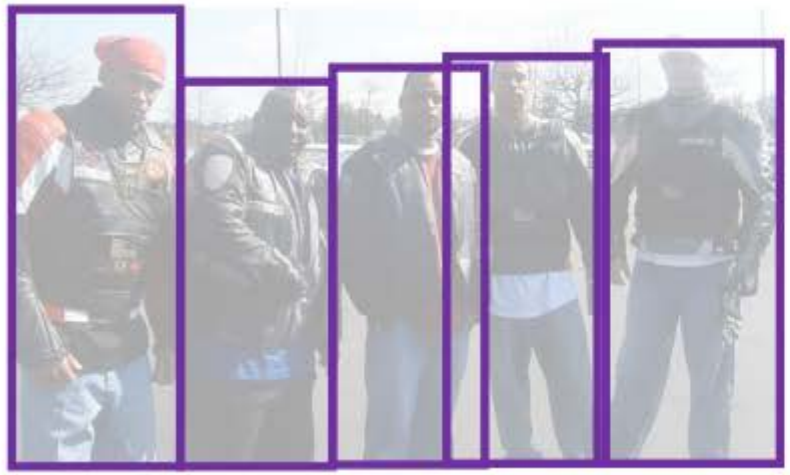
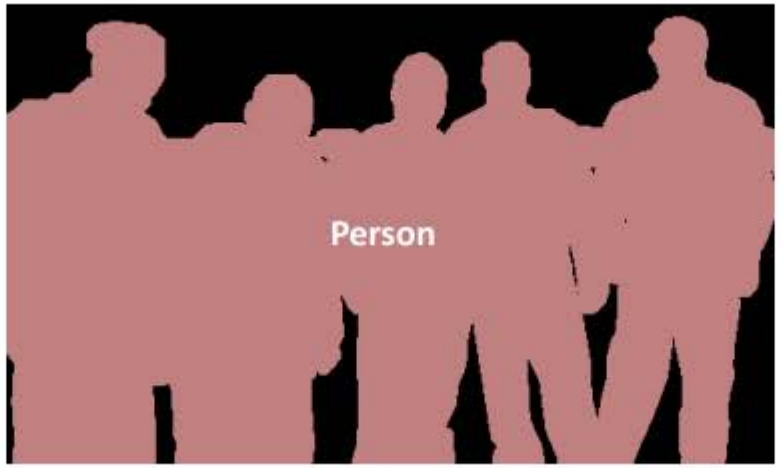
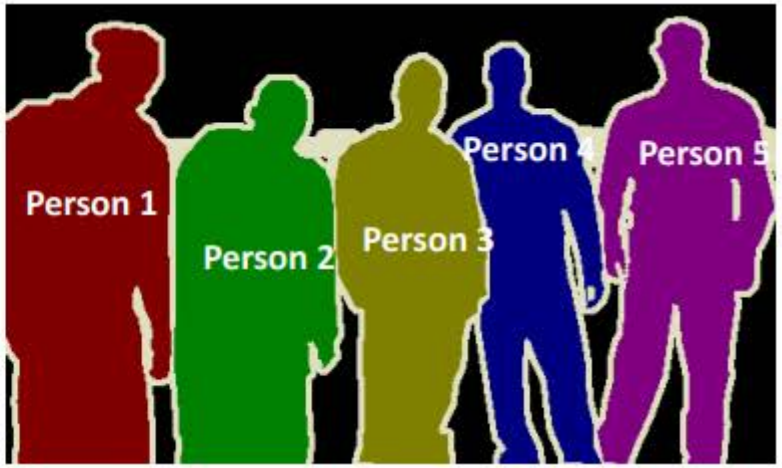






