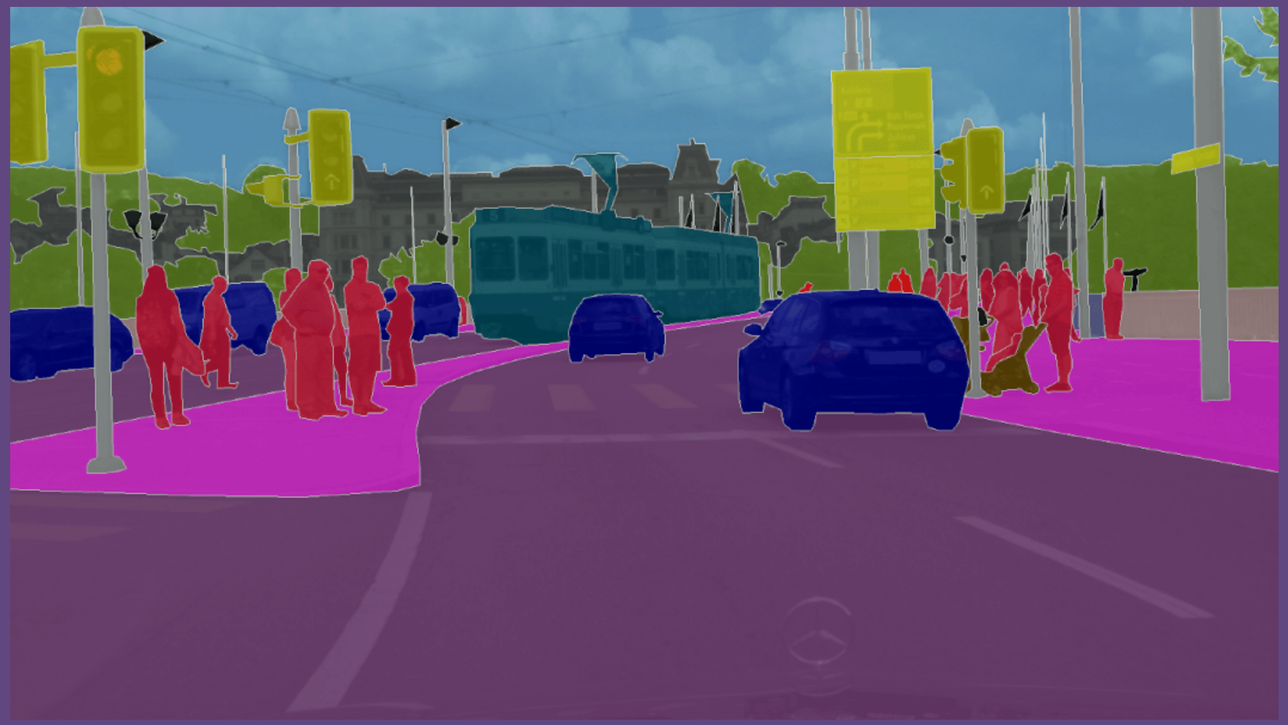
# Unet for image segmentation

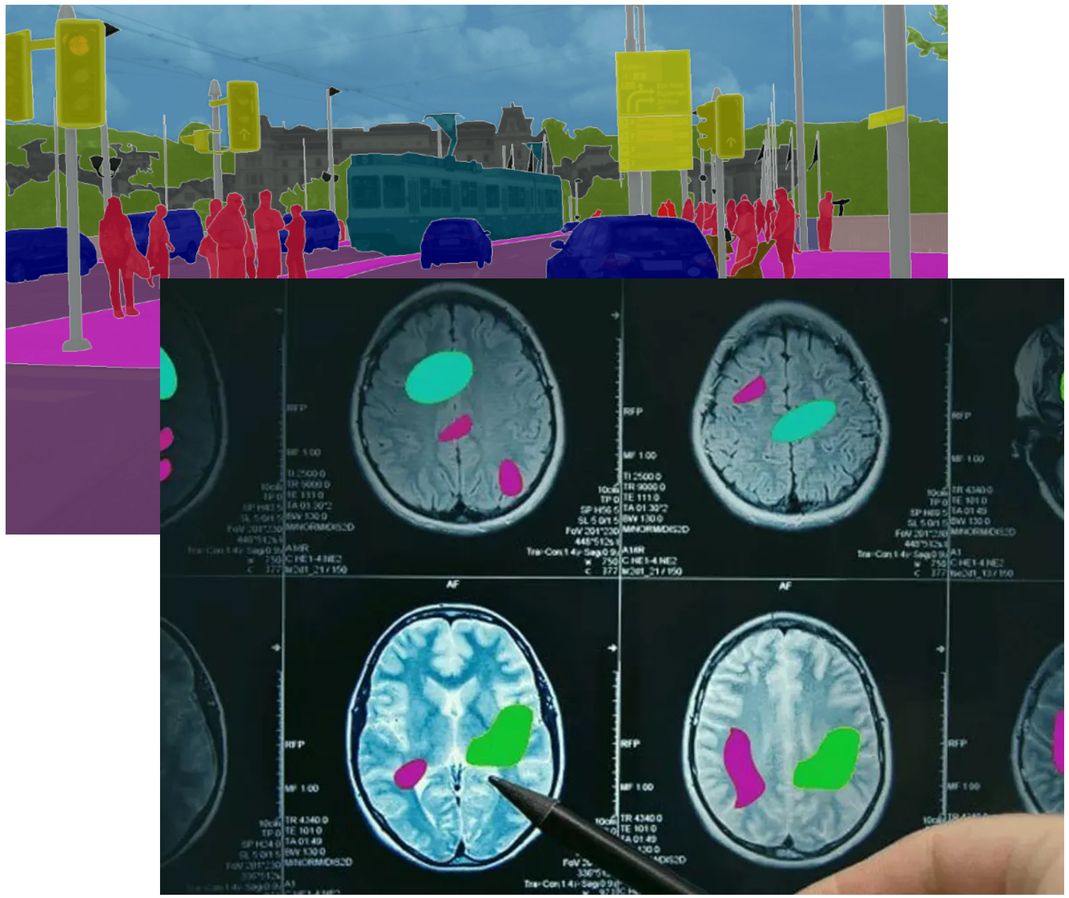


**Practical Deep Learning for Science**
**06 June, 2024**

