CSCI4622 Machine Learning 2020 Spring

Convolutional Neural Networks (3)

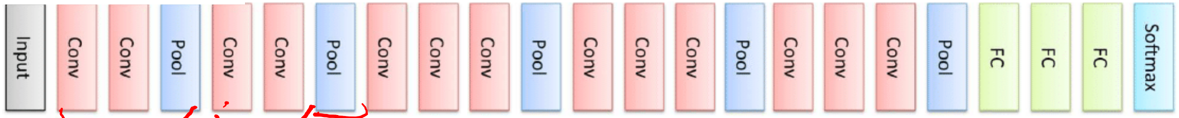
Geena Kim

University of Colorado Boulder

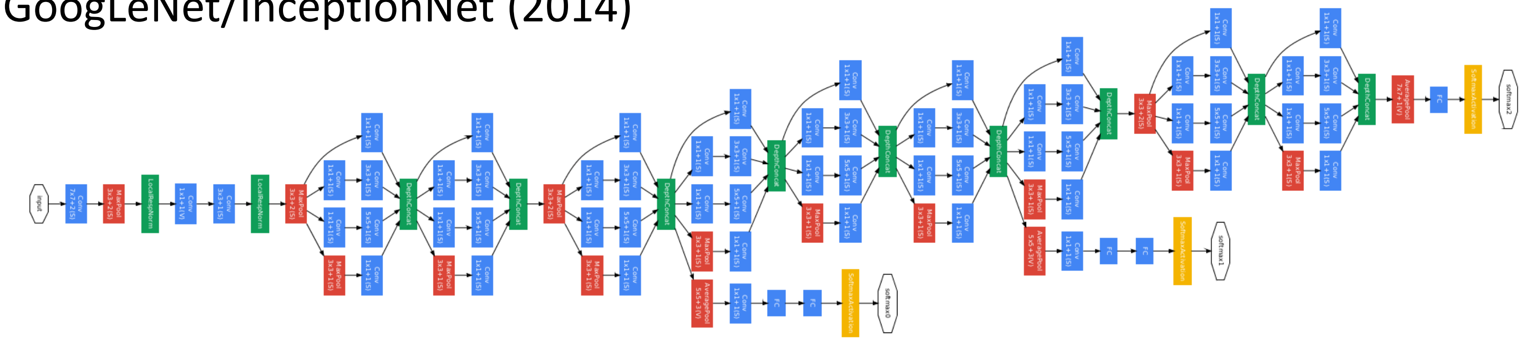
Most successful CNN models from ImageNet

Review

VGGNet (2014)

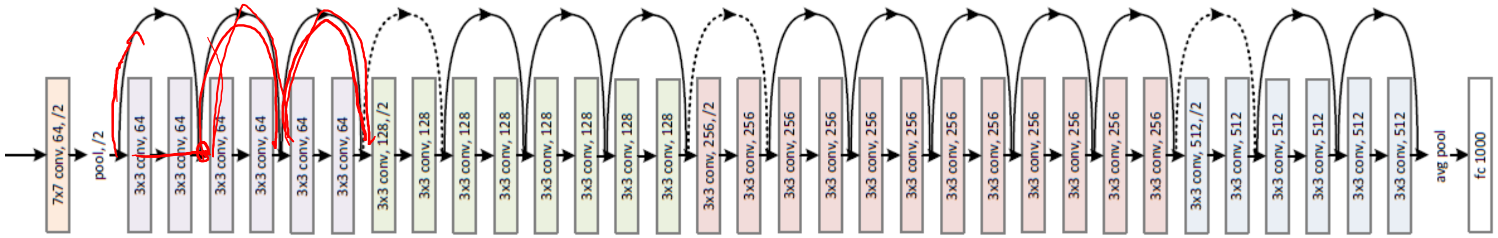


GoogLeNet/InceptionNet (2014)



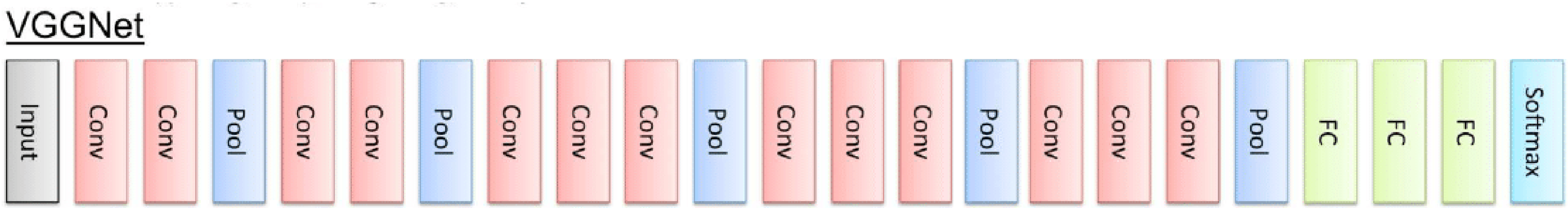
1x1

ResNet (2015)



VGGNet

Review



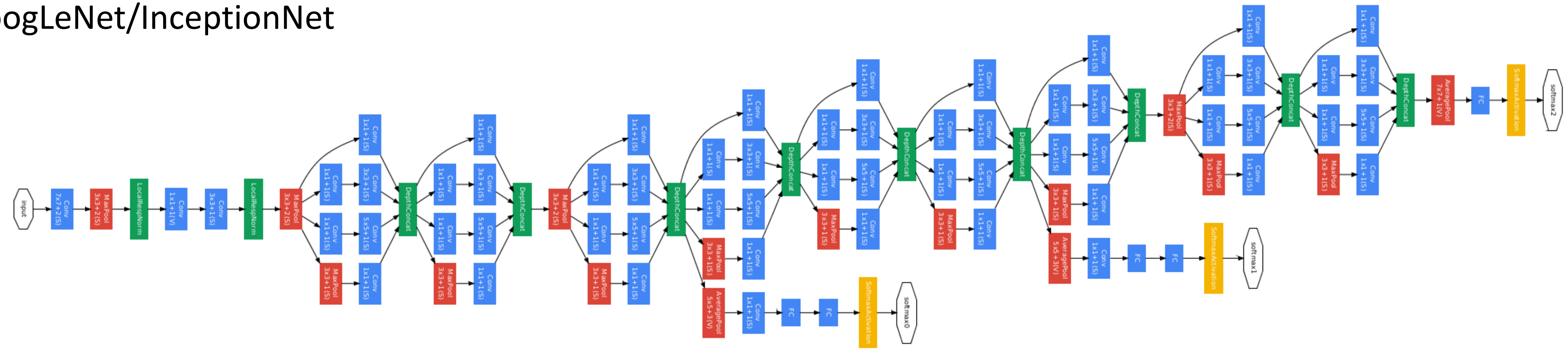
https://arxiv.org/abs/1409.1556

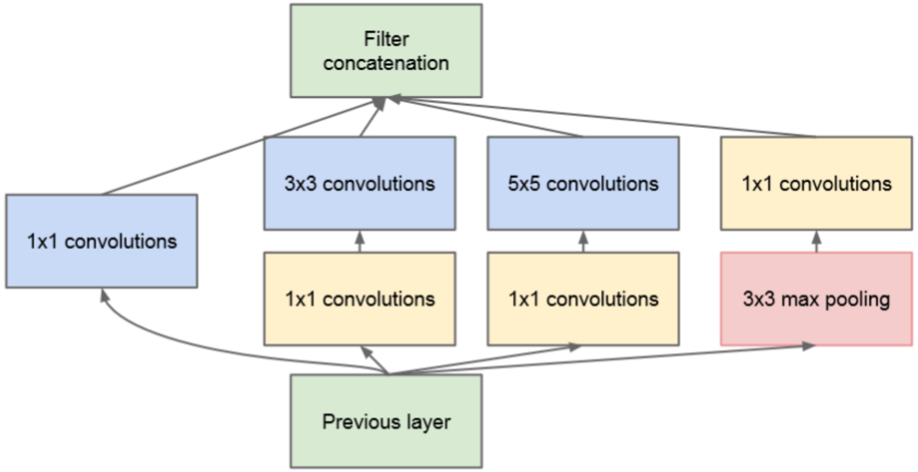
ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	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					

