RECOD
reasoning for complex data

Deep Neural Networks

Machine Learning and Pattern Recognition

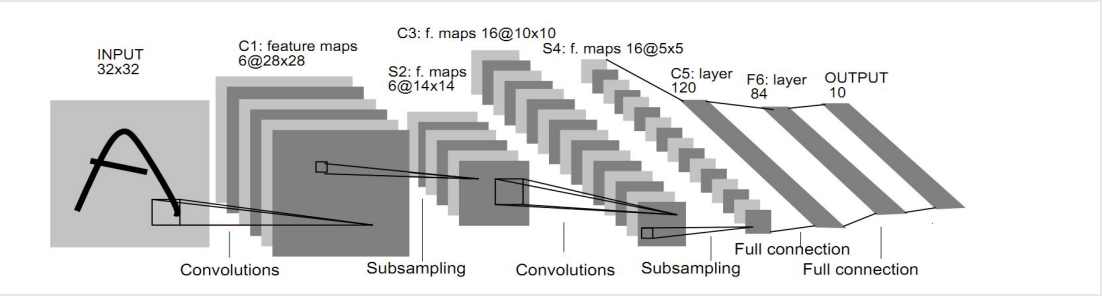
(Largely based on slides from Fei-Fei Li & Justin Johnson & Serena Yeung)

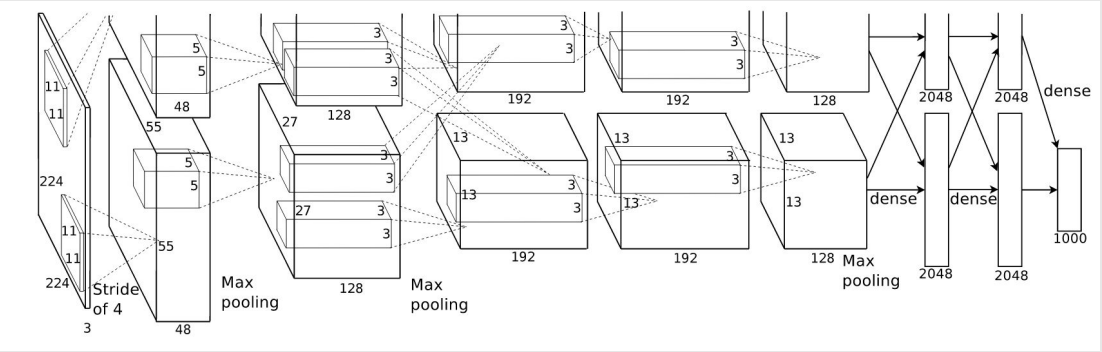
Prof. Sandra Avila
Institute of Computing (IC/Unicamp)