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

*- Nilotpal*

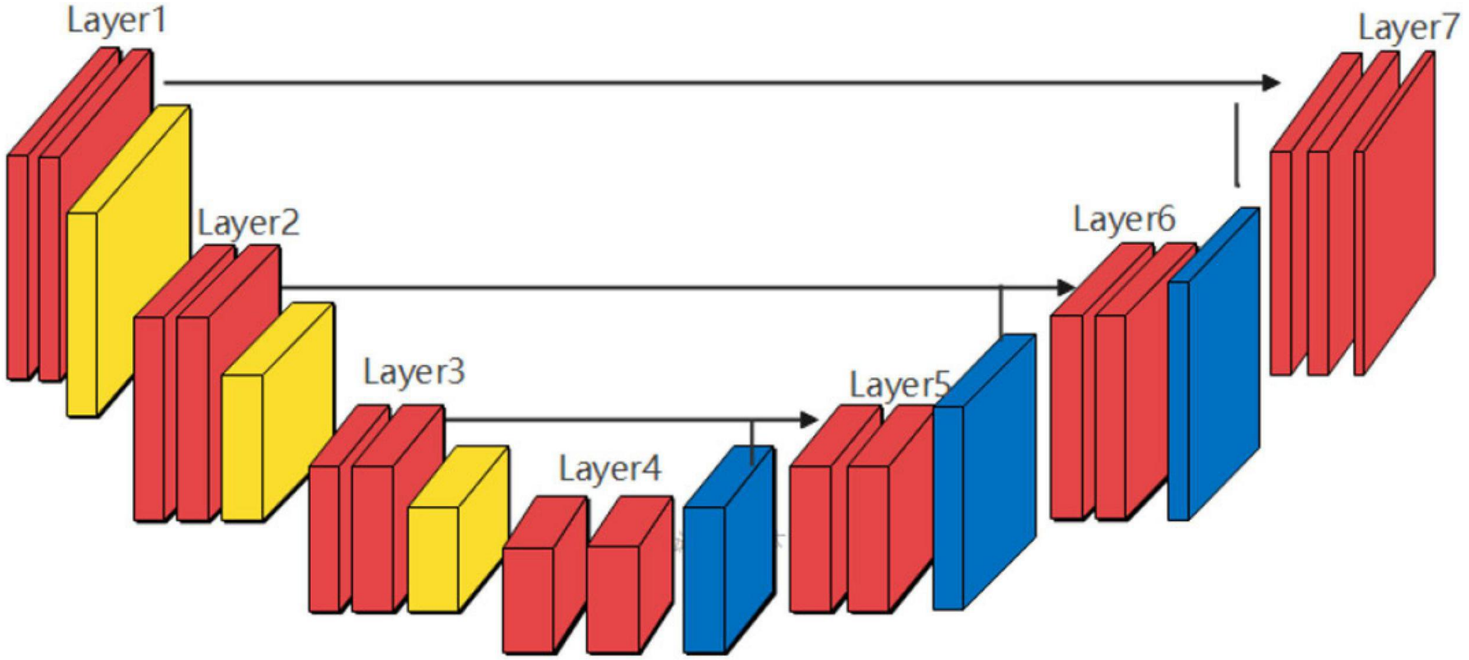
# Unet: Two main usage



segmentation


Diffusion (next lecture)

# Unet



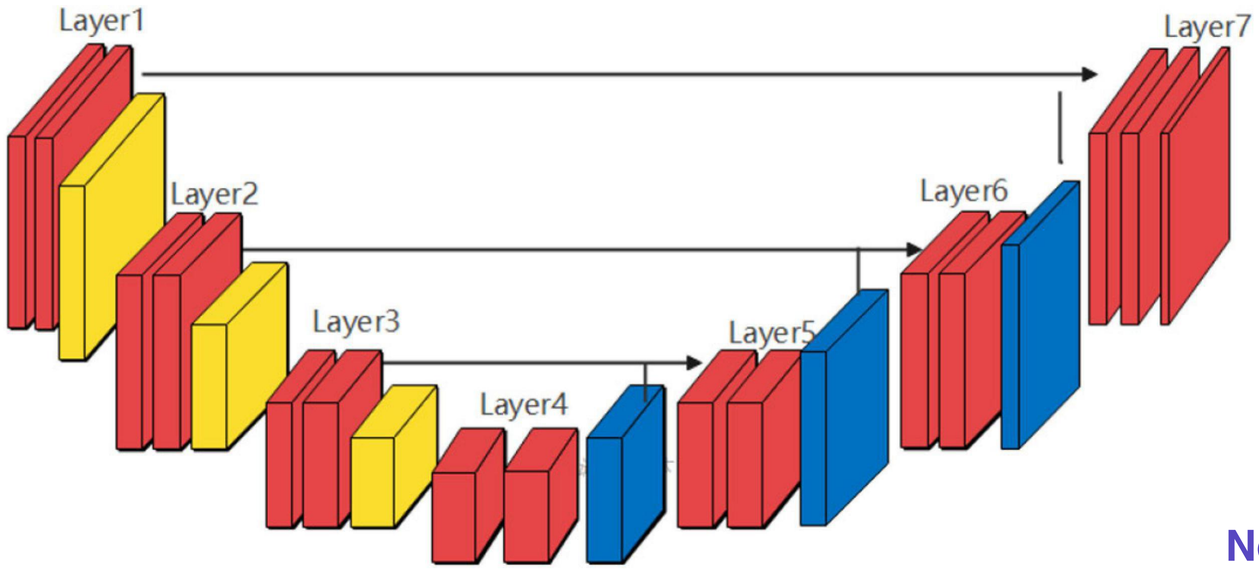

- U-shaped
- Very similar to the VAE
- Differences
  - Input is not the target
  - We won't be sampling. (No KLD in the middle)
  - Skip connections
    - Helps upsampling