GoogLeNet

Review

GoogLeNet/InceptionNet

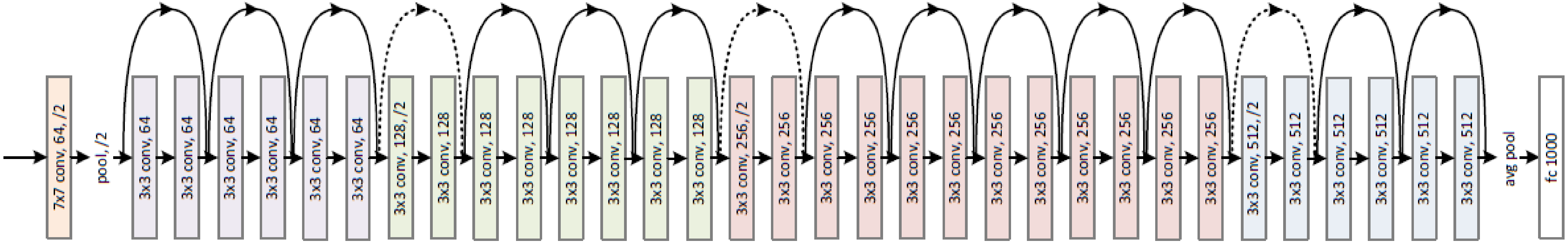




(b) Inception module with dimensionality reduction

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

ResNet



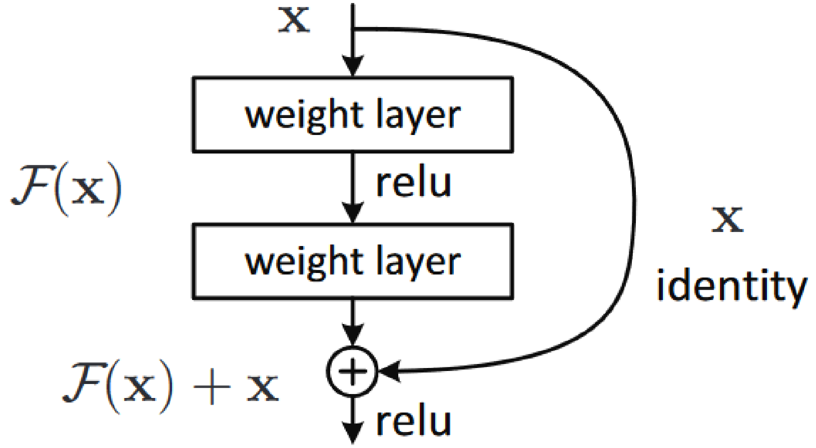
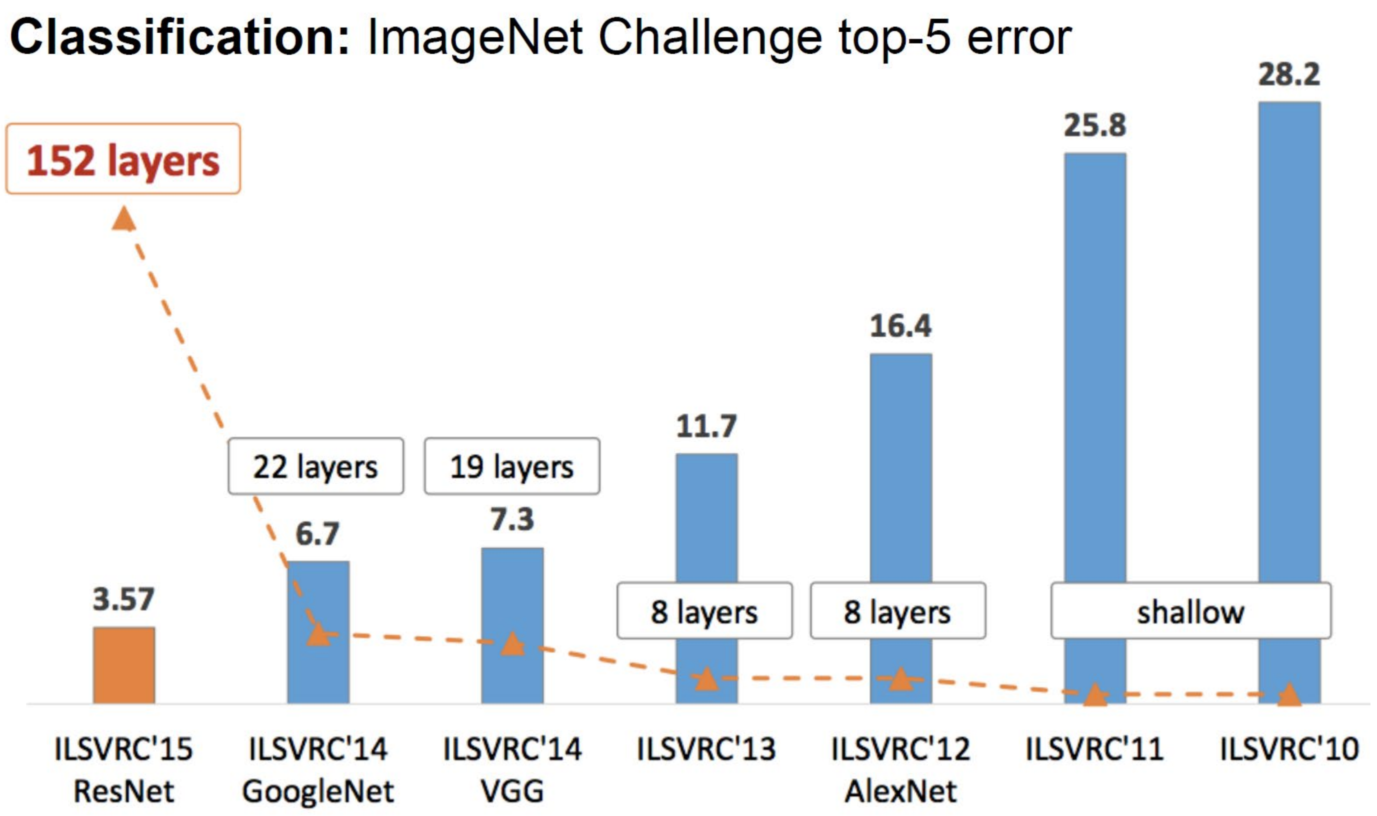


Figure 2. Residual learning: a building block.

http://arxiv.org/abs/1512.03385

ResNet



Credit: Kaiming He

Architecture summary

Review

- If the data is bigger and complex, deep network is beneficial

- Improve computational efficiency while keeping the same depth
- e.g. small filter size, 1x1 filter compression

- Designs to have a better gradient propagation (avoid vanishing or exploding gradients)

Dense

Design Choices or Training

Key Tuning Parameters in Training

Review

- Optimization method SGD
- Learning rate, momentum, etc
- Number of epochs
- Regularization

