# CNN Segmentation

Visual Computing and Medical Imaging Lab.,
Department of Software Convergence,
Seoul Women's University
Min Jin Lee

Saturday, August 14, 2021

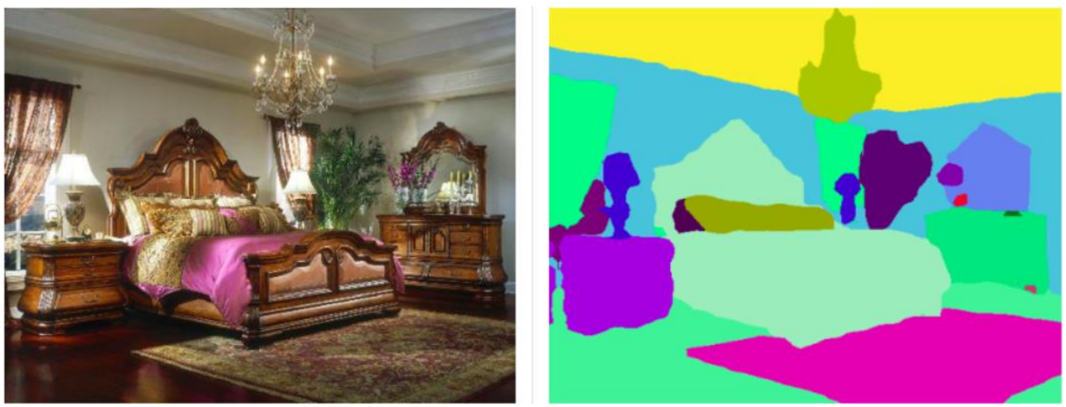
# Outline

- Segmentation Overview
- CNN Segmentation
  - Fully Convolutional Network (FCN)
  - DeconvNet
  - UNet
  - Vnet
  - Attention UNet
  - nnUNet
- Others
  - Loss function
  - Evaluation metrics
  - Public datasets

# Segmentation Overview

# Semantic Segmentation

- Partitioning image into meaningful instances




Input

Semantic Labels

# Semantic vs. Instance segmentation

원본 이미지

Semantic Segmentation

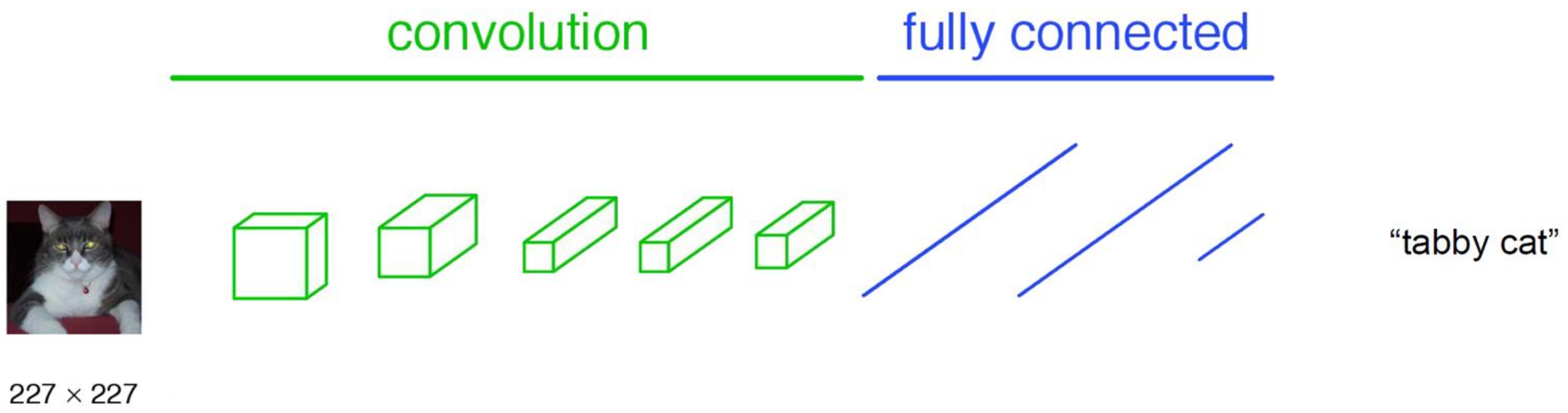
Instance Segmentation

# CNN Segmentation Models

[Long et al. CVPR 2015](#)

# Fully Convolutional Network (FCN)

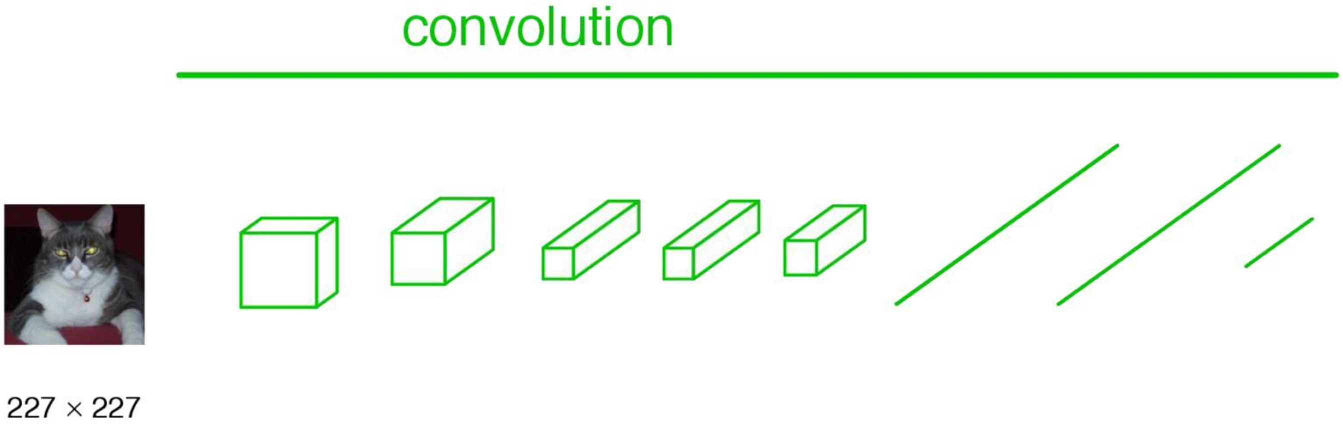
- From image classification to semantic segmentation




Slide credit: Jonathan Long

# Fully Convolutional Network (FCN)

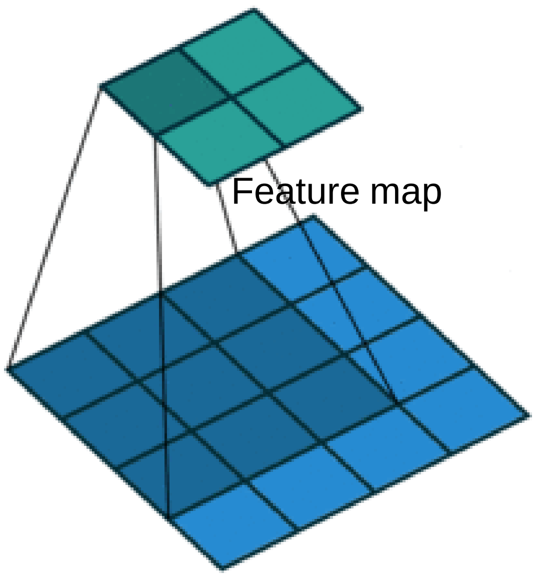
- Fully connected (FC) into fully convolutional layer




Slide credit: Jonathan Long

# Fully Convolutional Network (FCN)

- Convolution

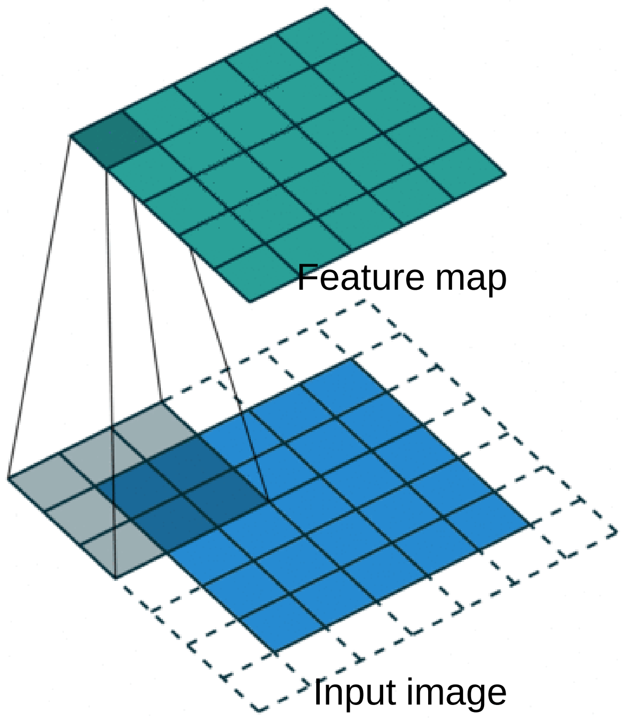



  - **Feature map**: by applying Filters to the input image
  - **Kernel(Filter)** : the field of view of the convolution — that is 3x3 pixels
  - **Padding**: how the border of a sample is handled
  - **Stride**: the step size of the kernel when traversing the image — usually 1, use a stride of 2 for downsampling

https://github.com/vdumoulin/conv_arithmetic/blob/master/README.md

# Fully Convolutional Network (FCN)

- Convolution

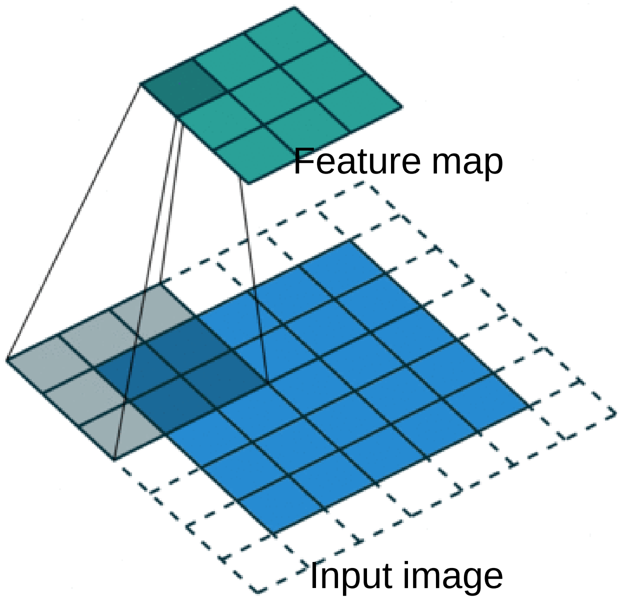



- **Feature map**: by applying Filters to the input image
- **Kernel(Filter)** : the field of view of the convolution — that is 3x3 pixels
- **Padding**: how the border of a sample is handled
- **Stride**: the step size of the kernel when traversing the image — usually 1, use a stride of 2 for downsampling