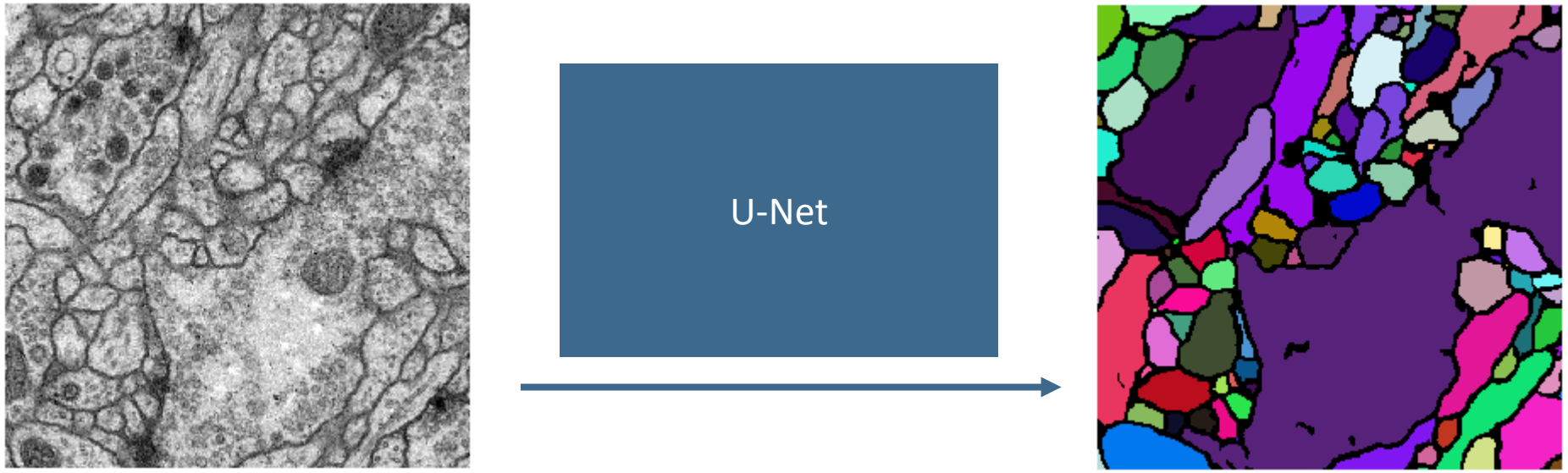
U-Net: Convolution Networks

Semantic Segmentation

Philip Wolschon 08.01.2020

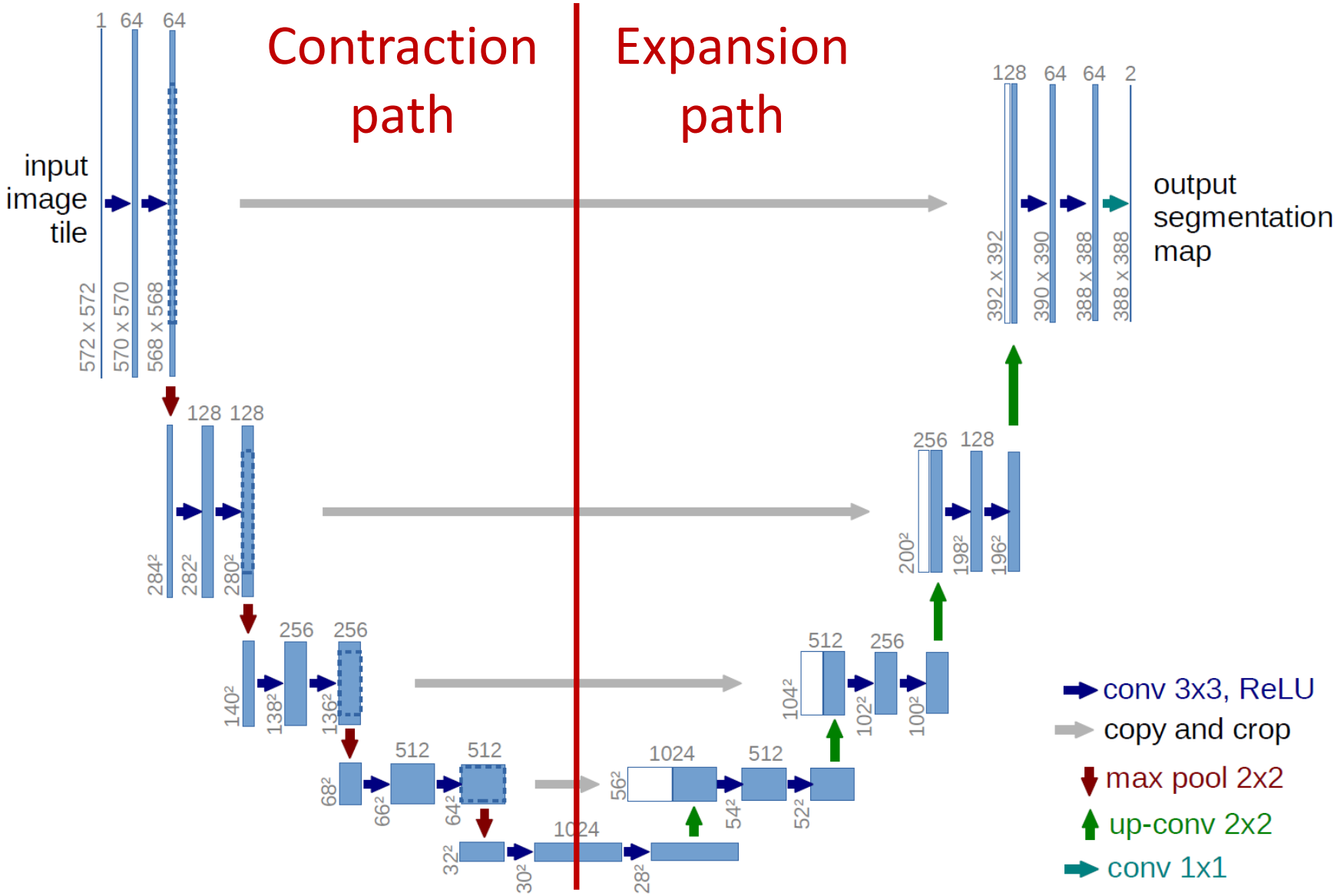
Table of contents

Definition



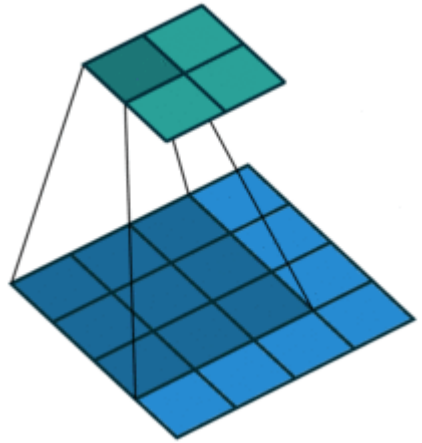
