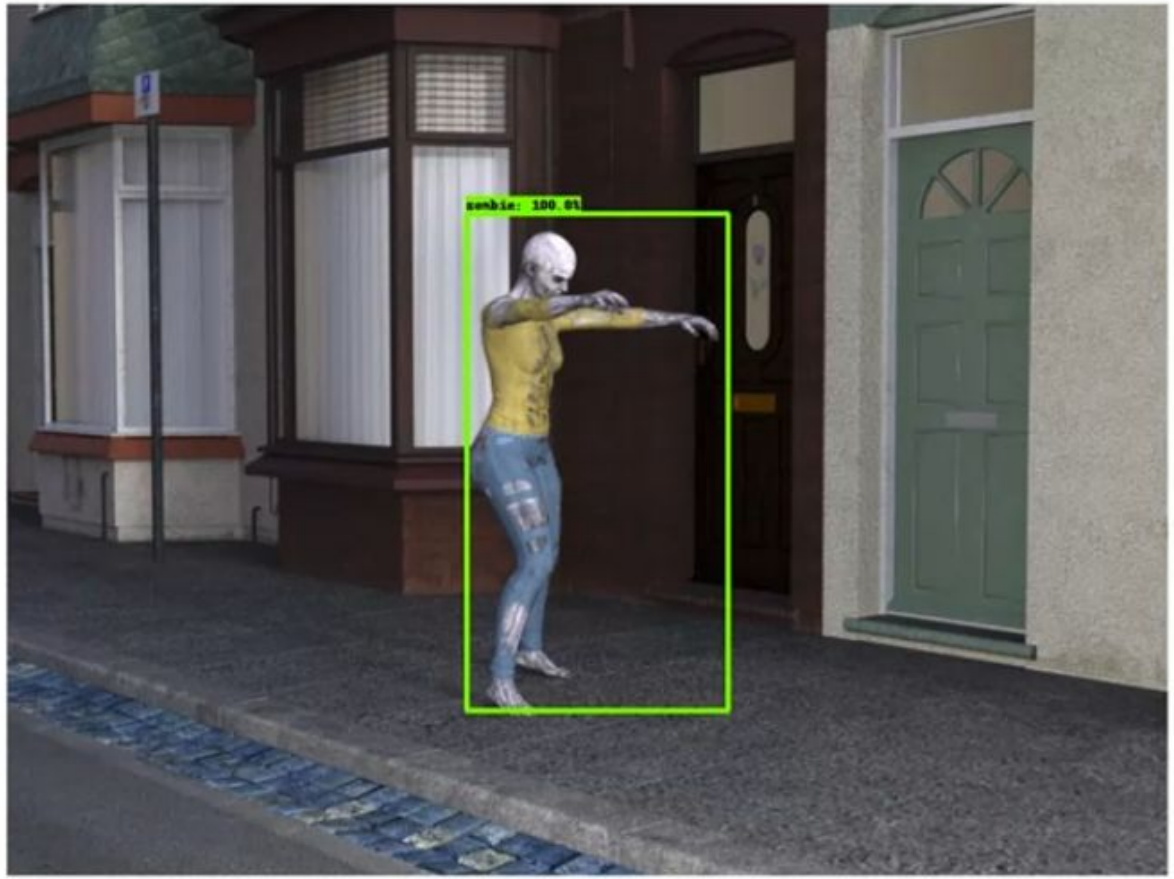
Image Segmentation

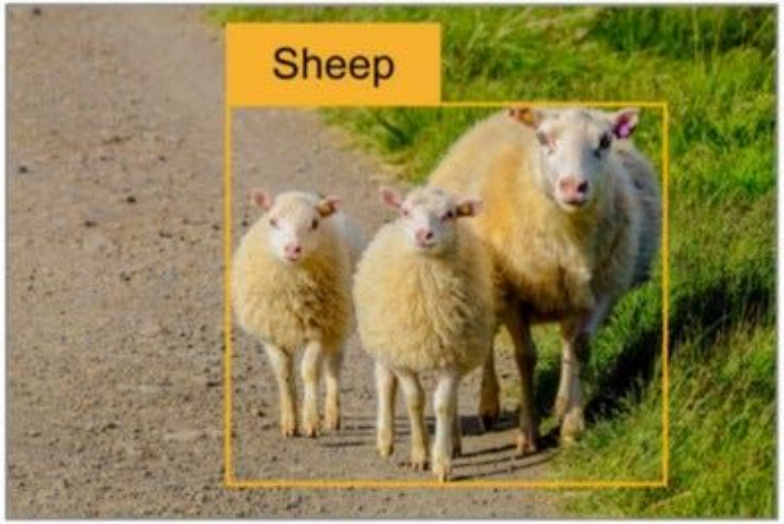
Segmentation VS Detection

BBox in Object Detection

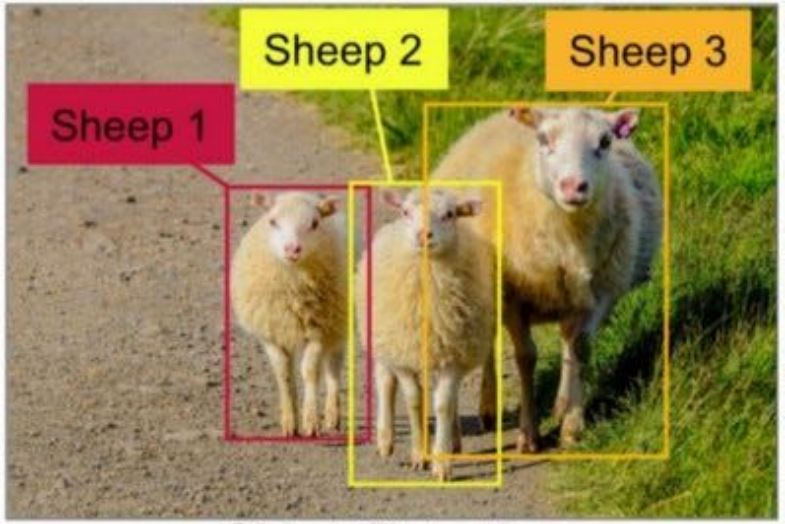


In Segmentation outline of the shape determine

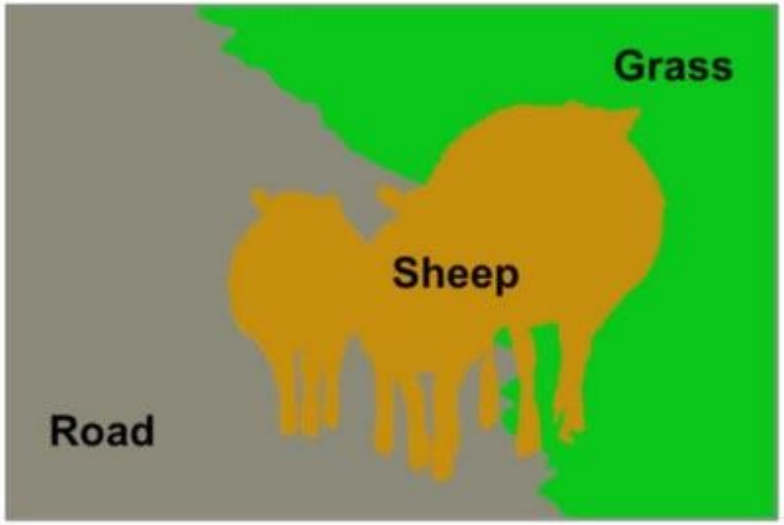
Segmentation VS Detection



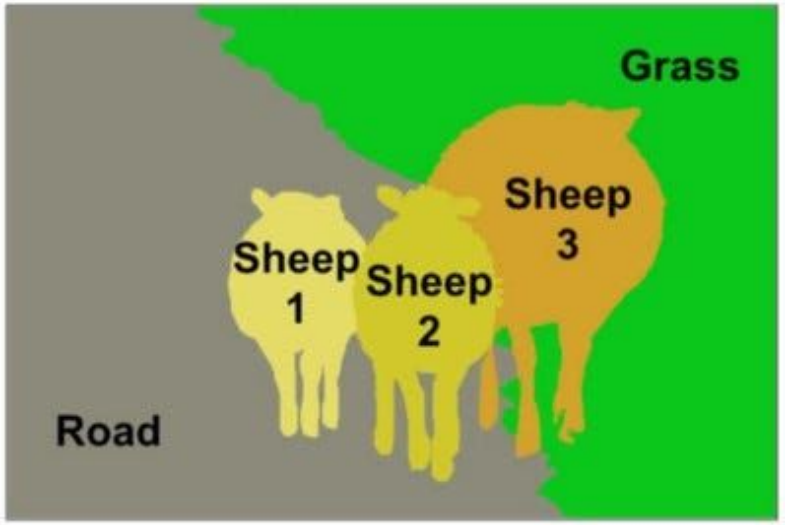
Classification + Localization



Object Detection

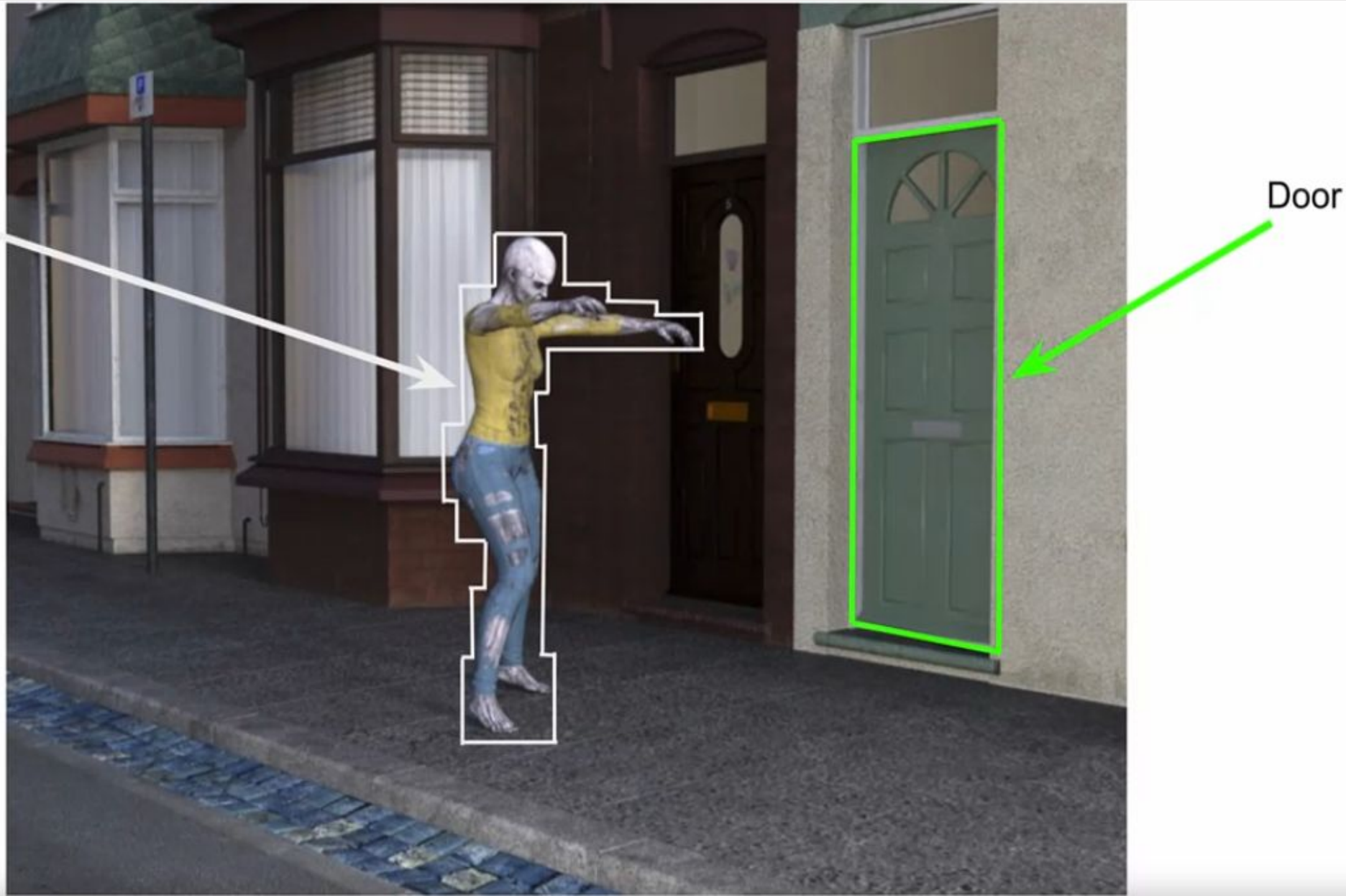


Semantic Segmentation



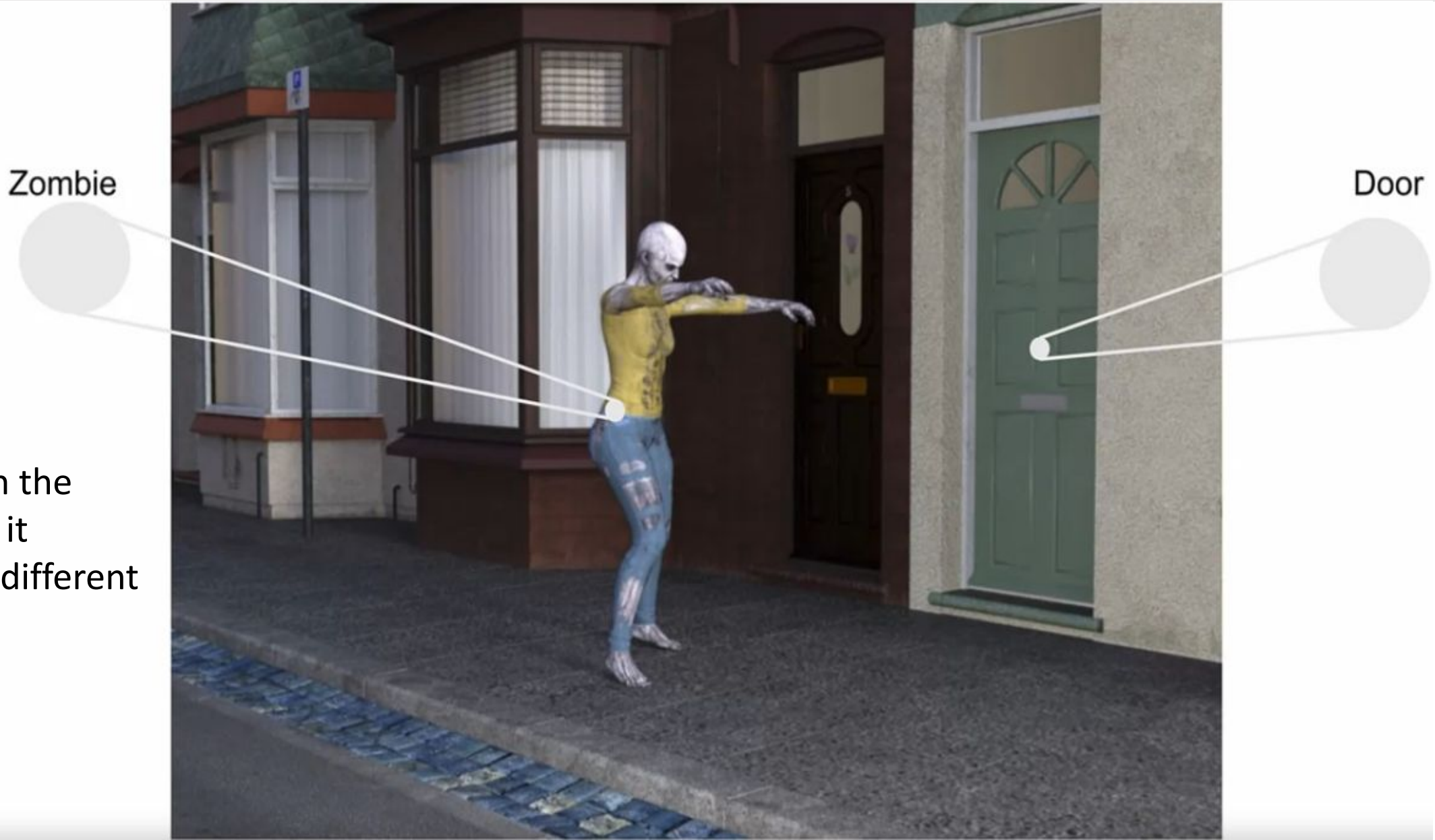
Instance Segmentation

What is image Segmentation?



Determine how to partition an image into multiple segments and have associated with an object.

What is image Segmentation?



Get each pixel in the image and have it classified into a different class.