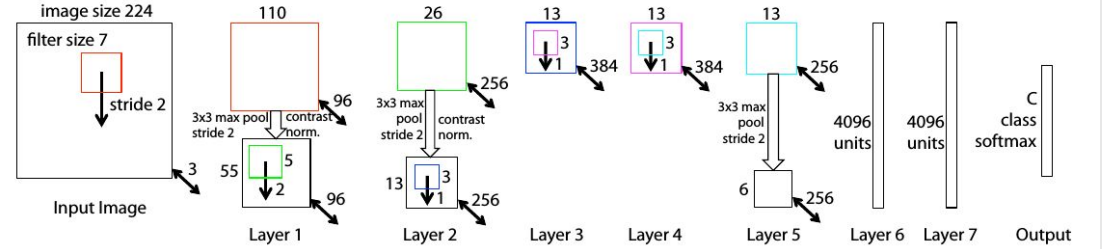
MC886/MO444, October 16, 2018

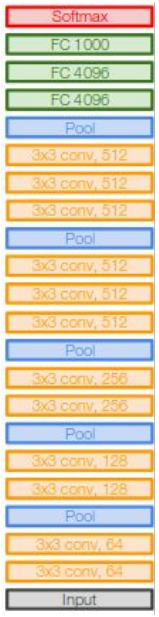
CNNs Architectures

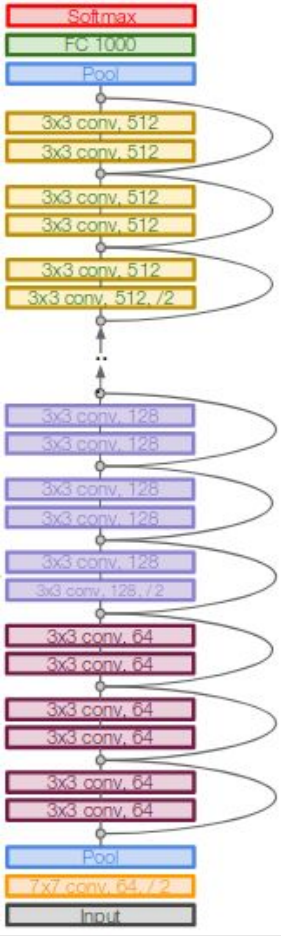
CNN-based Architectures



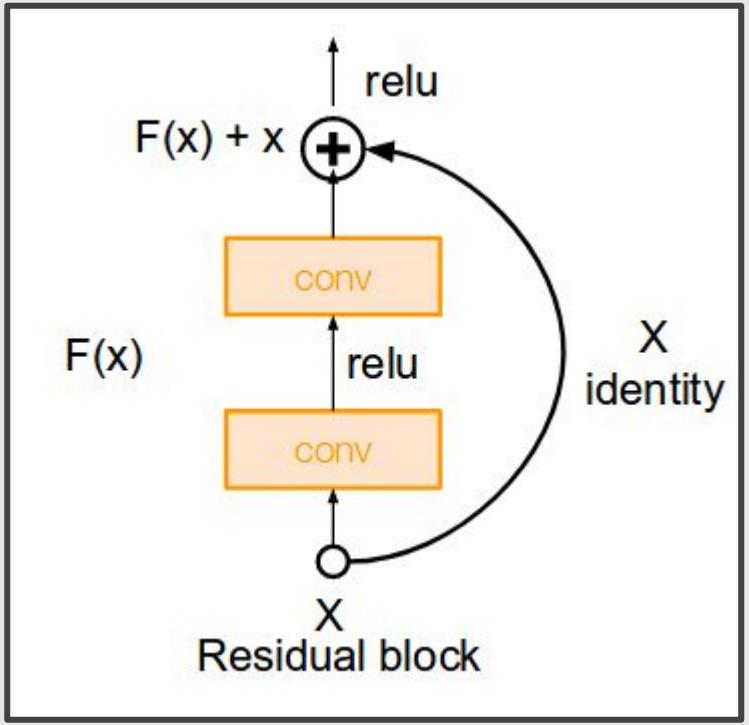


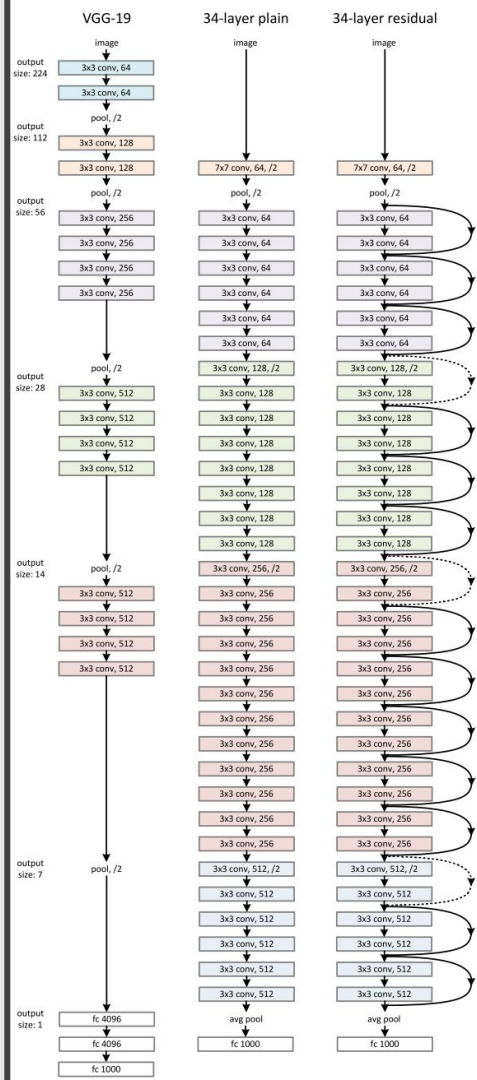






ResNet [He et al., 2015]





