

# U-Net Models

3 models for the segmentation task:

- **Classic U-Net:** *baseline U-Net model architecture*

# U-Net Models

3 models for the segmentation task:

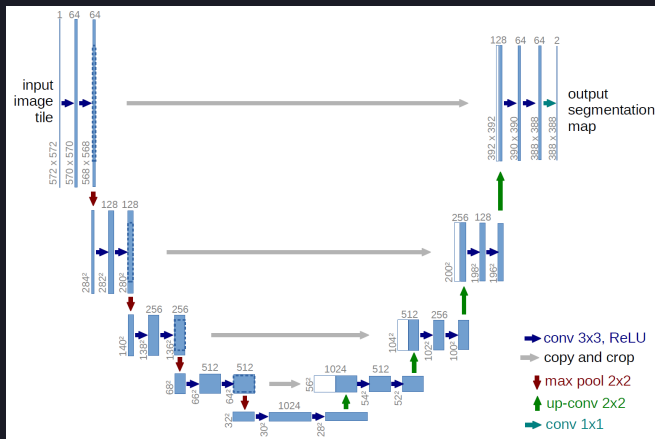
- **Classic U-Net:** *baseline U-Net model architecture*
- **Improved U-Net:** *small improvements, less parameters*

# U-Net Models

3 models for the segmentation task:

- **Classic U-Net:** *baseline U-Net model architecture*
- **Improved U-Net:** *small improvements, less parameters*
- **Attention U-Net:** *attention mechanism added*

## Classic U-Net

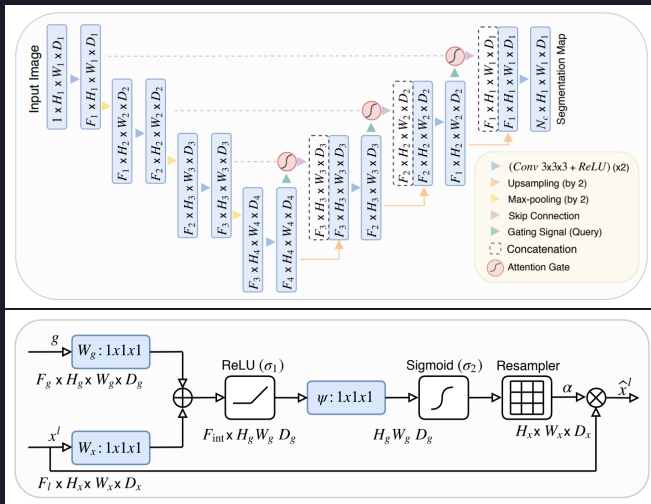


# Improved U-Net

Small improvements from previous → to reduce  $n^\circ$  of parameters and improve performance:

- **Separable Convolutions:** depthwise + pointwise convolutions
- **Batch Normalization:** to improve training and generalization
- **Larger Kernel Size:**  $7 \times 7$  kernels instead of  $3 \times 3$
- **Inverse Bottleneck:** expands + compresses channels
- **Additive Skip Connections:** instead of concatenated ones

## Attention U-Net



# Training Details

U-Net Models training parameters:

- **epochs:** 20
- **Optimizer:** Adam (with weight decay ( $1 \times 10^{-2}$ ))
- **Scheduler:** Exponential Decay ( $\gamma = 0.9$ )
- **Loss Function:** BCE with Logits Loss
- **learning rate:**  $2 \times 10^{-3}$
- **batch size:** 32 (both training and validation)
- **image size:**  $240 \times 240$
- **first encoder filters:** 32

# Performance Assessment

$$\text{Dice} = \frac{2 \times |X \cap Y|}{|X| + |Y|}$$

*Dice Coefficient*

*"overlap" metric*

$$\text{Precision} = \frac{TP}{TP + FP}$$

*Precision*

*prediction quality*

$$\text{Recall} = \frac{TP}{TP + FN}$$

*Recall*

*prediction quantity*



# Visualizing Attention Maps