Monitoring Overfitting in Training

Dataset split

Train / Validation / Test

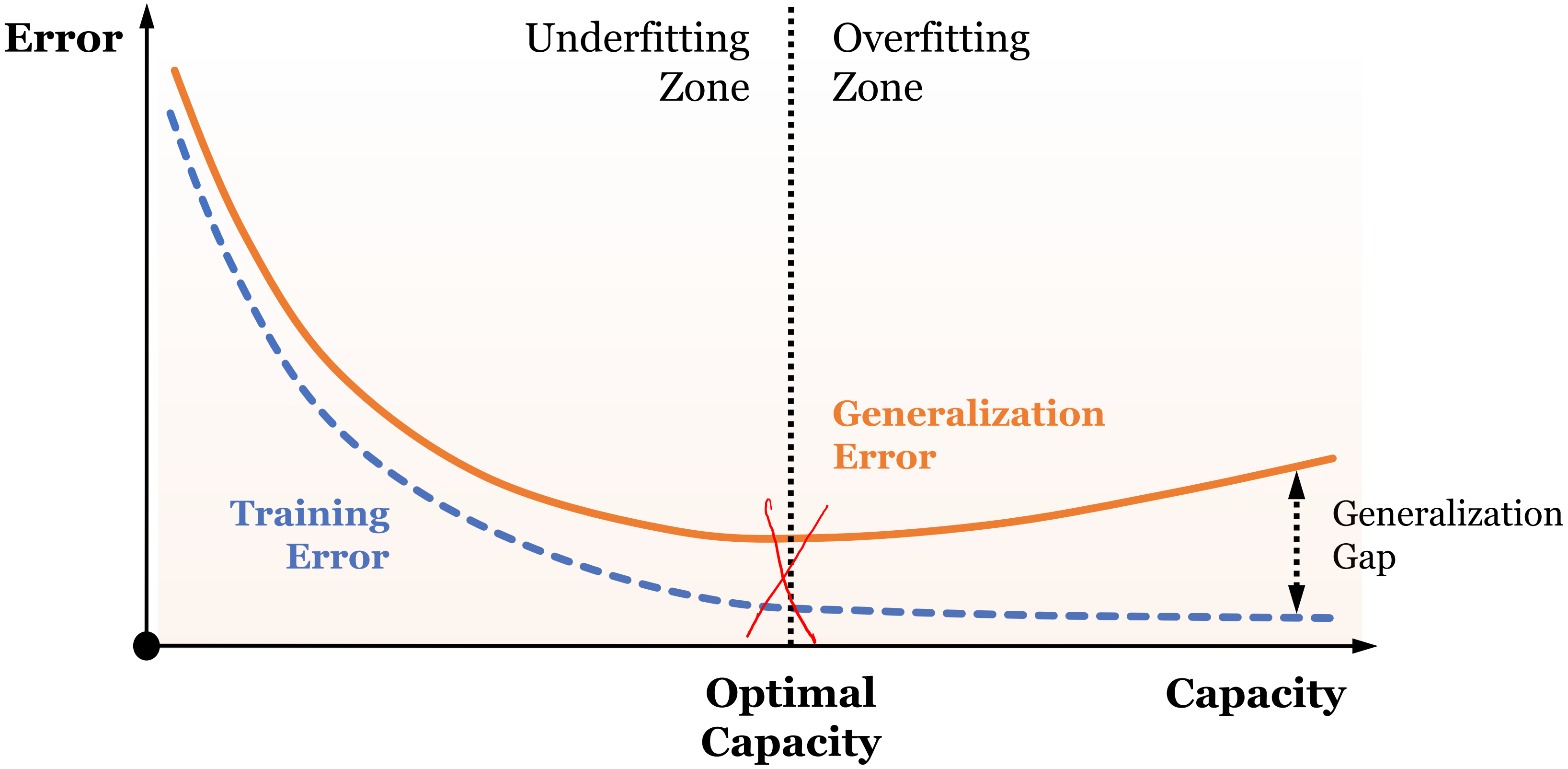
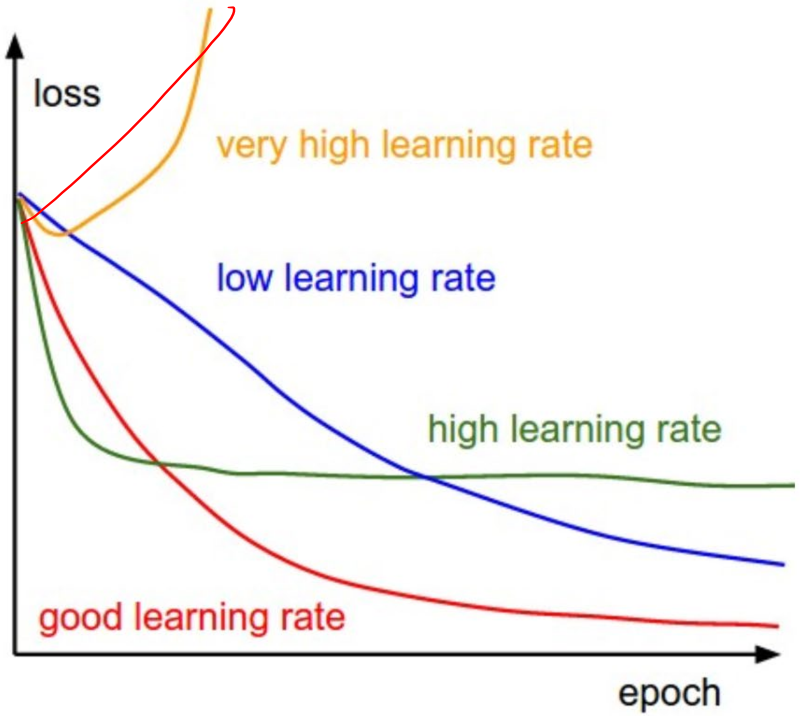
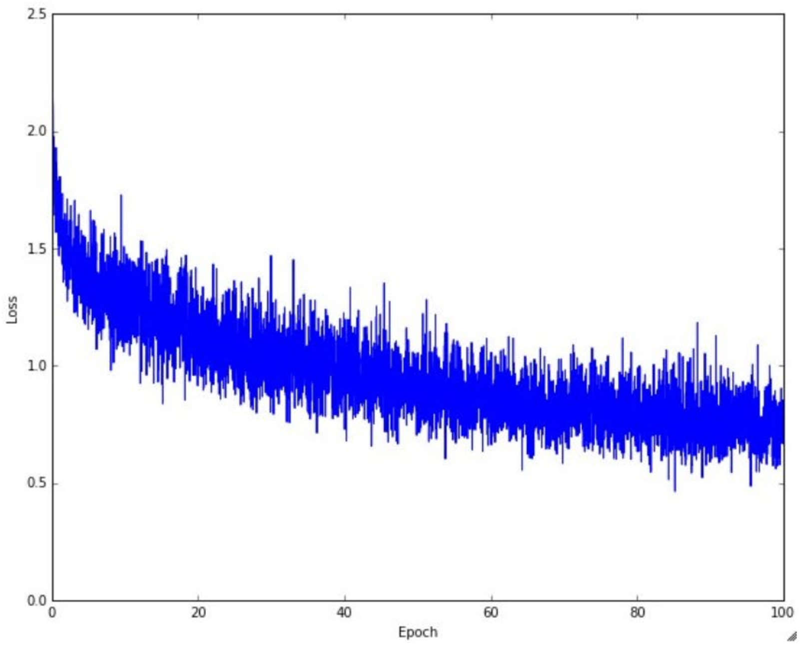
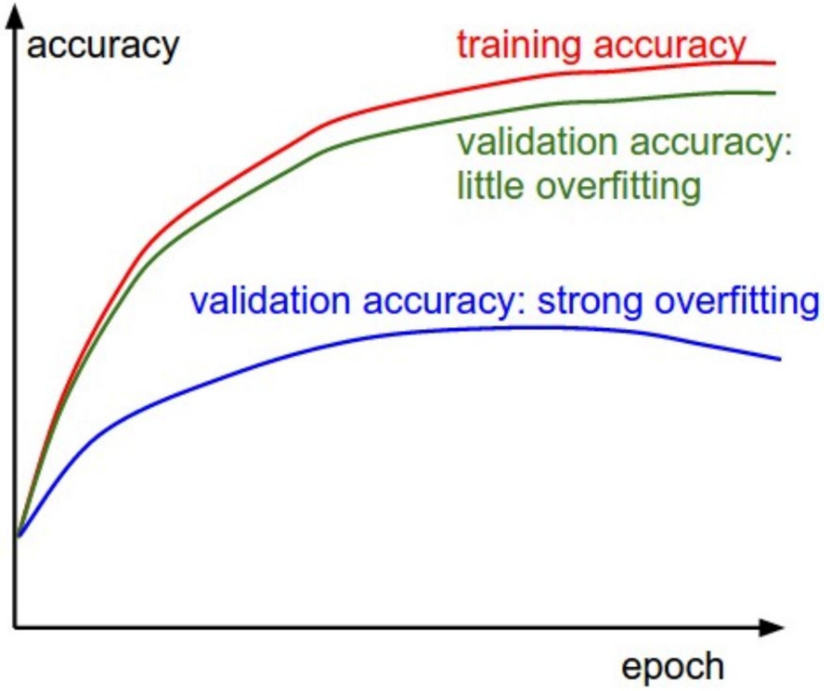


Diagram credit: Fei-Fei Li

Monitoring Overfitting in Training







http://cs231n.github.io/neural-networks-3/

SGD tuning parameters

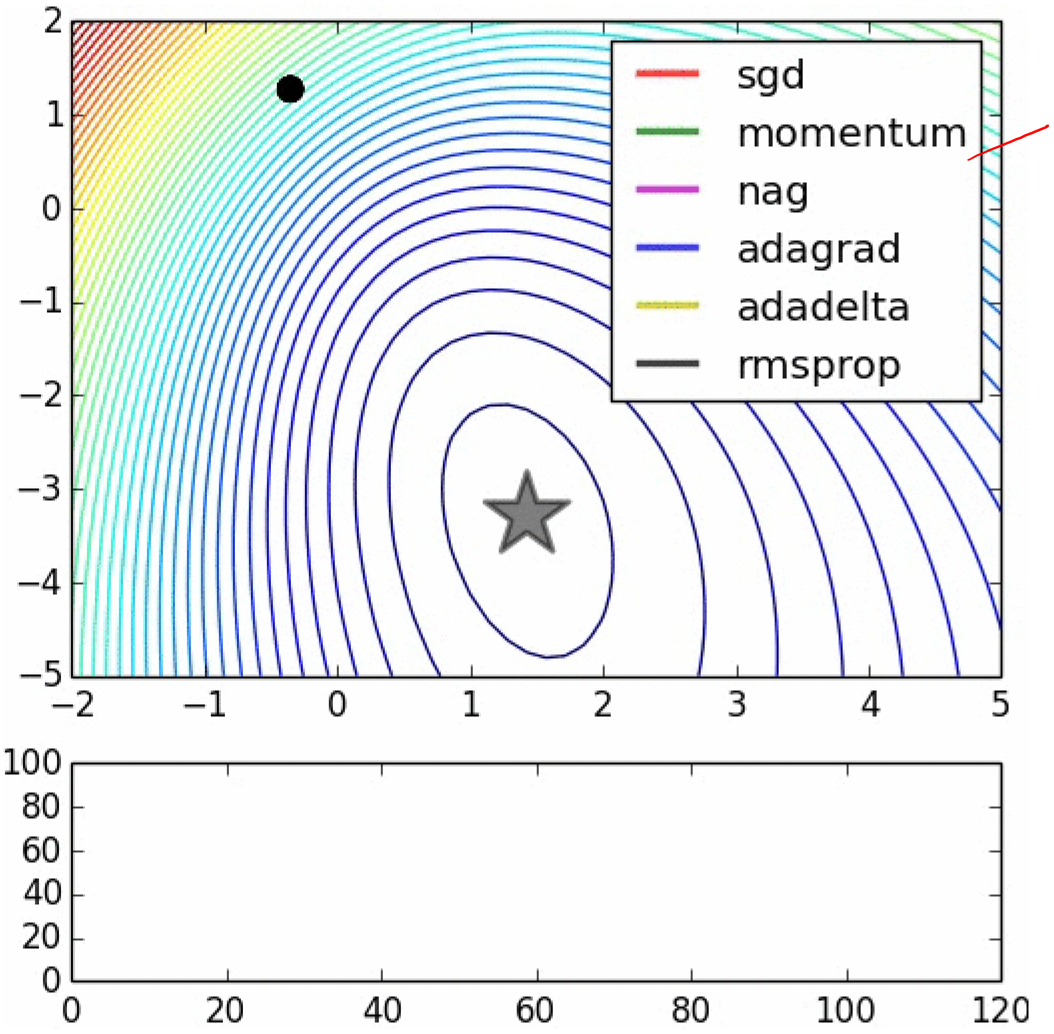

```
tf.keras.optimizers.SGD(
    learning_rate=0.01, momentum=0.0, nesterov=False, name='SGD', **kwargs
)
```

Popular options to tweak

- learning_rate: the base learning rate
- momentum
- decay
- nestrov
- (advanced) callback

