The U-net Model

A Fully Convolutional Neural Network Model

Roberto Souza
Assistant Professor
Electrical and Computer Engineering
Schulich School of Engineering

W2024

UNIVERSITY OF CALGARY

Outline

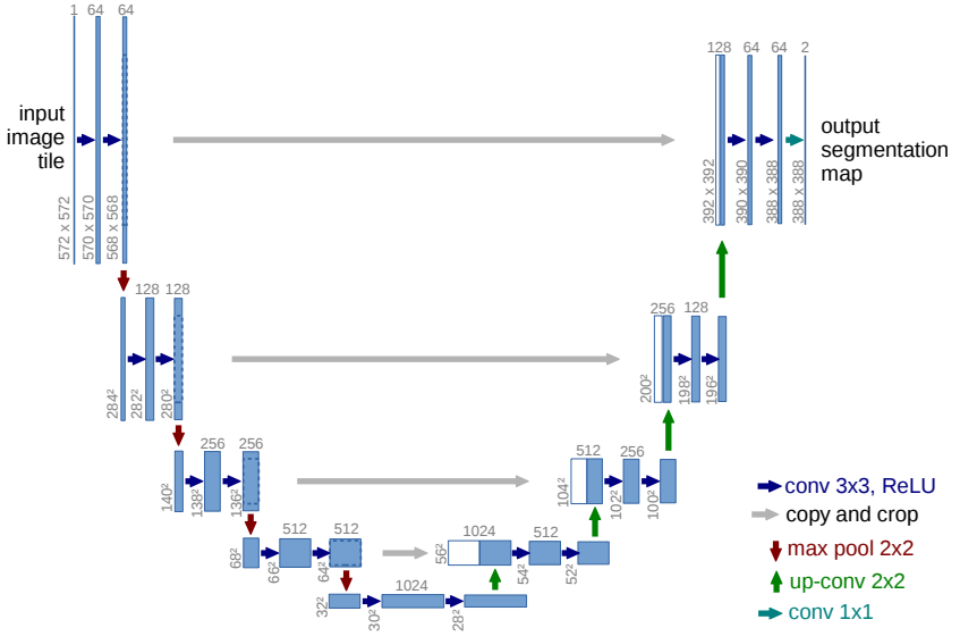
- Learning Goals
- The U-net Model
- Summary

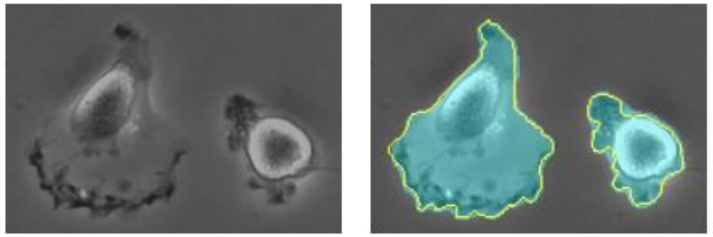
Learning Goals

- Understand the U-net architecture and its building blocks
- Discuss potential applications of the U-net model

The U-net Model

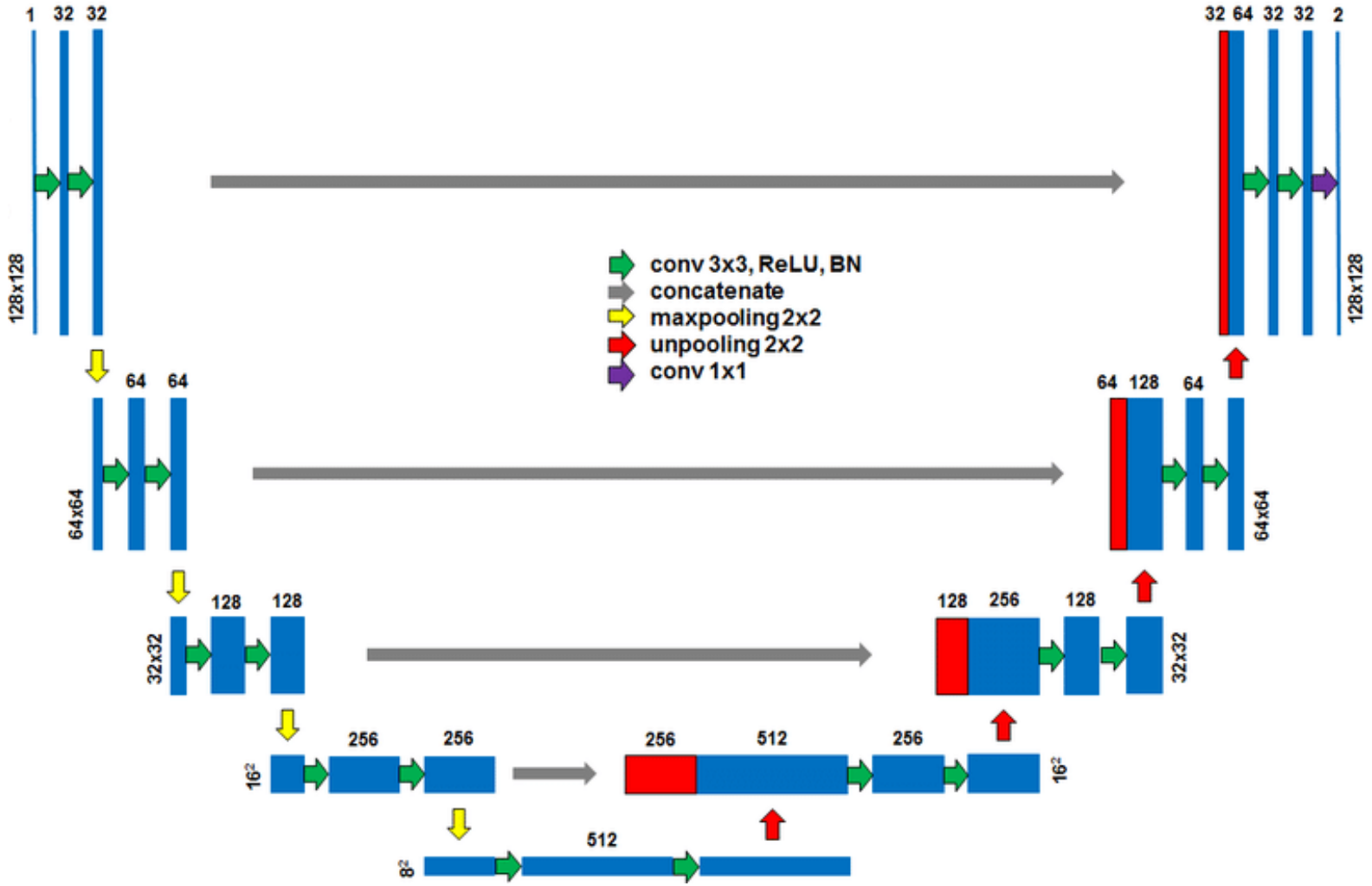
- The U-net is a fully convolutional neural network (i.e., no fully connected layers)
- Initially proposed for biomedical image segmentation problems
- It maps an input of size N into an output also of size N (if the convolutions are padded)