- U-Net learns semantic segmentation in an end-to-end setting
- each pixel is classified
- supervised learning

U-Net architecture

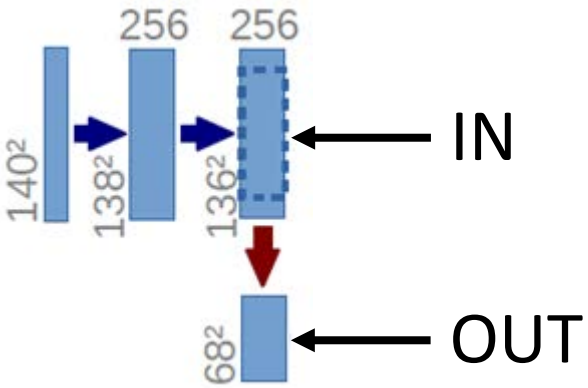


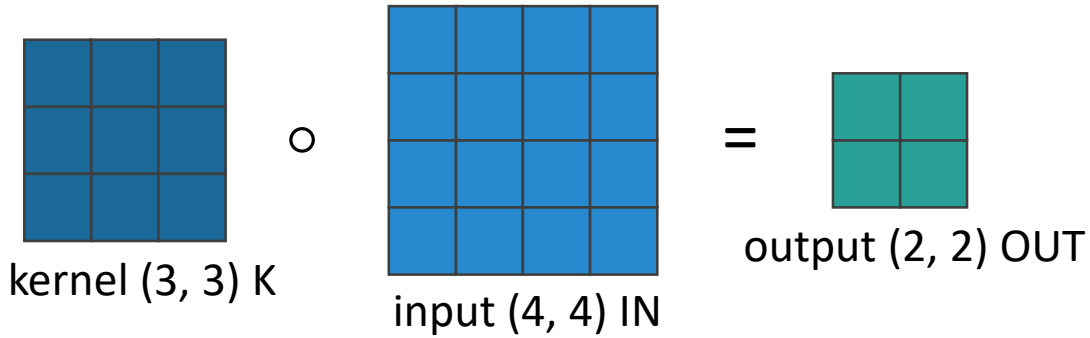
U-Net architecture

Contraction path (encoder):