https://github.com/vdumoulin/conv_arithmetic/blob/master/README.md

# Fully Convolutional Network (FCN)

- Convolution

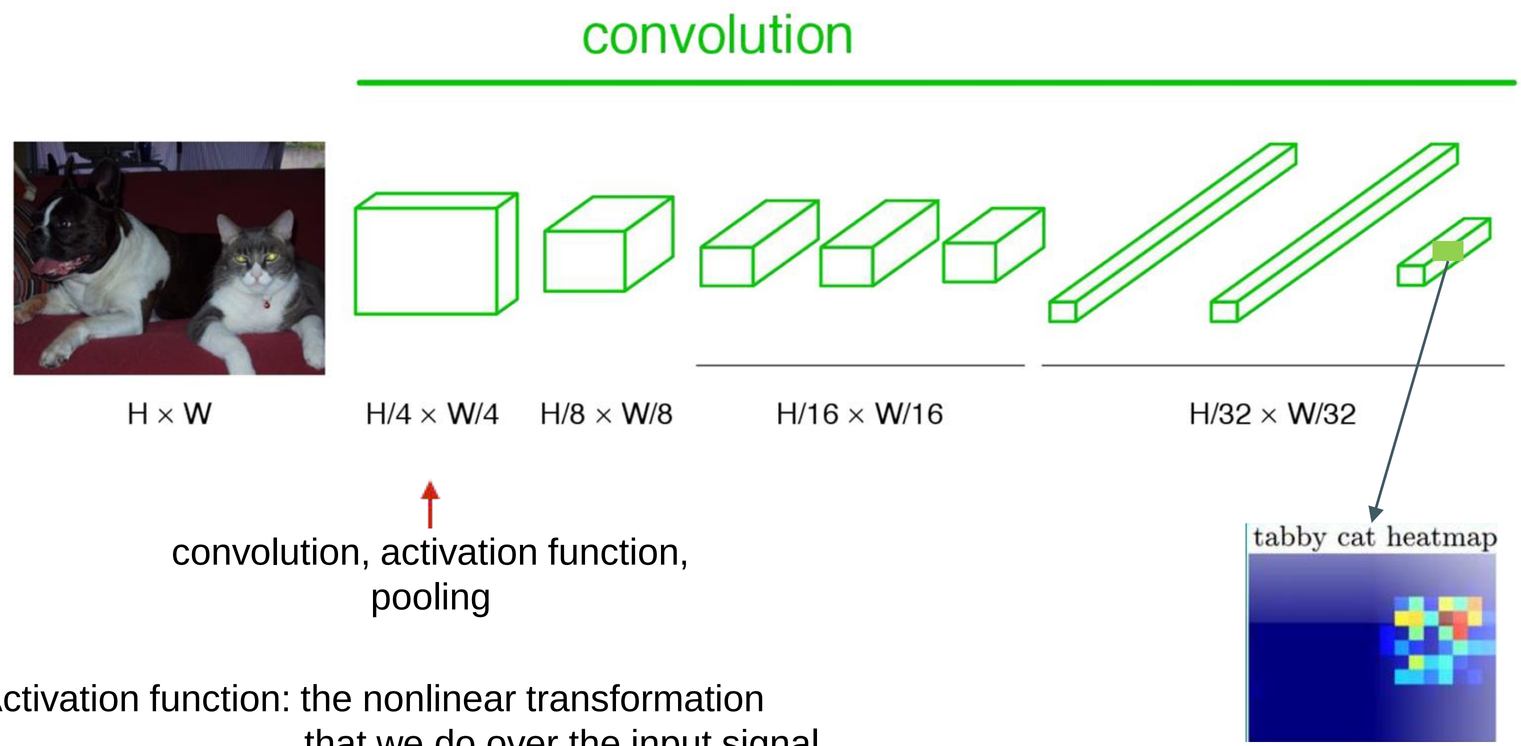



  - **Feature map**: by applying Filters to the input image
  - **Kernel(Filter)** : the field of view of the convolution — that is 3x3 pixels
  - **Padding**: how the border of a sample is handled
  - **Stride**: the step size of the kernel when traversing the image — usually 1, use a stride of 2 for downsampling

https://github.com/vdumoulin/conv_arithmetic/blob/master/README.md

# Fully Convolutional Network (FCN)

- Fully connected (FC) into fully convolutional la



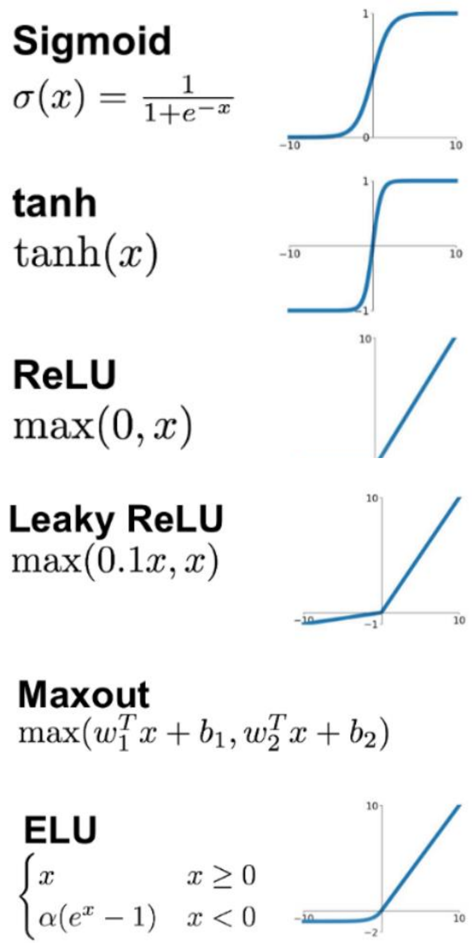

Activation function: the nonlinear transformation that we do over the input signal

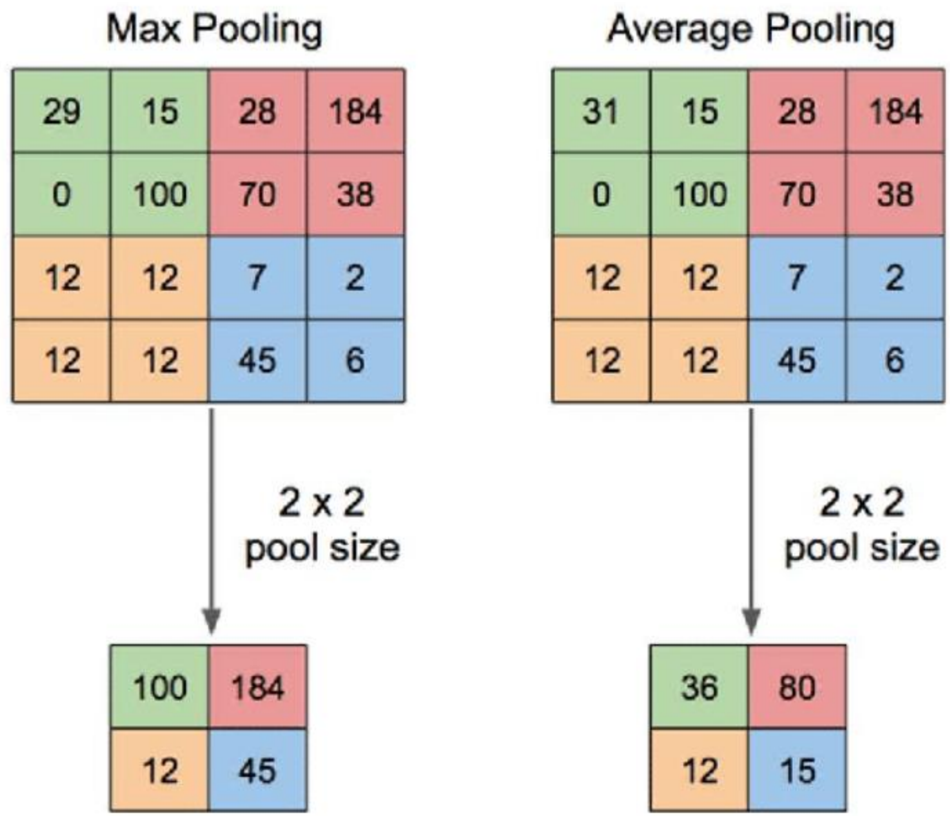
Pooling: reduce the dimensions of the feature maps