UNIVERSITY OF CALGARY

The U-net Model



UNIVERSITY OF CALGARY

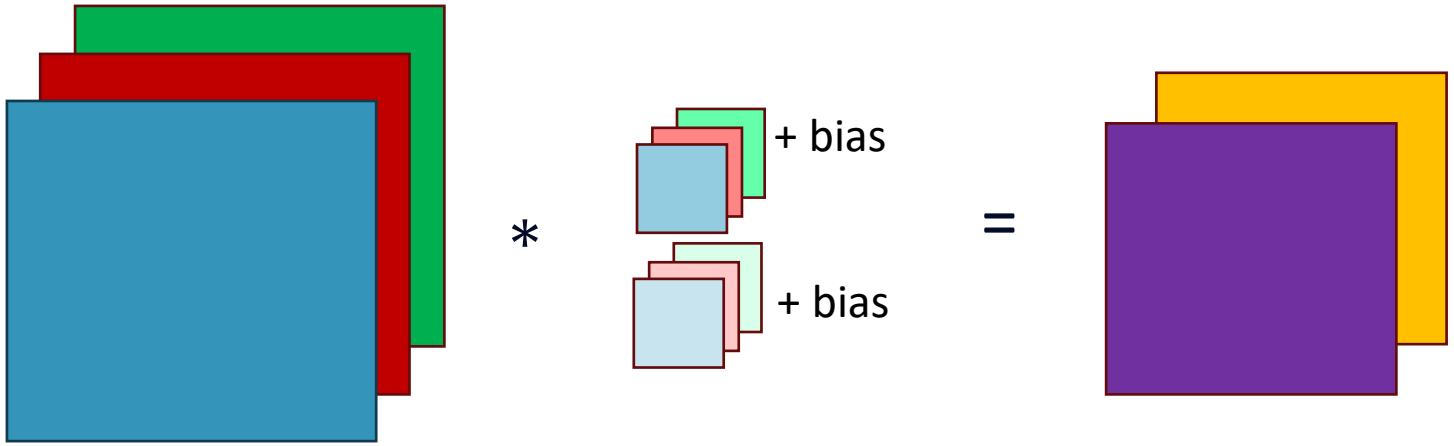
Number of parameters

For one layer:

params =

weights + biases =

[input maps x filter size x output maps] + output maps



Number of parameters

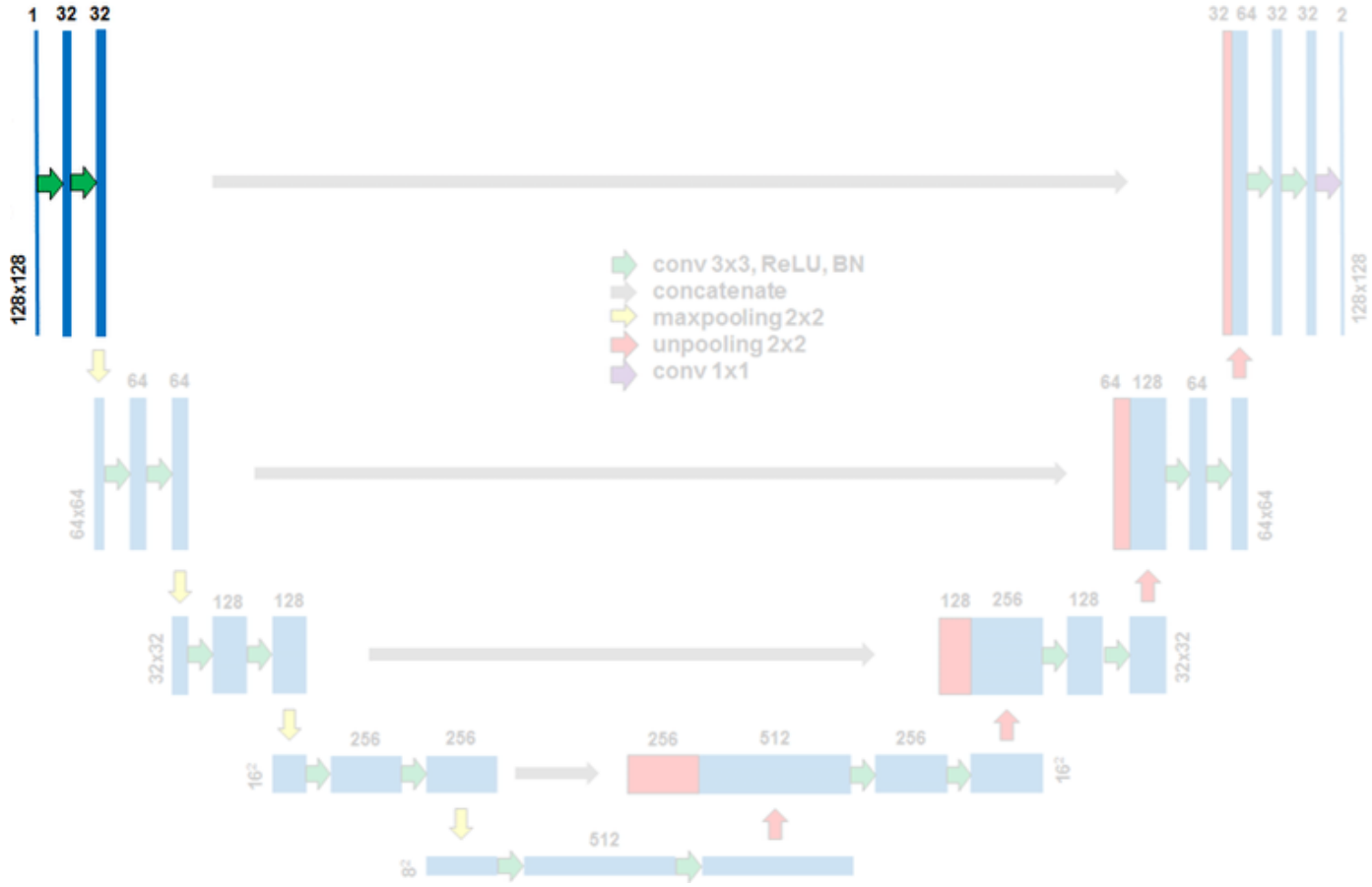
Layer 1 params:

1 x (3 x 3) x 32 + 32 = 320

Layer 2 params:

32 x (3 x 3) x 32 + 32 = 9248

Total: 9568



UNIVERSITY OF CALGARY

Number of parameters

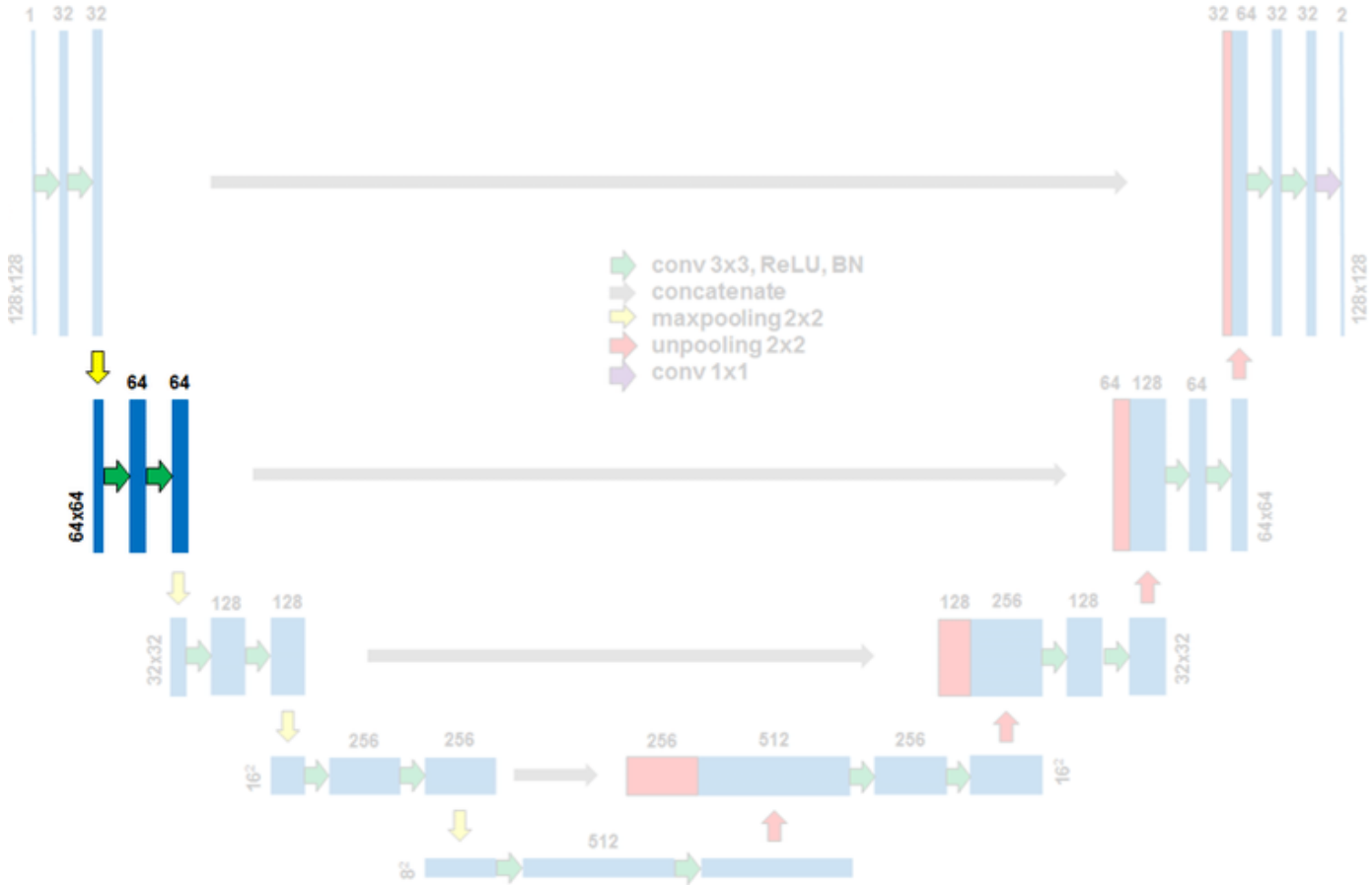
Layer 1 params:

32 x (3 x 3) x 64 + 64 = 18,496

Layer 2 params:

64 x (3 x 3) x 64 + 64 = 36,928

Total: 55,424



UNIVERSITY OF CALGARY

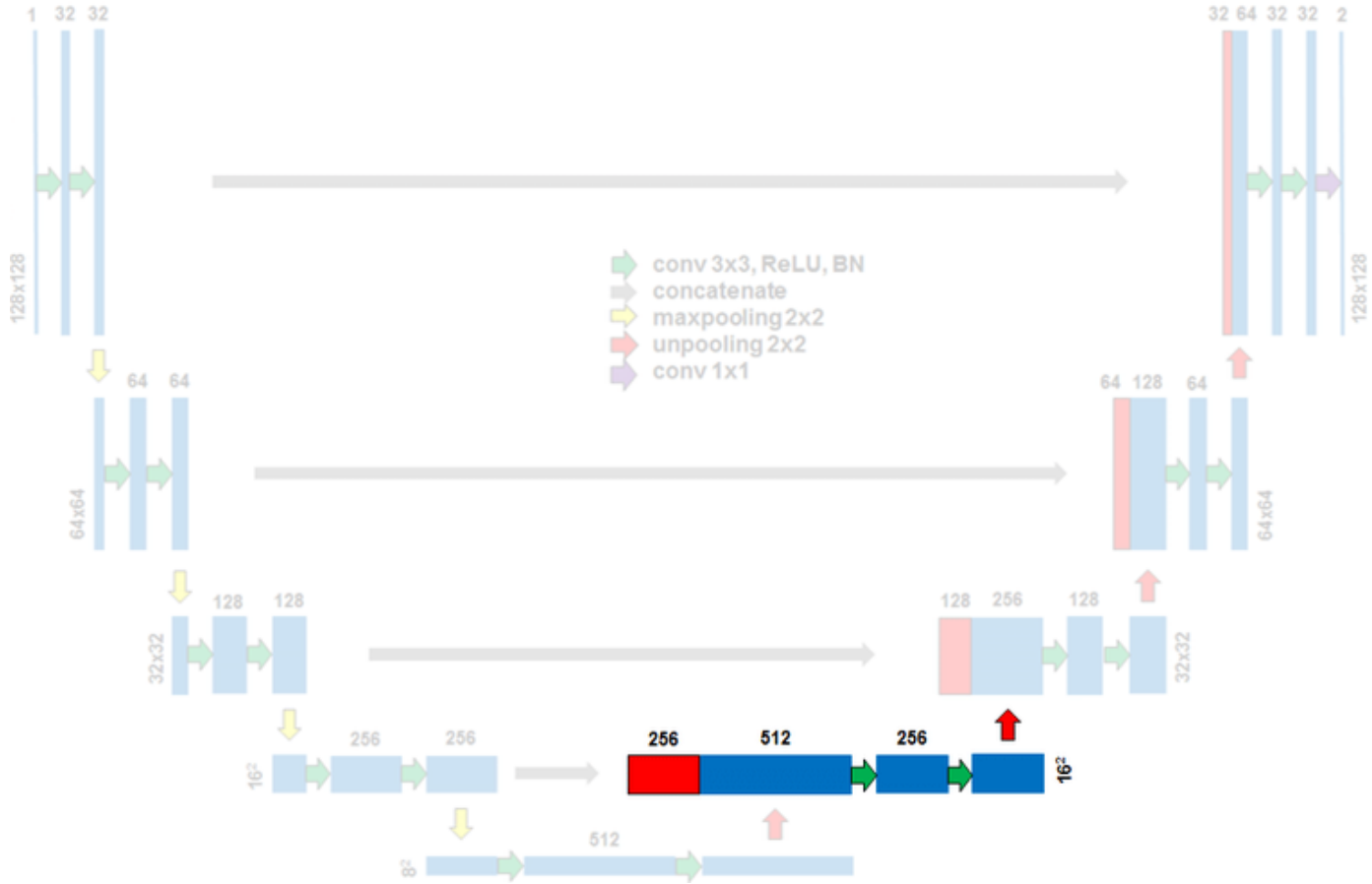
Number of parameters

Layer 1 params:
768 x (3 x 3) x 256 + 256 = 1,769,728