Optimization: Momentum

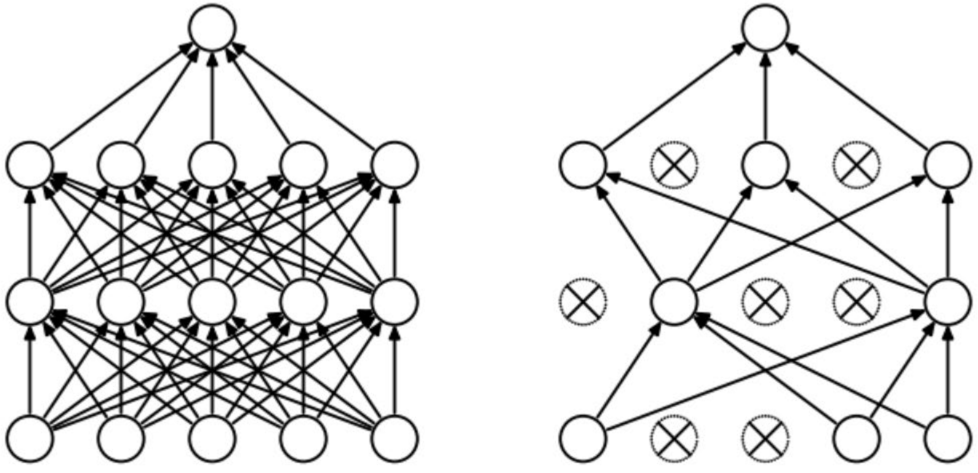
Review



Ways to reduce overfitting

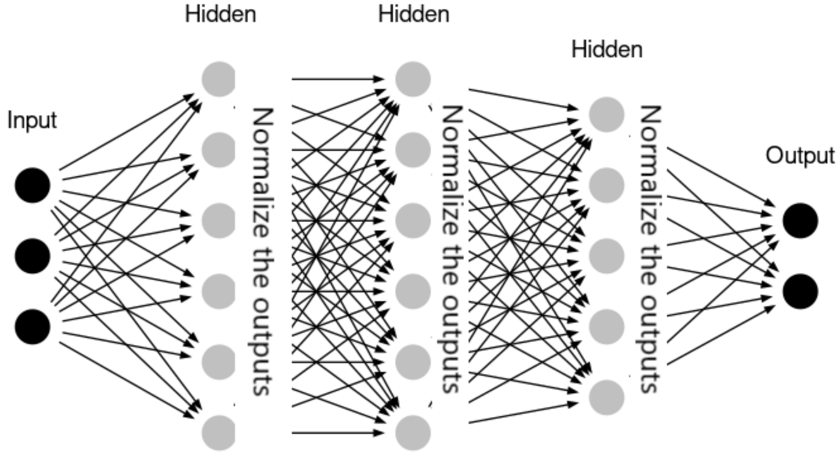
Review

Dropout



(a) Standard Neural Net (b) After applying dropout.

Batch normalization



Practical Tips for CNN

Optimization

- Learning rate (0.01~0.0001)
- Optimization method: Adam or RMSProp

Architecture

- ReLU/PReLU for hidden layers, Sigmoid/Softmax/Tanh/PReLU for the output layer
- 3x3 filters
- $[\text{Conv-Conv-MaxPool}]_n$ structure

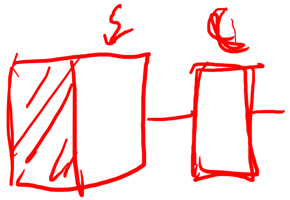
Regularization

- L2 regularization
- Dropouts
- Batch Normalization

Transfer Learning

How long it takes to train on a large dataset?

Multi GPU + Deep CNN + Large Dataset = 2-3 Weeks Training time



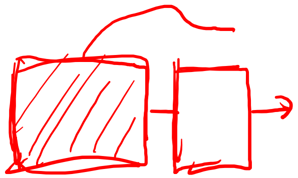
Transfer learning

- As fixed feature extractor: remove the output layer, weights frozen
- Fine-tuning the CNN: also let weights updated
- Use part of layers

Transfer Learning

Review

When should I use a pre-trained network?



- New dataset is small and similar to original dataset (X)
- New dataset is large and similar to the original dataset (O)
- New dataset is small but very different from the original dataset (X)
- New dataset is large and very different from the original dataset (O)

Various models pre-trained on ImageNet

Keras https://keras.io/applications/

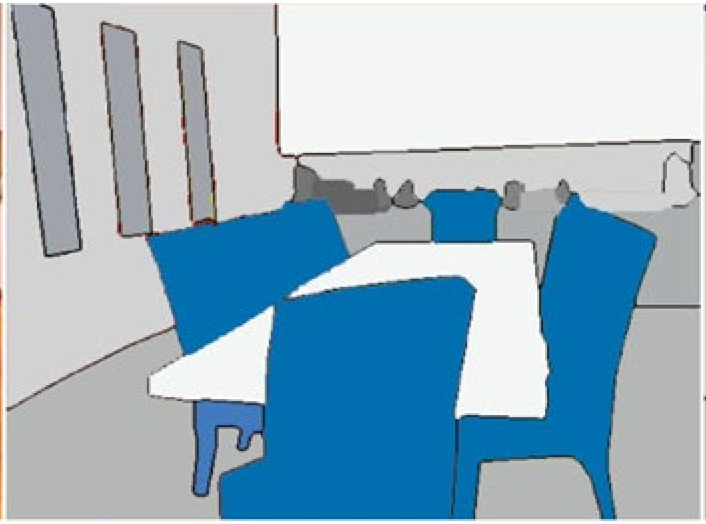
Image Segmentation

Why Segmentation?

