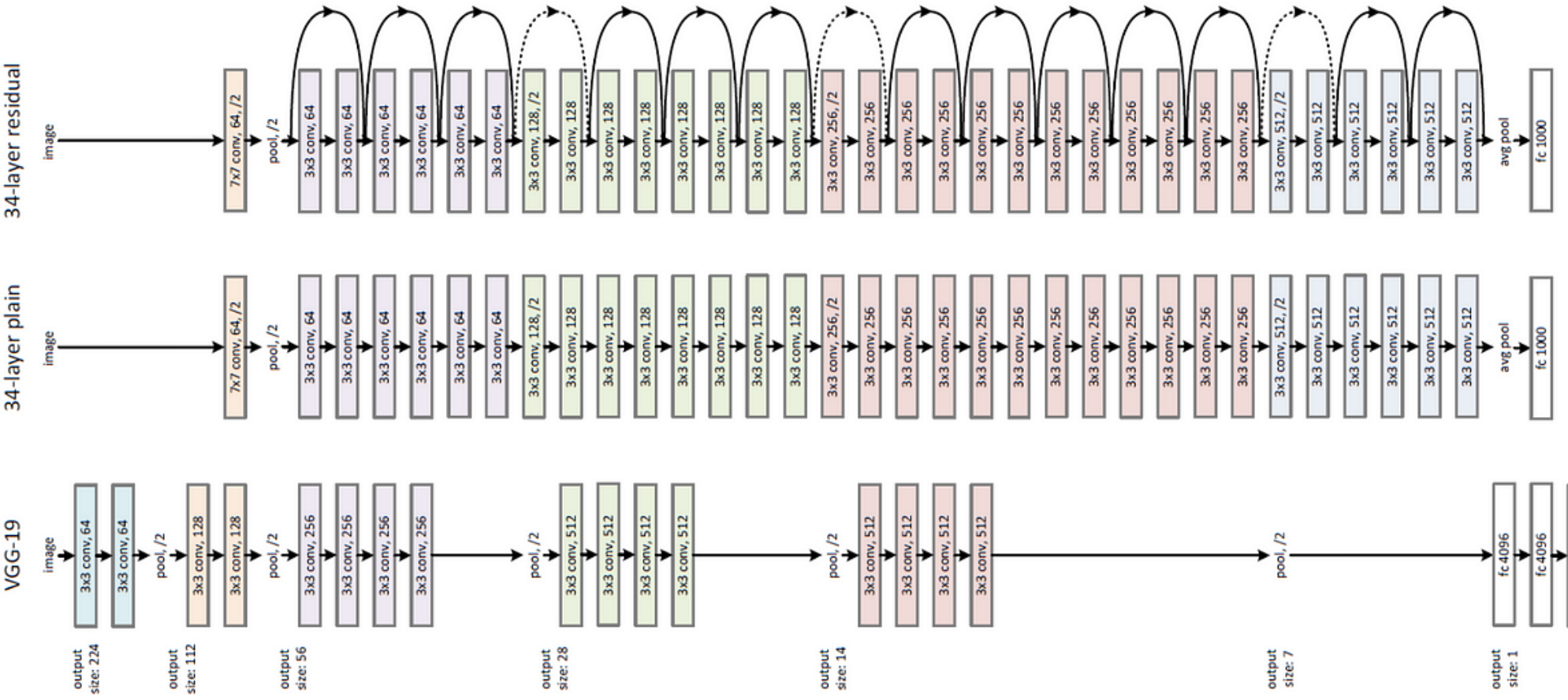


Figure 2. Residual learning: a building block.