Layer 2 params:
256 x (3 x 3) x 256 + 256 = 590,080

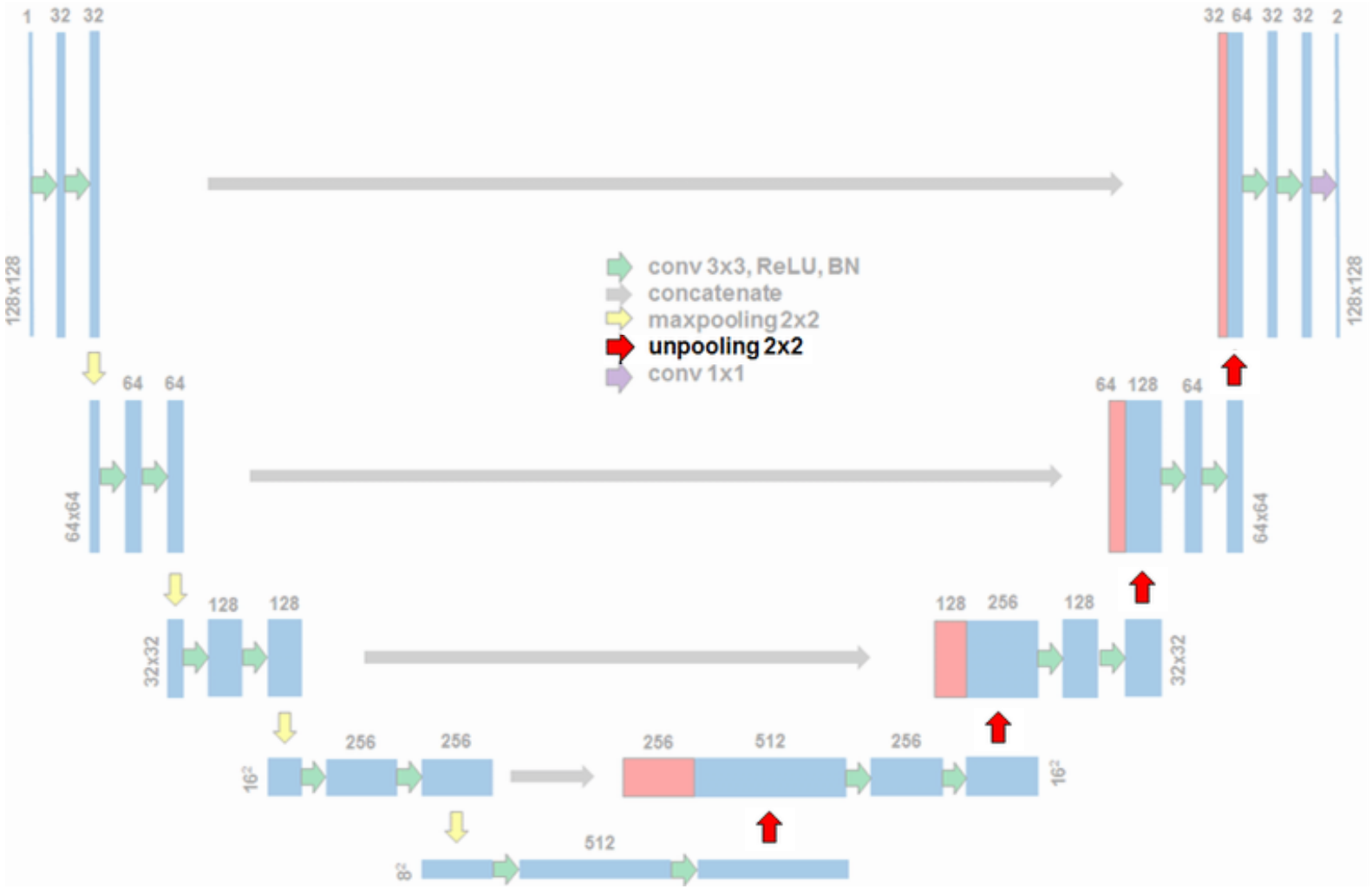
Total: 2,359,808

Grand total: ~7.8 million



UNIVERSITY OF CALGARY

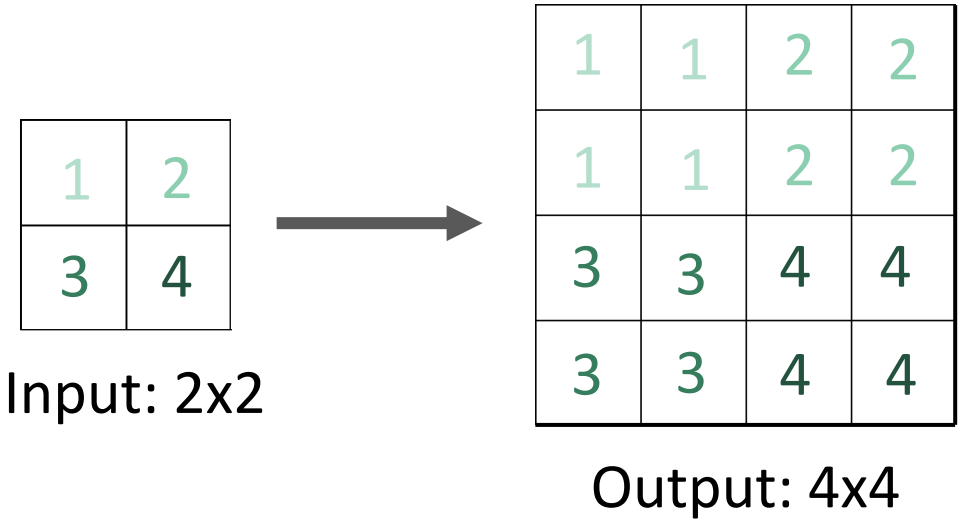
Upsampling