Segmentation types

Semantic Segmentation

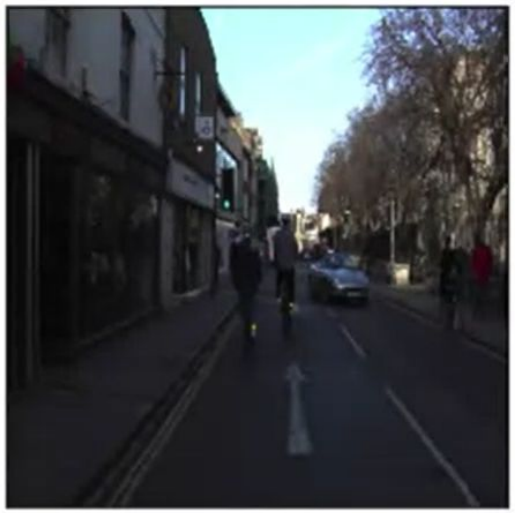
Instance Segmentation

Segmentation masks/levels/groundtruth

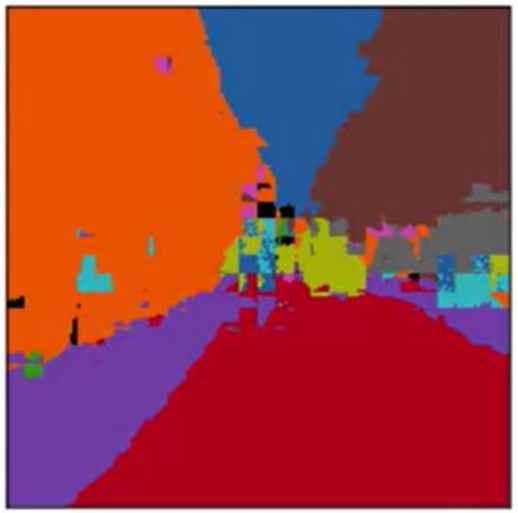
Classes indices = [0 = Background 1 = People]

Segmentation prediction

Sample Visualization of Predicted Segments



Original Image



Predicted Segments

Ground Truth Segments

Image Segmentation

- Encoder
  - CNN without fully connected layers
  - Aggregates low level features to high level features
- Decoder
  - Replaces fully connected layers in a CNN
  - Up samples image to original size to generate a pixel mask

Popular Architectures

- Fully Convolutional Neural Networks.
  - SegNet
  - UNet
  - PSPNet

Used for semantic segmentation

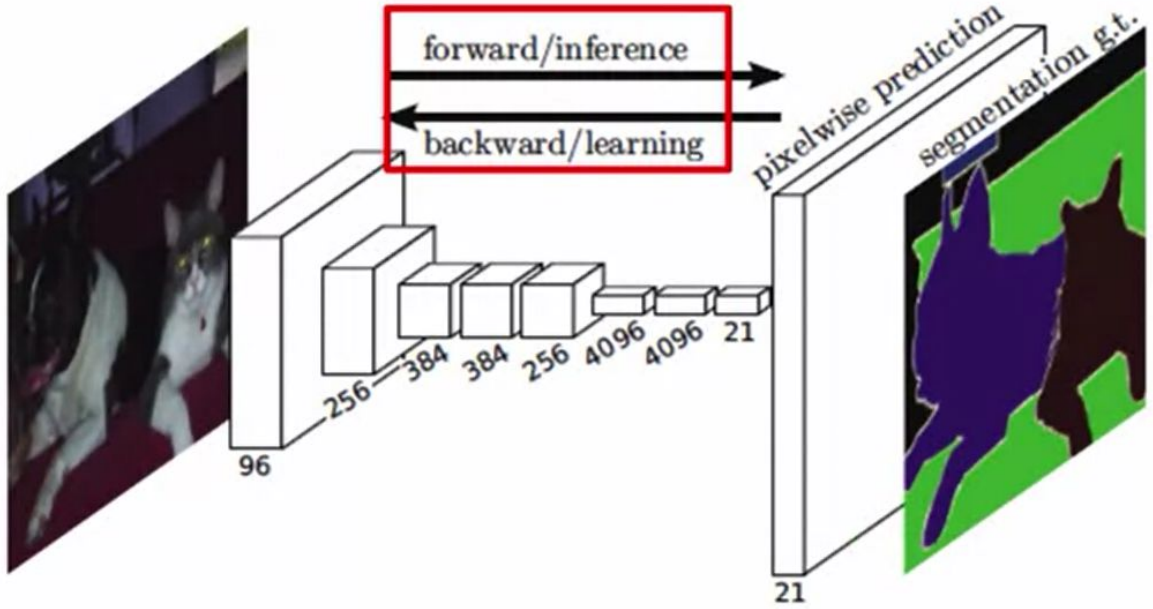
- Mask-RCNN → Used for Instance segmentation

Fully Convolutional Neural Networks

- "Fully Convolutional Networks for Semantic Segmentation" https://arxiv.org/abs/1411.4038
- Replace the fully connected layers with convolutional layers
- Earlier conv layers: Feature extraction and down sampling
- Later conv layers: up sample and pixel-wise labelmap.

Semantic Segmentation Model: FCN

Fully Convolutional Neural Networks



https://people.eecs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf

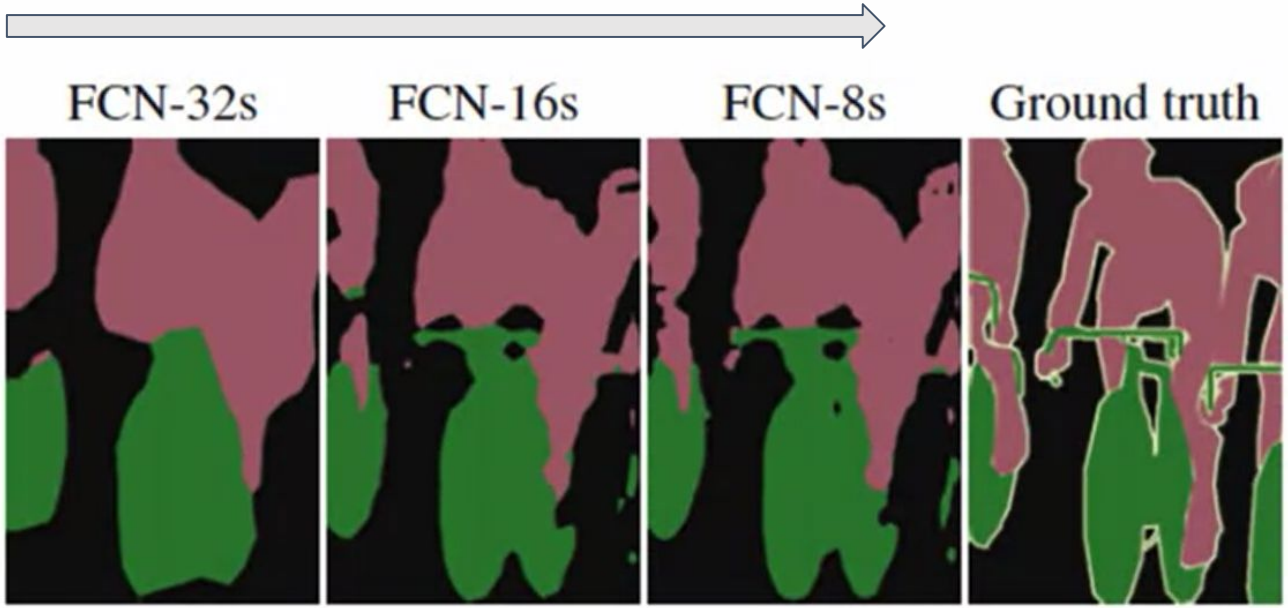
Encoder

- Popular encoder architectures:
  - VGG-16
  - ResNet-50
  - MobileNet
- Reuse convolutional layers for feature extraction.
  - Do not reuse fully connected layers

Semantic Segmentation Model: FCN

Comparison of Different FCNs

Increase the performance based on way of upsampling

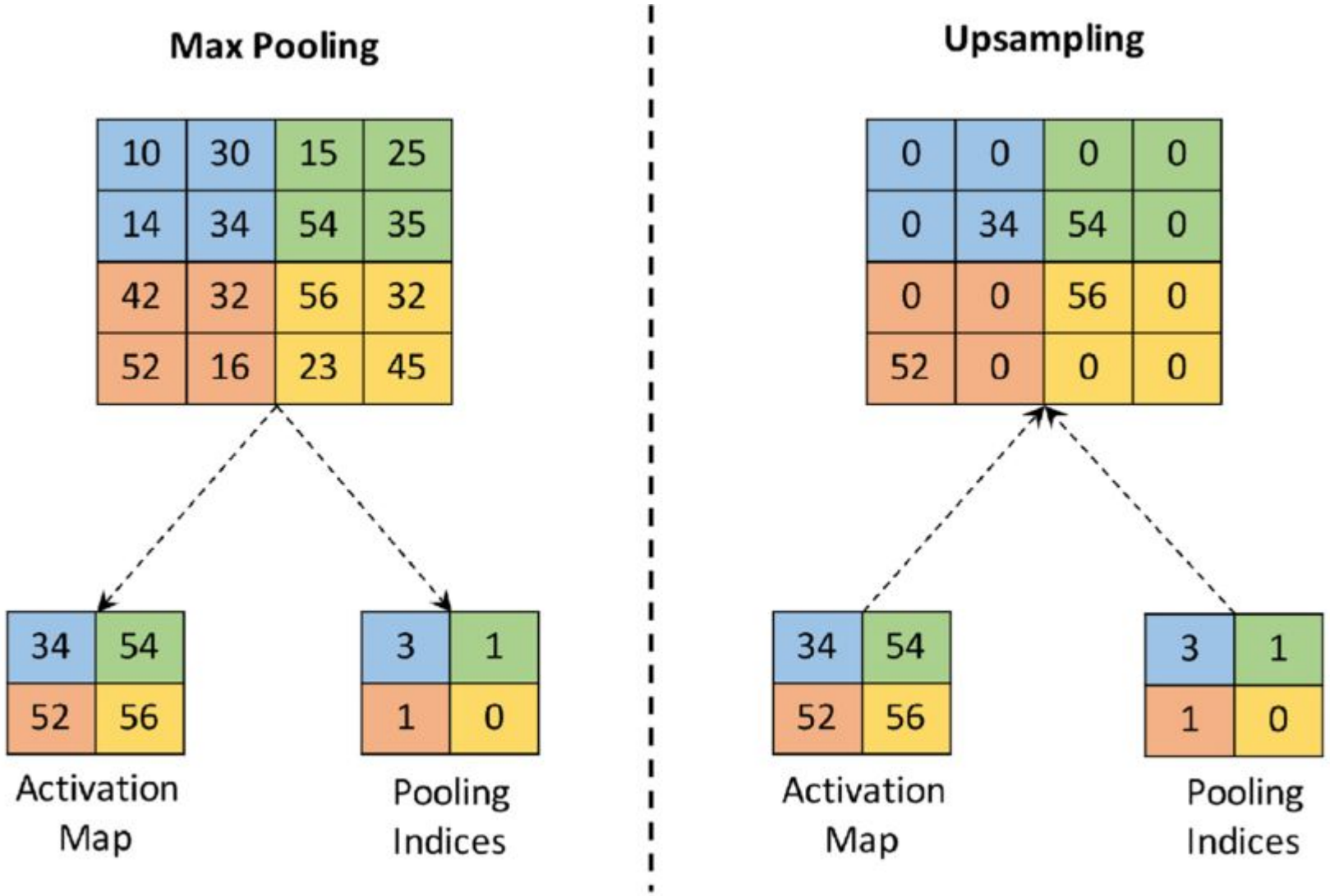


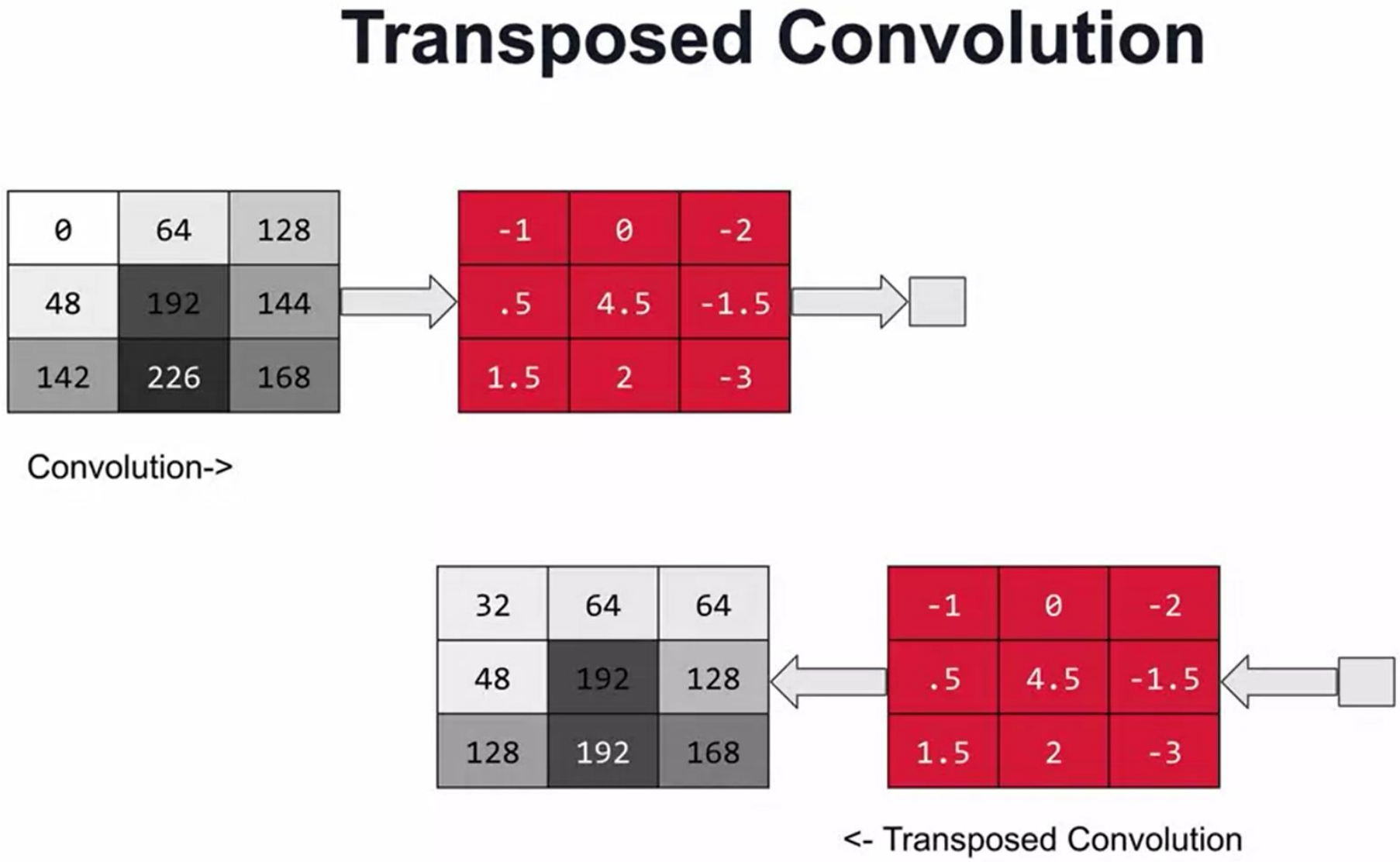
https://arxiv.org/pdf/1411.4038.pdf

Simple Scaling - UpSampling2D

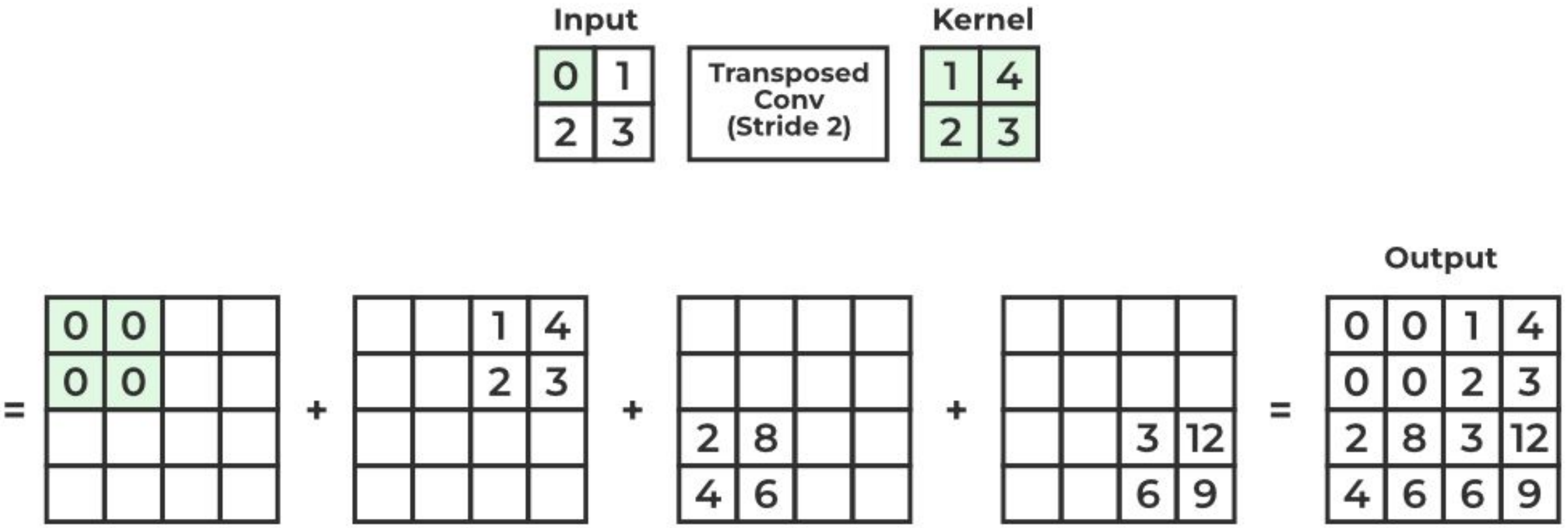
- Upsampling2D scales up the image
- Two Types of scaling:
  - Nearest
    - Copies value from nearest pixel.
  - Bilinear
    - linear interpolation from nearby pixels.