Slide credit: Jonathan Long

# Fully Convolutional Network (FCN)

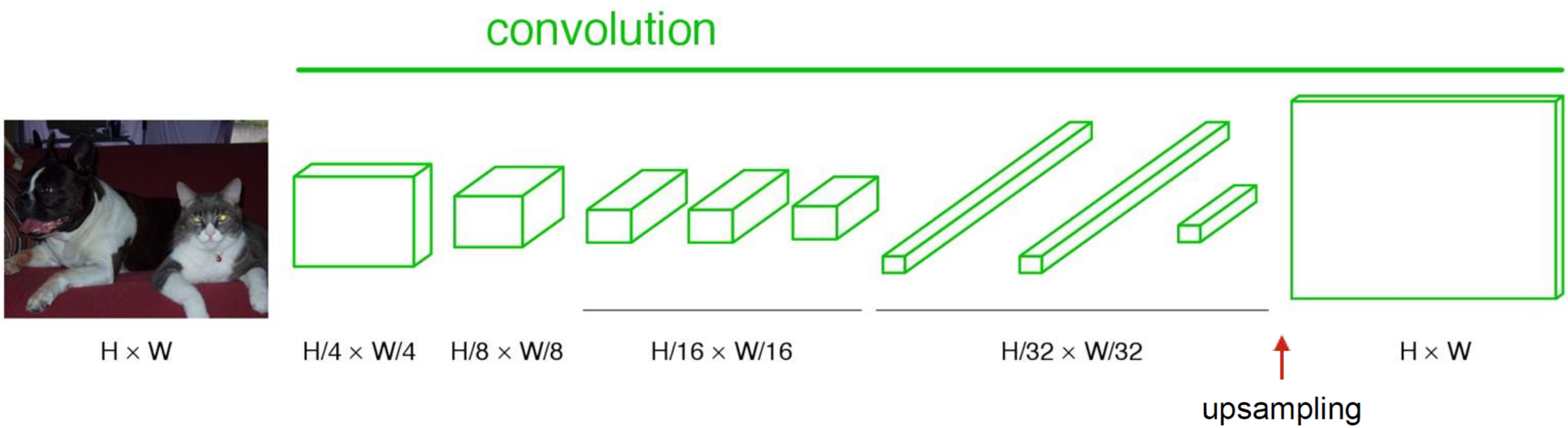
- Activation function and pooling layer



# Fully Convolutional Network (FCN)

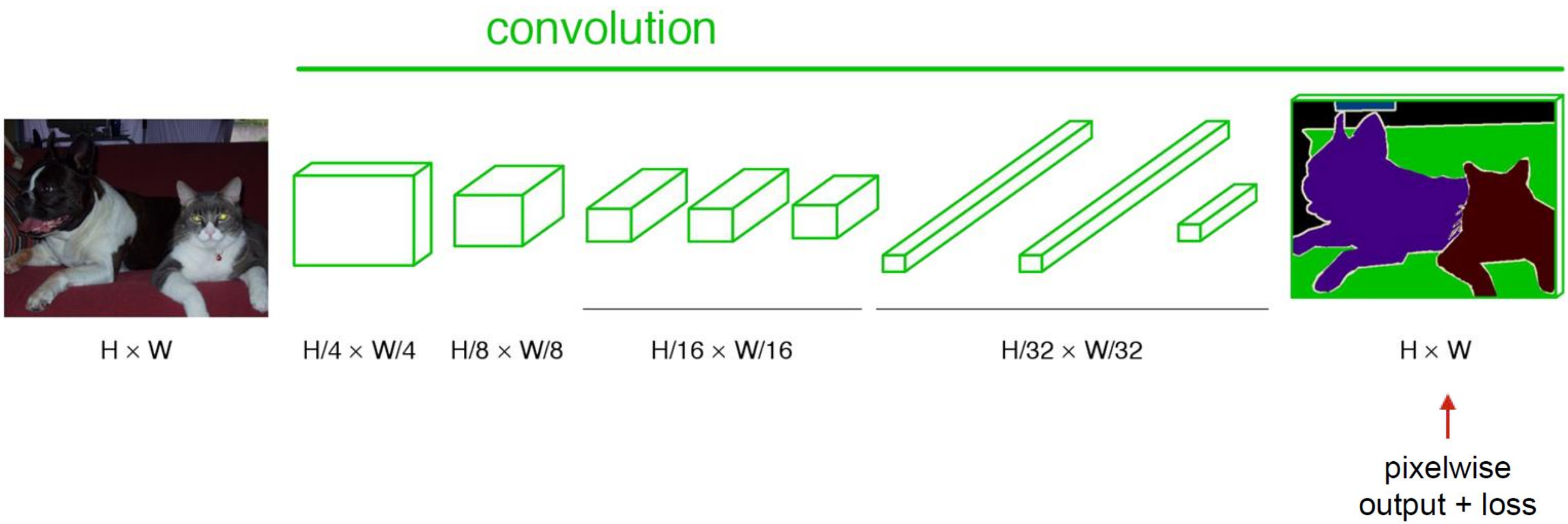
- Pixelwise prediction




Slide credit: Jonathan Long

# Fully Convolutional Network (FCN)

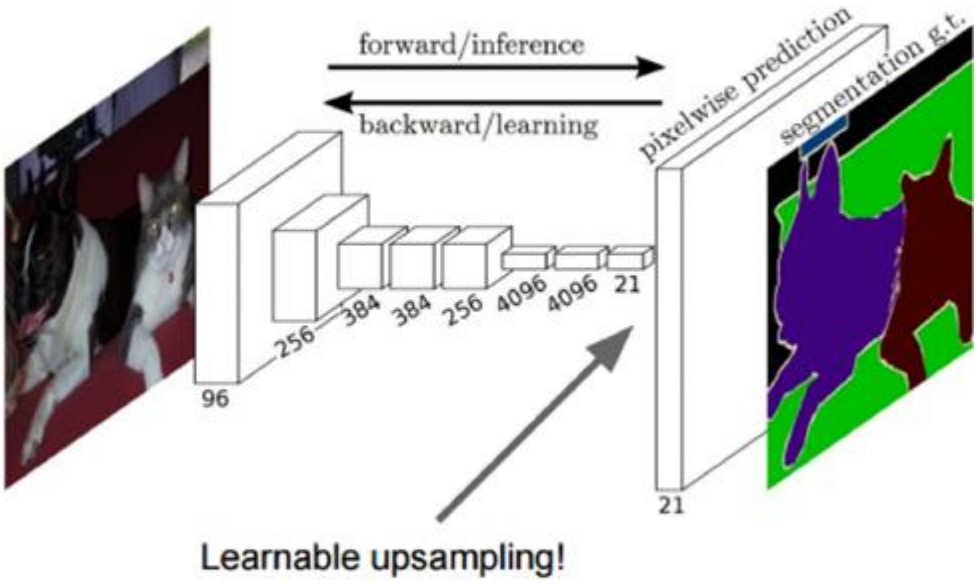
- Pixels-to-pixels output

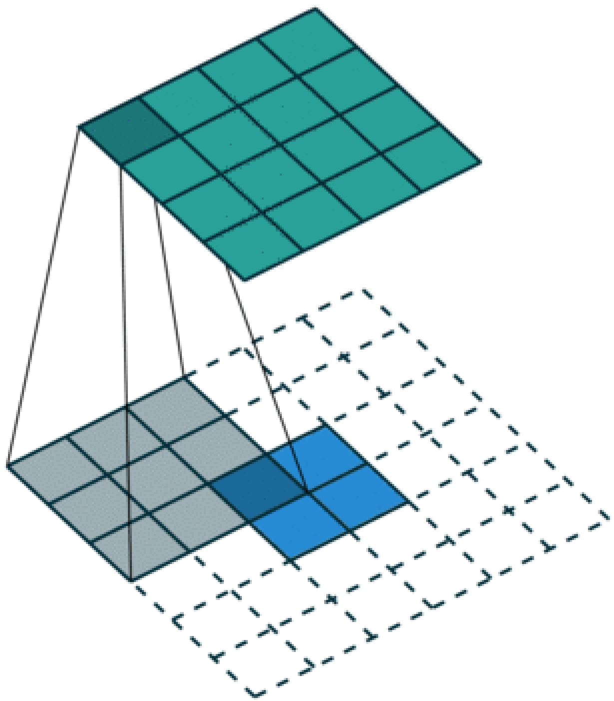



Slide credit: Jonathan Long

# Fully Convolutional Network (FCN)

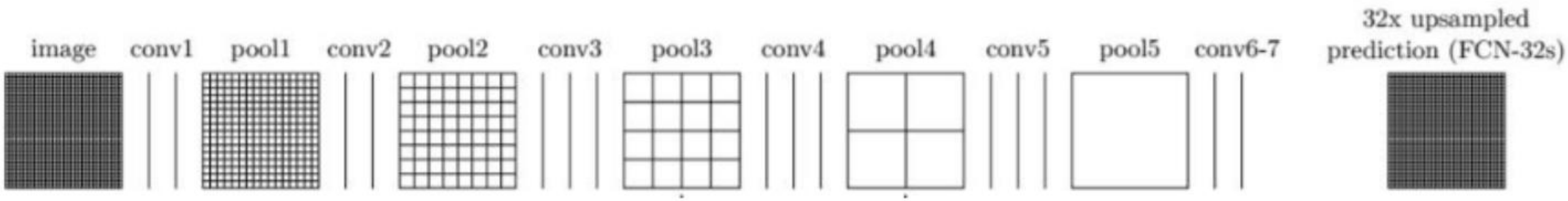
- Upsampling via transposed convolution



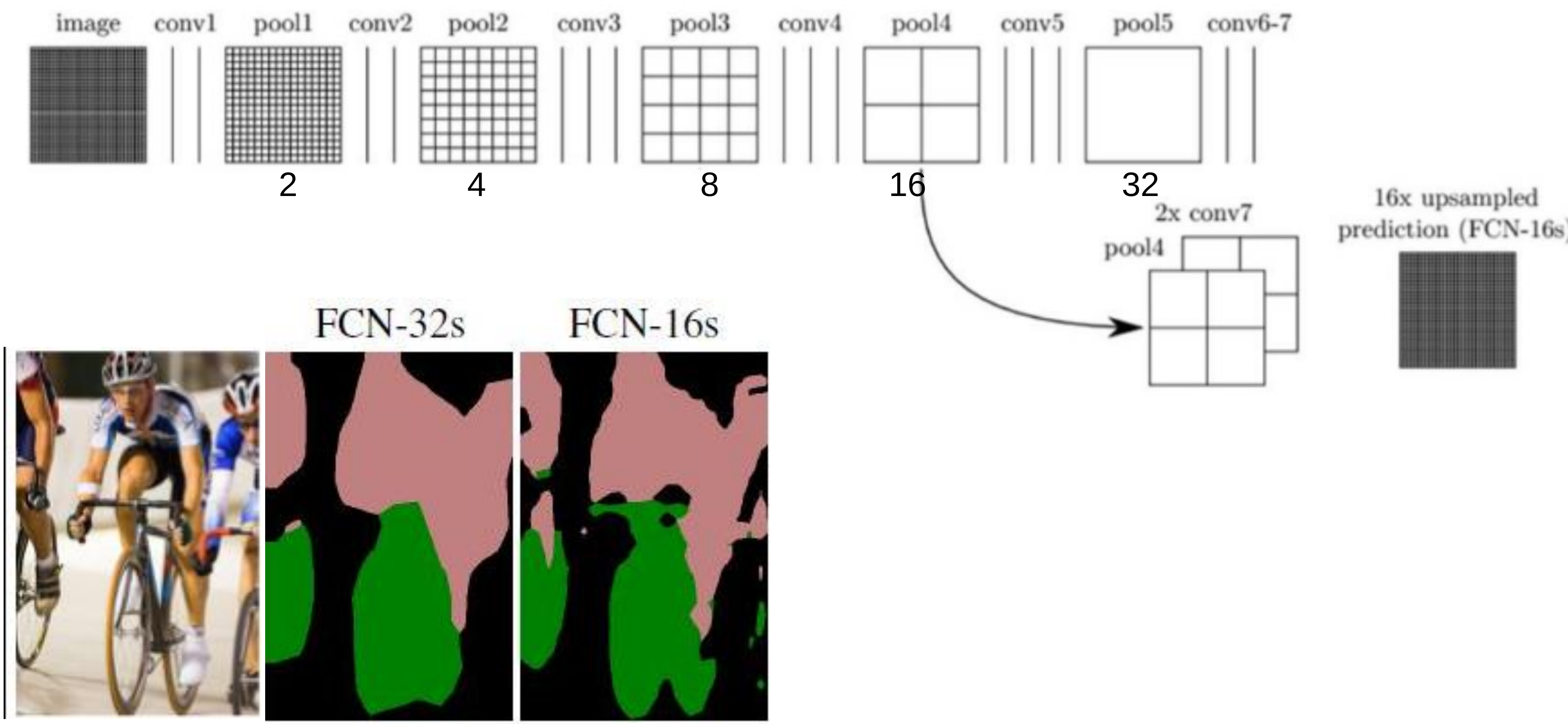
# Fully Convolutional Network (FCN)