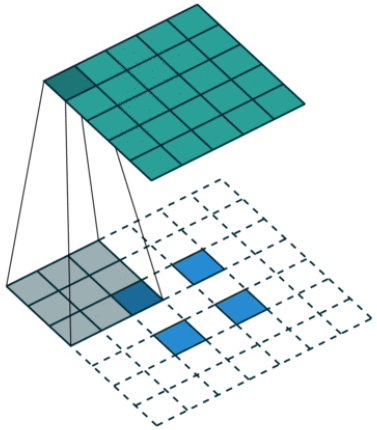
Convolution → Convolution → Max Pooling





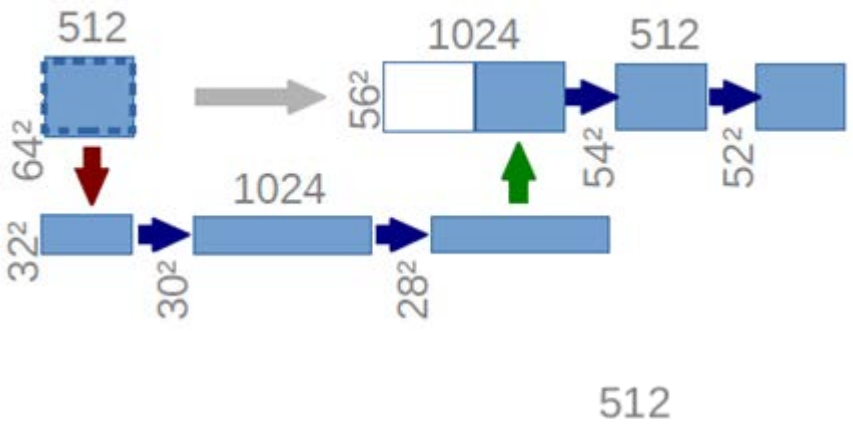
$$K \circ IN = OUT$$

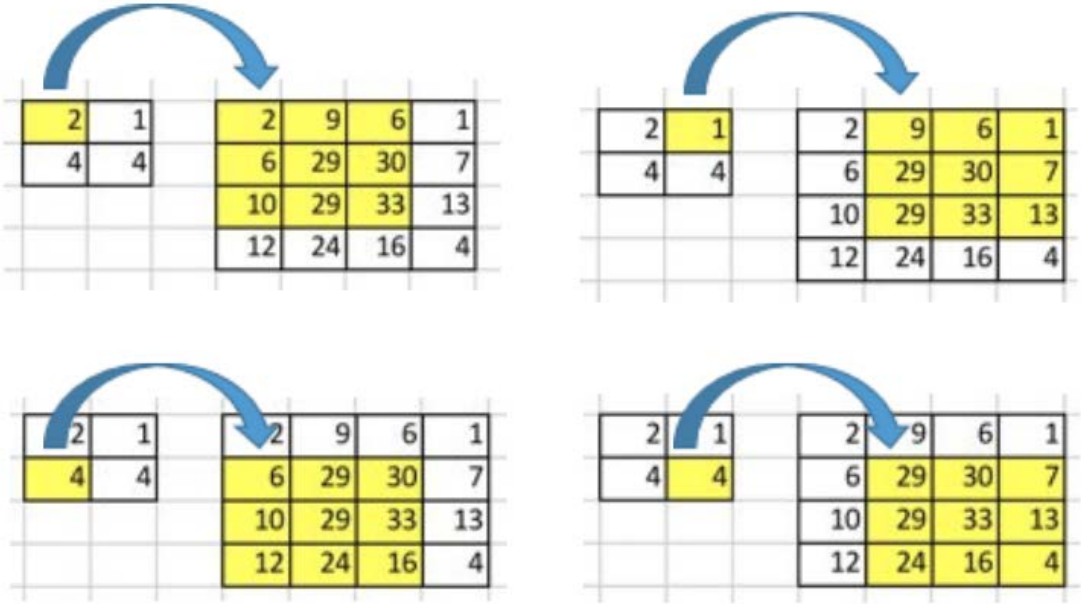
U-Net architecture



Expansion path (decoder):

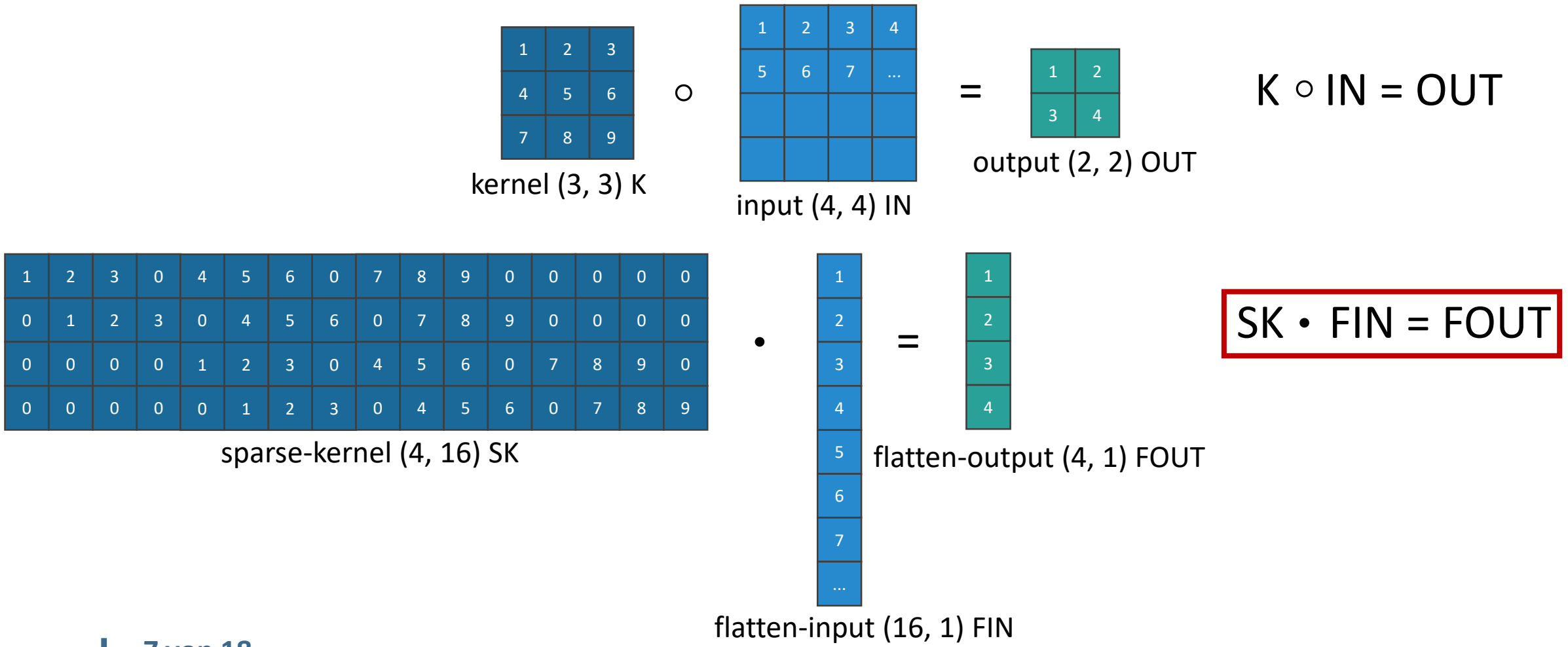
Transposed Convolution → Concatenate → Convolution → Convolution