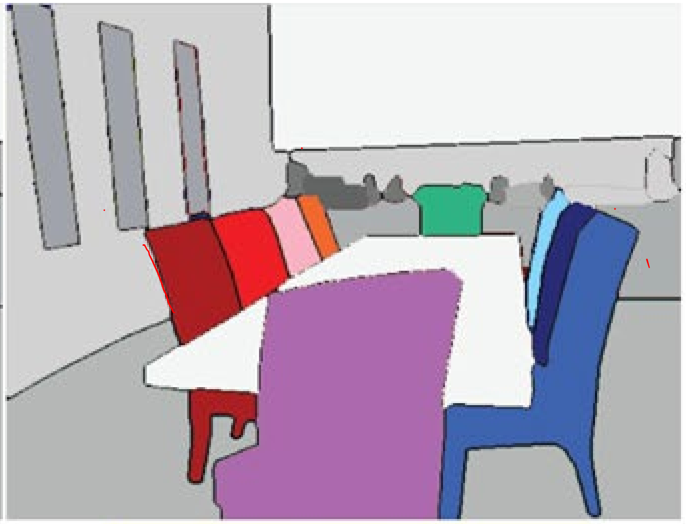
Application in self-driving car

Segmentation Task

Input Image



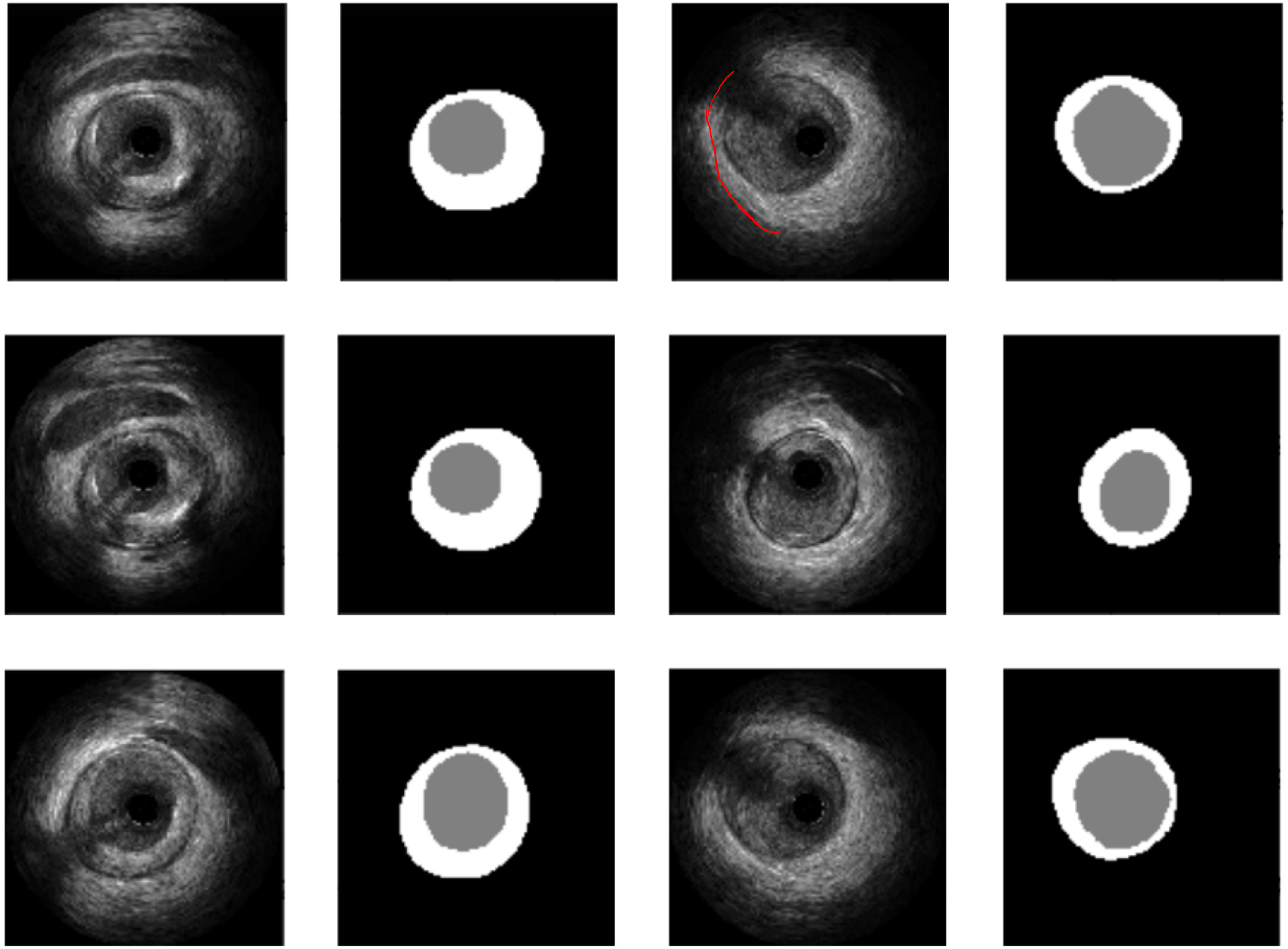
Semantic Segmentation



Instance Segmentation

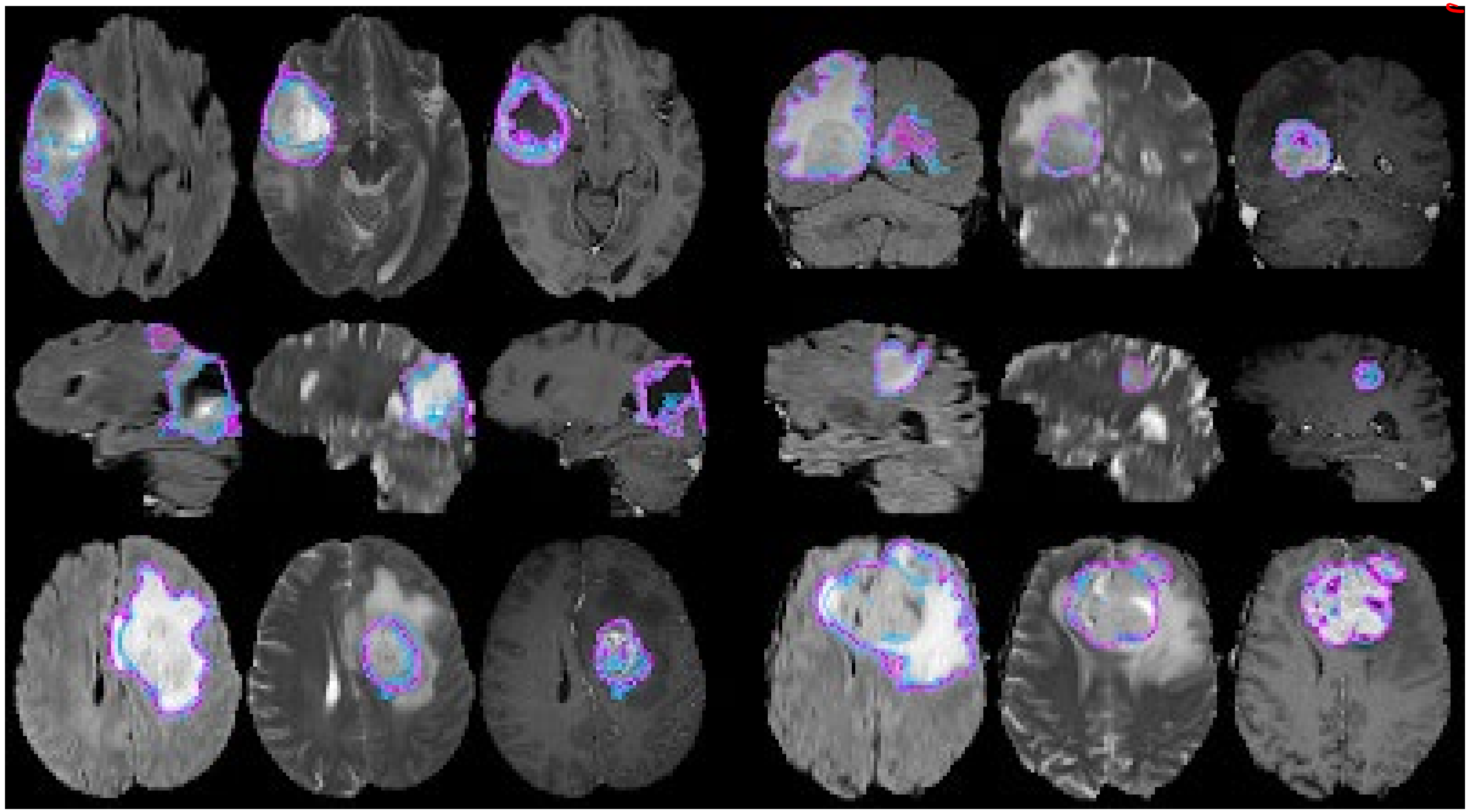
Why Segmentation?

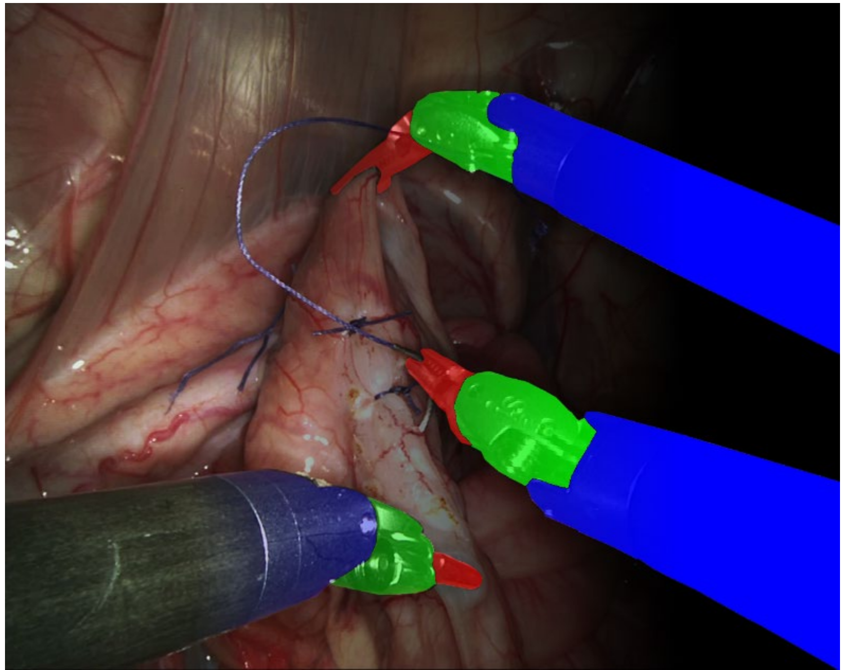
Application in medical images



Why Segmentation?

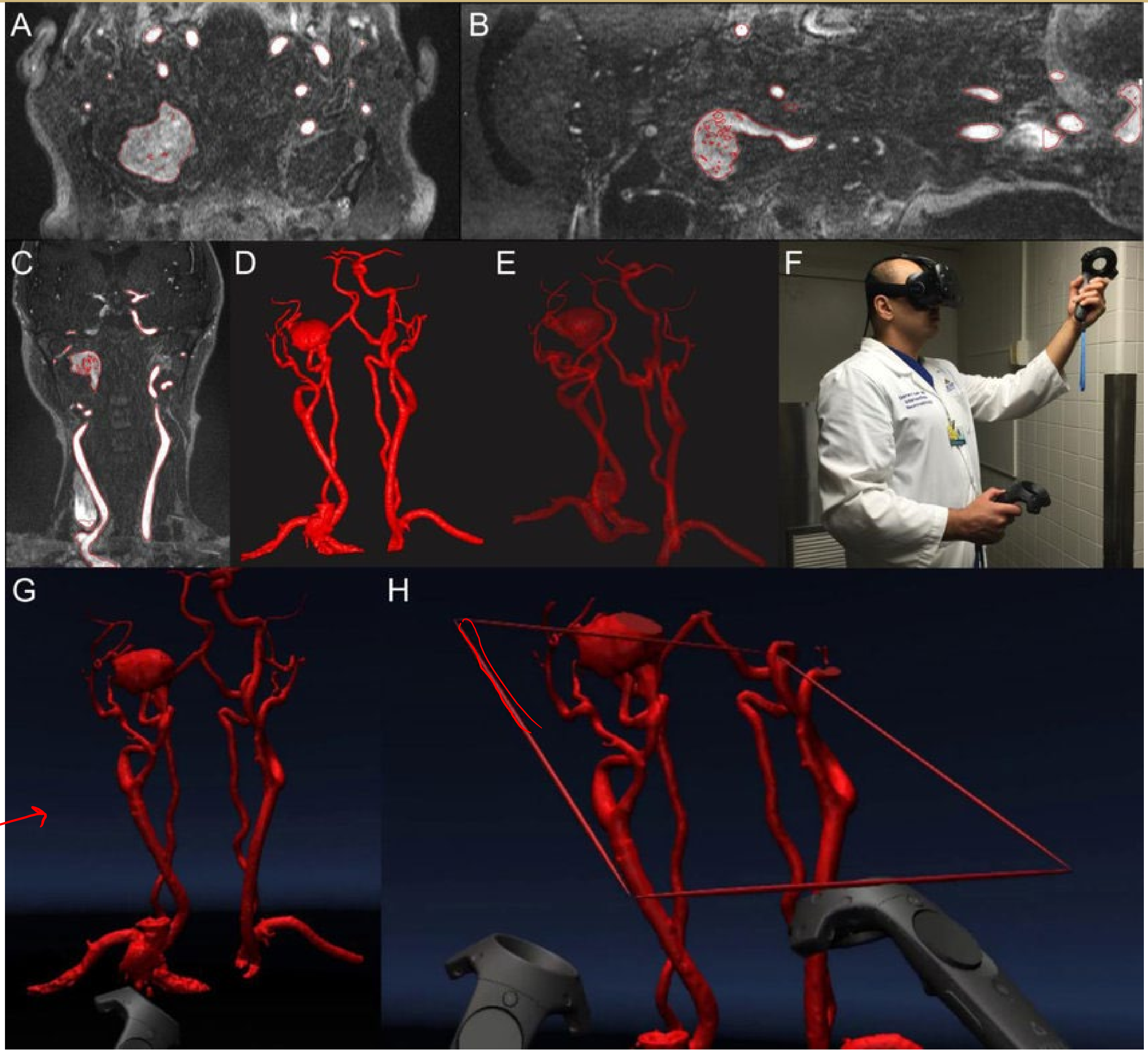
Application in medical images





Why Segmentation?