- Fusing the output



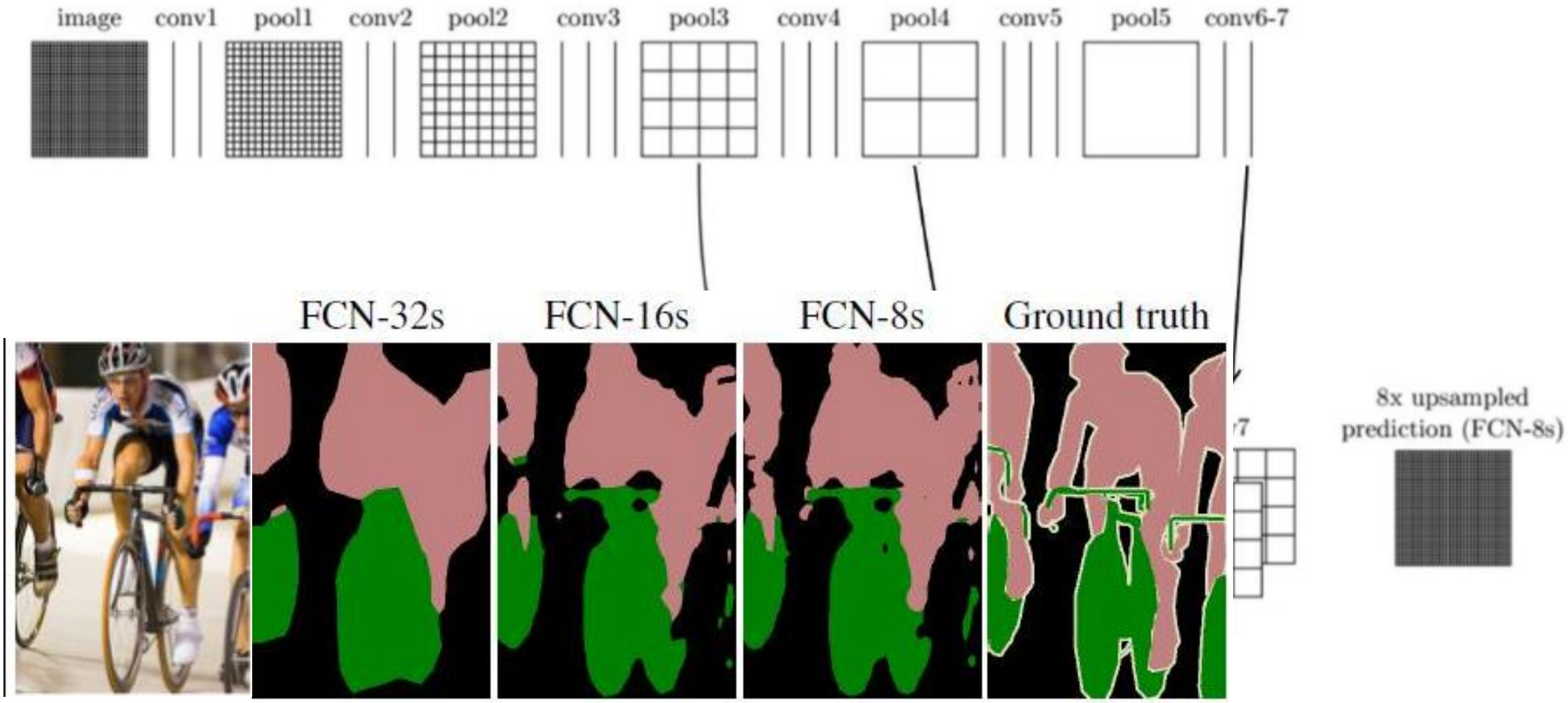

# Fully Convolutional Network (FCN)

- Fusing the output



# Fully Convolutional Network (FCN)

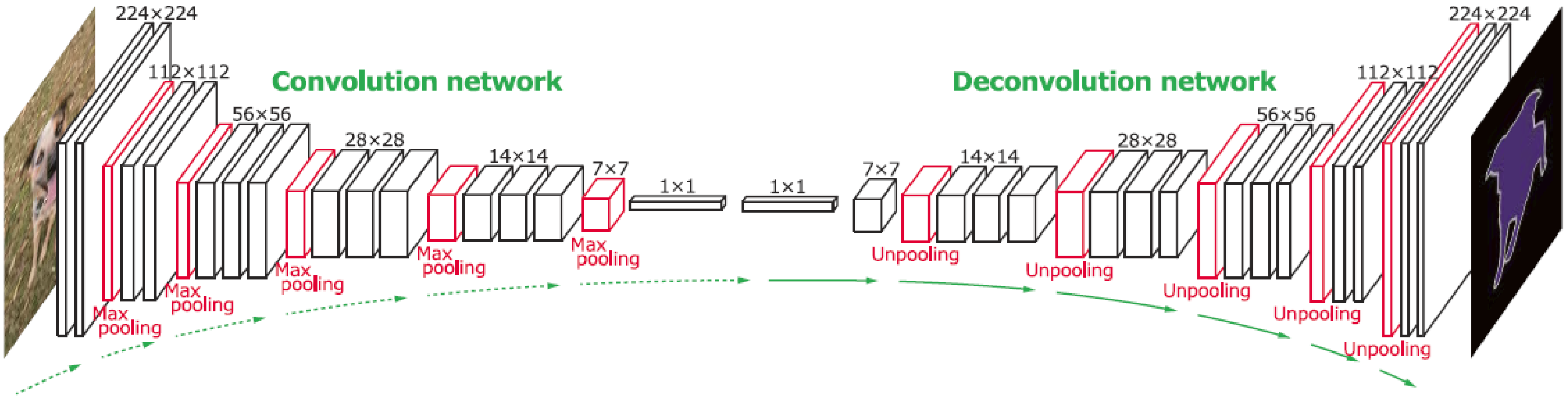
- Fusing the output



Noh et al. ICCV 2015

# DeconvNet

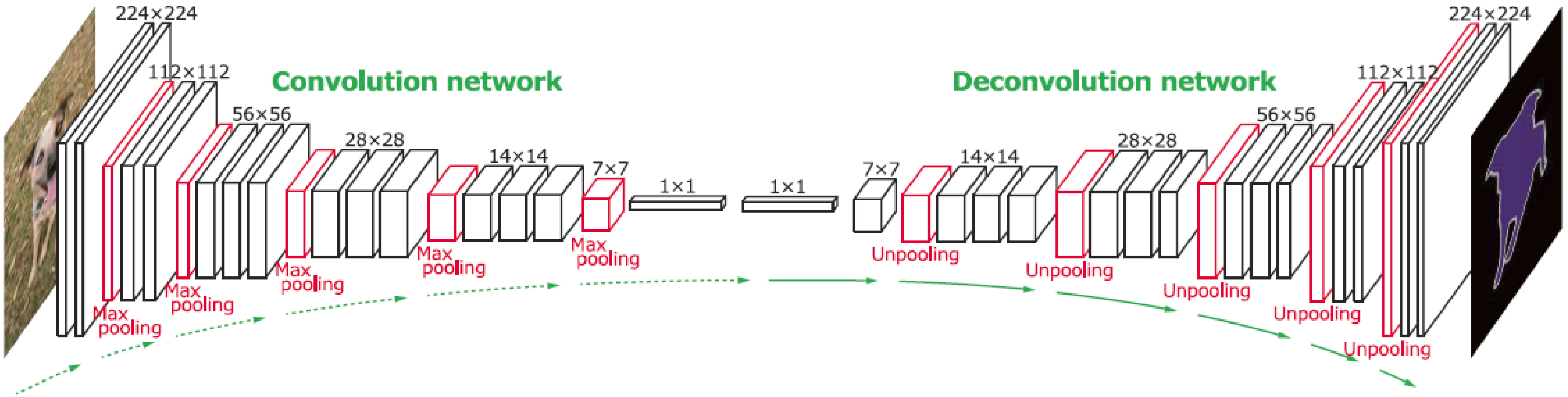
- Convolution + deconvolution network

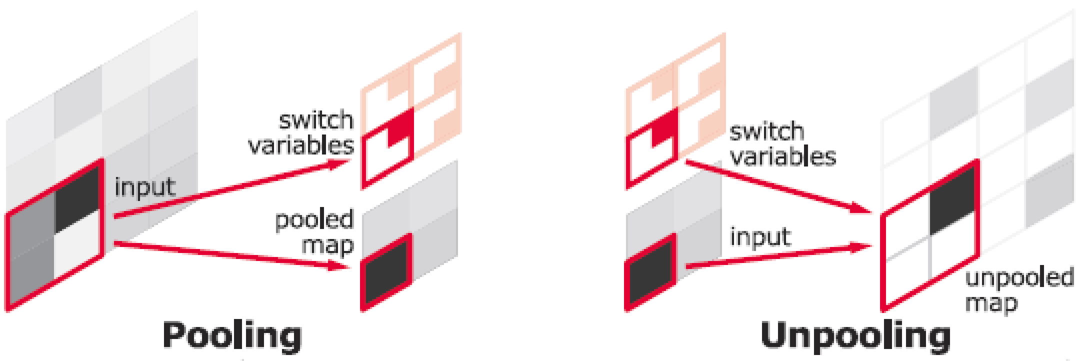


Noh et al. ICCV 2015

# DeconvNet

- Unpooling

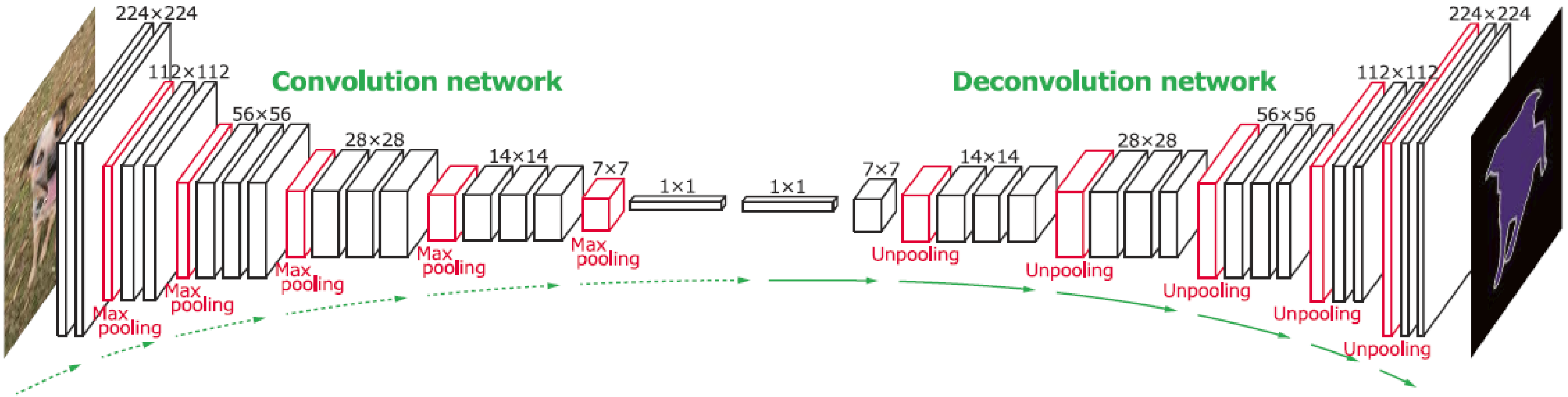


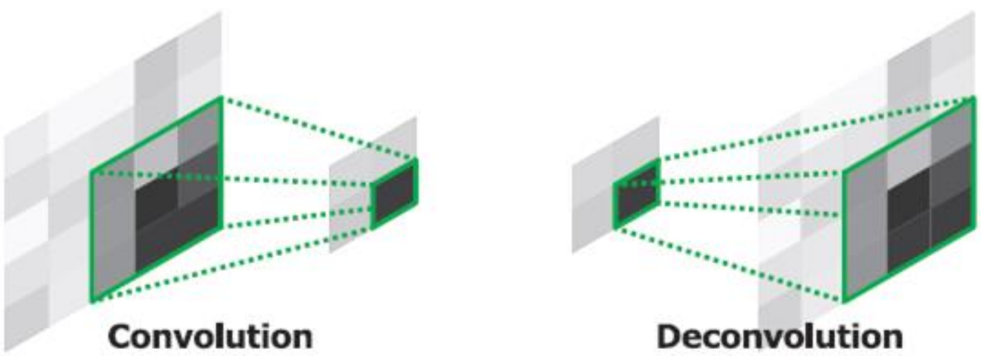



[Noh et al. ICCV 2015]