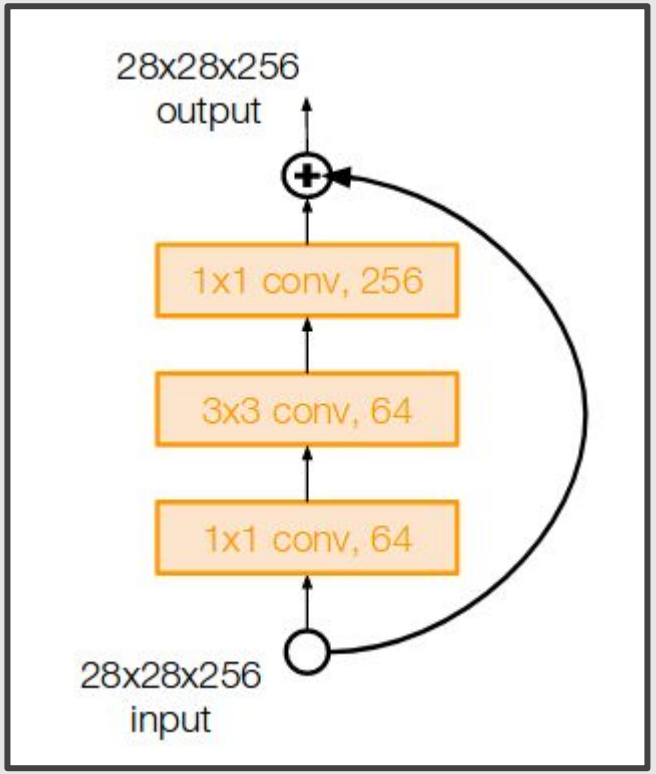
ResNet [He et al., 2015]

For deeper networks (**ResNet-50+**), use "bottleneck" layer to improve efficiency (similar to GoogLeNet)

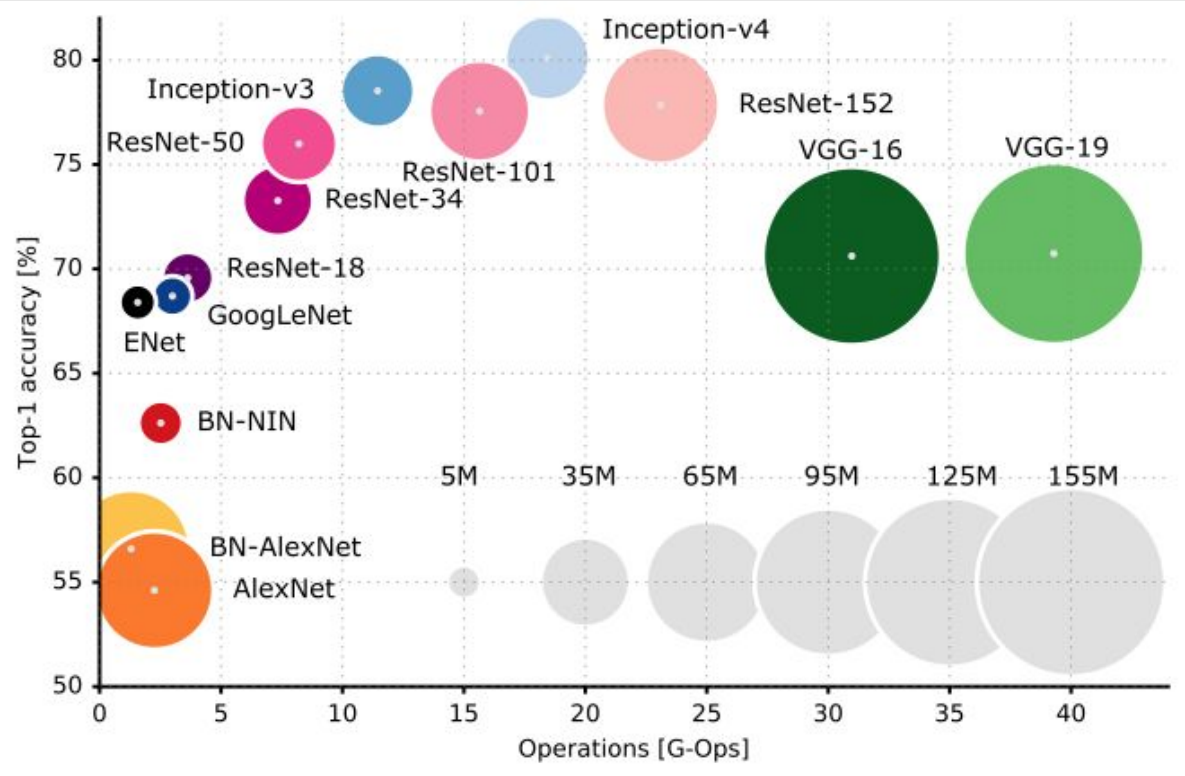
1x1 conv, 256 filters projects back to 256 feature maps (28x28x256)