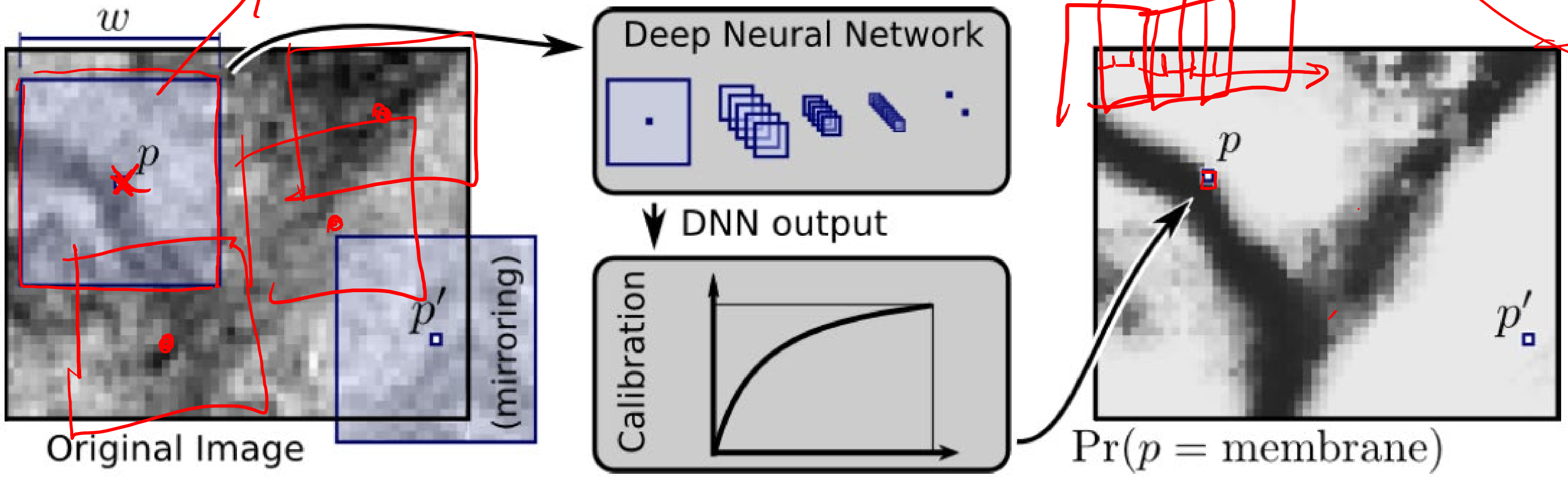
Application in medical images

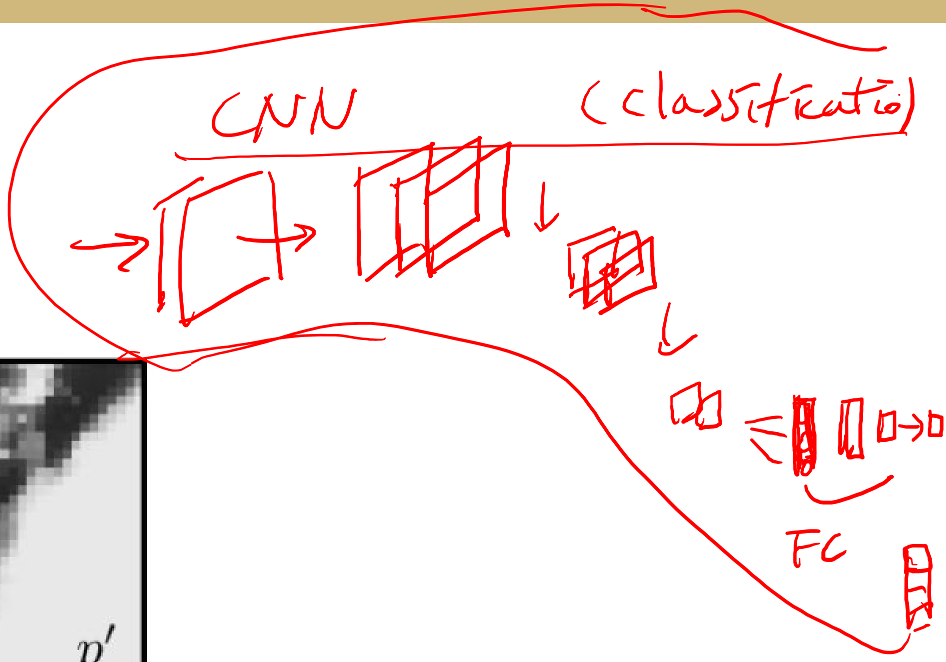


Ong, Chin et al. Virtual Reality in Neurointervention. Journal of vascular and interventional neurology. 10 (2018)

How to build a segmentation model with CNN?

Patch-based method

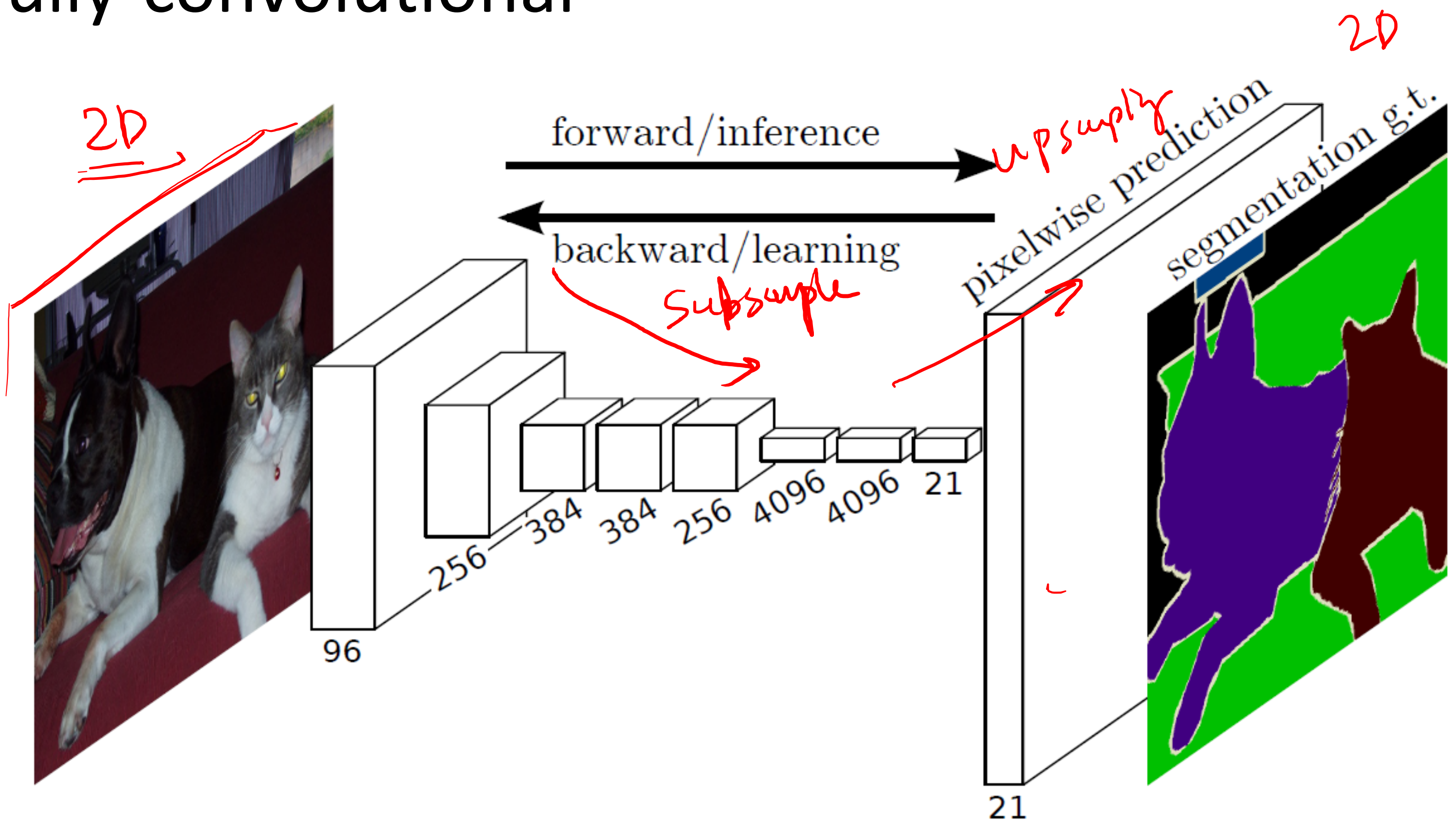




http://papers.nips.cc/paper/4741-deep-neural-networks-segment-neuronal-membranes-in-electron-microscopy-images.pdf

How to build a segmentation model with CNN?