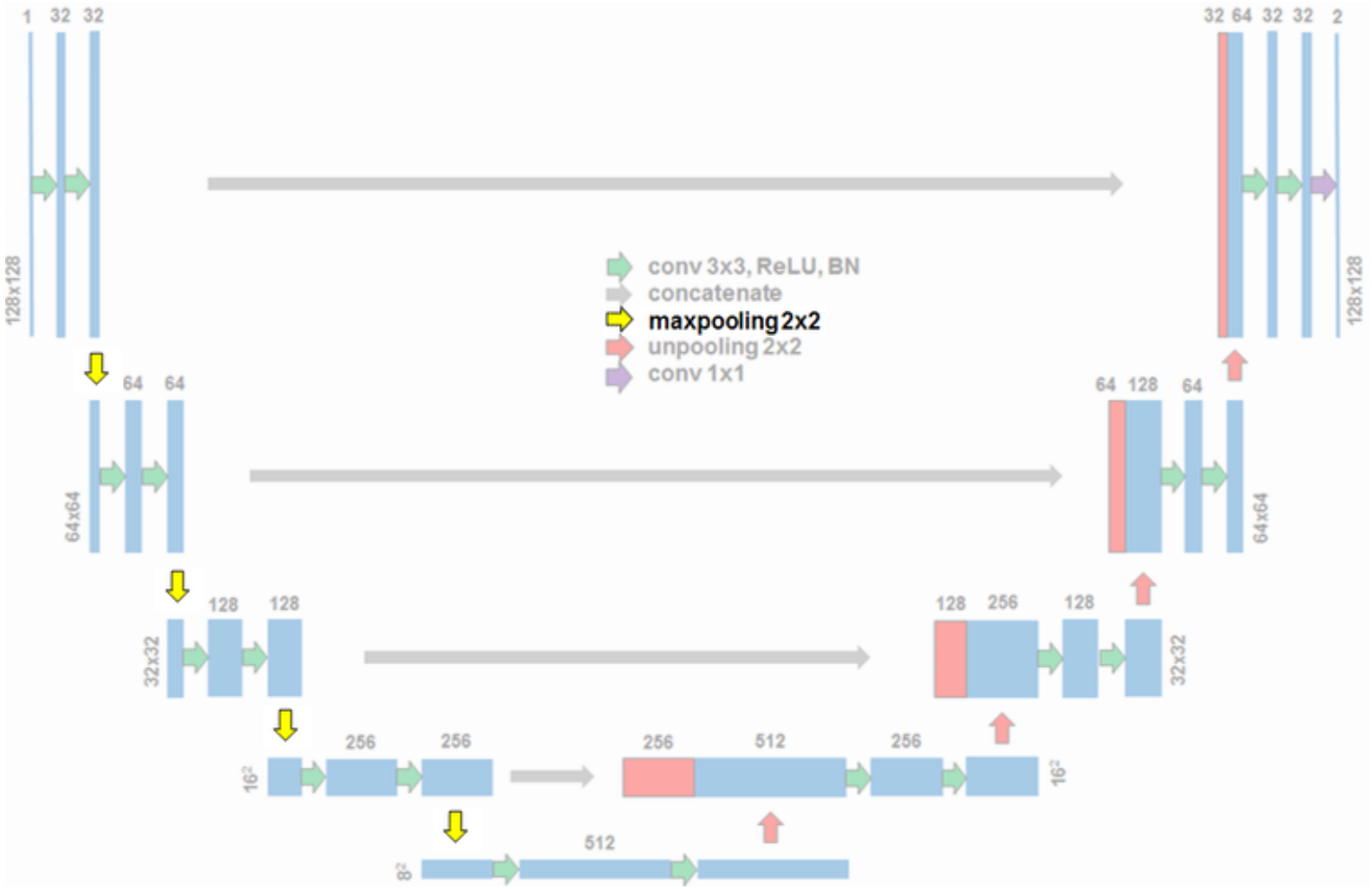
UNIVERSITY OF CALGARY

Upsampling

- Opposite effect of max-pooling
- Many ways to do it
  - Transpose convolution
  - Nearest neighbor interpolation
  - Linear interpolation
  - …