Upsampling

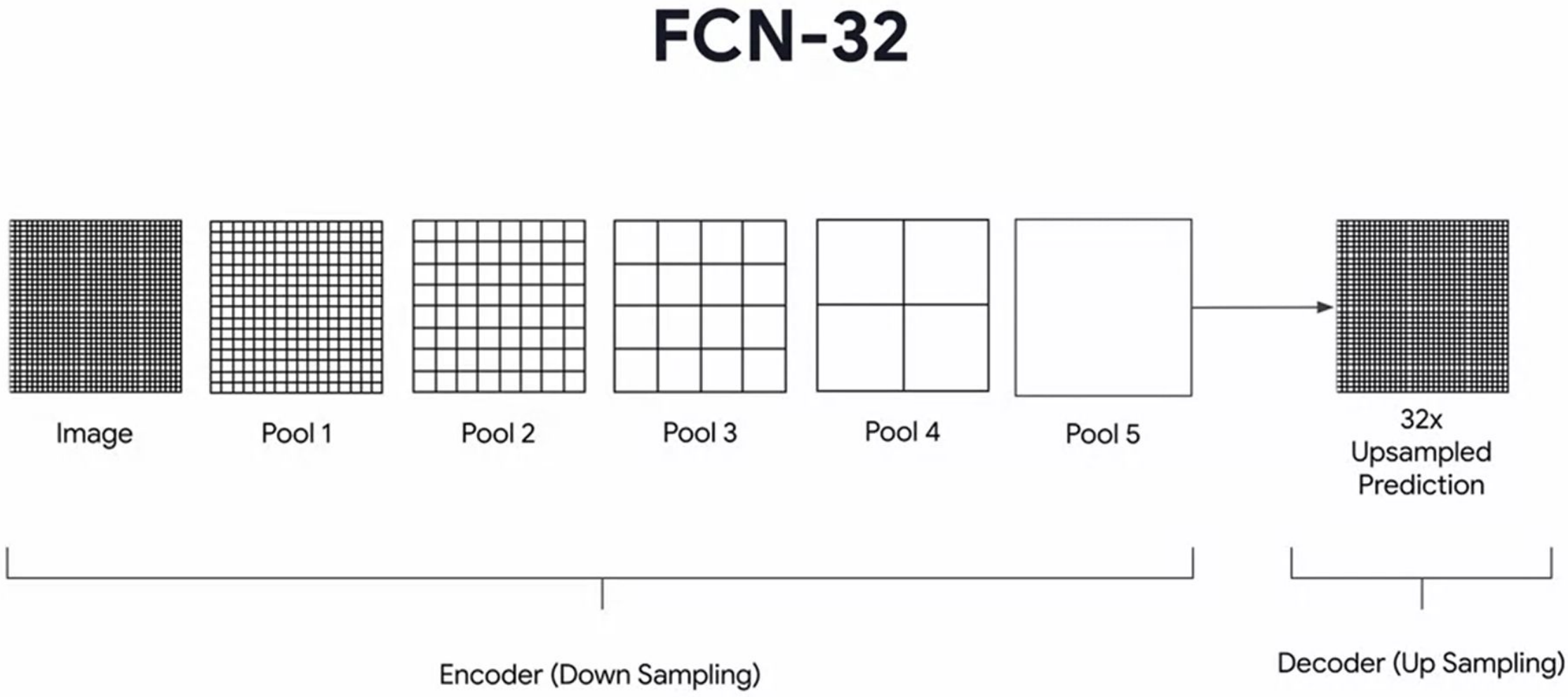

Max Pooling
10 30 15 25
14 34 54 35
42 32 56 32
52 16 23 45
34 54
52 56
Activation Map
3 1
1 0
Pooling Indices
Upsampling
0 0 0 0
0 34 54 0
0 0 56 0
52 0 0 0
34 54
52 56
Activation Map
3 1
1 0
Pooling Indices


Transposed Convolution
0
64
128
48
192
144
142
226
168
-1
0
-2
.5
4.5
-1.5
1.5
2
-3
Convolution->
32
64
64
48
192
128
128
192
168
-1
0
-2
.5
4.5
-1.5
1.5
2
-3
<- Transposed Convolution