# Upsampling

- For VAE, we used transposeConv2D
- Today we'll try (linear upsampling + conv2D)
  - Conceptually easier
- Linear upsampling => insert new pixel between already existing pixel
  - The new pixel value can be computed using different types of interpolation
  - Eg. Bilinear interpolation: new pixel value is the weighted average of the surrounding pixels

# DoubleConv

- (conv2D + batchNorm + ReLU) x2 + upsampling/downsampling




(conv2D + BN + ReLU)

Downsampling

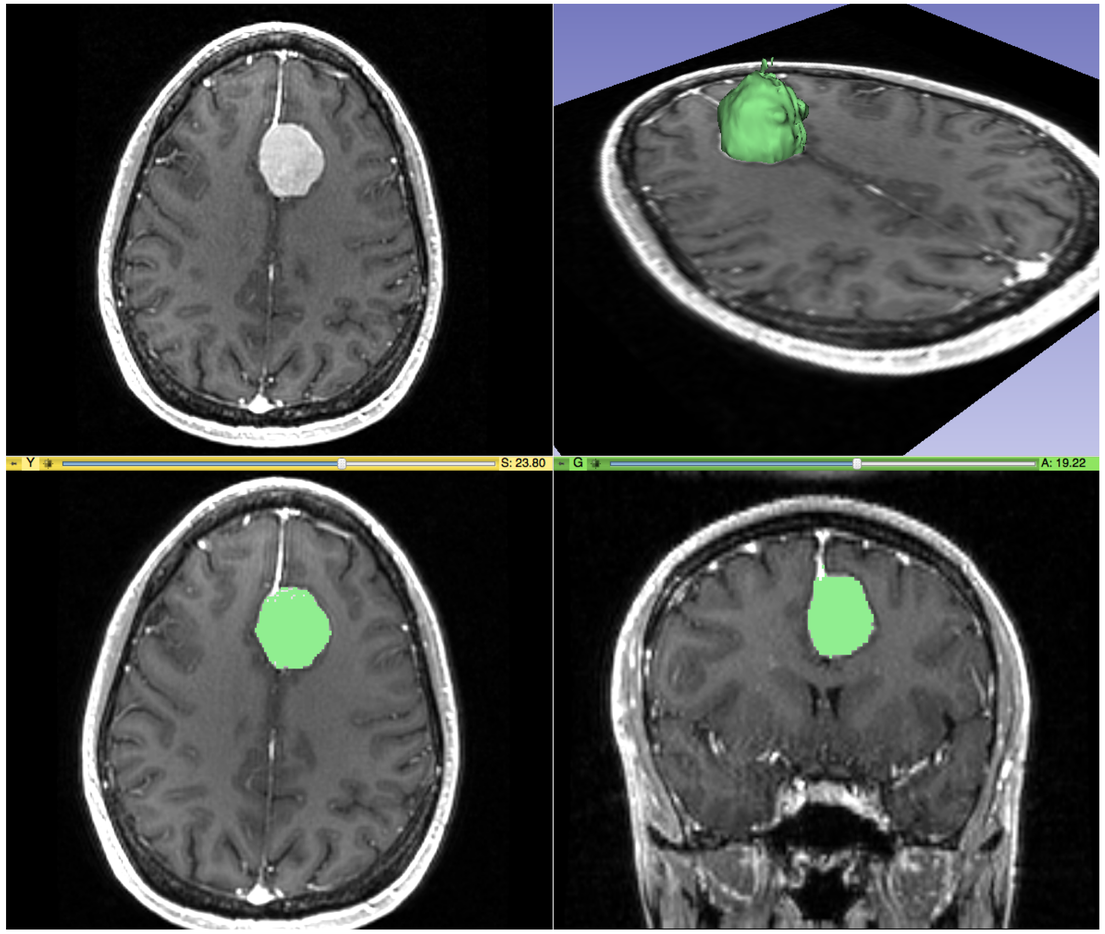
MaxPool

Upsampling

Linear upsampling

**No learnable parameter in up/down sampling**

# Dice loss: why?




- Classification on each pixel
- Way more background (0) pixels than signal (1)
- Class imbalance!
- Network will biased to predict 0