Fully-convolutional

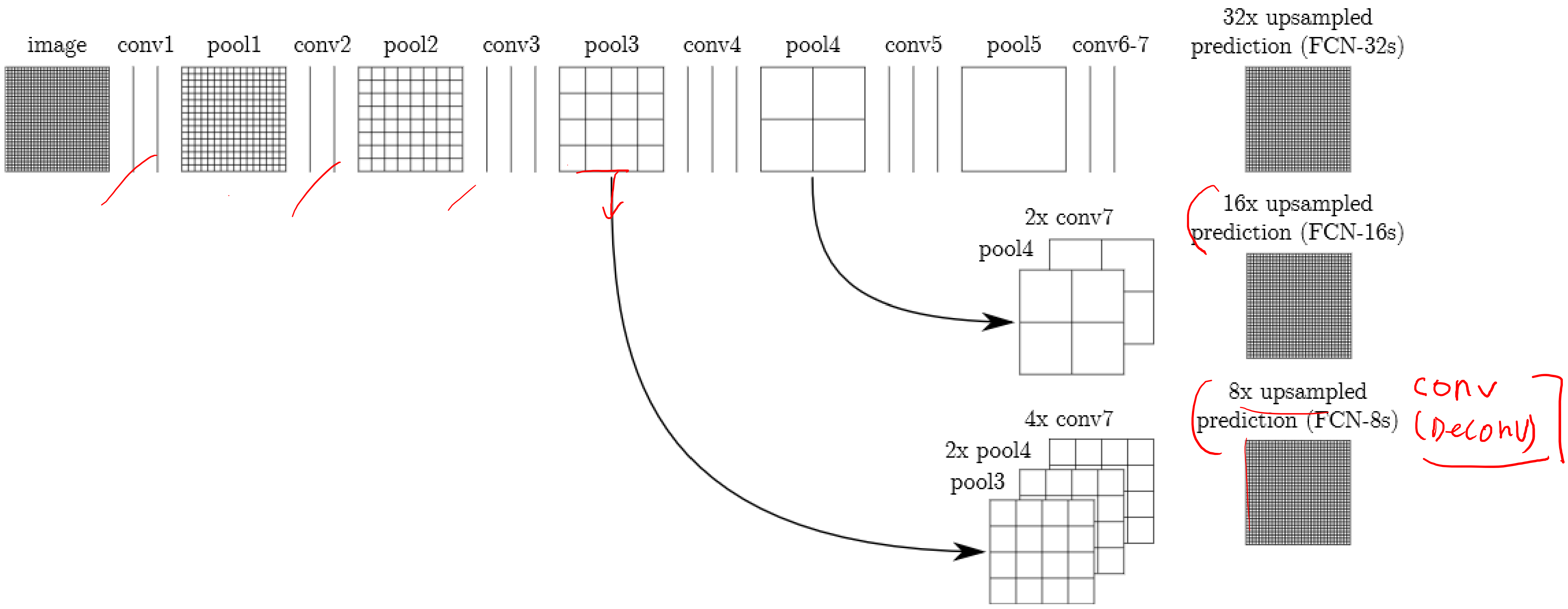


https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Long_Fully_Convolutional_Networks_2015_CVPR_paper.pdf

How to build a segmentation model with CNN?

Fully-convolutional

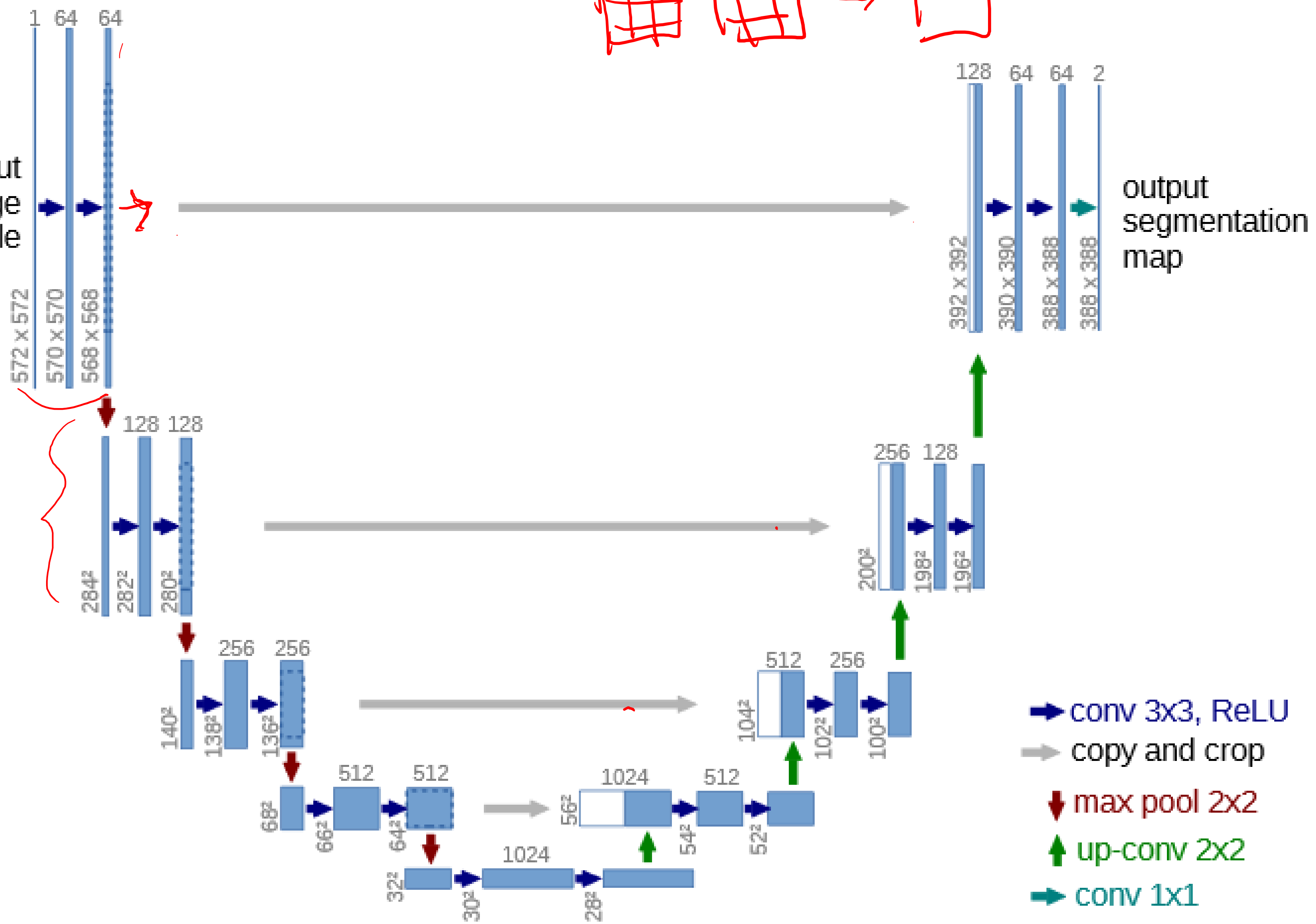
FCNN

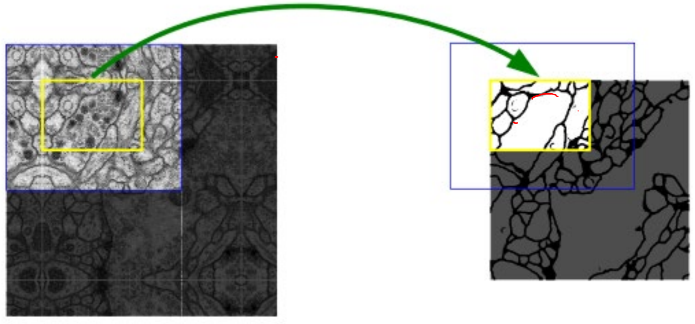


https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Long_Fully_Convolutional_Networks_2015_CVPR_paper.pdf

How to build a segmentation model with CNN?