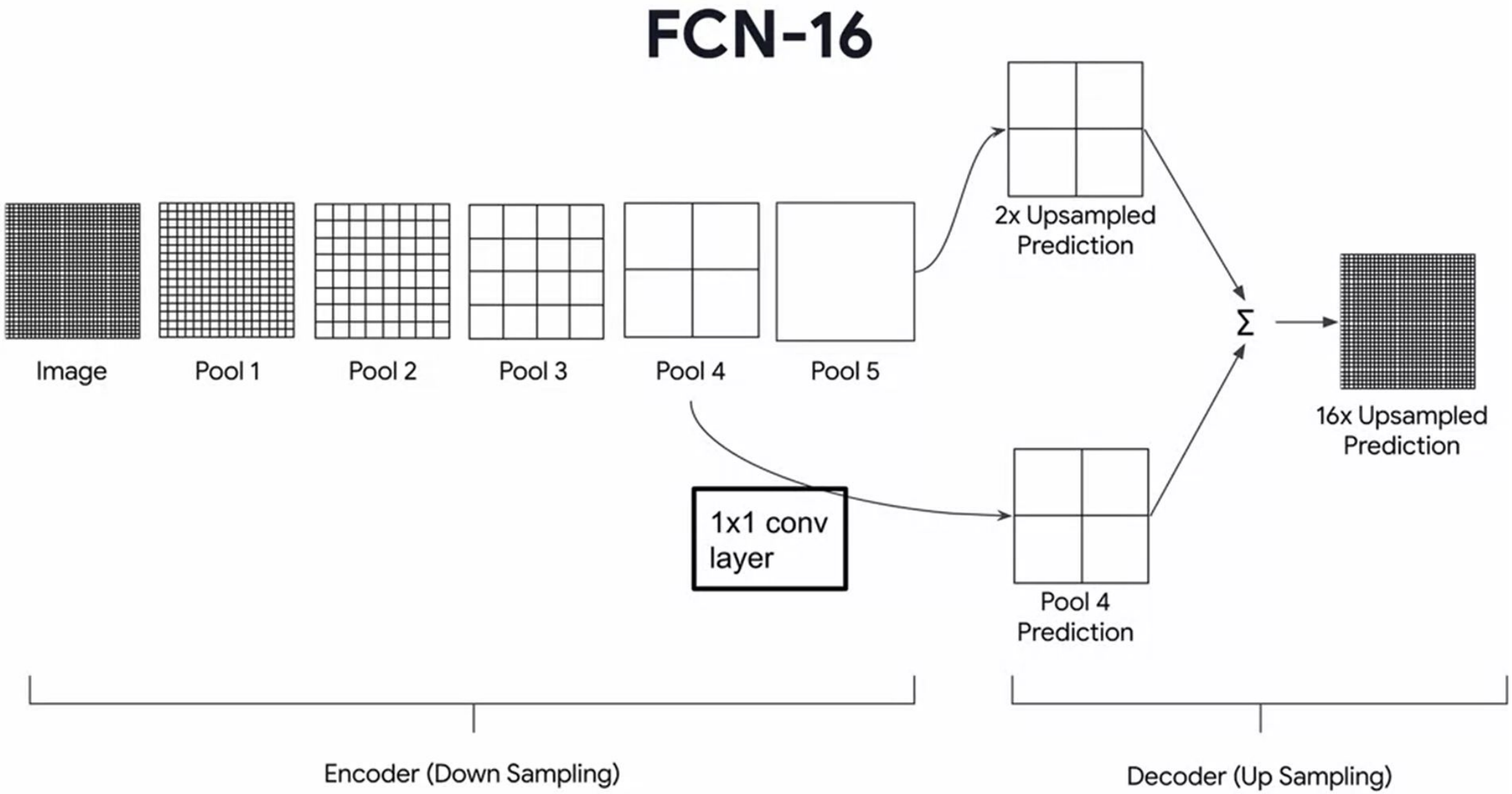
Transposed Convolution



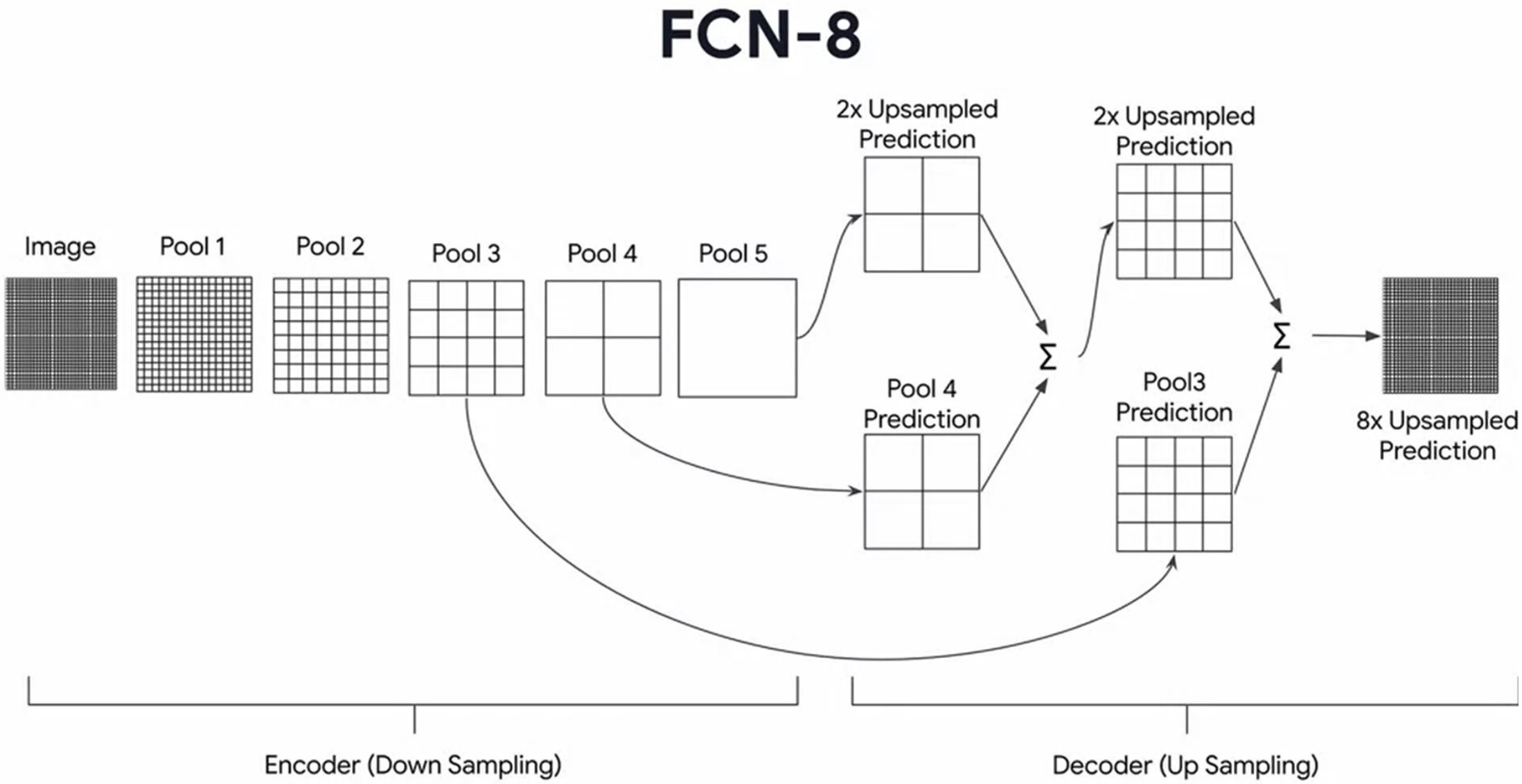
Semantic Segmentation Model: FCN Types



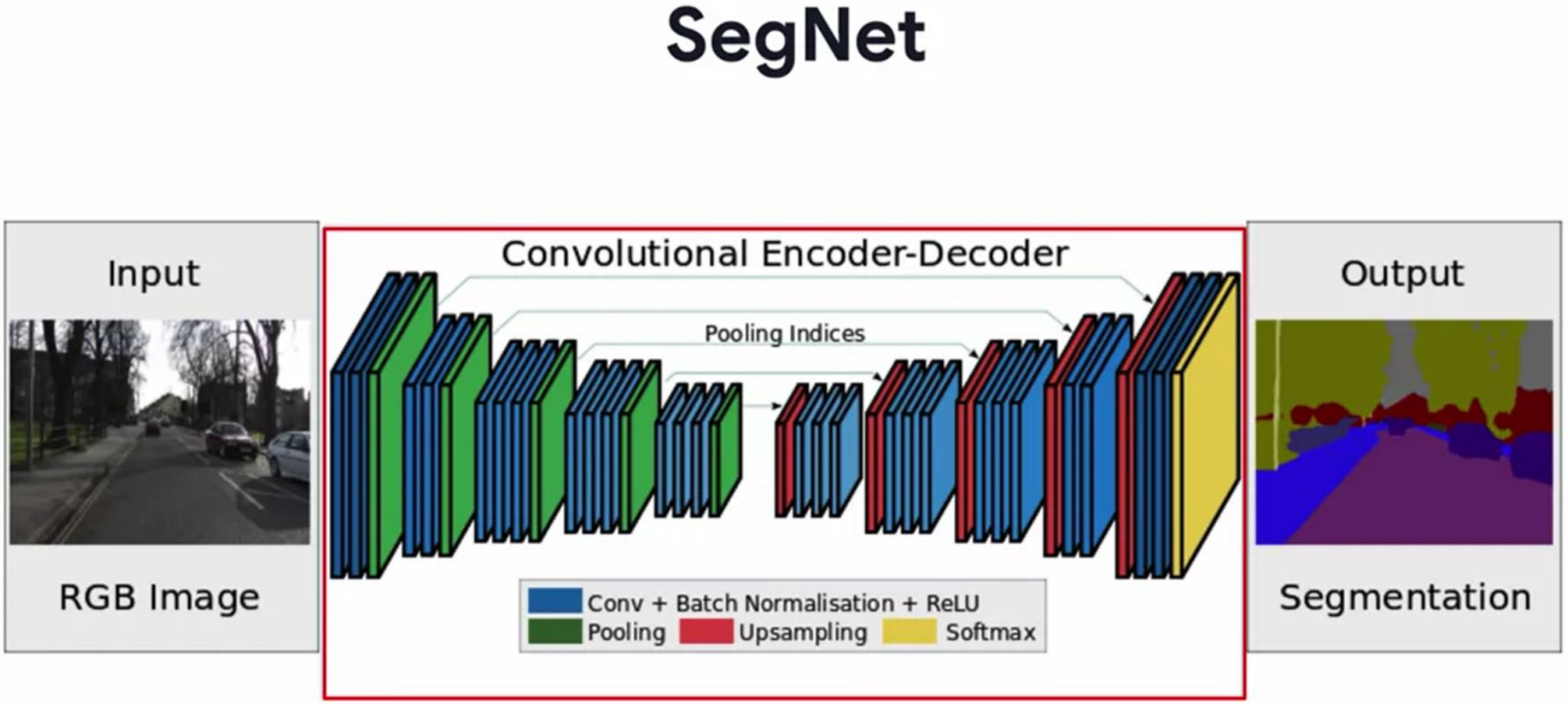
Semantic Segmentation Model: FCN Types



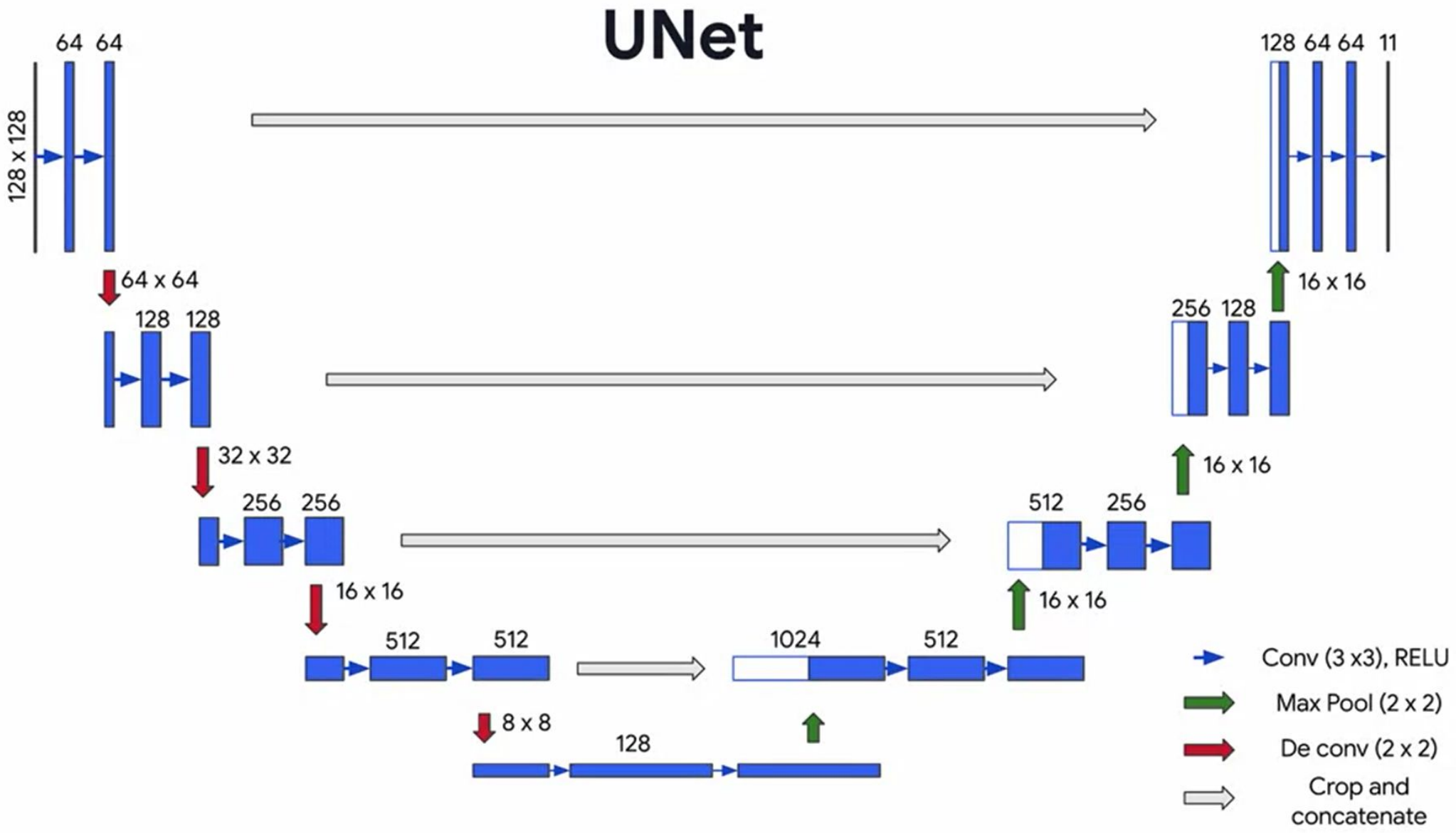
Semantic Segmentation Model: FCN Types



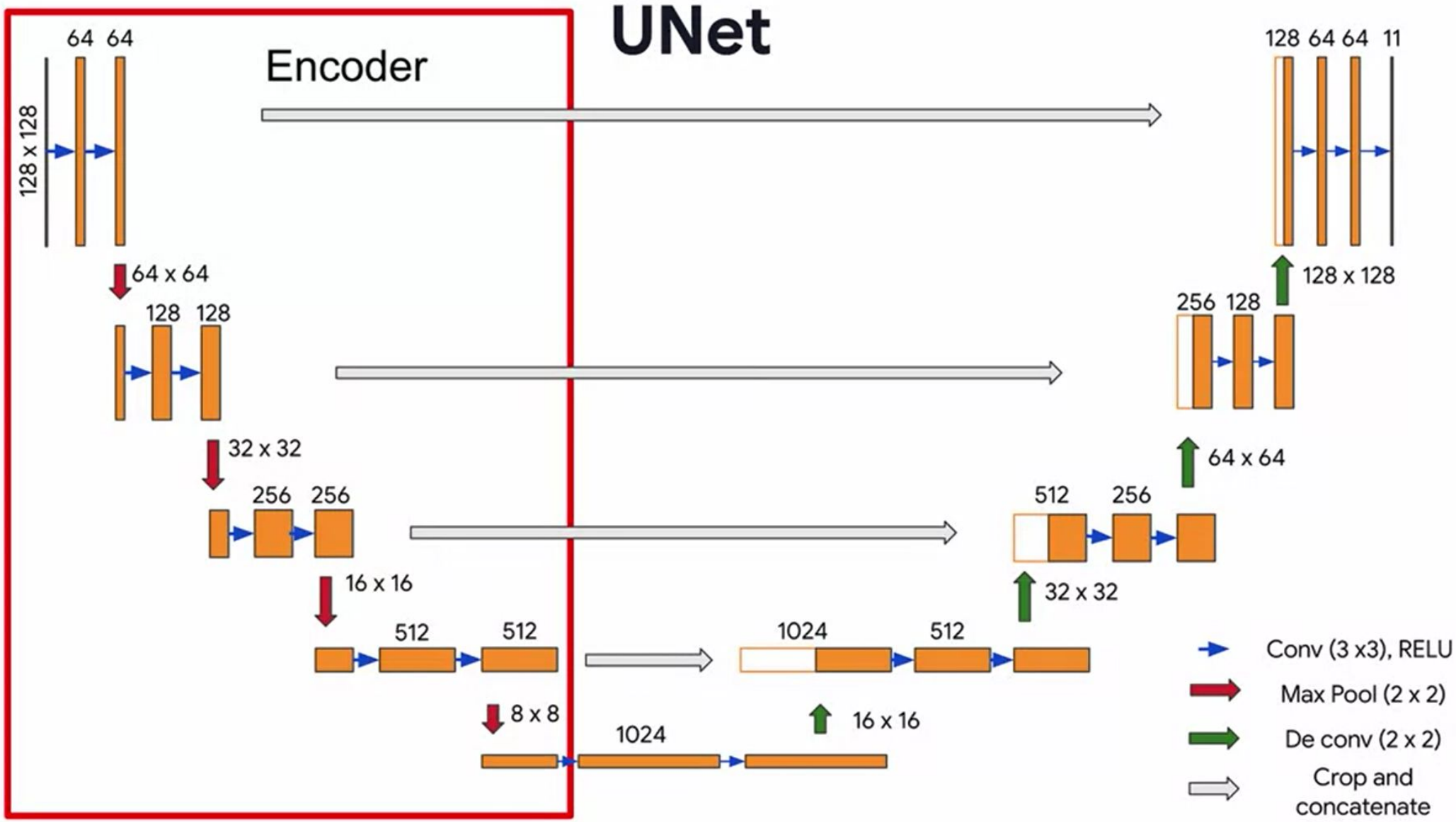
Semantic Segmentation Model : SegNet



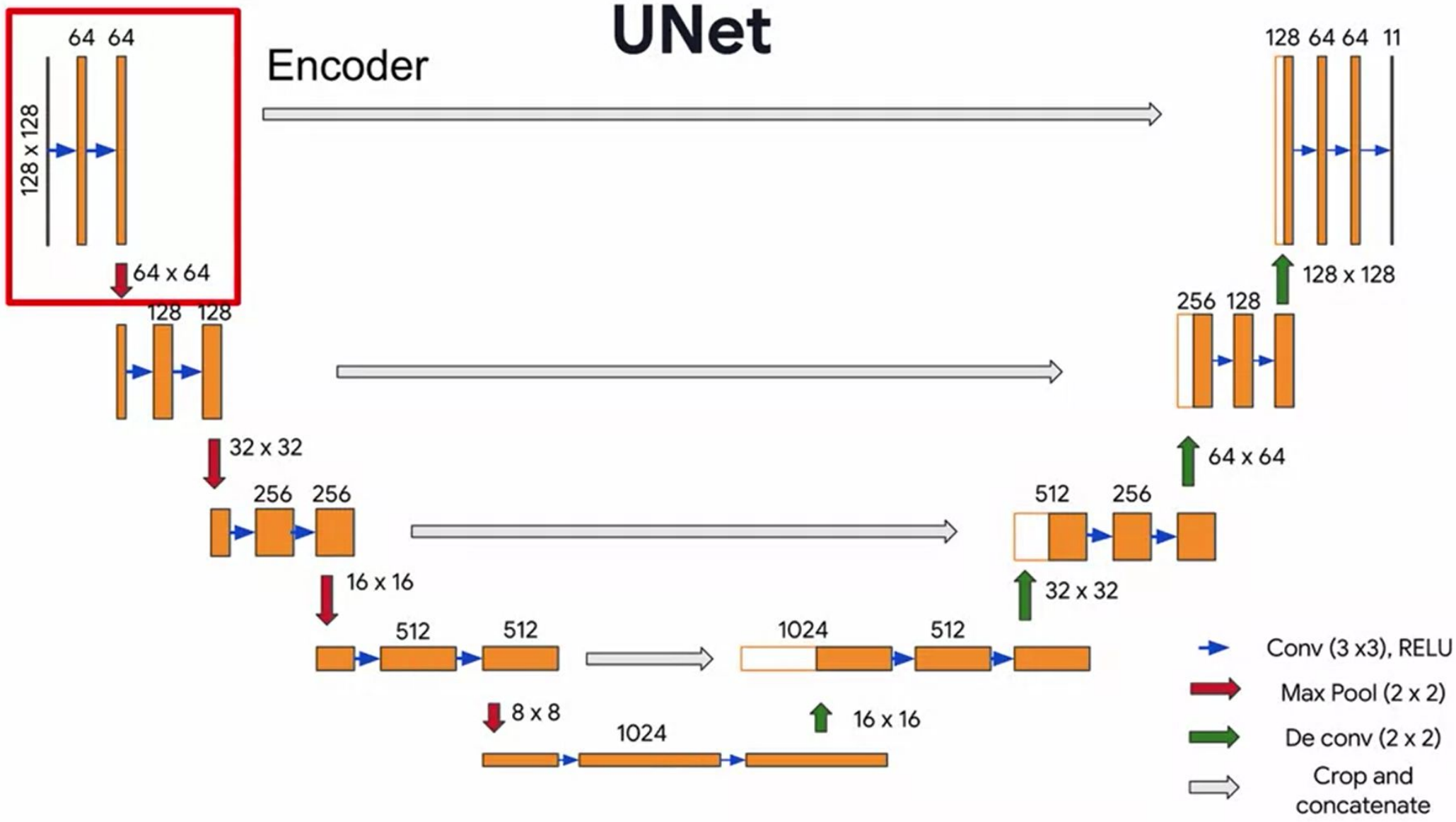
Semantic Segmentation Model: Unet



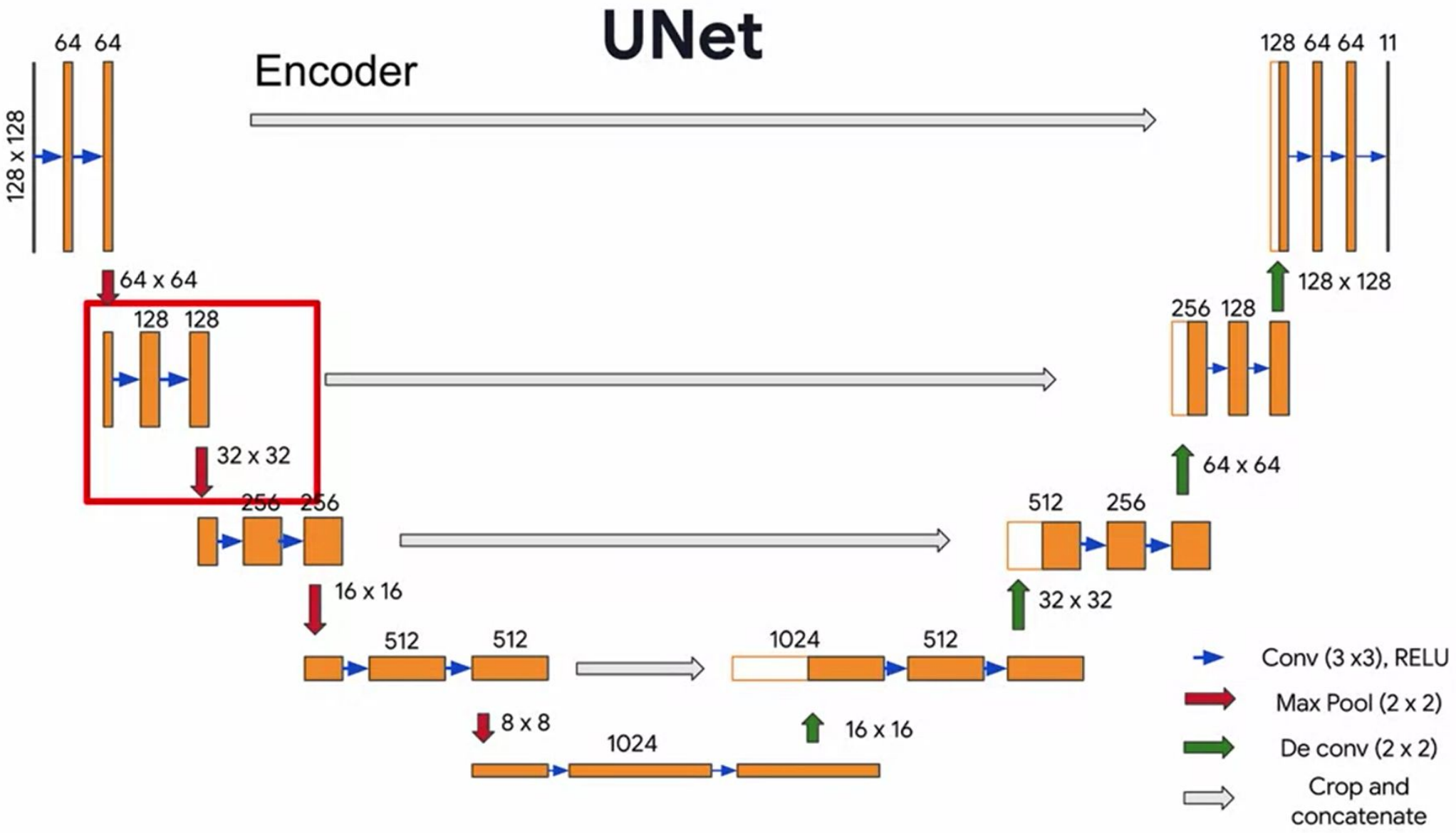
Semantic Segmentation Model: Unet



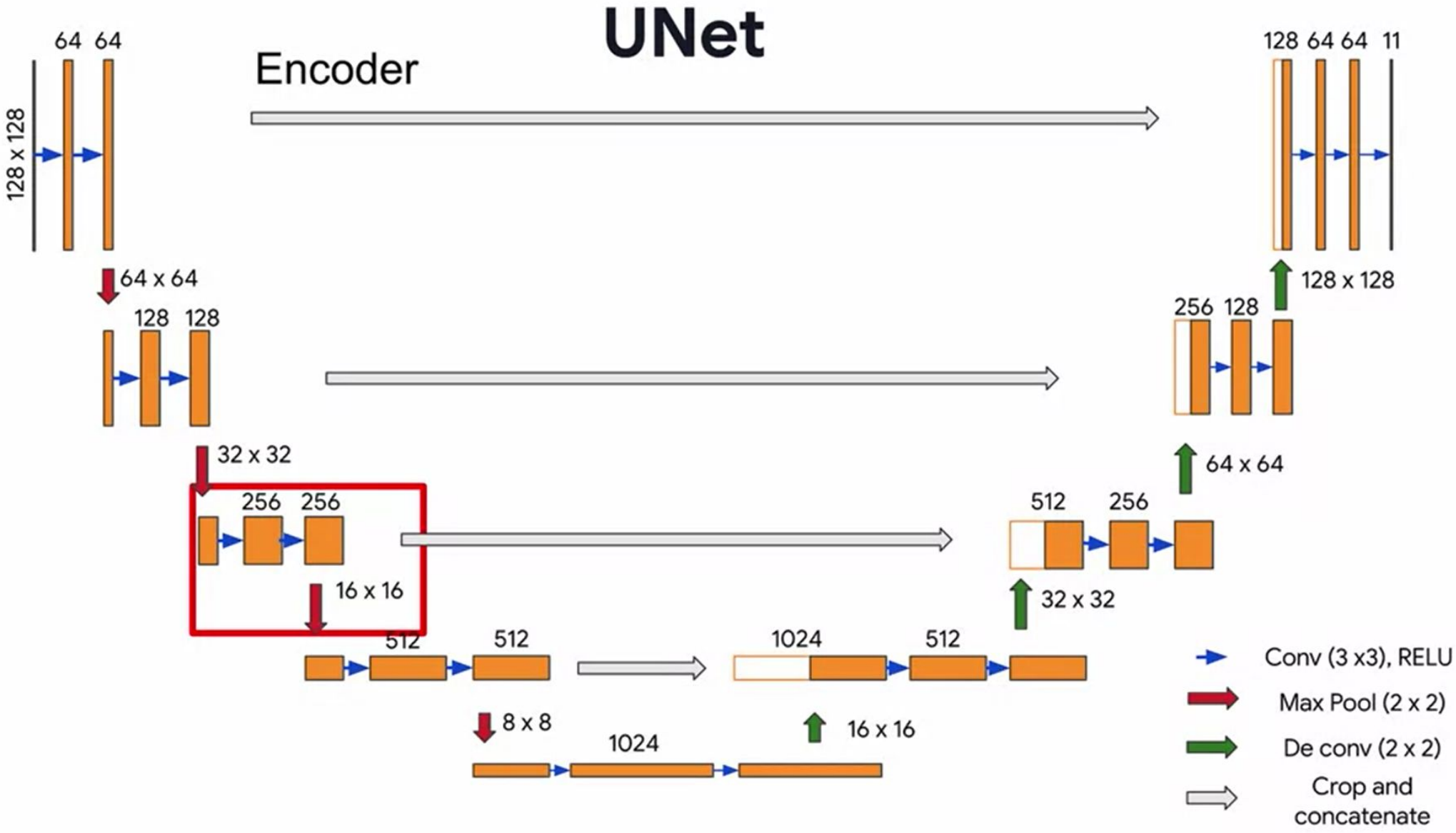