3x3 conv operates over only 64 feature maps

1x1 conv, 64 filters to project to 28x28x64



The size of the blobs is proportional to the number of network parameters.



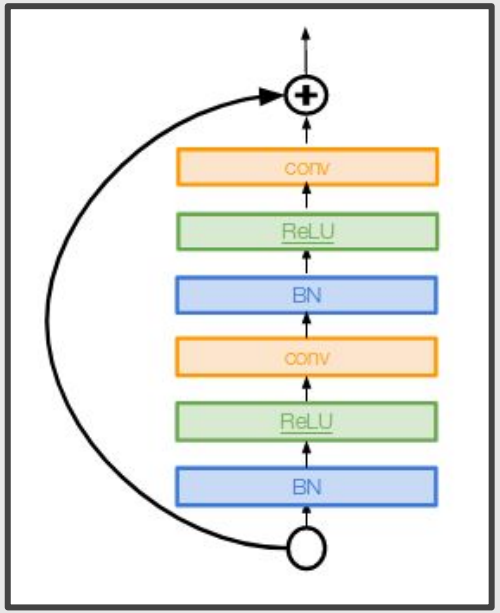
https://medium.com/towards-data-science/neural-network-architectures-156e5bad51ba

Other CNNs Architectures

Improving ResNet ...

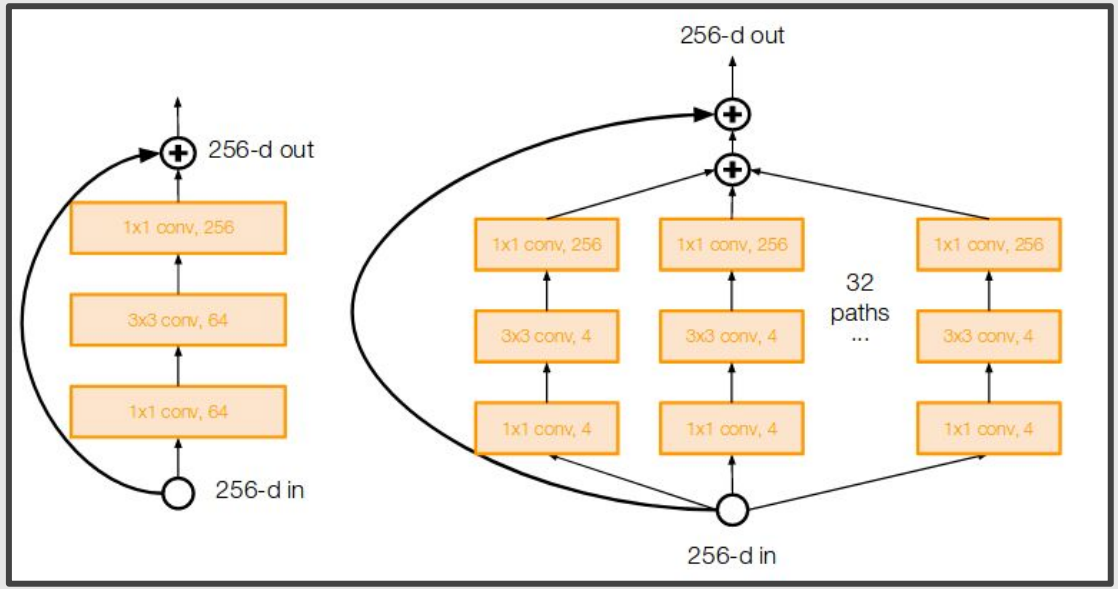
Identity Mappings in Deep Residual Networks [He et al., 2016]

- Improved ResNet block design from creators of ResNet
- Creates a more direct path for propagating information throughout network (moves activation to residual mapping pathway)
- Gives better performance



Improving ResNet ...