U-Net architecture

Expansion path (decoder):



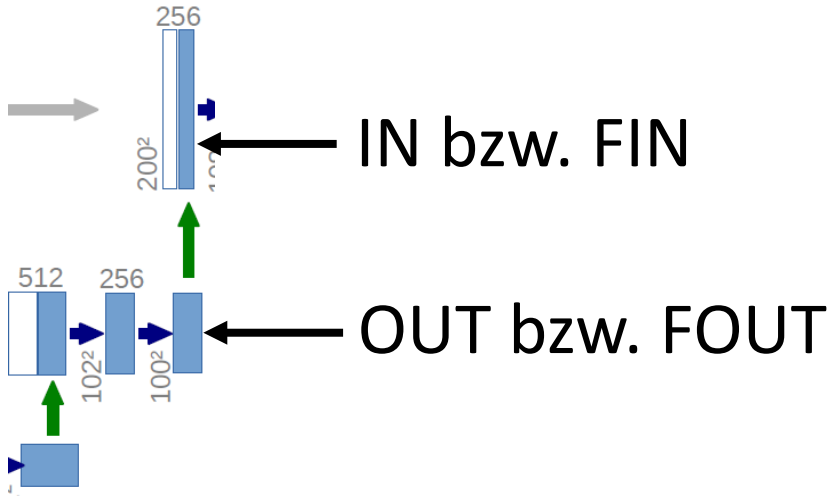
U-Net architecture

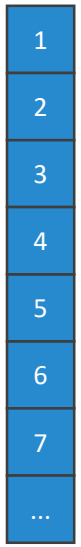
Expansion path (decoder):

$$SK \cdot FIN = FOUT$$

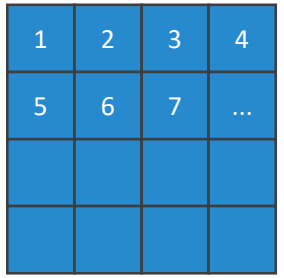
$$SK^T \cdot SK \cdot FIN = SK^T \cdot FOUT$$