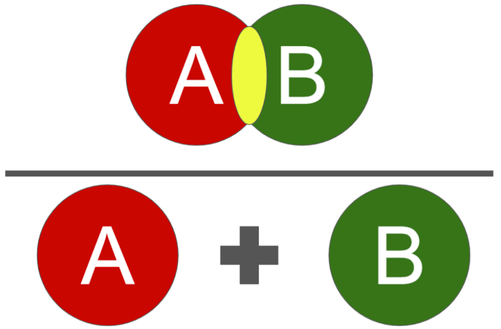
https://mphy0026.readthedocs.io/en/latest/segmentation/segmentation_methods.html

# Dice loss

Traget mask


Pred mask


Dice coeff = $2 \times \frac{A \cap B}{A + B}$



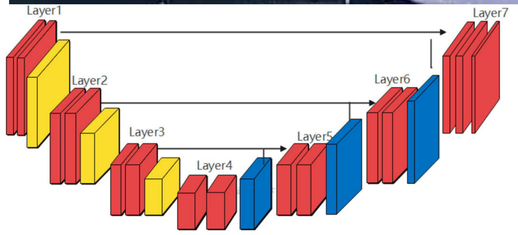

**A and B have zero overlap => 0 (minimum)**

**A and B are same => 1 (maximum)**

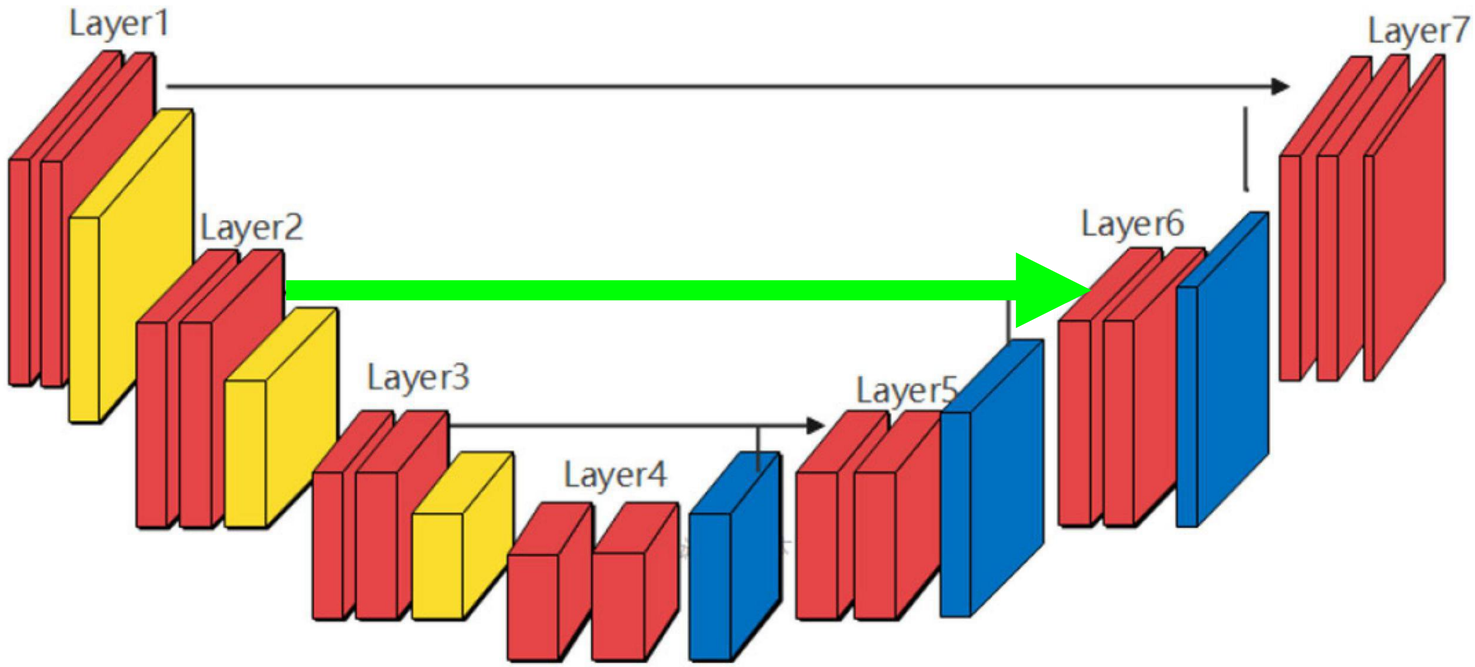
**Dice loss = 1 - dice coefficient**

https://pycad.co/the-difference-between-dice-and-dice-loss/


Transformer
Transformer
Transformer
Transformer