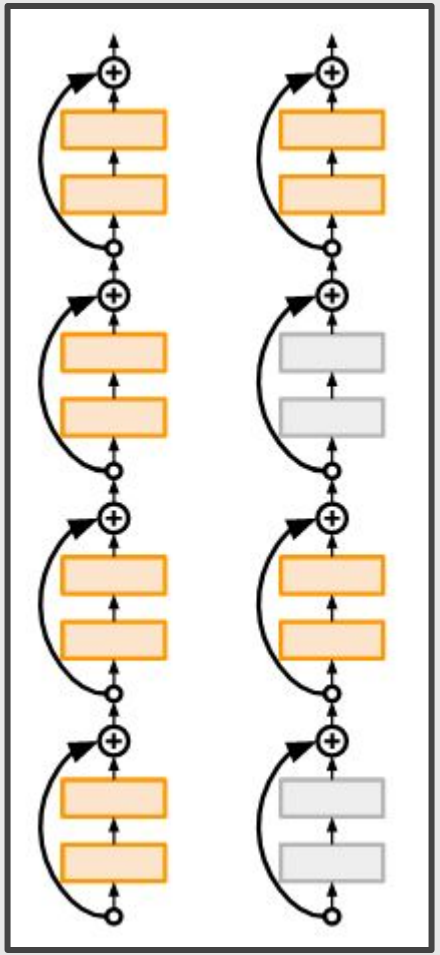
Aggregated Residual Transformations for Deep Neural Networks (**ResNeXt**) [Xie et al., 2016]



Improving ResNet ...

Deep Networks with Stochastic Depth [Huang et al., 2016]

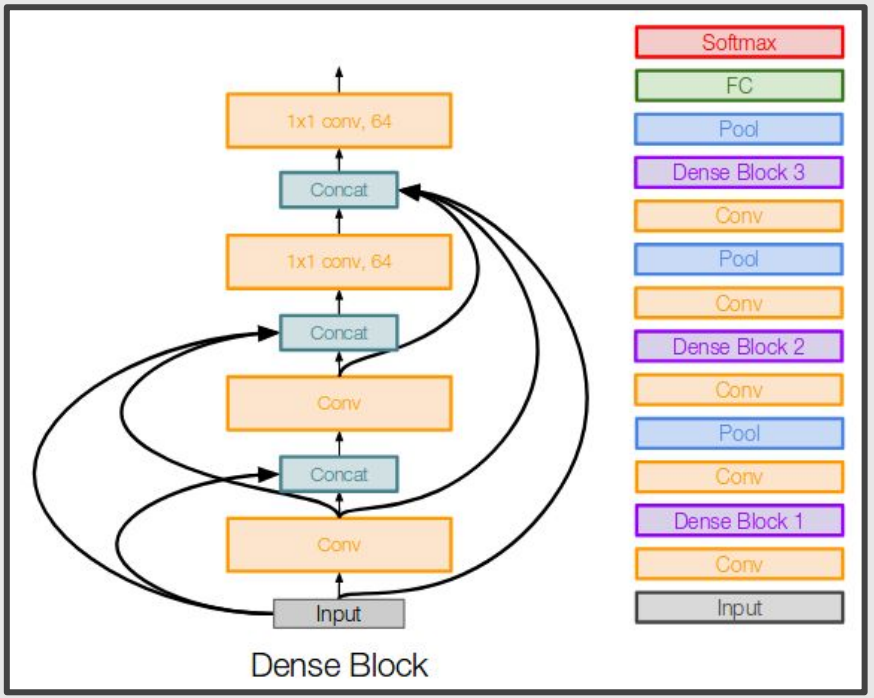
- Motivation: reduce vanishing gradients
- Randomly drop a subset of layers during each training pass
- Bypass with identity function



Beyond ResNet ...