https://arxiv.org/pdf/1512.03385.pdf

ResNet 2015



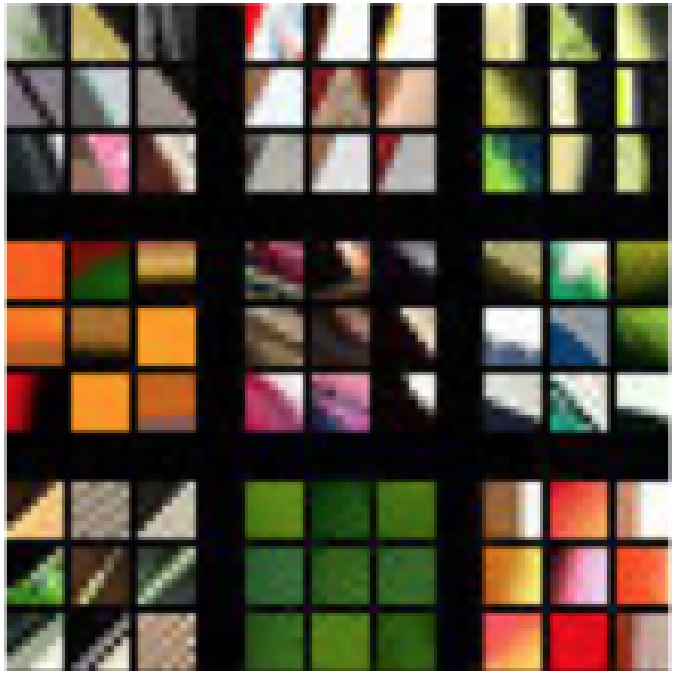
https://arxiv.org/pdf/1512.03385.pdf

ResNet 2015

- Residual connections
- Low error with up to 1000 layers
- About 270,000,000 – 19,000,000,000 parameters
- Test error 4.5%