$$FIN = SK^T \cdot FOUT \rightarrow IN$$



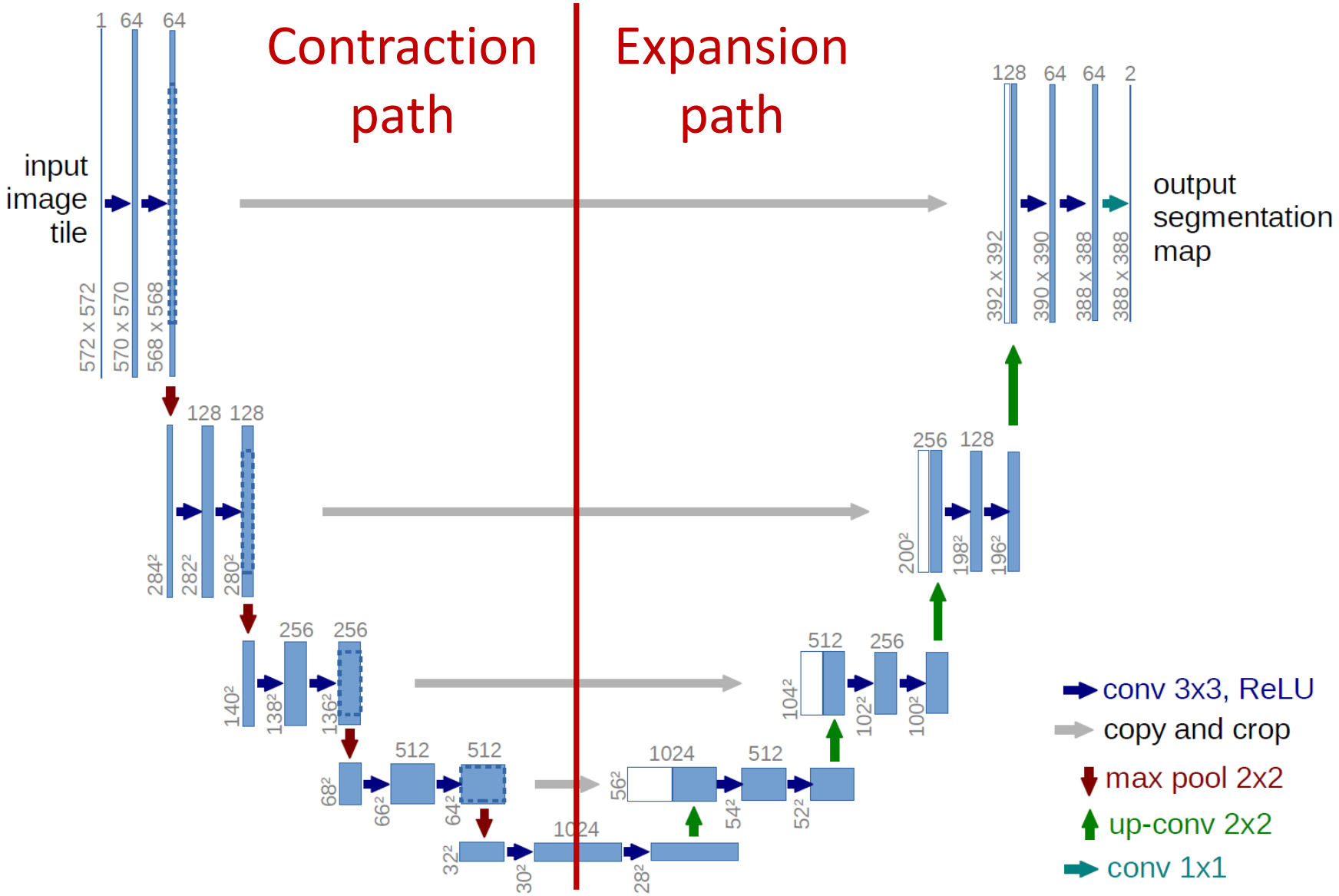


flatten-input (16, 1) FIN



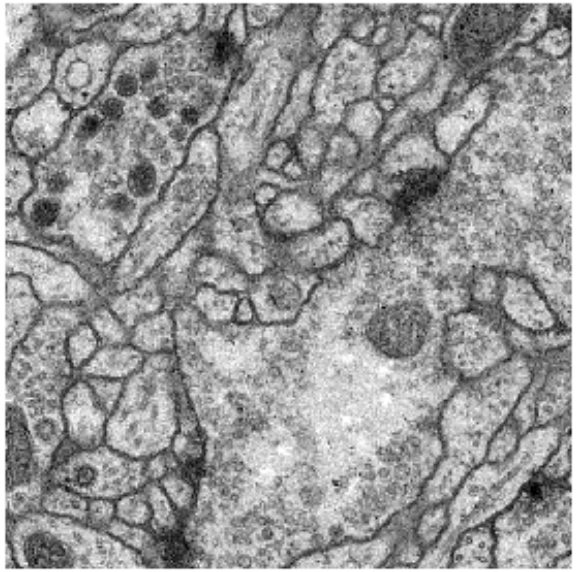
input (4, 4) IN

U-Net architecture



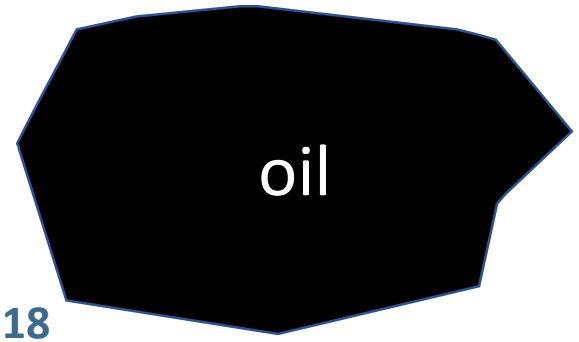
History

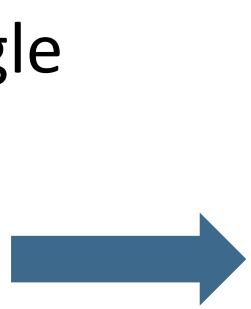
- Created in 2015 at the University of Freiburg
- Extended and modified network of the FCN (fully convolutional network)
- work with fewer training images and yield more precise segmentations