Semantic Segmentation Model: Unet



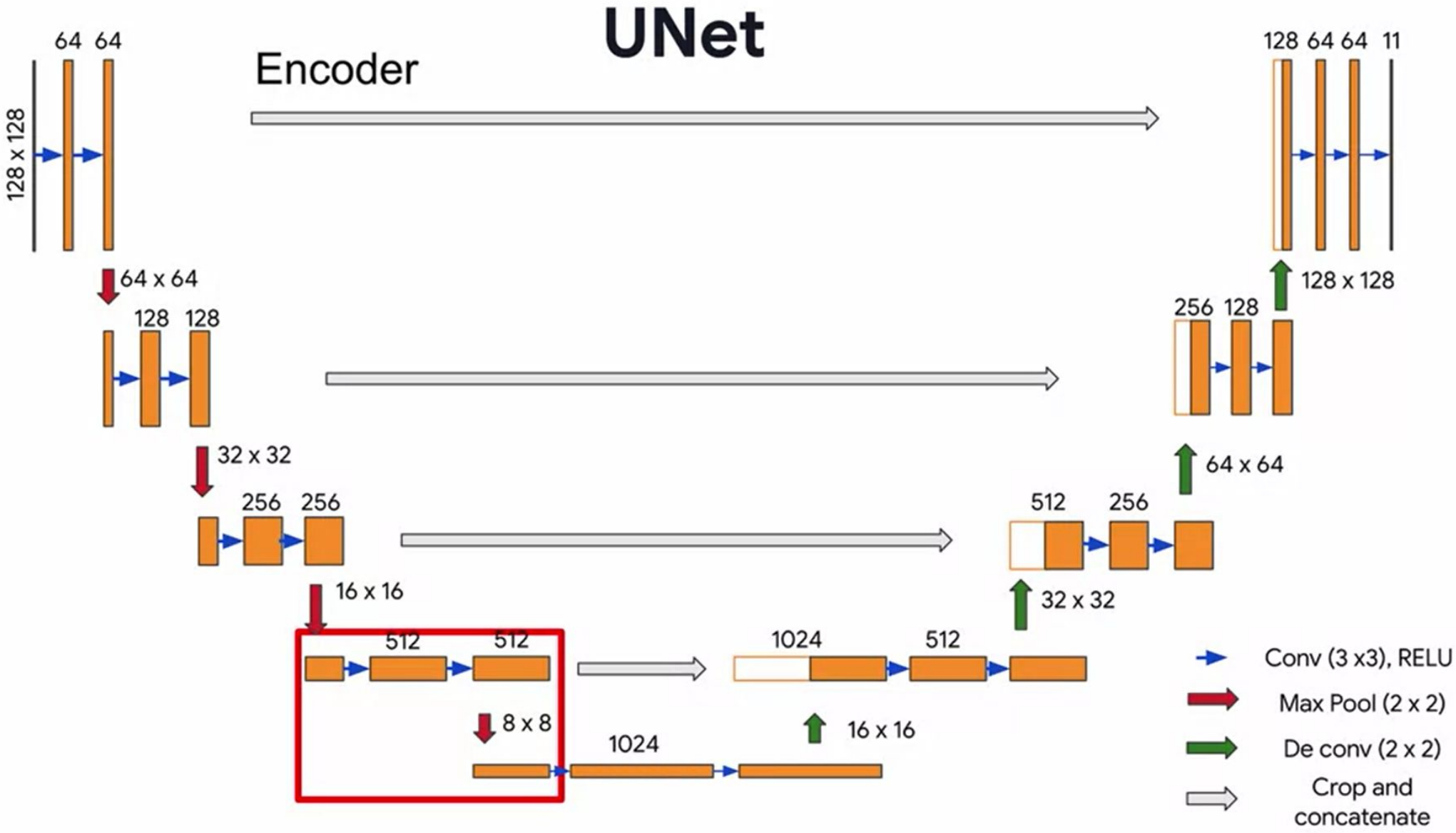
Semantic Segmentation Model: Unet



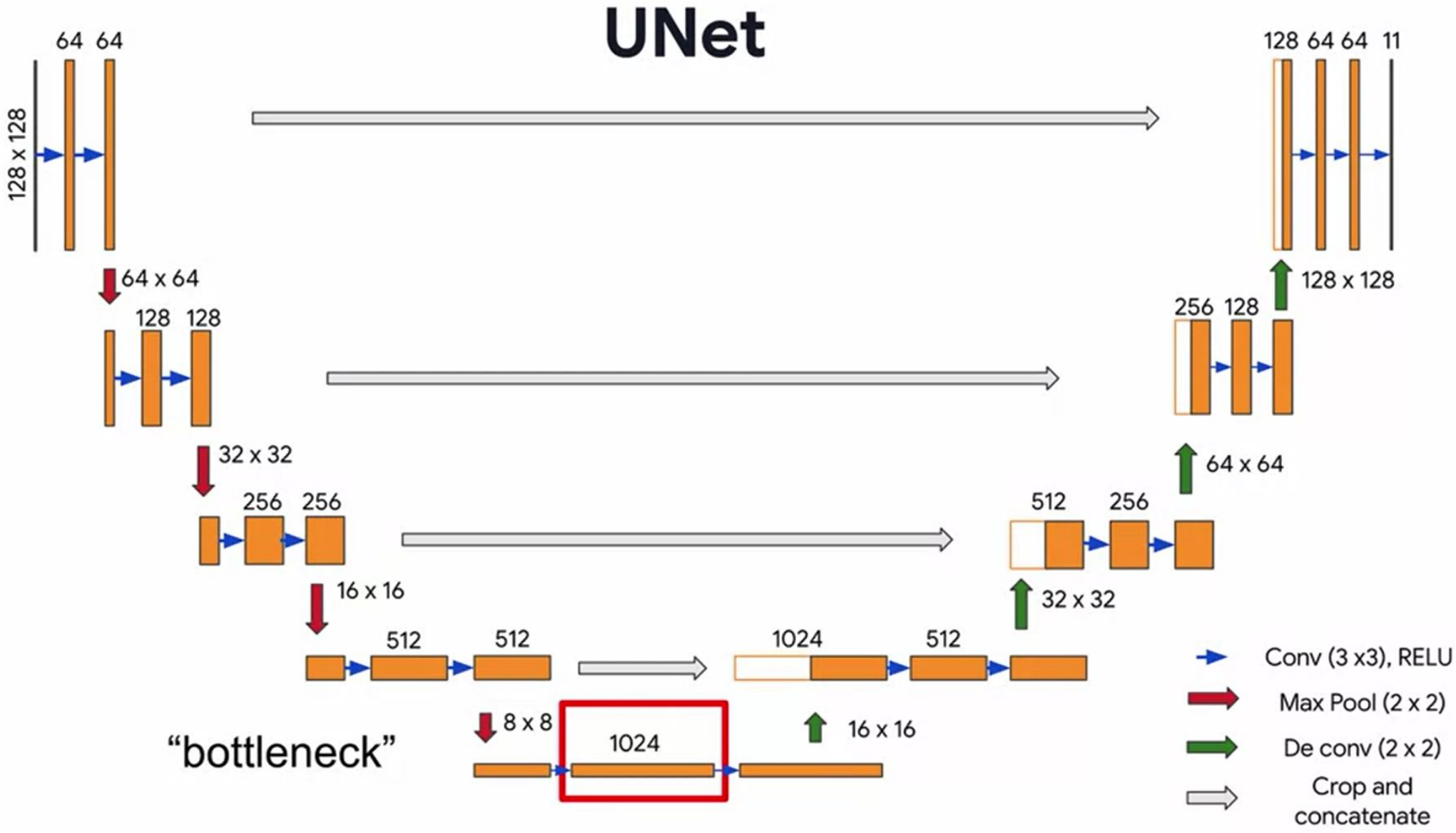
Semantic Segmentation Model: Unet



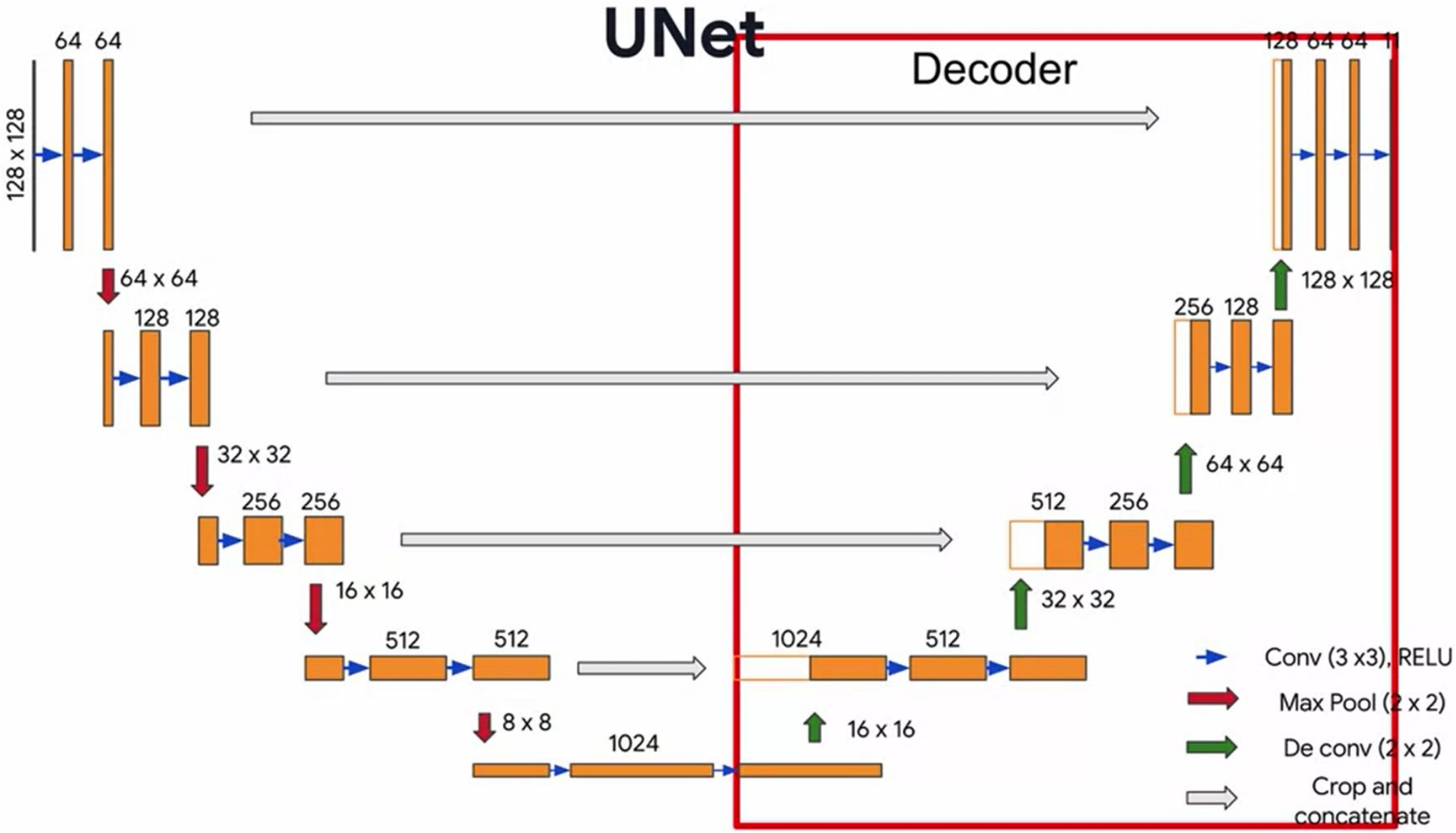
Semantic Segmentation Model: Unet



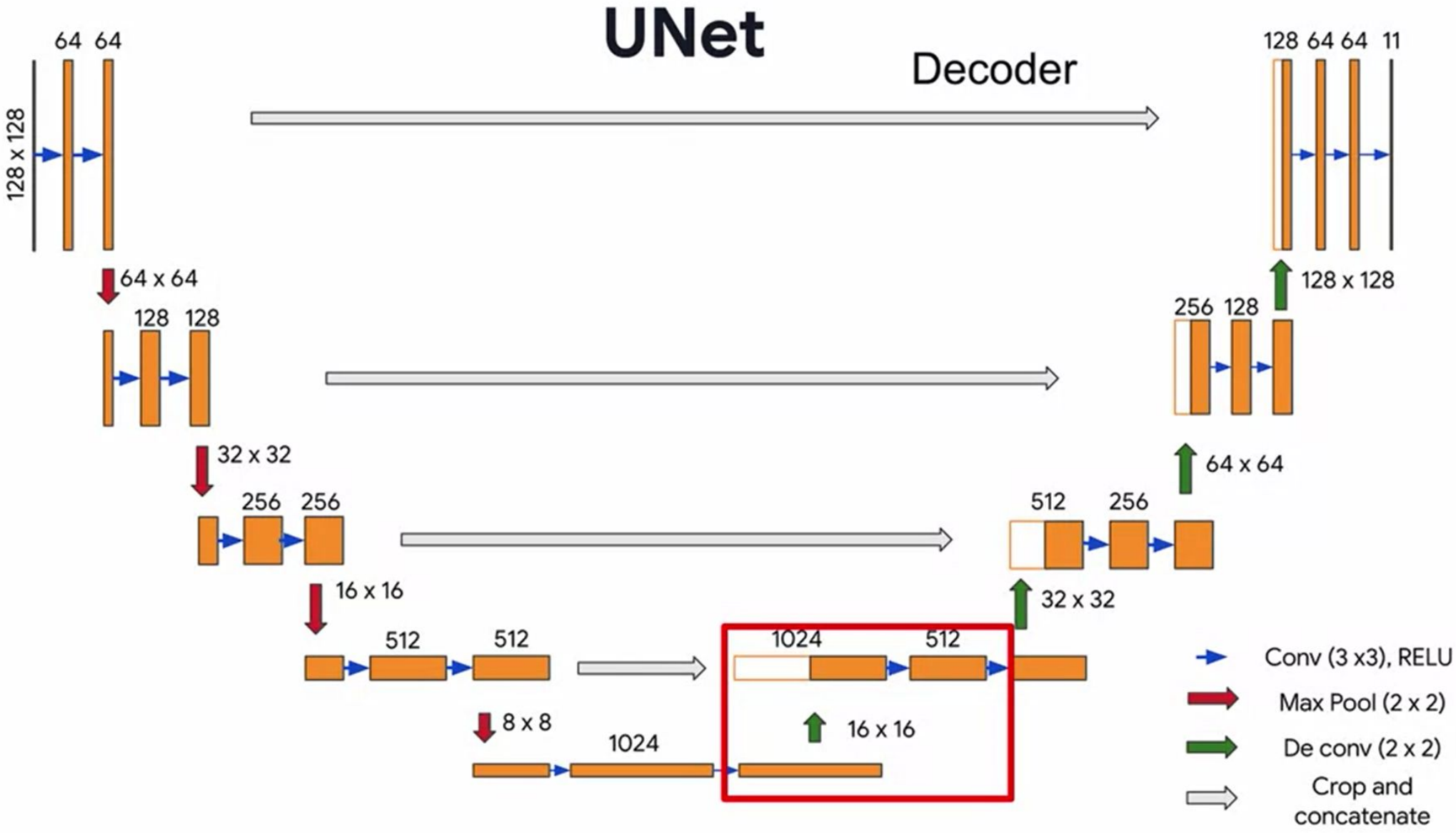
Semantic Segmentation Model: Unet



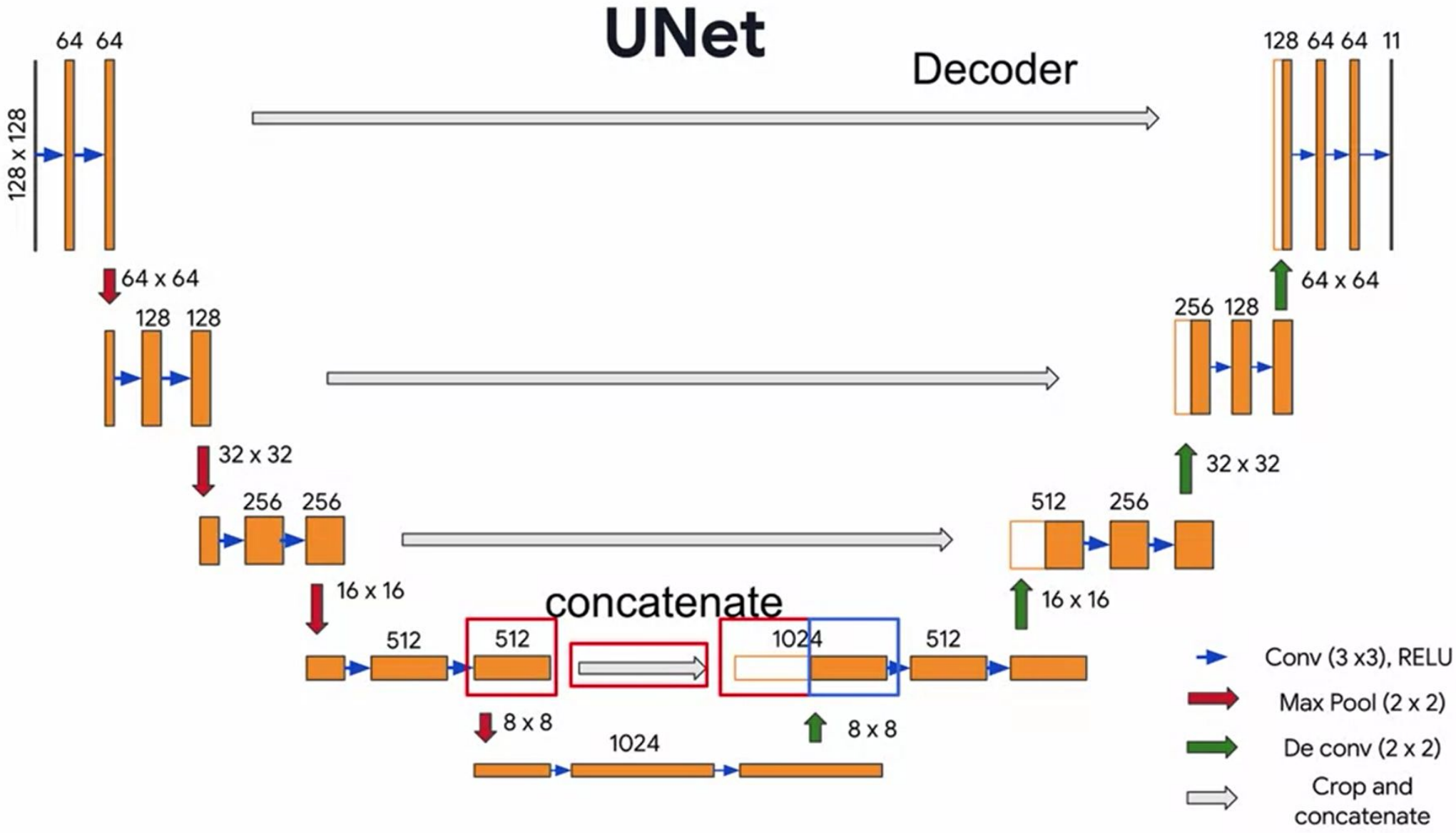
Semantic Segmentation Model: Unet



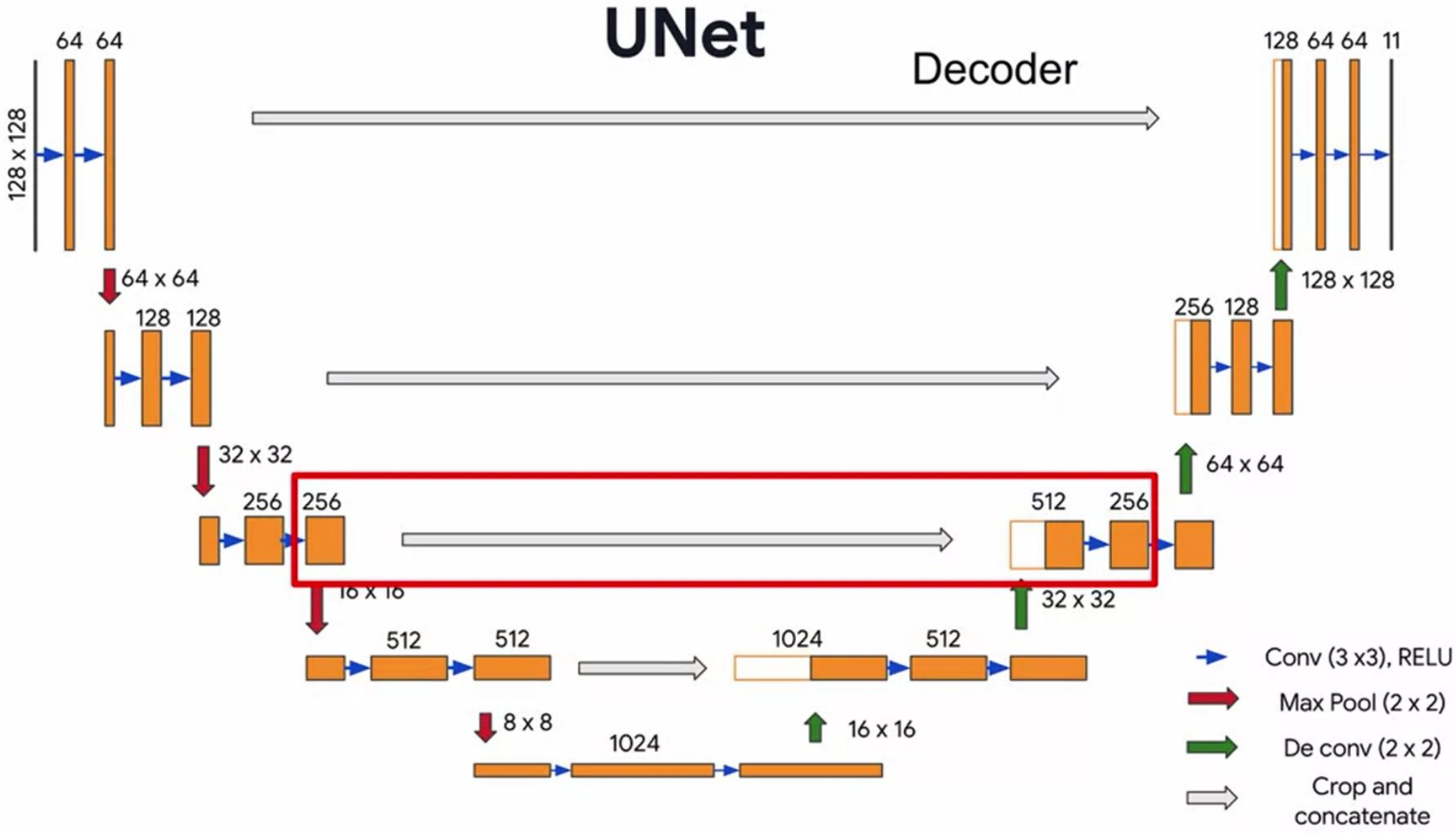
Semantic Segmentation Model: Unet



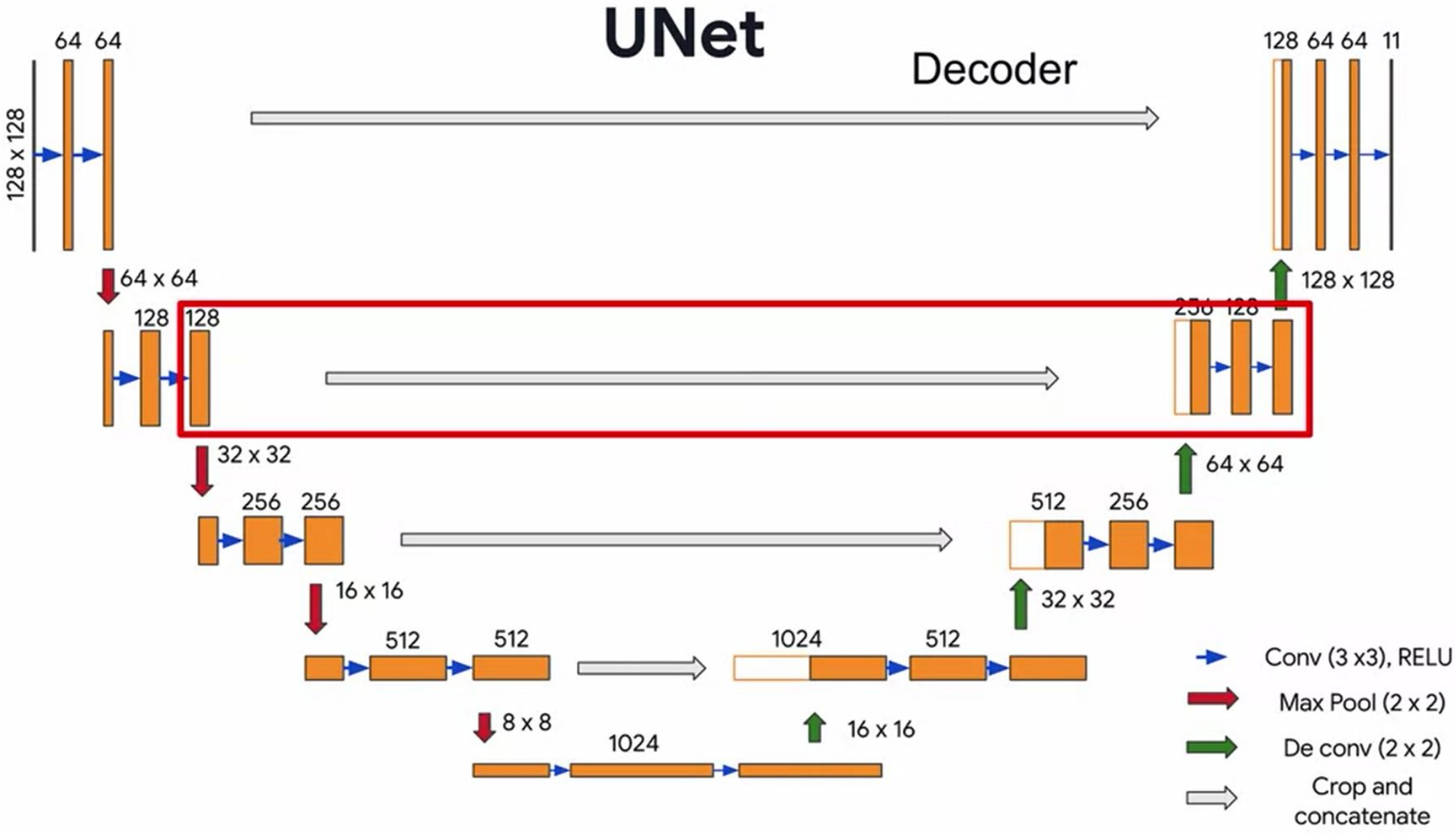
Semantic Segmentation Model: Unet