# DeconvNet

- Deconvolution

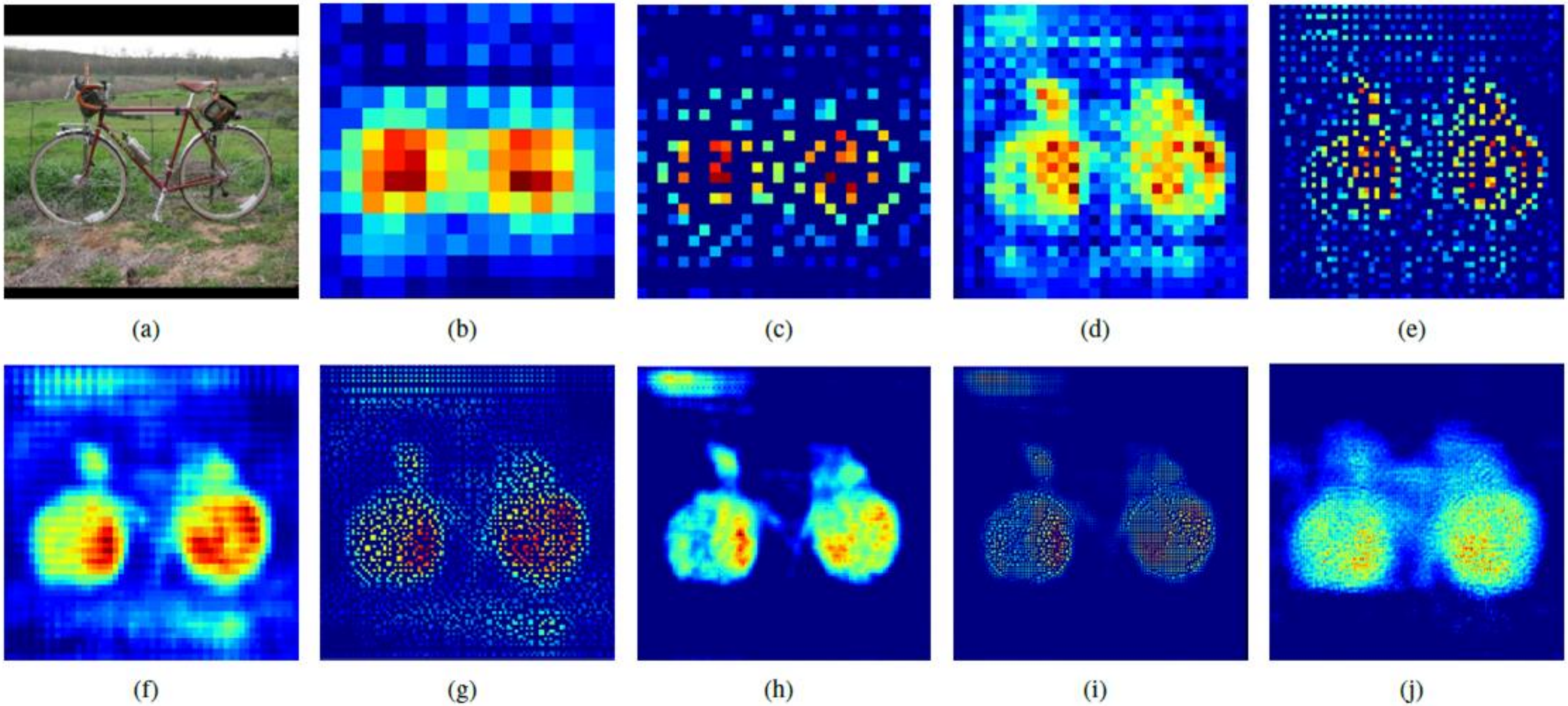


Noh et al. ICCV 2015

# DeconvNet

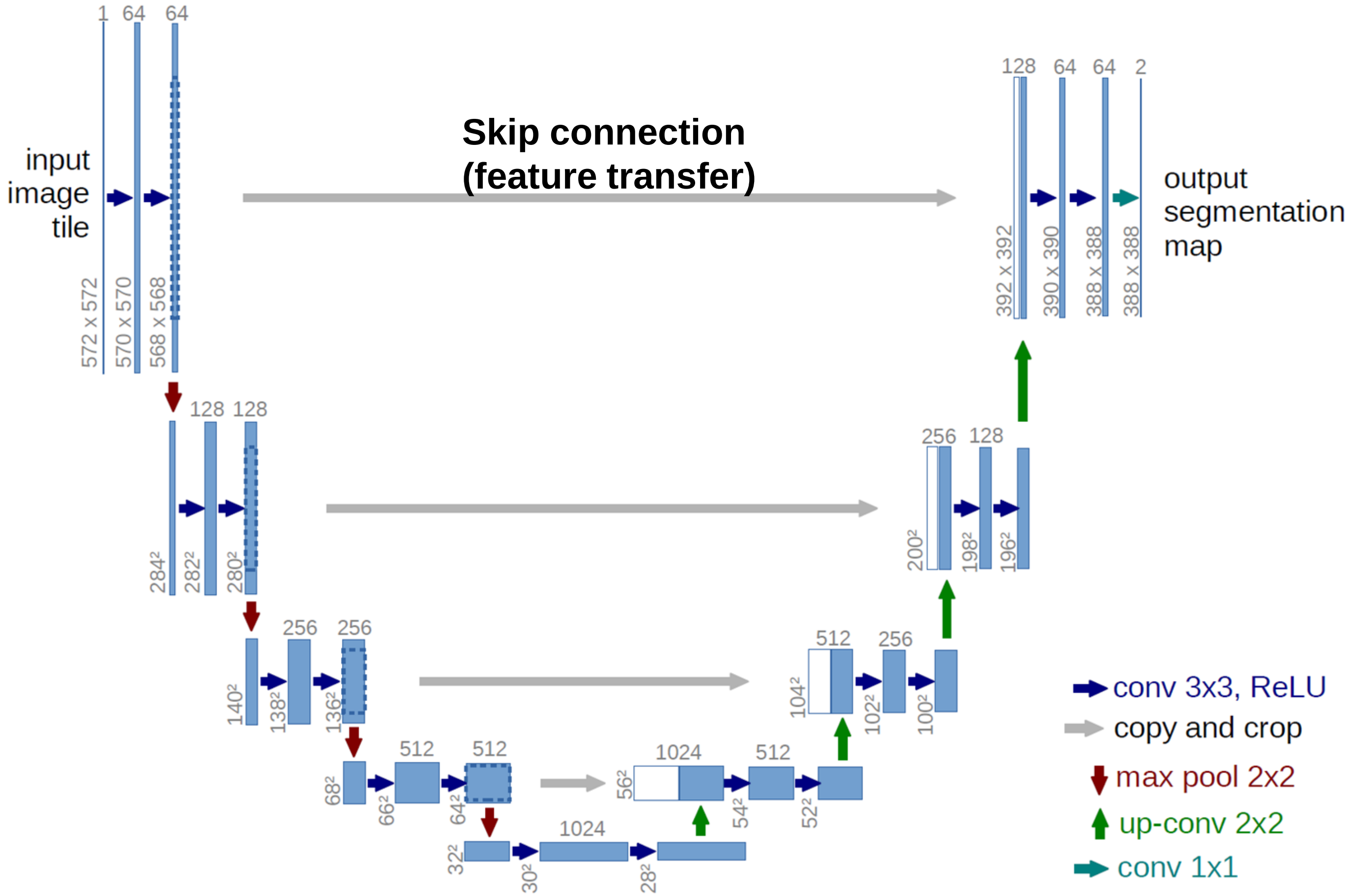
- Deconvolution and unpooling



Ronneberger et al. MICCAI 2015

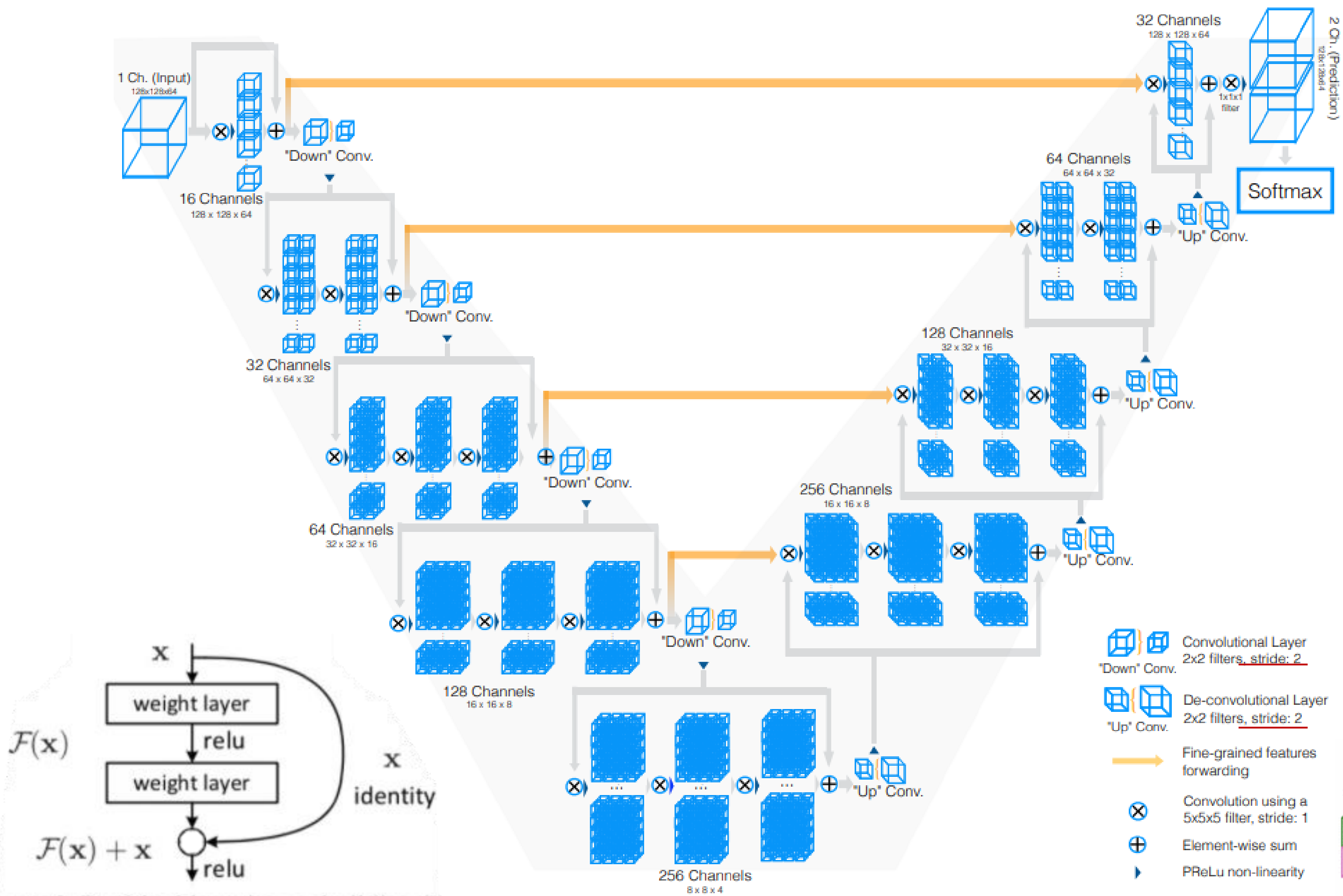
# UNet

- Encoder-Decoder + skip connection



# VNet

- 3D UNet + residual learning - pooling operation



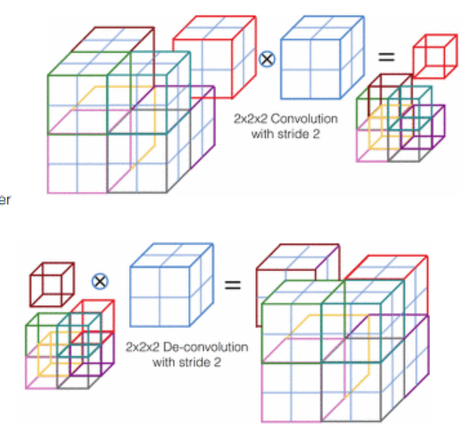



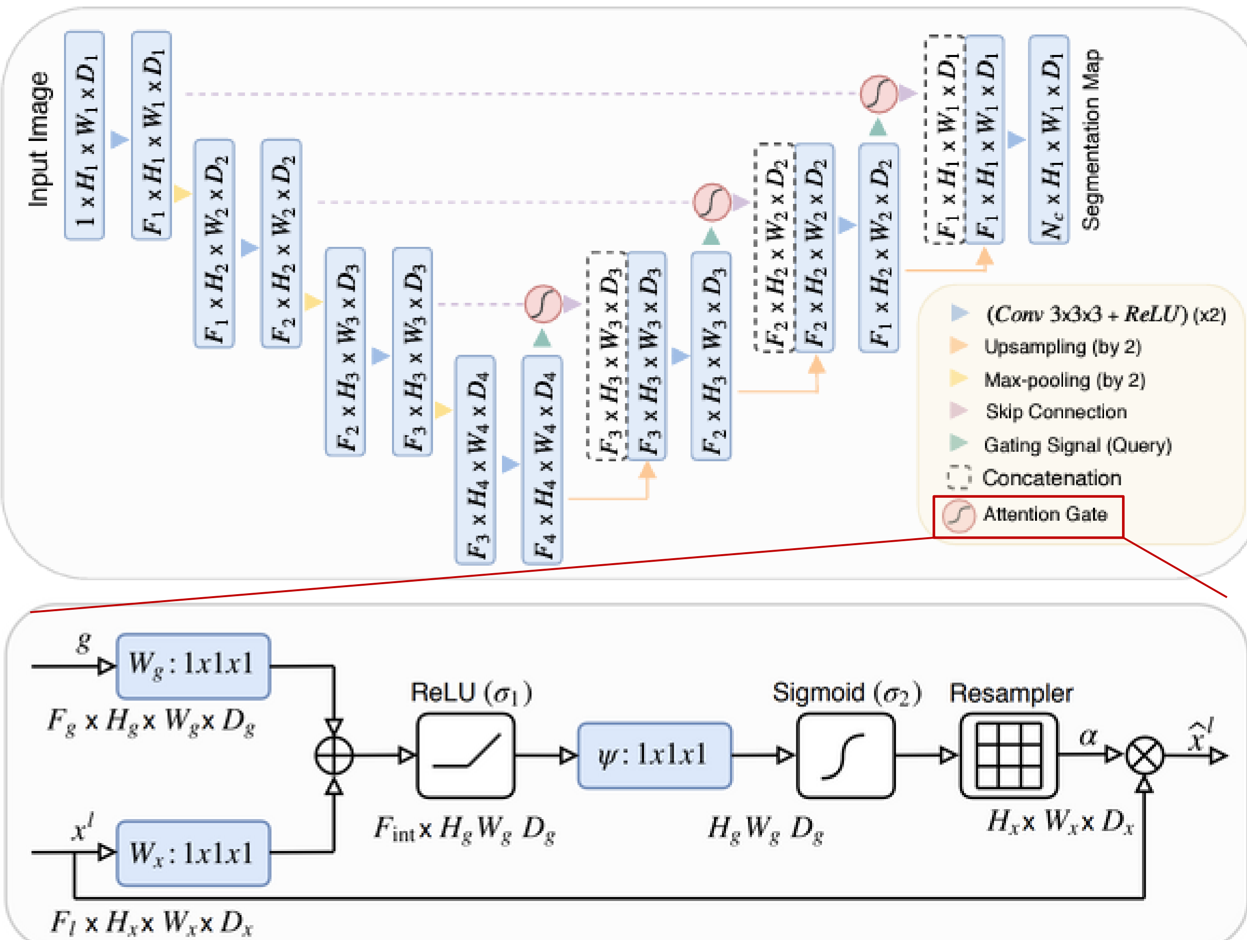



Figure 2. Residual learning: a building bl

Oktay et al. MIDL 2018

# Attention UNet

- 3D UNet + attention



# Attention UNet

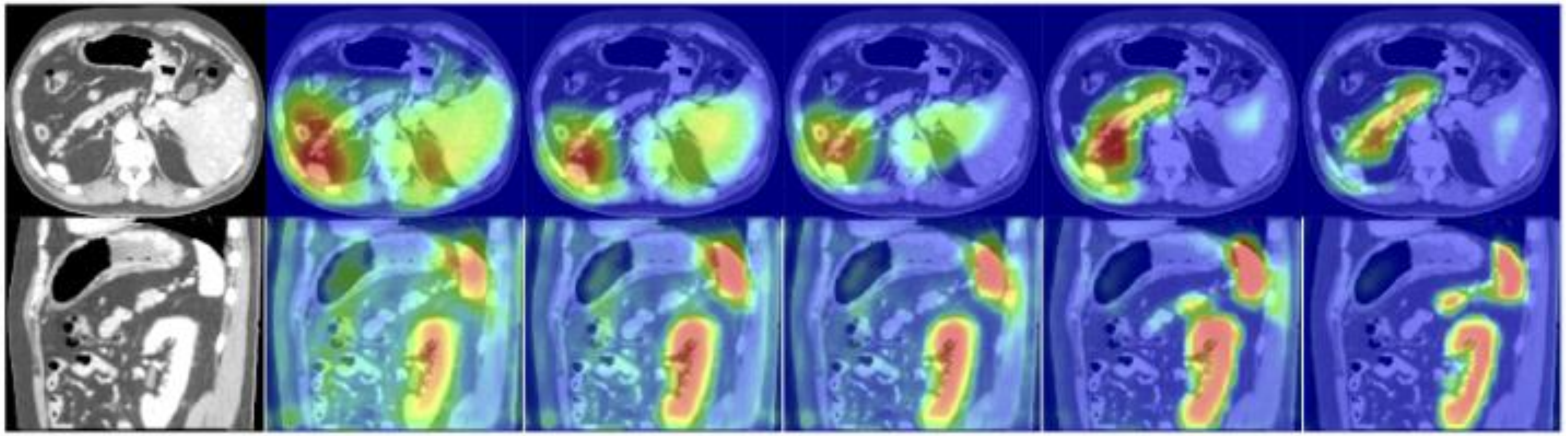
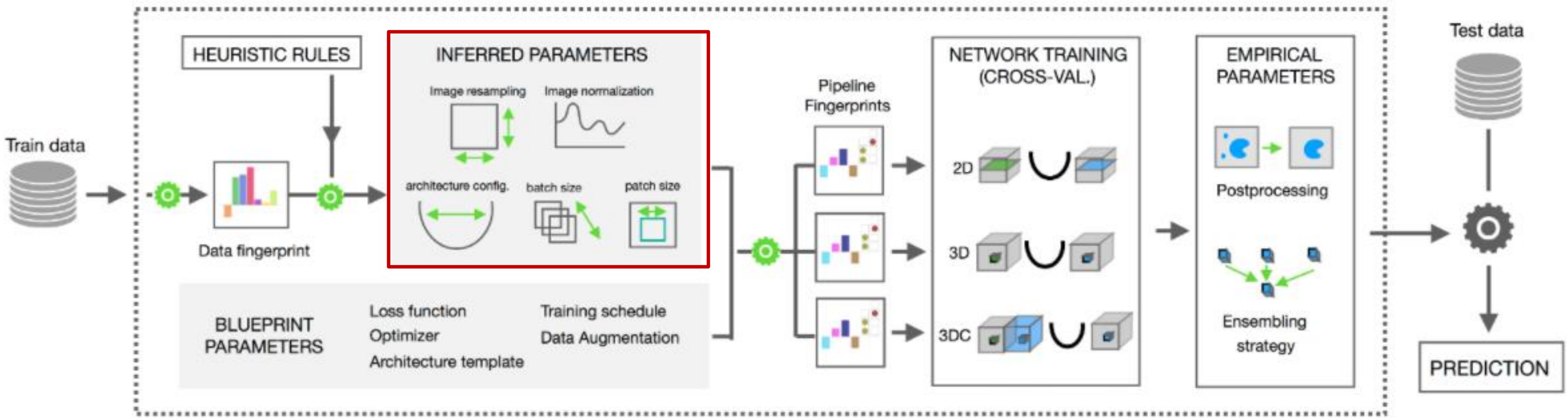


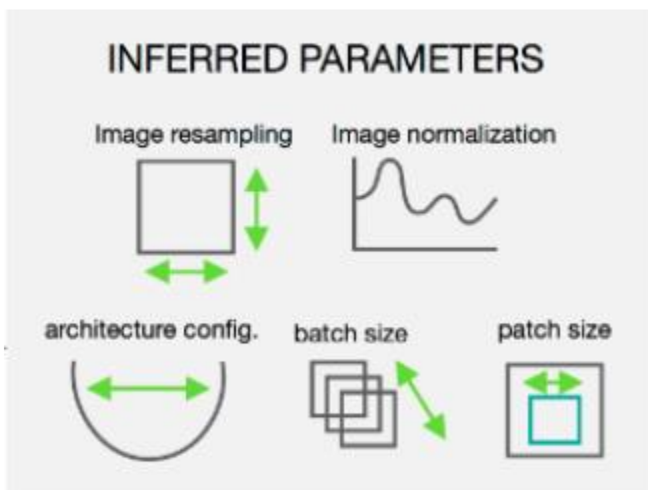