Densely Connected Convolutional Networks (**DenseNet**) [Huang et al., 2017]

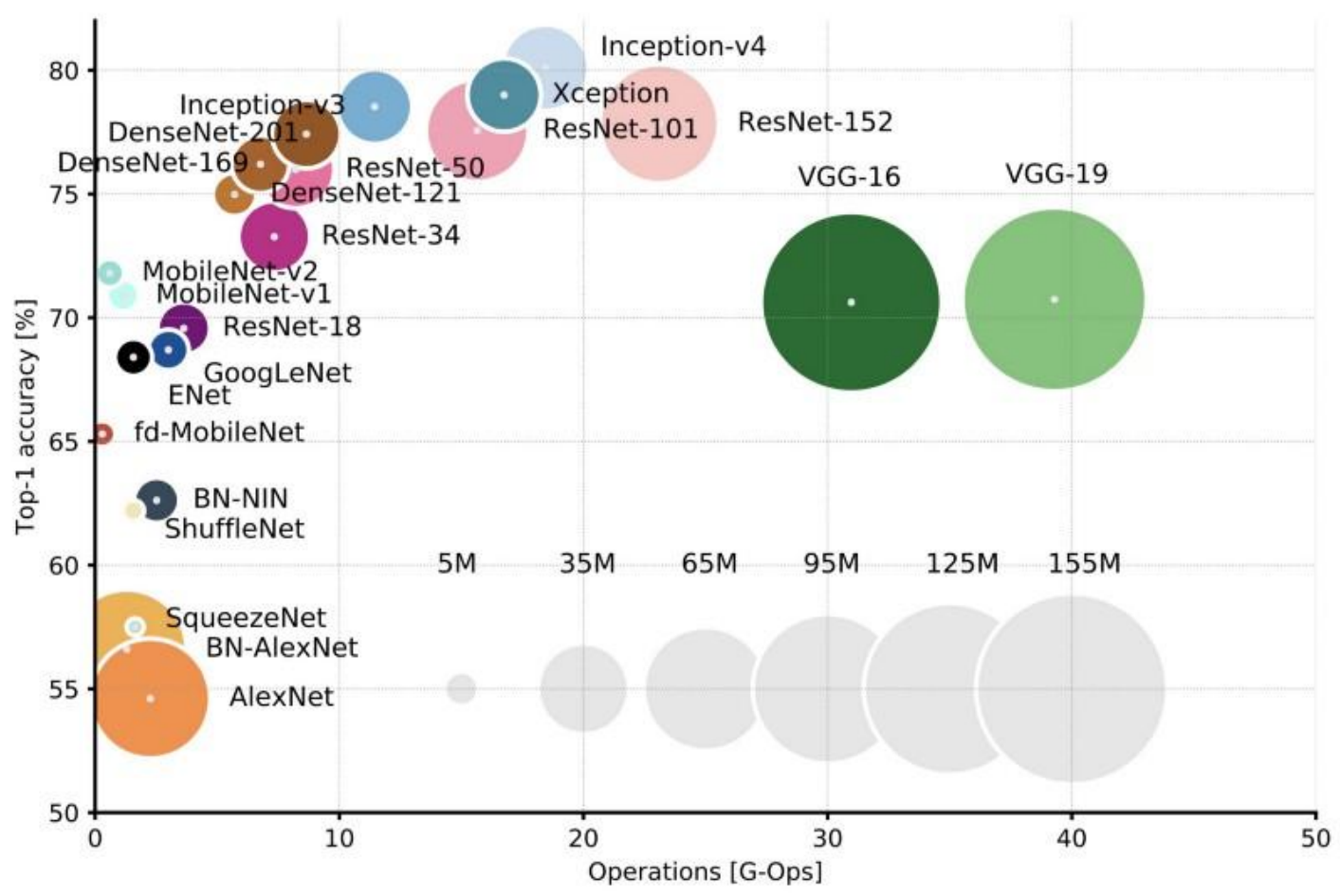
- Dense blocks where each layer is connected to every other layer in feedforward fashion
- Alleviates vanishing gradient, strengthens feature propagation, encourages feature reuse



Summary

- VGG, GoogLeNet, ResNet all in wide use, available in model zoos
- ResNet current best default
- Trend towards extremely deep networks
- Significant research centers around design of layer / skip connections and improving gradient flow
- Even more recent trend towards examining necessity of depth vs. width and residual connections

The size of the blobs is proportional to the **number of network parameters**.