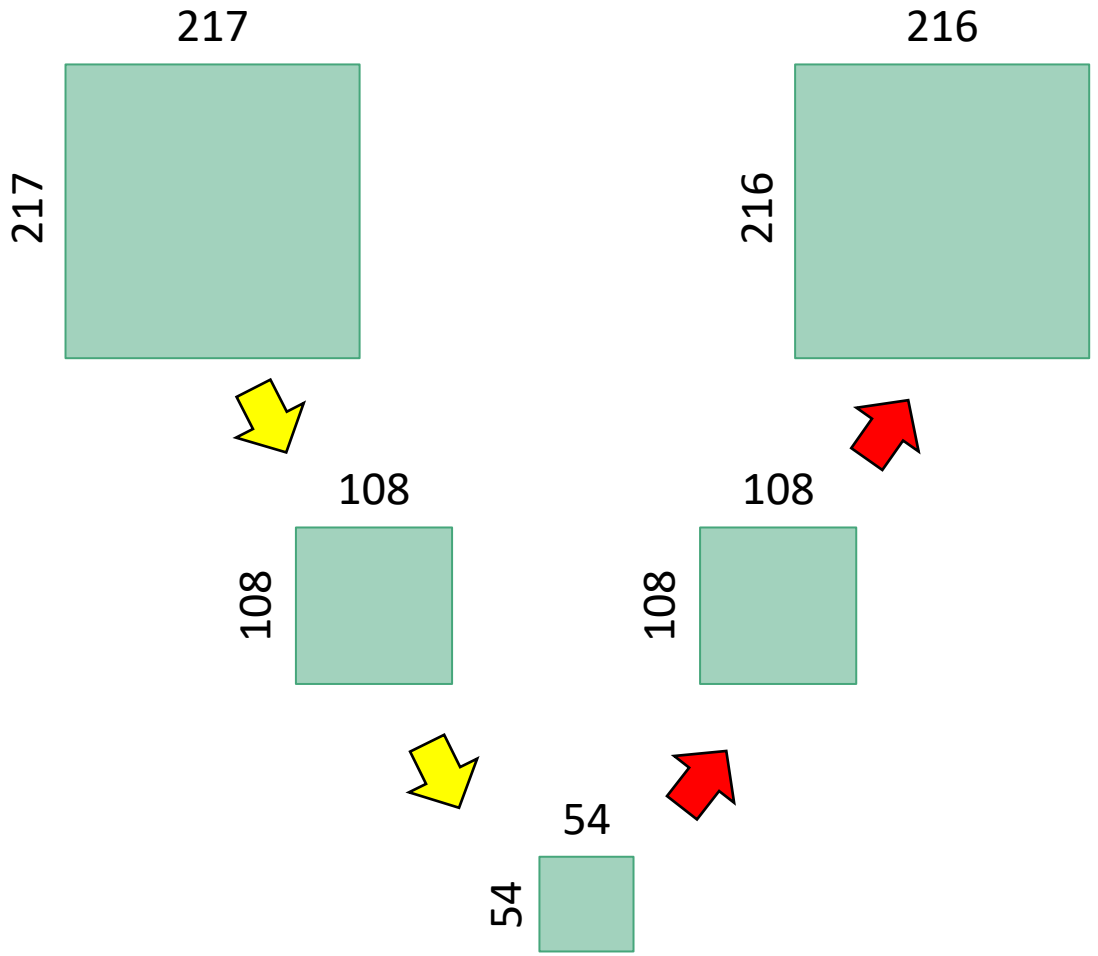
Max pooling



UNIVERSITY OF CALGARY

Maxpooling

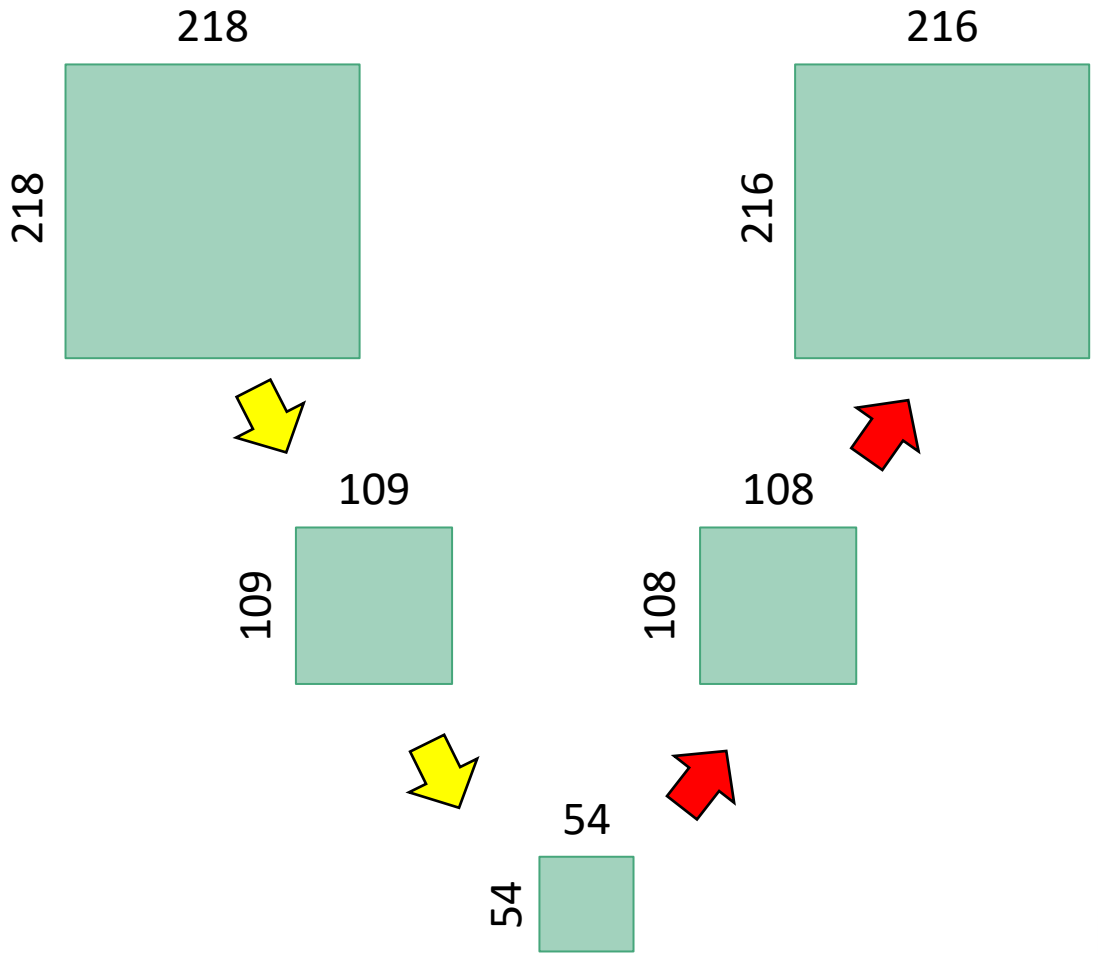
- U-net uses 2D max pooling
- What happens when we have
  - Odd image dimensions?