Figure 4: The figure shows the attention coefficients ($\alpha^{l_{s2}}$, $\alpha^{l_{s3}}$) across different training epochs (3, 6, 10, 60, 150). The images are extracted from sagittal and axial planes of a 3D abdominal CT scan from the testing dataset. The model gradually learns to focus on the pancreas, kidney, and spleen.

| Method (Train/Test Split) | U-Net (120/30) | Att U-Net (120/30) | U-Net (30/120) | Att U-Net (30/120) |
|---|---|---|---|---|
| Pancreas DSC | 0.814±0.116 | **0.840±0.087** | 0.741±0.137 | **0.767±0.132** |
| Pancreas Precision | 0.848±0.110 | 0.849±0.098 | 0.789±0.176 | **0.794±0.150** |
| Pancreas Recall | 0.806±0.126 | **0.841±0.092** | 0.743±0.179 | **0.762±0.145** |
| Pancreas S2S Dist (mm) | 2.358±1.464 | **1.920±1.284** | 3.765±3.452 | 3.507±3.814 |
| Spleen DSC | 0.962±0.013 | 0.965±0.013 | 0.935±0.095 | **0.943±0.092** |
| Kidney DSC | 0.963±0.013 | 0.964±0.016 | 0.951±0.019 | 0.954±0.021 |
| Number of Params | 5.88 M | 6.40 M | 5.88 M | 6.40 M |
| Inference Time | 0.167 s | 0.179 s | 0.167 s | 0.179 s |

Isensee et al. nnU-Net 2018

# no-new-UNet (nnUNet)



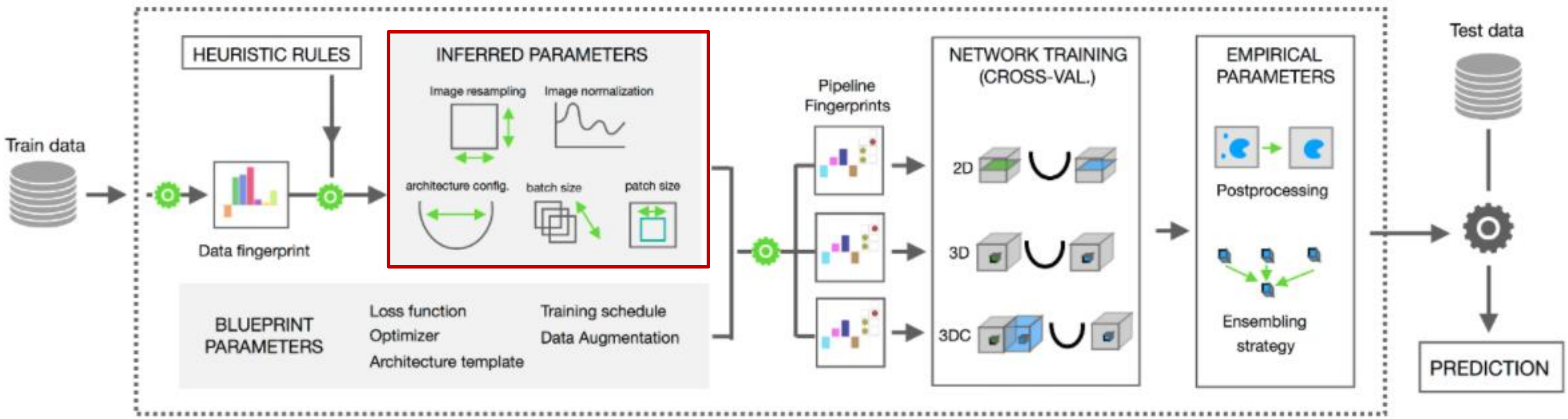

**Intensity normalization**: normalized by clipping to the [0.5, 99.5] percentiles of these intensity values, followed by a z-score normalization based on the mean and standard deviation of all collected intensity values

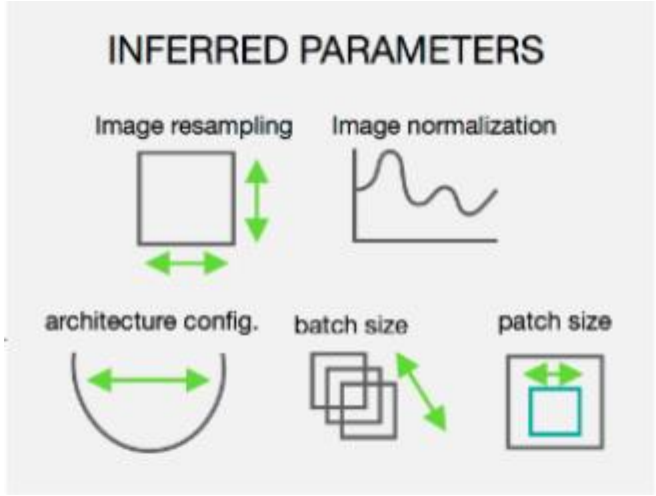
**Resampling**: resampled to the median voxel spacing of their respective dataset

**Cropping**: cropped to the region of nonzero values

https://medium.com/miccai-educational-initiative/nnu-net-the-no-new-unet-for-automatic-segmentation-8d655f3f6d2a