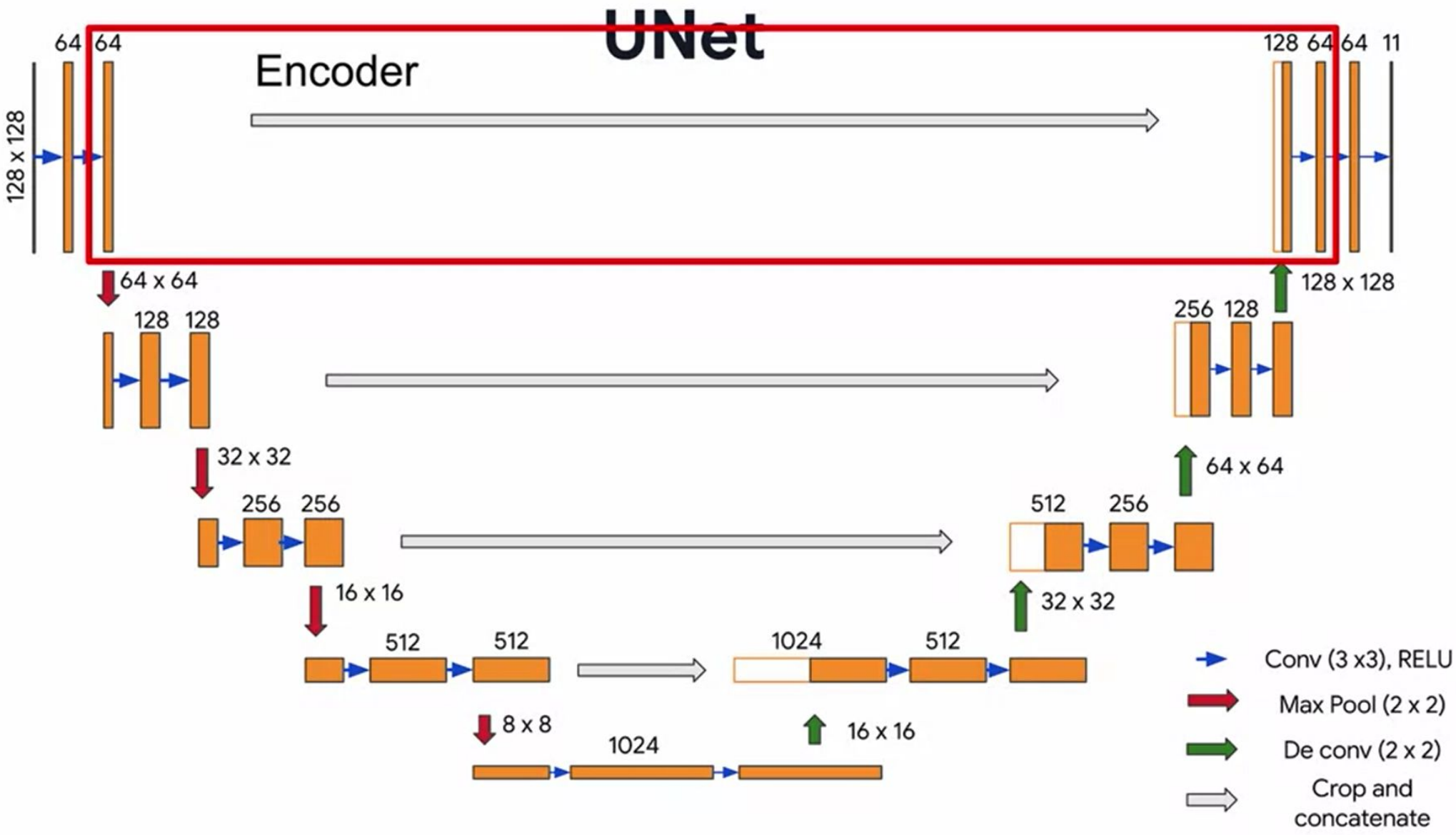
Semantic Segmentation Model: Unet



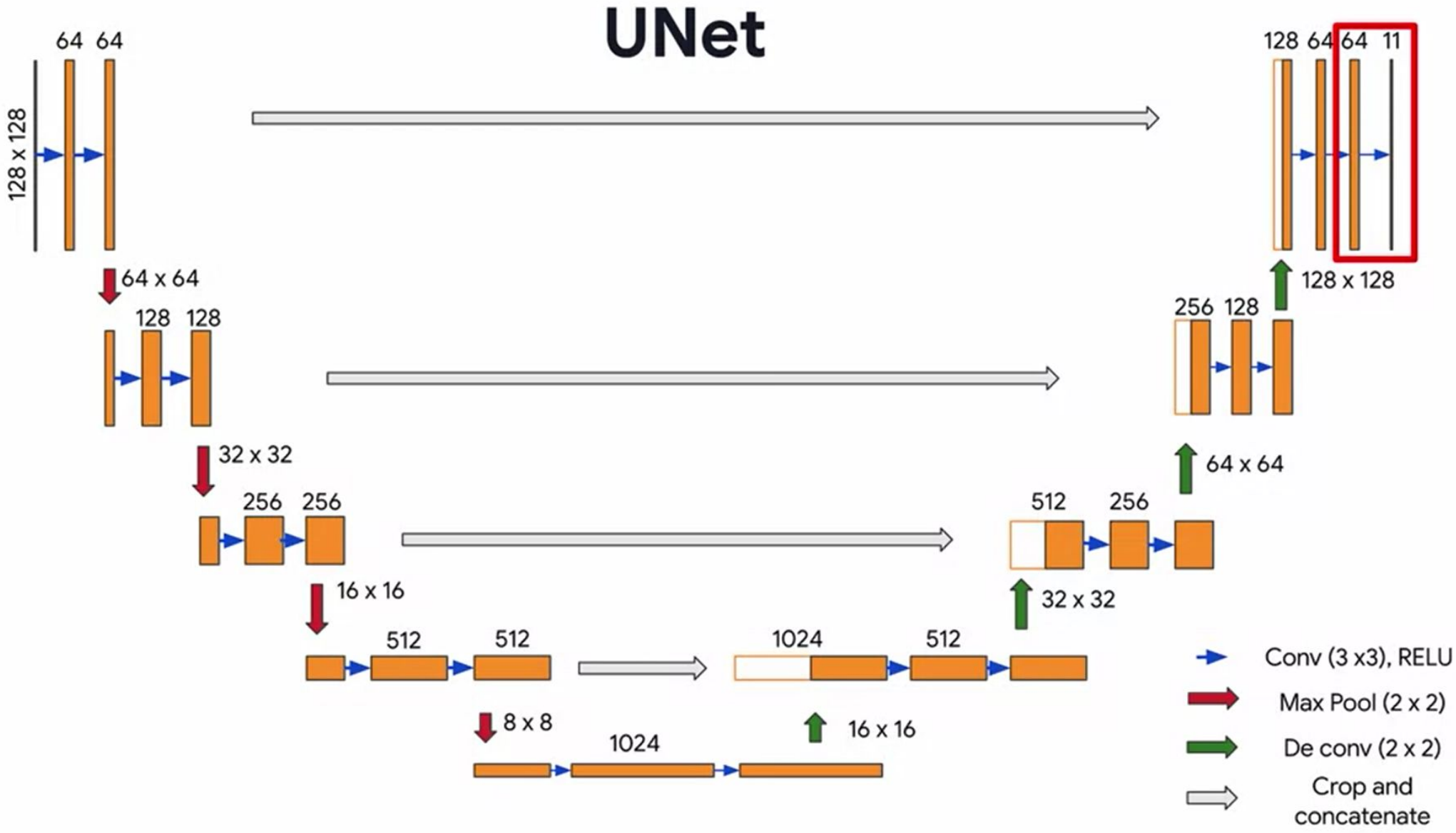
Semantic Segmentation Model: Unet



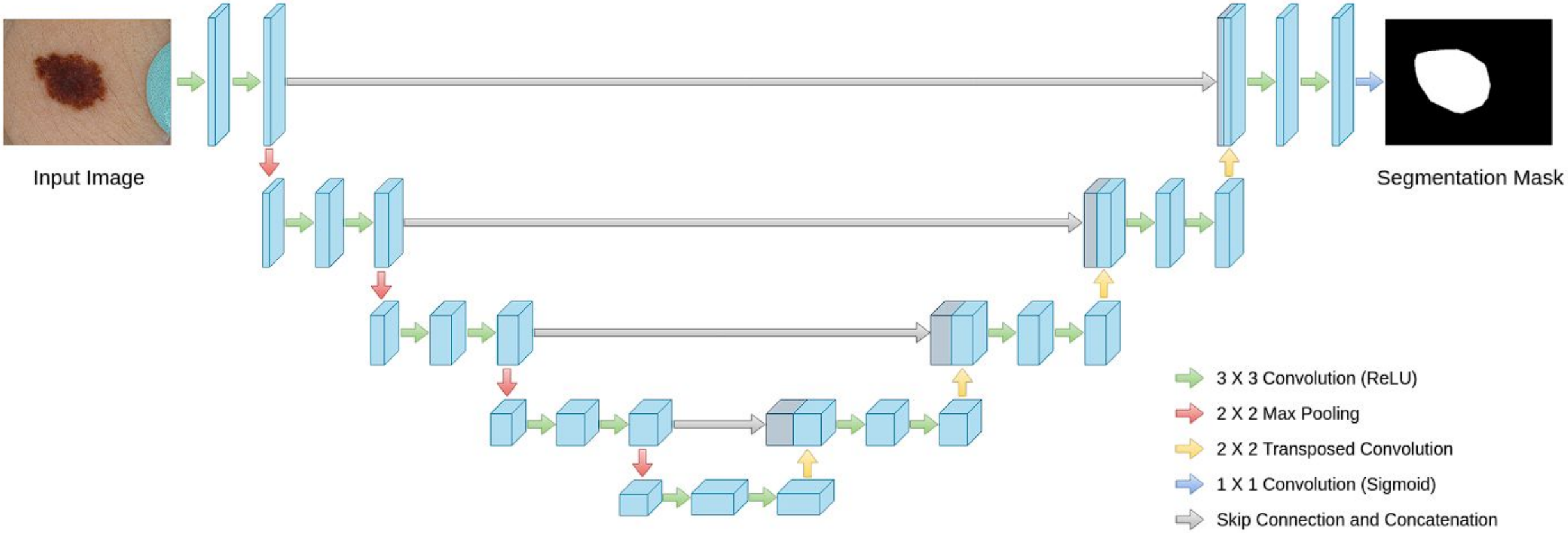
Semantic Segmentation Model: Unet



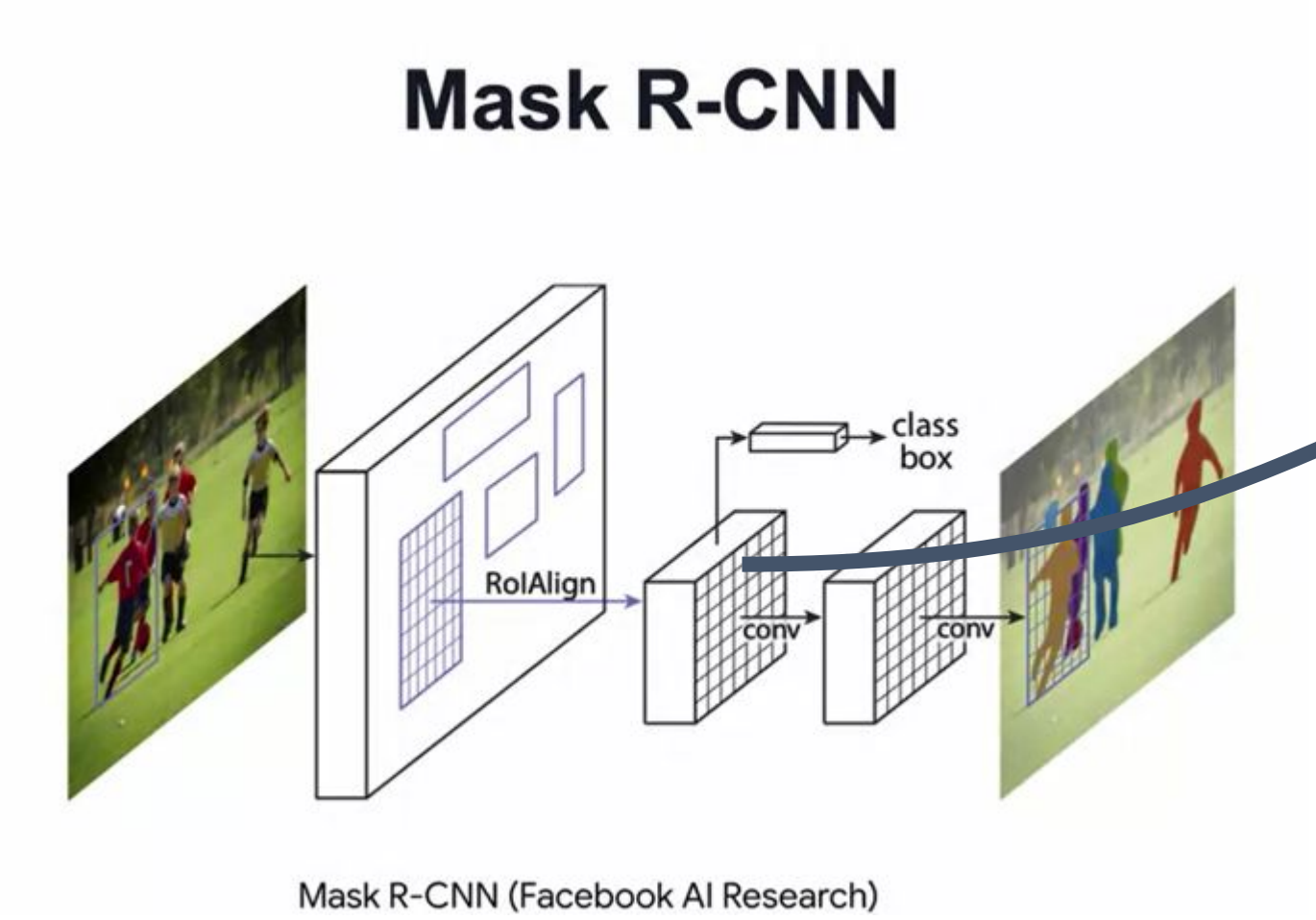
Semantic Segmentation Model: Unet

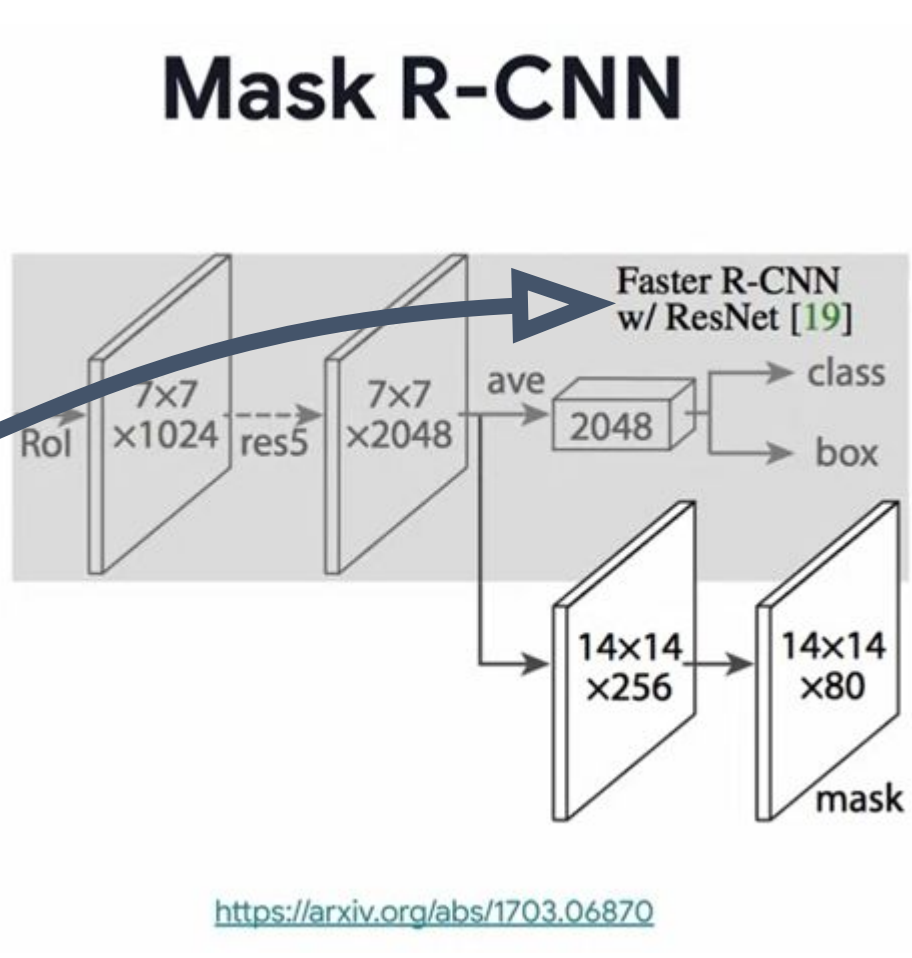


Semantic Segmentation Model: Unet

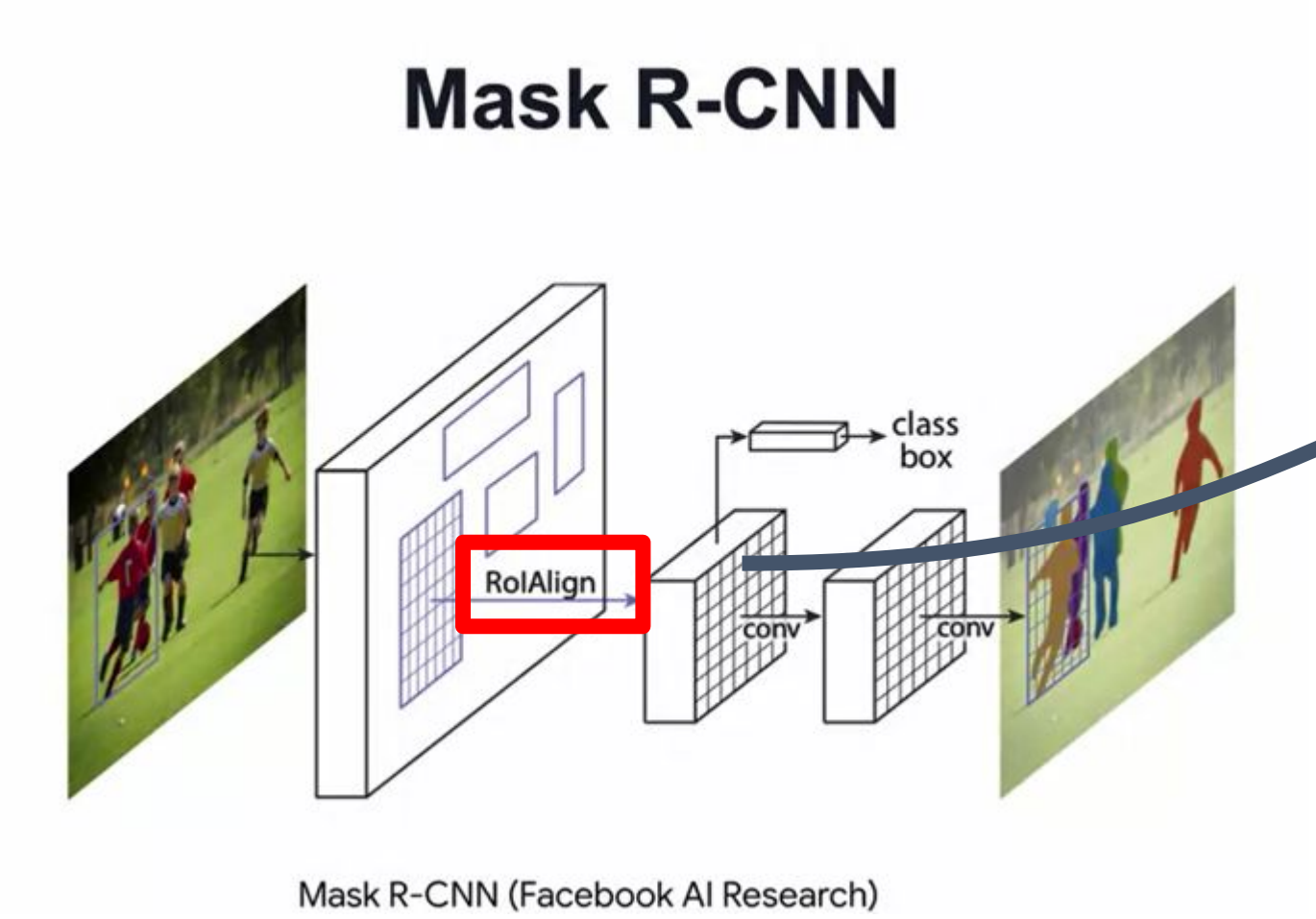


Segmentation model: Mark R-CNN

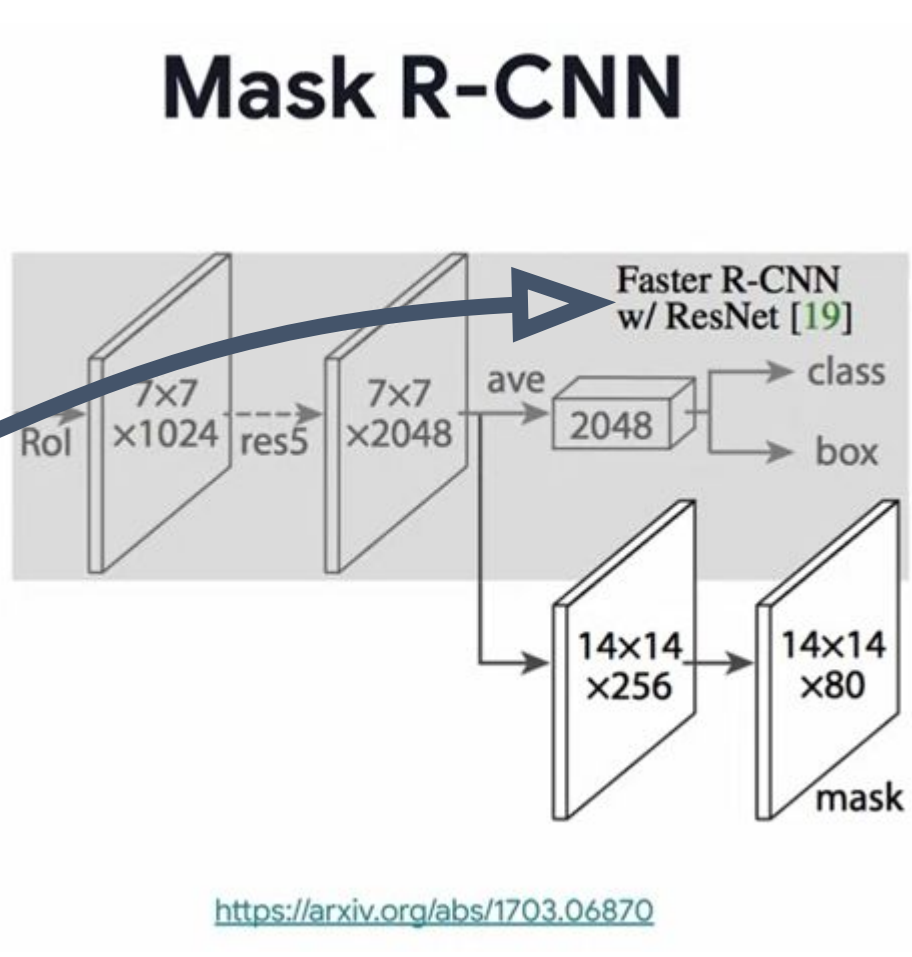




Segmentation model: Mark R-CNN

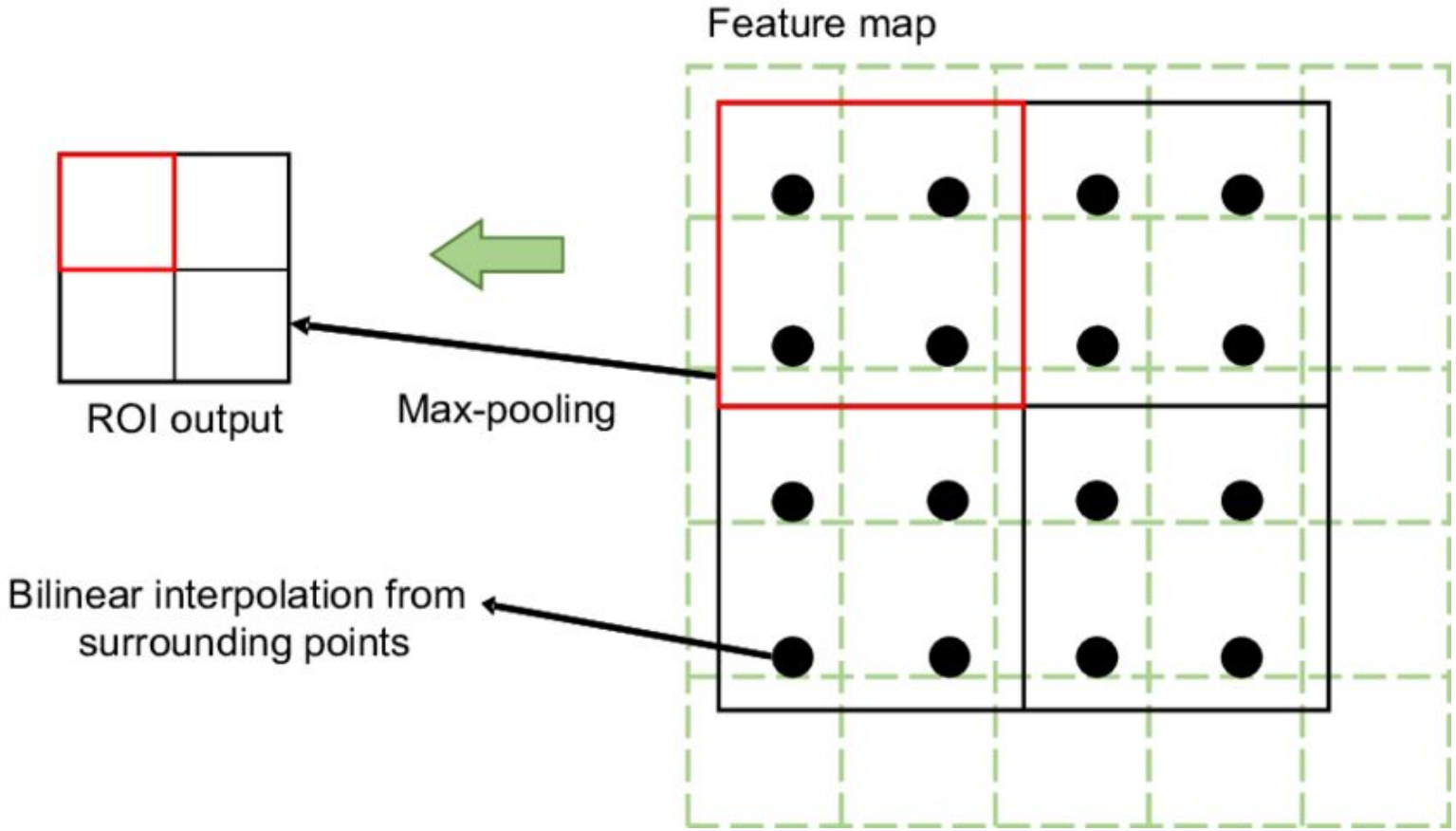


Mask R-CNN (Facebook AI Research)