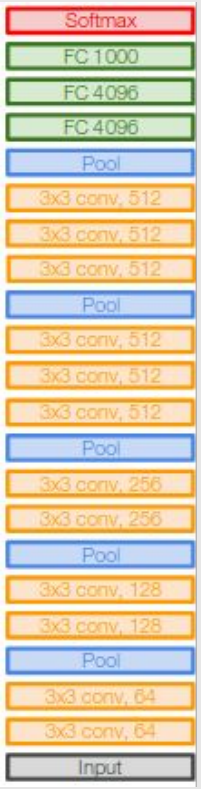
https://medium.com/@culurciello/analysis-of-deep-neural-networks-dcf398e71aae

Transfer Learning

“You need a lot of a data if you want to train/use CNNs”

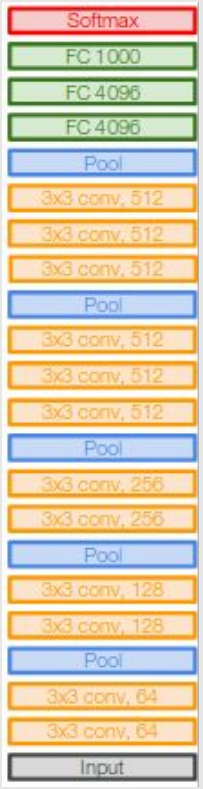
Transfer Learning

Transfer Learning with CNNs

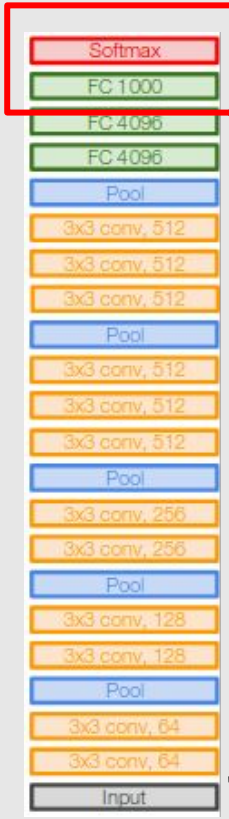


1. Train on ImageNet

Transfer Learning with CNNs



1. Train on ImageNet

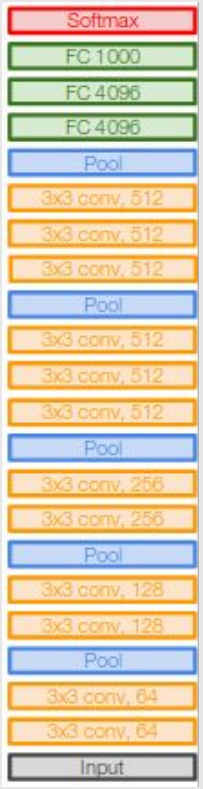


Train this

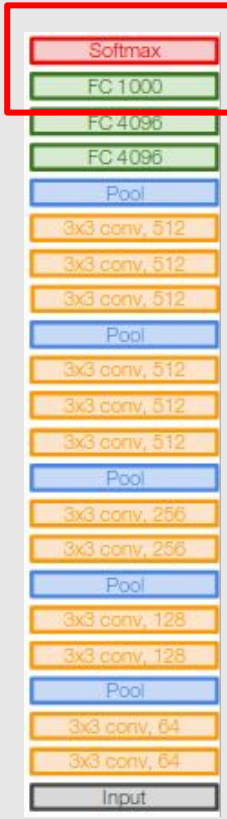
Freeze these

2. Small dataset: **Feature extractor**

Transfer Learning with CNNs



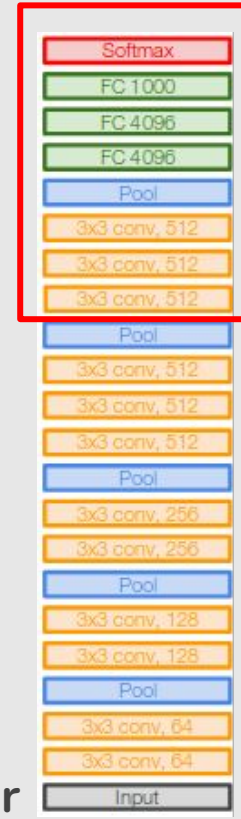
1. Train on ImageNet



Train this

Freeze these

2. Small dataset: **Feature extractor**



Train this

Freeze these

3. Medium dataset: **Finetuning**

References

Machine Learning Books

- Hands-On Machine Learning with Scikit-Learn and TensorFlow, Chap. 11 & 13

Machine Learning Courses

- https://www.coursera.org/learn/neural-networks
- “The 3 popular courses on Deep Learning”: https://medium.com/towards-data-science/the-3-popular-courses-for-deeplearning-ai-ac37d4433bd