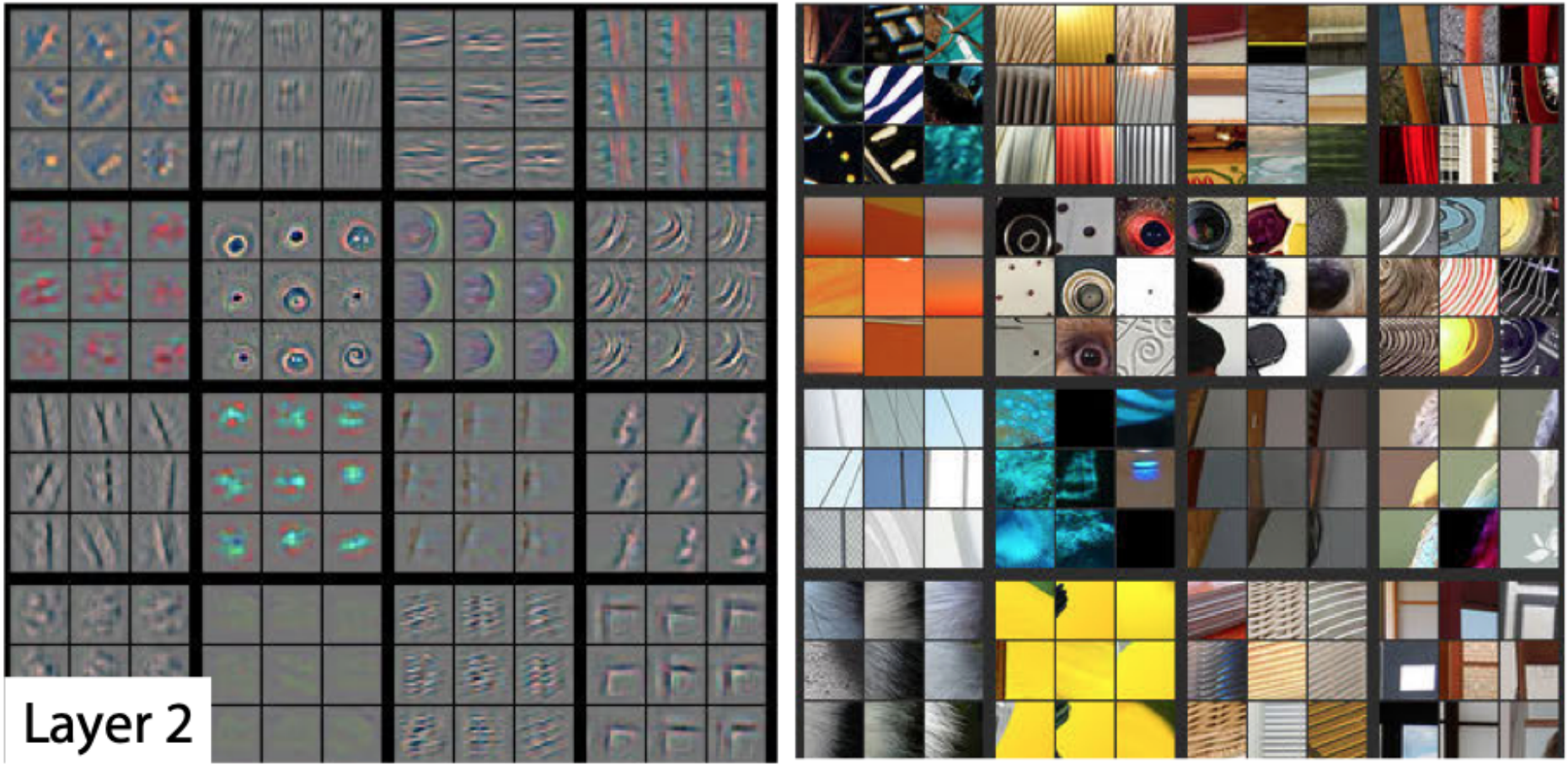
Representation of the filters

Layer 1



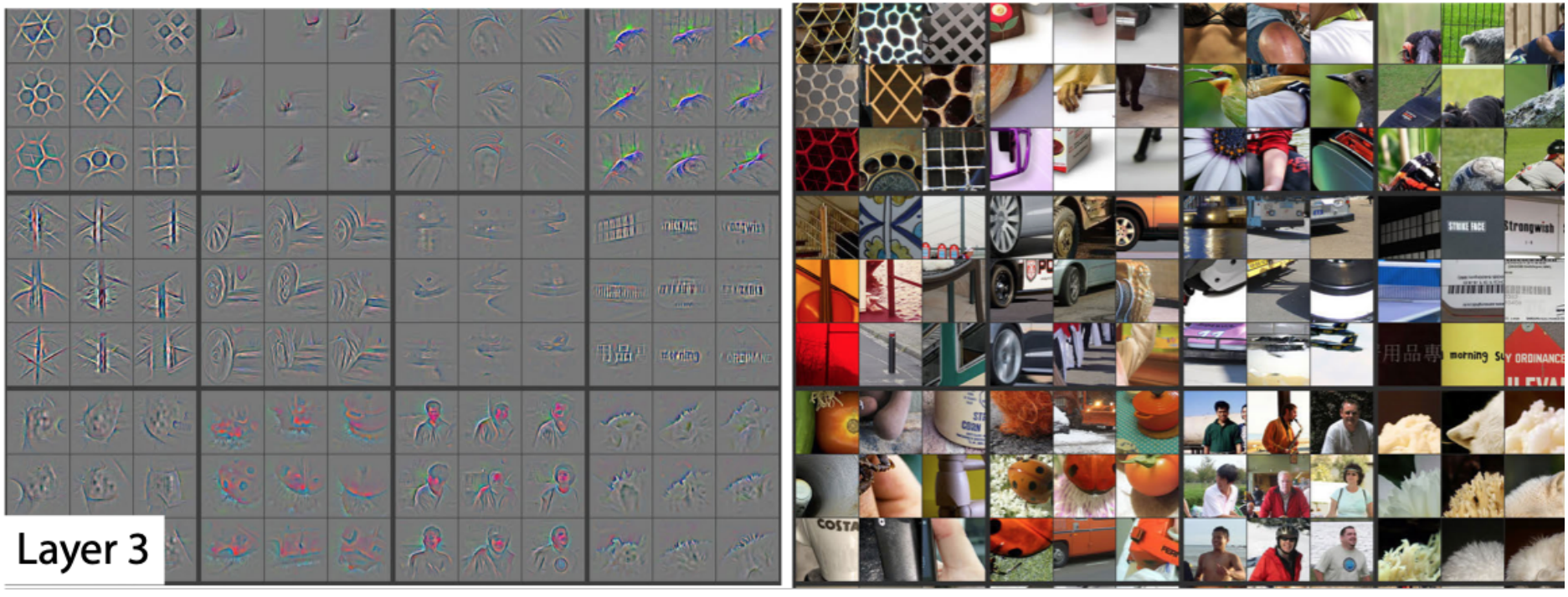
https://arxiv.org/pdf/1311.2901.pdf

Representation of the filters



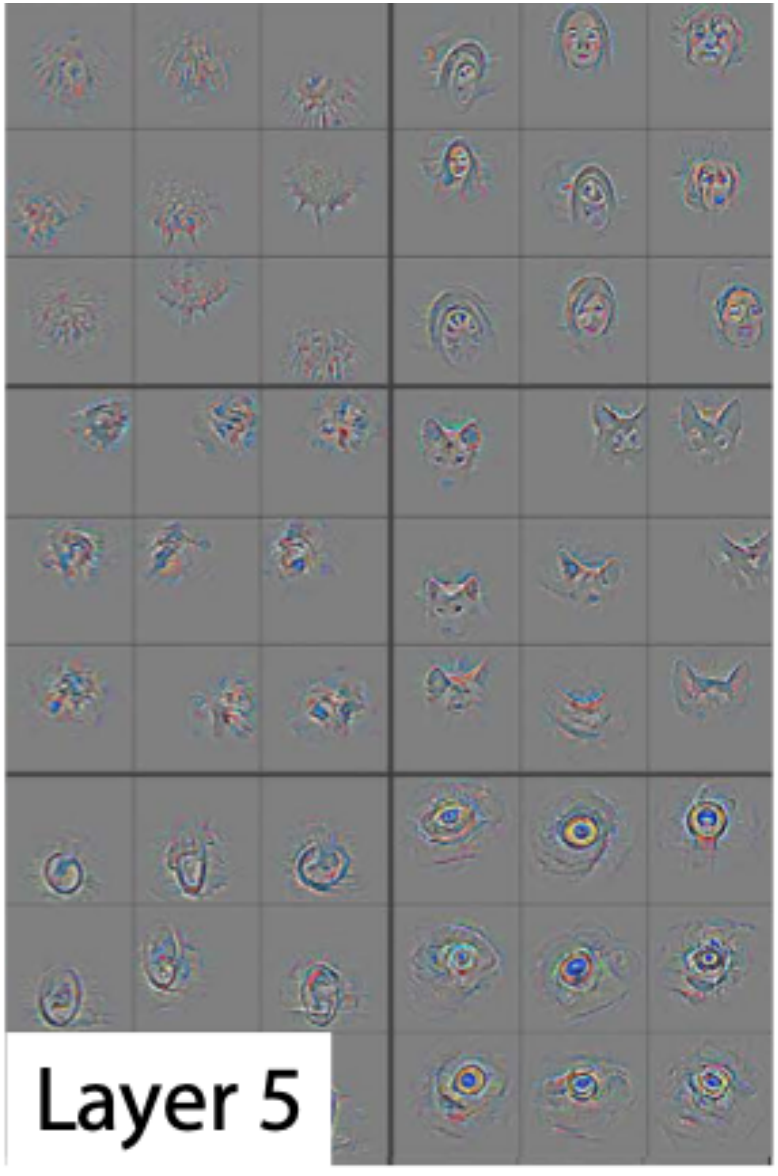