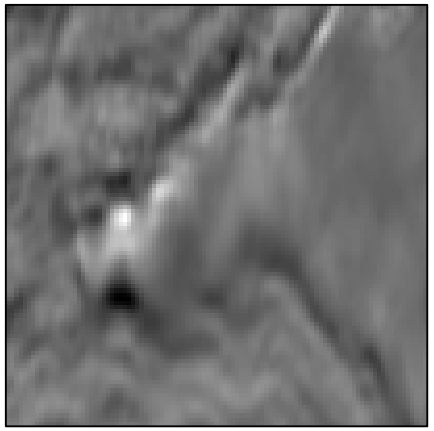


Used dataset

Overview: TGS salt dataset from kaggle





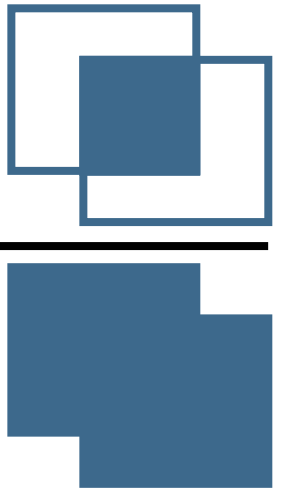


Implementation

Implementation:

- U-Net-Model (Standard)
  - with skip-connections
  - without skip-connections
- U-Net-Model with data augmentation
  - horizontal flip
  - vertical flip
- Evaluation metric: Intersection over union