# no-new-UNet (nnUNet)




**Patch size, batch size**: automatically adapted on the basis of the median image size after resampling
**Architecture config.**
**The num of down-sampling**: until the feature maps have size 8 and each axis is down-sampled individually

https://medium.com/miccai-educational-initiative/nnu-net-the-no-new-unet-for-automatic-segmentation-8d655f3f6d2a

# no-new-UNet (nnUNet)

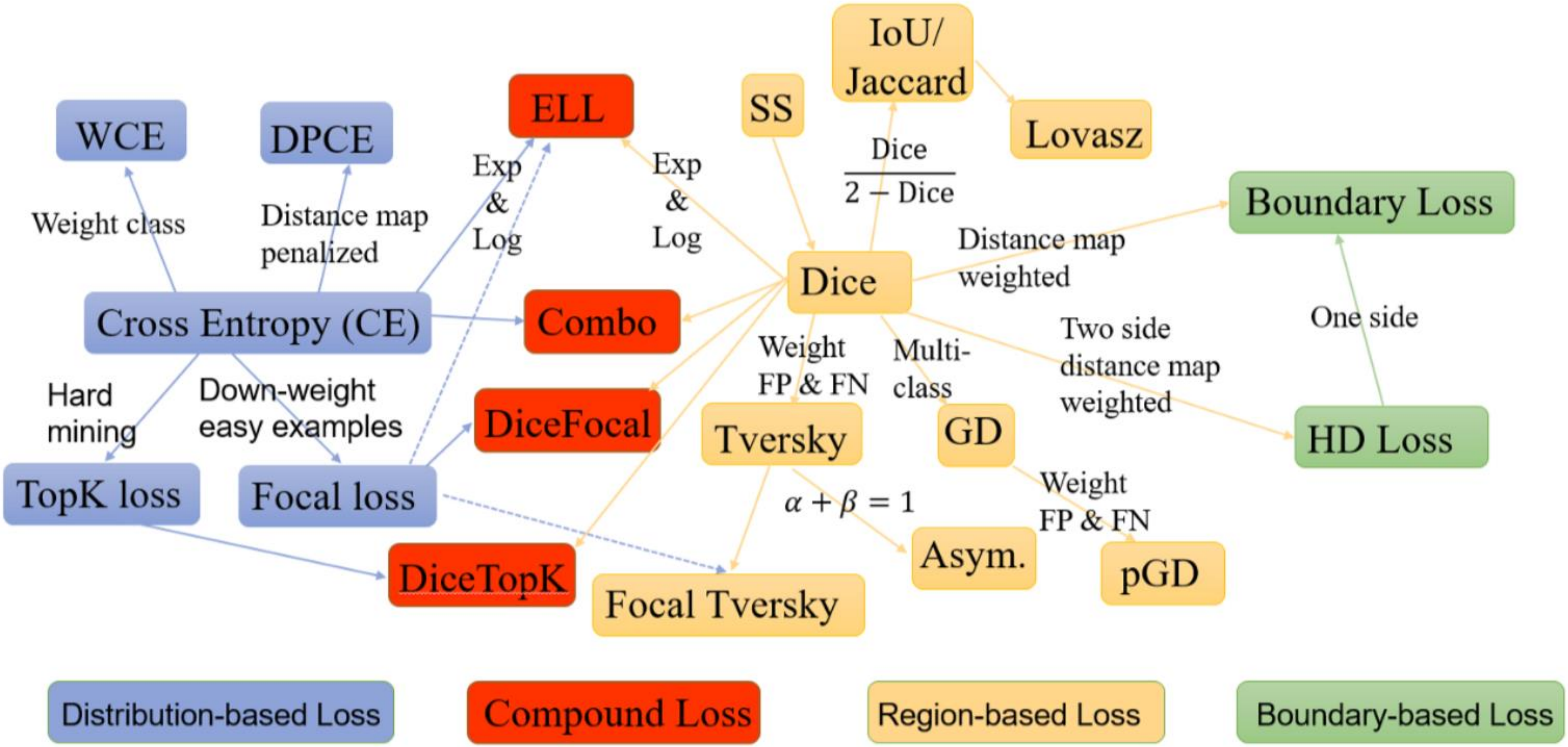
Table 3: Rankings for the development phase and the mystery phase, median and interquartile range (IQR) of the Dice Similarity Coefficient (DSC) values of all team. The ranking was computed as described in Section 2.4.1.

| The development phase | | | | The mystery phase | | | |
|---|---|---|---|---|---|---|---|
| **Rank** | **Team ID** | **Median DSC** | **IQR DSC** | **Rank** | **Team ID** | **Median DSC** | **IQR DSC** |
| 1 | ***nnU-Net*** | 0.79 | (0.61,0.88) | 1 | ***nnU-Net*** | 0.71 | (0.58,0.82) |
| 2 | *K.A.V.athlon* | 0.77 | (0.58,0.86) | 2 | *NVDLMED* | 0.69 | (0.54,0.79) |
| 3 | *NVDLMED* | 0.78 | (0.57,0.87) | 3 | *K.A.V.athlon* | 0.67 | (0.49,0.79) |
| 4 | *Lupin* | 0.75 | (0.52,0.86) | 4 | *LS Wang's Group* | 0.64 | (0.46,0.78) |
| 5 | *CerebriuDIKU* | 0.76 | (0.52,0.88) | 5 | *CerebriuDIKU* | 0.56 | (0.15,0.71) |
| 6 | *LS Wang's Group* | 0.75 | (0.51,0.88) | 6 | *MIMI* | 0.65 | (0.45,0.75) |
| 7 | *MIMI* | 0.73 | (0.51,0.86) | 7 | *Whale* | 0.55 | (0.20,0.68) |
| 8 | *Whale* | 0.65 | (0.28,0.83) | 8 | *UBIlearn* | 0.55 | (0.05,0.69) |
| 9 | *UBIlearn* | 0.72 | (0.40,0.85) | 9 | *LfB* | 0.49 | (0.16,0.63) |
| 10 | *VST* | 0.69 | (0.39,0.84) | 10 | *Jiafucang* | 0.48 | (0.04,0.67) |
| 11 | *BCVuniandes* | 0.70 | (0.42,0.86) | 11 | *A-REUMI01* | 0.51 | (0.14,0.65) |
| 12.5 | *BUT* | 0.72 | (0.39,0.84) | 12 | *AI-MED* | 0.33 | (0.01,0.52) |
| 12.5 | *A-REUMI01* | 0.70 | (0.42,0.85) | 13 | *Lupin* | 0.57 | (0.19,0.69) |
| 14 | *Jiafucang* | 0.49 | (0.11,0.81) | 14 | *VST* | 0.41 | (0.00,0.64) |
| 15 | *LfB* | 0.68 | (0.43,0.82) | 15 | *Lesswire1* | 0.40 | (0.08,0.52) |
| 16 | *AI-Med* | 0.63 | (0.29,0.79) | 16 | *BUT* | 0.38 | (0.01,0.60) |
| 17 | *Lesswire1* | 0.65 | (0.33,0.79) | 17 | *BCVuniandes* | 0.10 | (0.01,0.38) |
| 18 | *EdwardMa12593* | 0.31 | (0.01,0.69) | 18 | *RegionTec* | 0.29 | (0.00,0.50) |
| 19 | *RegionTec* | 0.57 | (0.19,0.73) | 19 | *EdwardMa12593* | 0.08 | (0.01,0.18) |

https://arxiv.org/abs/2106.05735: The Medical Segmentation Decathlon

# Others

# Loss function




https://github.com/JunMa11/SegLoss

# Evaluation metrics

- Overlap based metrics
    - Dice index, Jaccard index

$$DICE = \frac{2\,|S_g^1 \cap S_t^1|}{|S_g^1| + |S_t^1|} = \frac{2TP}{2TP + FP + FN} \qquad JAC = \frac{|S_g^1 \cap S_t^1|}{|S_g^1 \cup S_t^1|} = \frac{TP}{TP + FP + FN}$$

    - Accuracy,  Recall(=sensitivity), Specificity, Precision,

$$Recall = Sensitivity = TPR = \frac{TP}{TP + FN} \qquad Precision = PPV = \frac{TP}{TP + FP}$$

$$Specificity = TNR = \frac{TN}{TN + FP}$$

- Distance based metrics
    - Average Hausdorff Distance

$$AVD(A, B) = max(d(A, B), d(B, A)) \quad d(A, B) = \frac{1}{N} \sum_{a \in A} \min_{b \in B} \|a - b\|$$

# Public datasets for medical segmentation

| Dataset | Objects | URL |
|---|---|---|
| LiTS [137] | Liver | https://competitions.codalab.org/competitions/17094 |
| Sliver07 [138] | Liver | http://www.sliver07.org/ |
| 3Dircadb [139] | Liver | https://www.ircad.fr/research/3dircadb/ |
| NIH Pancreas [140] | Pancreas | http://academictorrents.com/details/80ecfefcabede760cdbdf63e38986501f7becd49 |
| COLONOGRAPHY [141] | Colon cancer | https://wiki.cancerimagingarchive.net/display/Public/CT+COLONOGRAPHY#dc149b9170f54aa29e88f1119e25ba3e |
| AMRG Cardiac Atlas [142] | Heart | http://www.cardiacatlas.org/studies/amrg-cardiac-atlas/ |
| LIDC-IDRI [143] | Lung | https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI# |
| PROMISE12 [144] | Prostate | https://promise12.grand-challenge.org/ |
| OASIS [145] | Brain | http://www.oasis-brains.org/ |
| BRATS [3] [146] [147] | Brain | https://www.med.upenn.edu/sbia/brats2018/registration.html |
| ISLES [148] | Brain | http://www.isles-challenge.org/ |
| mTOP [149] | Brain | https://www.smir.ch/MTOP/Start2016 |
| KITS [150] | Kidney | https://kits19.grand-challenge.org |
| CHAOS [151] | Spleen, Liver, Kidneys | https://chaos.grand-challenge.org |
| Medical Segmentation Decathlon [152] | Spleen, Liver, Pancreas, Brain Tumor, Heart, Hippocampus, prostate, Lung, Hepatic Vessel, Colon | http://medicaldecathlon.com/index.html |

Medical Image Segmentation Using Deep Learning: A Survey (2020)

# Thank you