https://arxiv.org/pdf/1311.2901.pdf

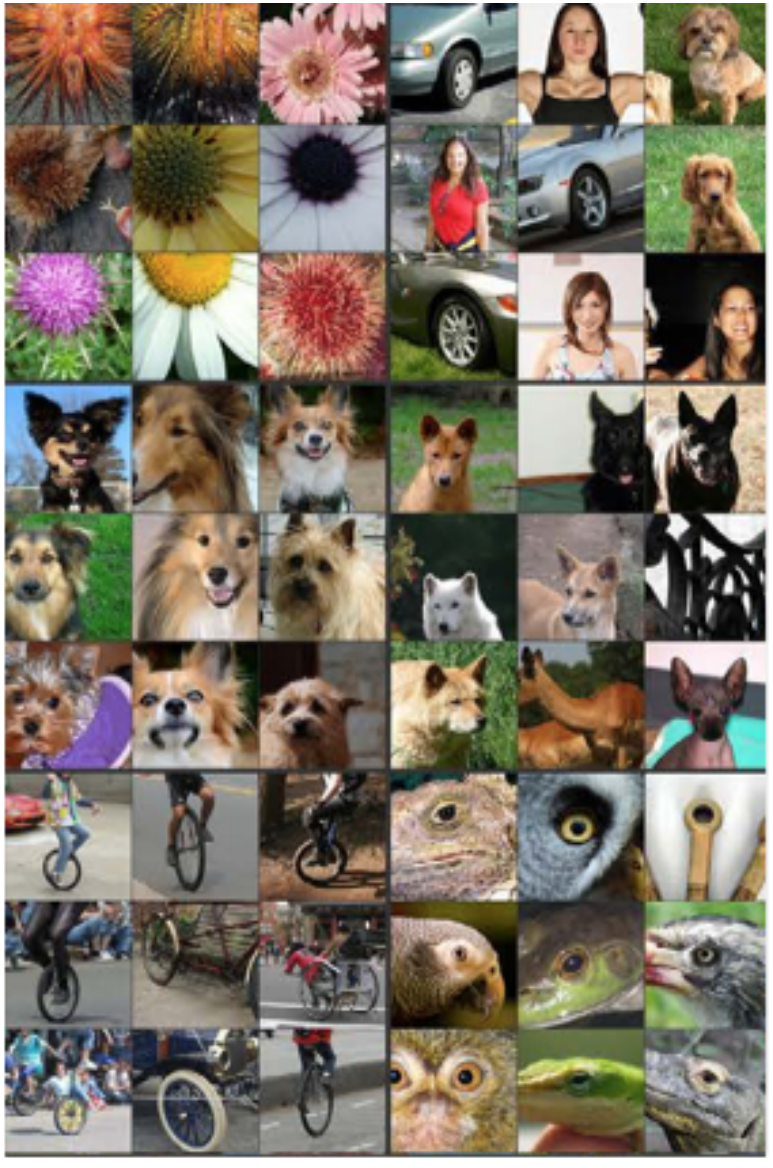
Representation of the filters



https://arxiv.org/pdf/1311.2901.pdf

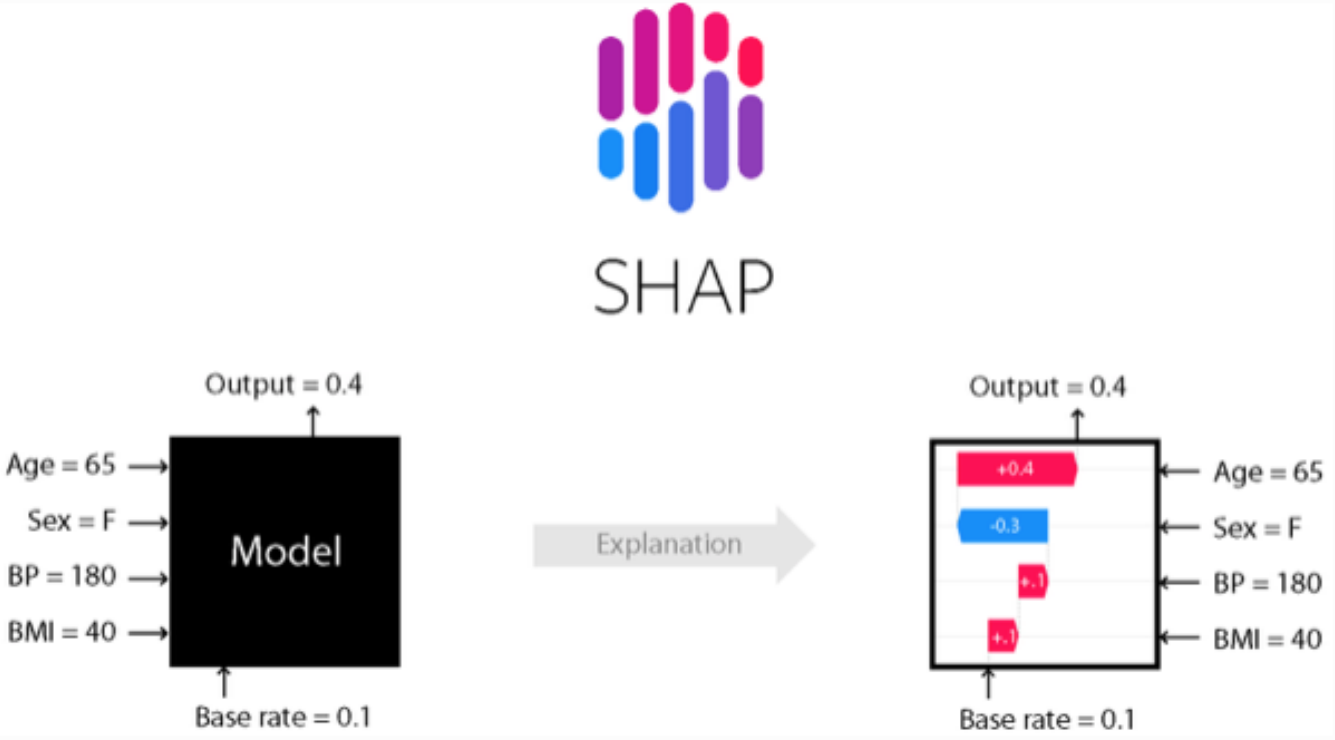
Representation of the filters





https://arxiv.org/pdf/1311.2901.pdf

Shapley Values



SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the classic Shapley values from game theory and their related extensions (see papers for details and citations).

`https://shap.readthedocs.io/en/latest/`