$$\text{IoU} = \frac{\text{Area of intersection}}{\text{Area of union}}$$

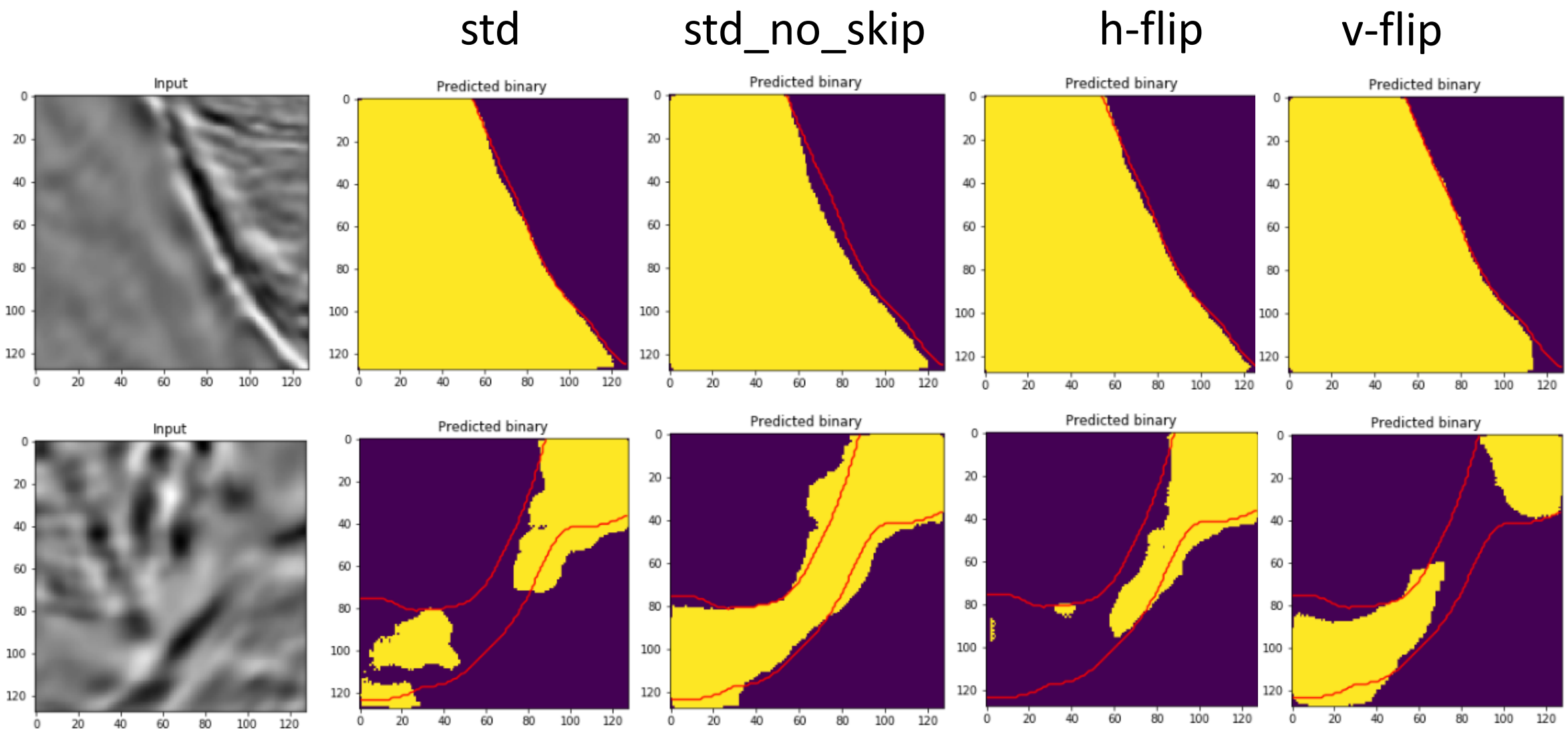


Result

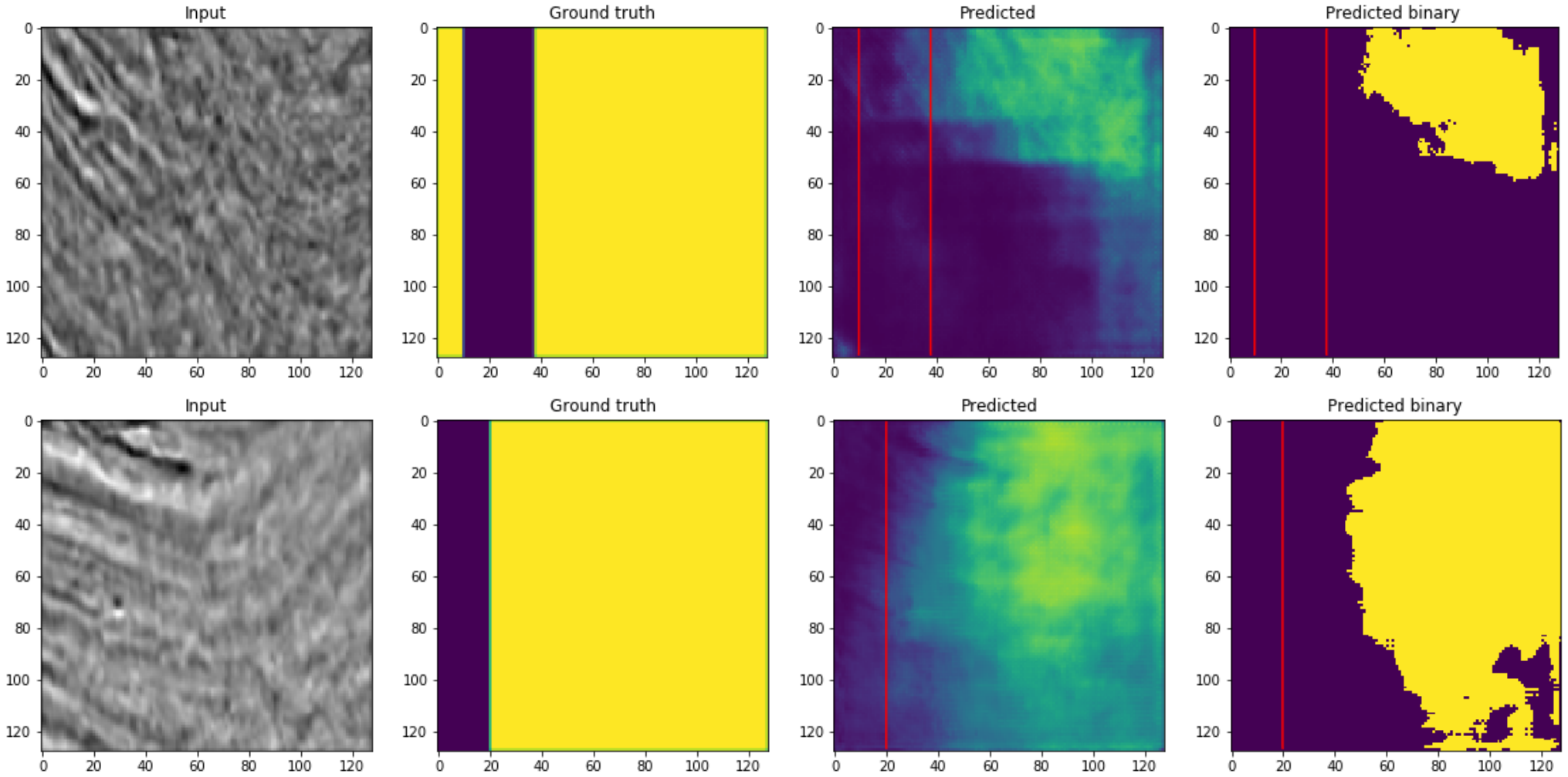
Comparison (test data):

U-Net-Model	**Loss**	**Mean IoU**
standard	0.152	0.788
standard without skip-connections	0.169	0.775
horizontal flip	**0.146**	**0.793**
vertical flip	0.177	0.749

Result


std
std_no_skip
h-flip
v-flip
Input
Predicted binary
Input
Predicted binary
Predicted binary
Predicted binary
Predicted binary

Challenges


Input
Ground truth
Predicted
Predicted binary
Input
Ground truth
Predicted
Predicted binary

Conclusion

- U-Net = Semantic segmentation (pixelwise classification)
- Contraction and expansion path
- Skip connections
- Challenges

The end

Sources

- https://github.com/vdumoulin/conv_arithmeticChallenges
- https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5
- https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/
- https://medium.com/activating-robotic-minds/up-sampling-with-transposed-convolution-9ae4f2df52d0
- https://www.codeastar.com/u-net-object-detection-iou/