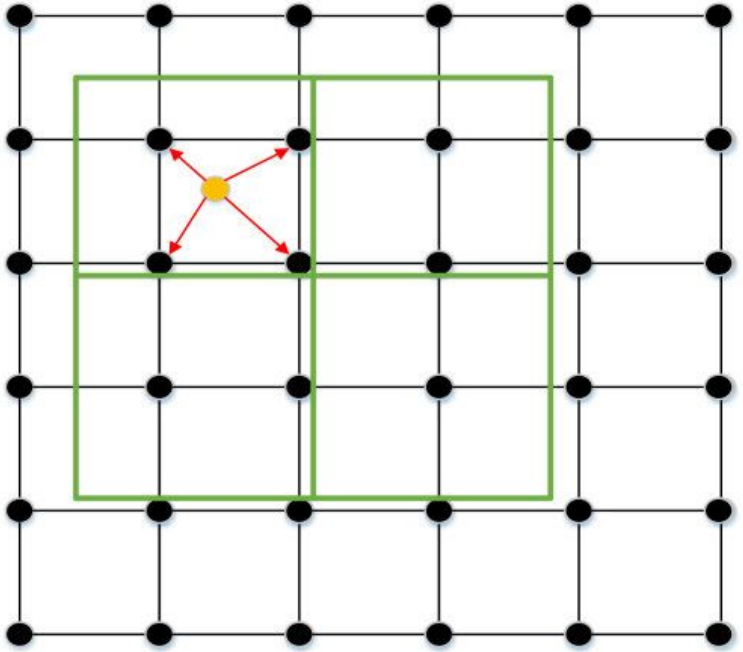


https://arxiv.org/abs/1703.06870

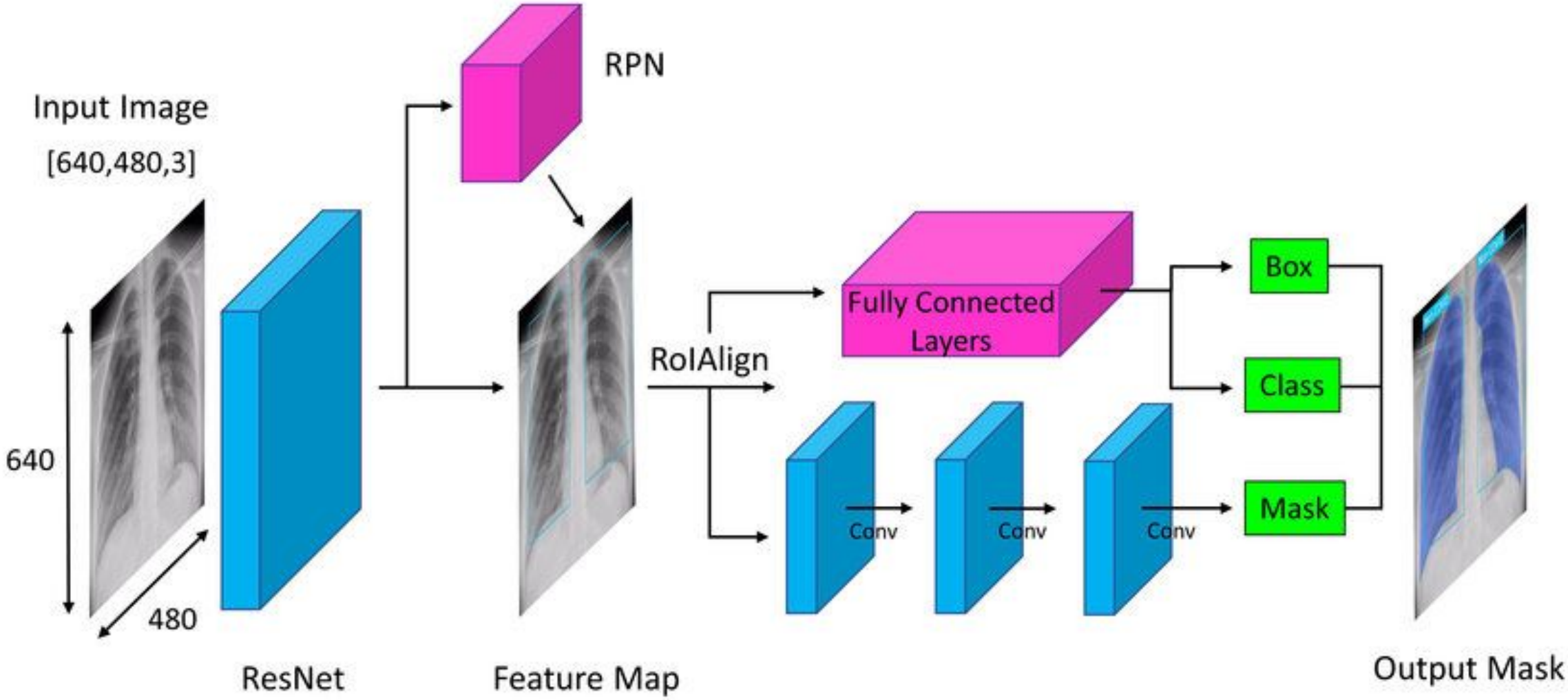
Segmentation model: Mark R-CNN



RoI Align



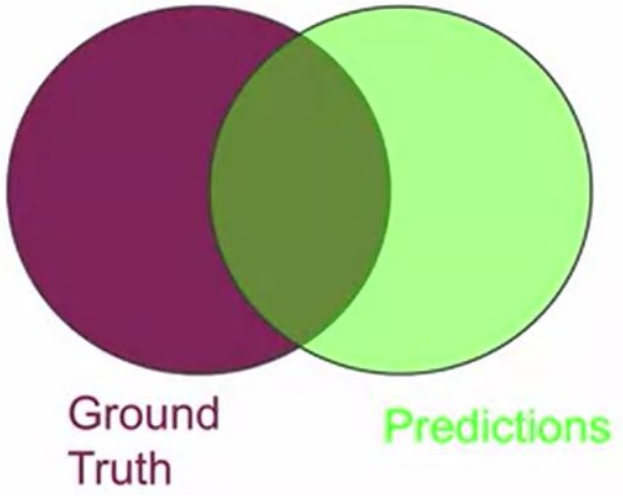
Segmentation model: Mark R-CNN



Segmentation evaluation metrics idea

Area of Overlap

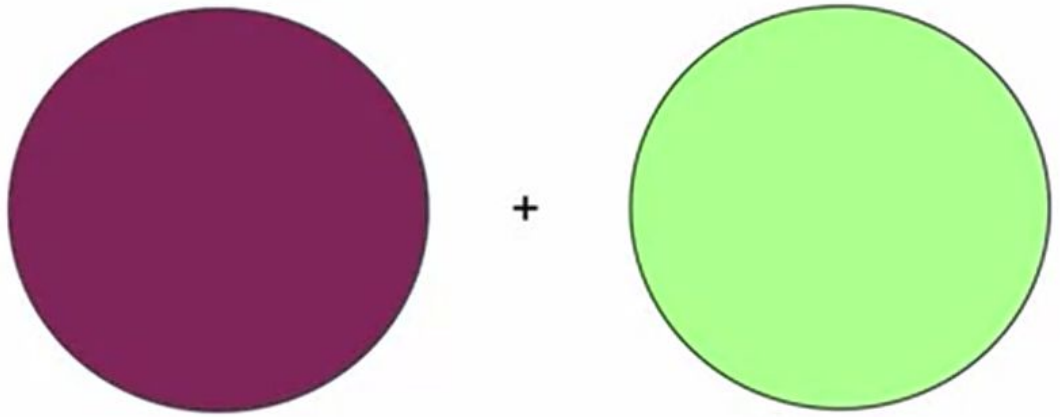
Area of Overlap = sum(True Positives)



Segmentation evaluation metrics idea

Combined Area

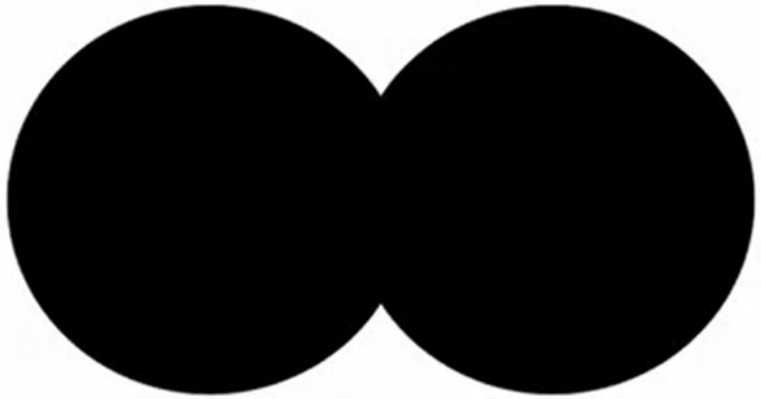
Combined Area = Total Pixels in predicted segmentation mask + Total Pixels in True Segmentation mask