Transformer

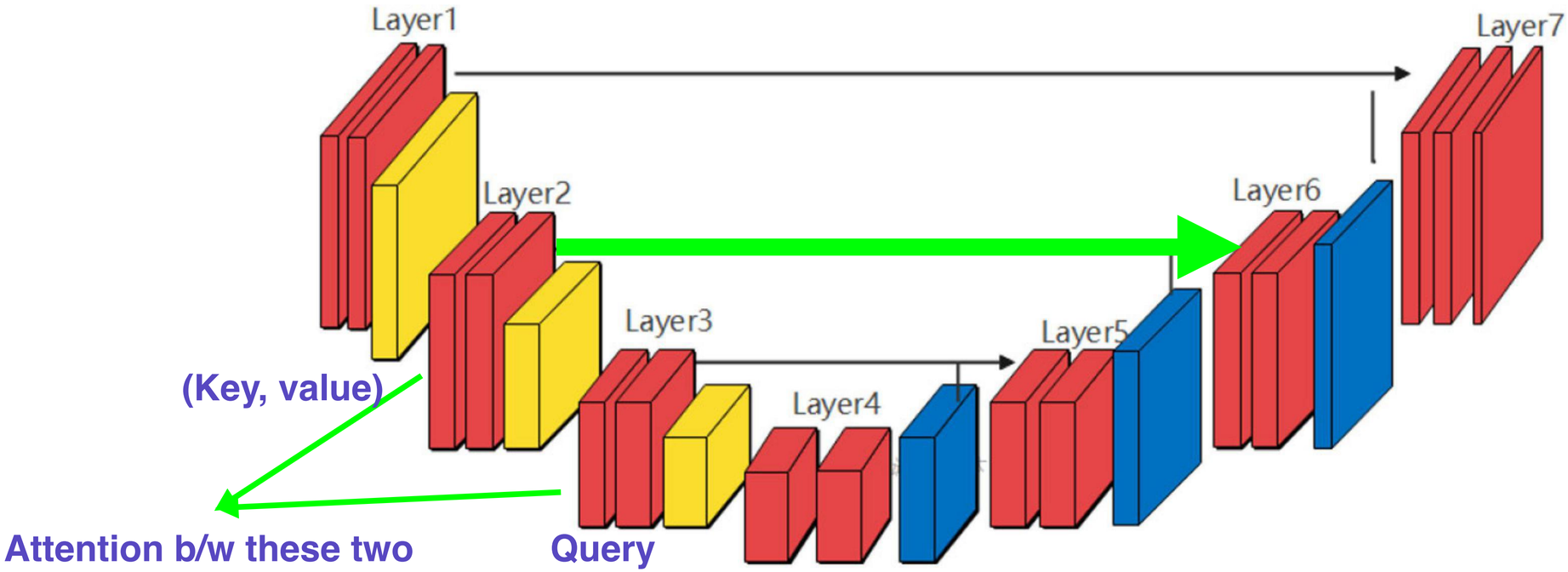

Layer1
Layer2
Layer3
Layer4
Layer5
Layer6
Layer7

# Attention Unet




- Skip connection helps with upsampling
- But also can bring redundant information
  - Counterproductive considering we wanted to extract important information

# Attention Unet




- Skipconnect only the information that was useful when downsampling

# COCO data format

- Common Objects in COntext
- Pretty common in object detection, instance segmentation...
- Contains - images, annotations, categories