U-Net

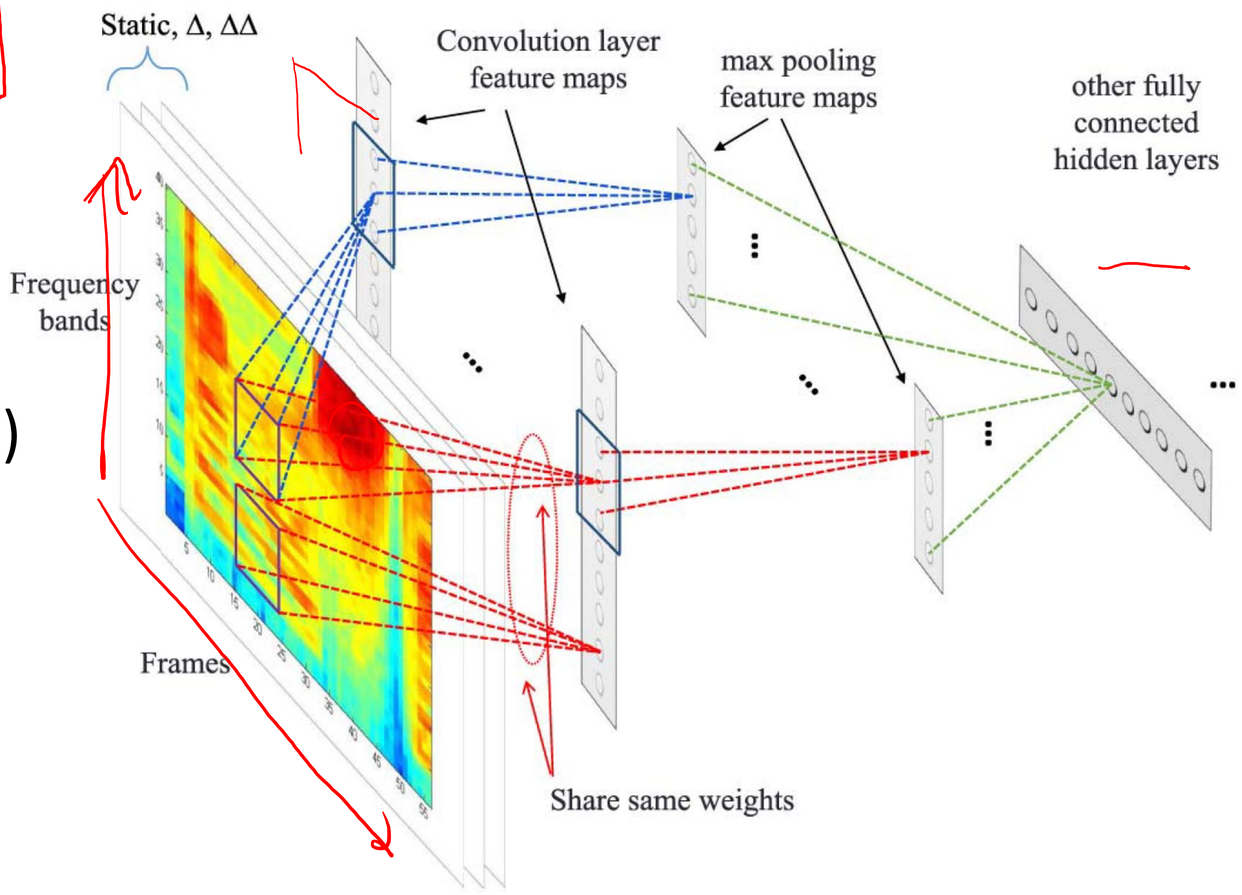




https://arxiv.org/pdf/1505.04597.pdf

Other data type- Sound

- Spectrogram(FFT)
- Mel-Frequency Cepstral Coefficient (MFCC)
- MFCC-derivatives
- Mel-Frequency Cepstral Spectrum
- Log-engergy



https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/CNN_ASLPTrans2-14.pdf

Other data type- Text

Word Embeddings:

Frequency-based

Count Vector

TF-IDF Vector

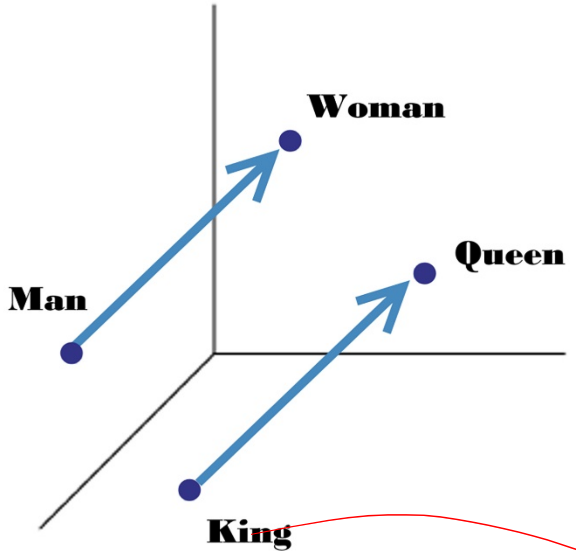
Co-Occurrence Vector

Prediction-based

Neural probabilistic model

Word2Vec

GloVe



Word embedding is a technique that treats words as vectors whose relative similarities correlate with semantic similarity.

TF-IDF(

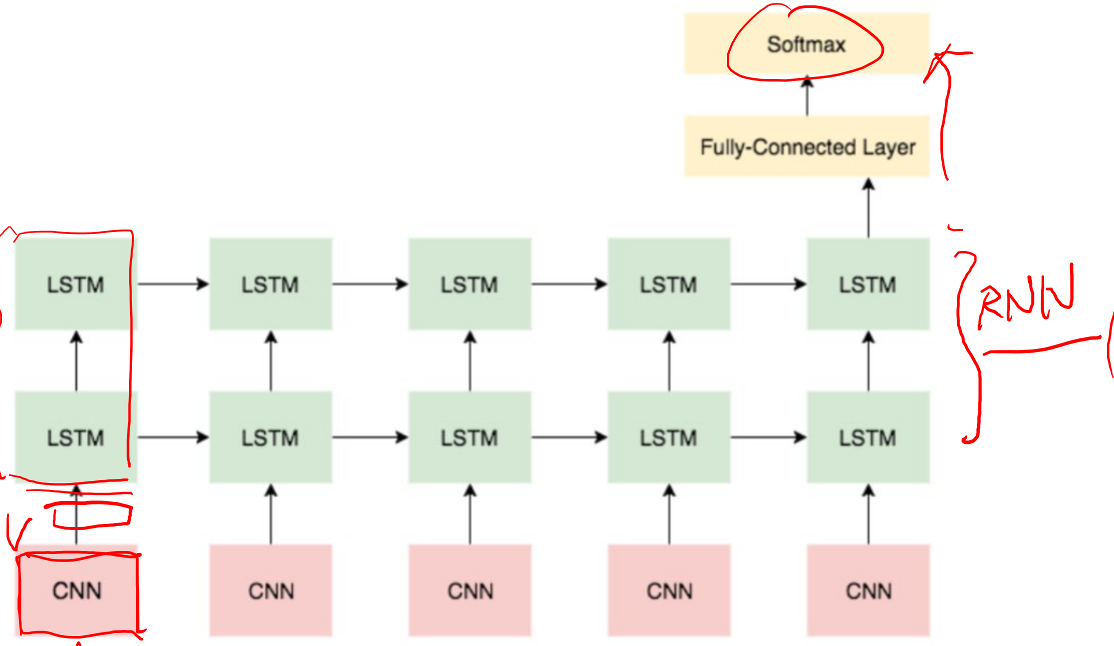
$$W : \text{words} \to \mathbb{R}^n$$

$$W(\text{“cat”}) = (0.2,\ \text{-}0.4,\ 0.7,\ \ldots)$$

$$W(\text{“mat”}) = (0.0,\ 0.6,\ \text{-}0.1,\ \ldots)$$

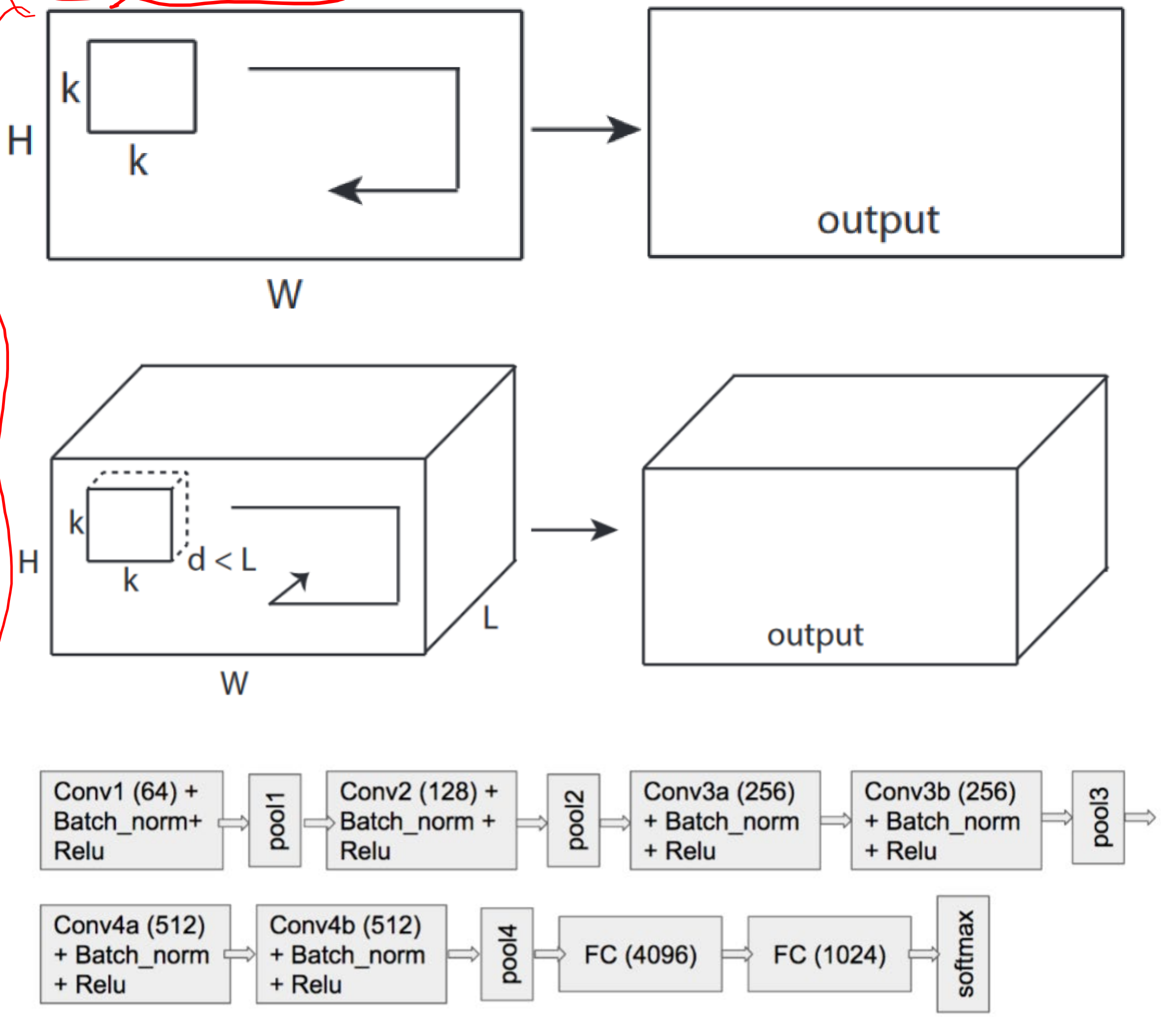
Other data type- Video

LSTM+CNN



https://arxiv.org/pdf/1503.08909.pdf

3D CNN



https://arxiv.org/pdf/1412.0767.pdf