Segmentation evaluation metrics idea

Area of Union

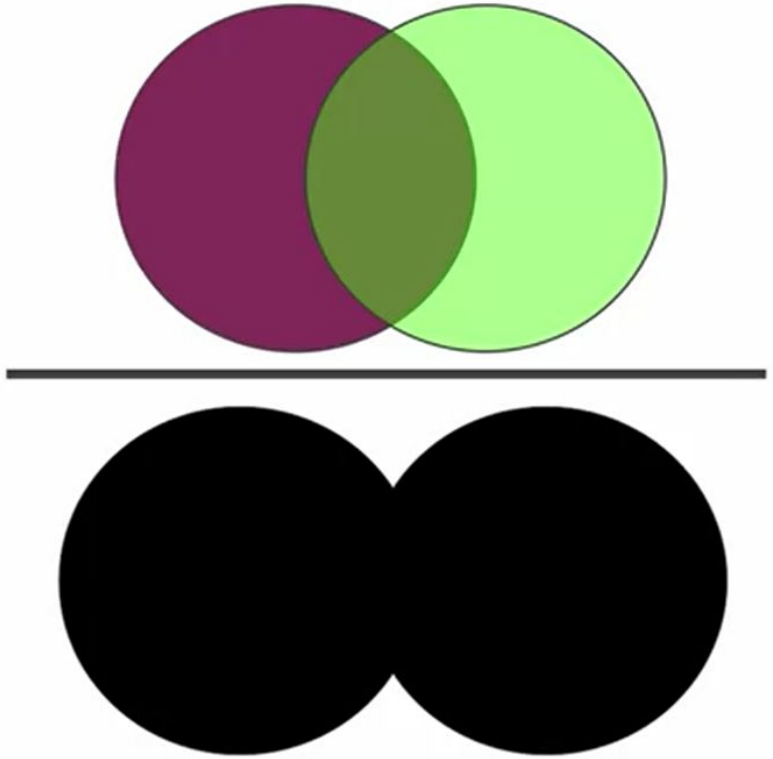
Area of Union = Total Pixels in predicted segmentation mask + Total Pixels in True Segmentation mask - Area of Overlap



Segmentation evaluation metrics idea

Intersection Over Union

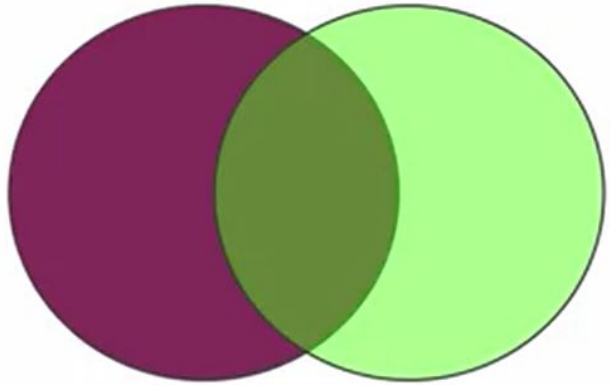
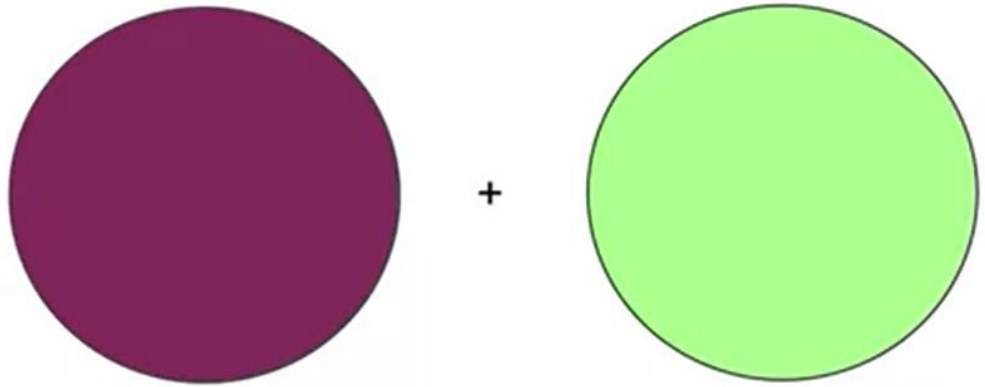
$$IOU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Segmentation evaluation metrics idea

Dice Score

$$\text{Dice Score} = 2 \times \frac{\text{Area of Overlap}}{\text{Combined Area}}$$

$$2 \times \frac{\text{  }}{\text{  }}$$