UNIVERSITY OF CALGARY

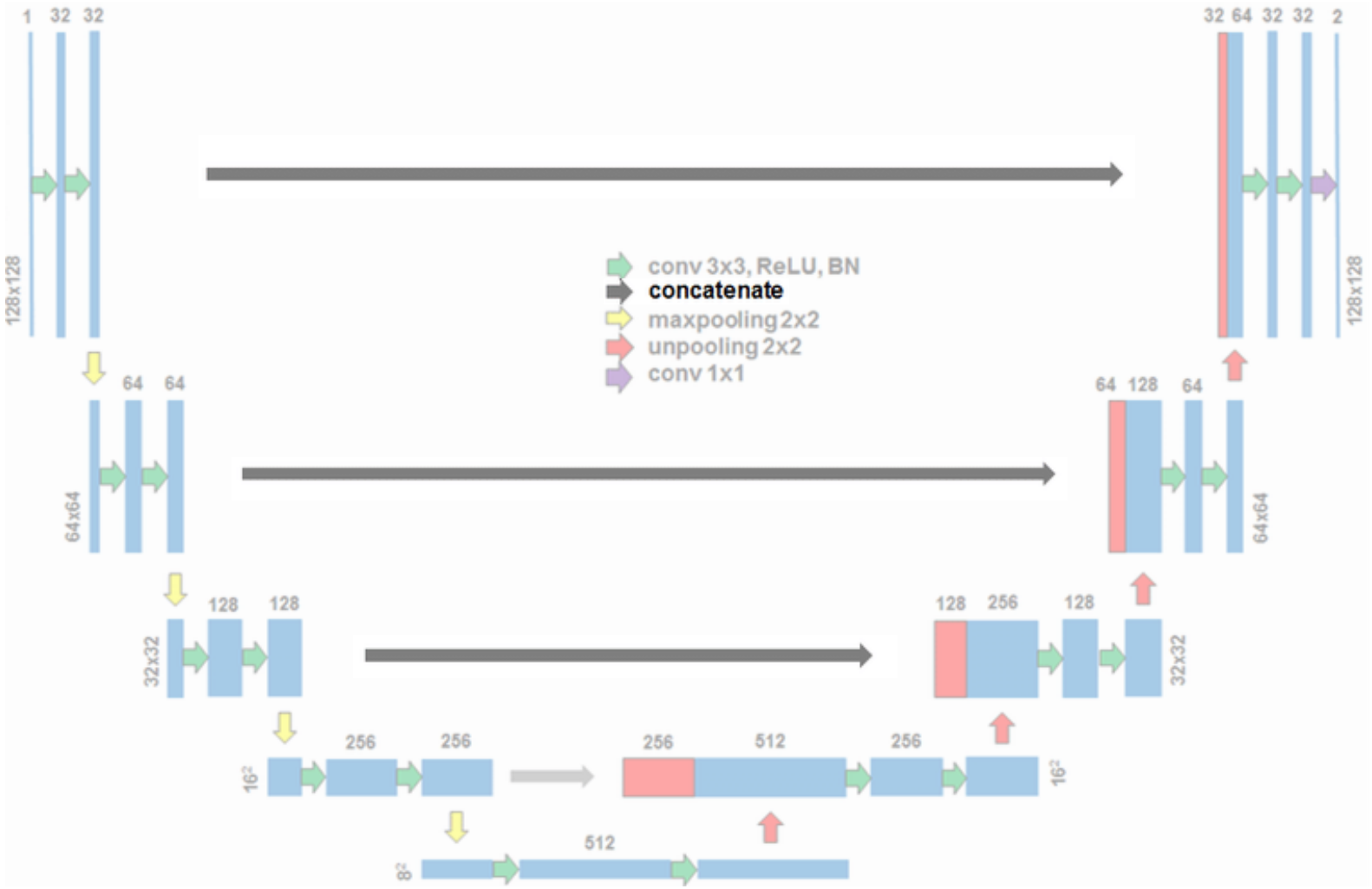
Maxpooling

- U-net uses 2D max pooling
- What happens when we have
  - Odd image dimensions?
  - Other even image dimensions?
- Solution: pad input to nearest dimension divisible by 2 # max pool layers



UNIVERSITY OF CALGARY

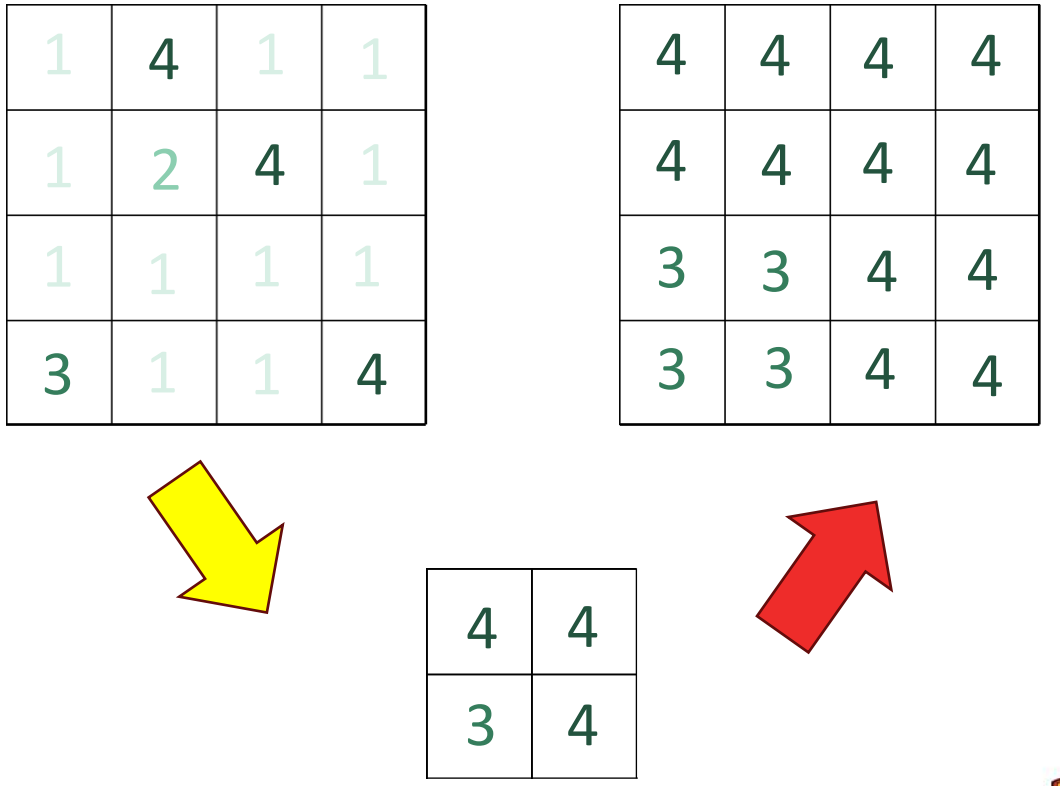
Skip connections



UNIVERSITY OF CALGARY

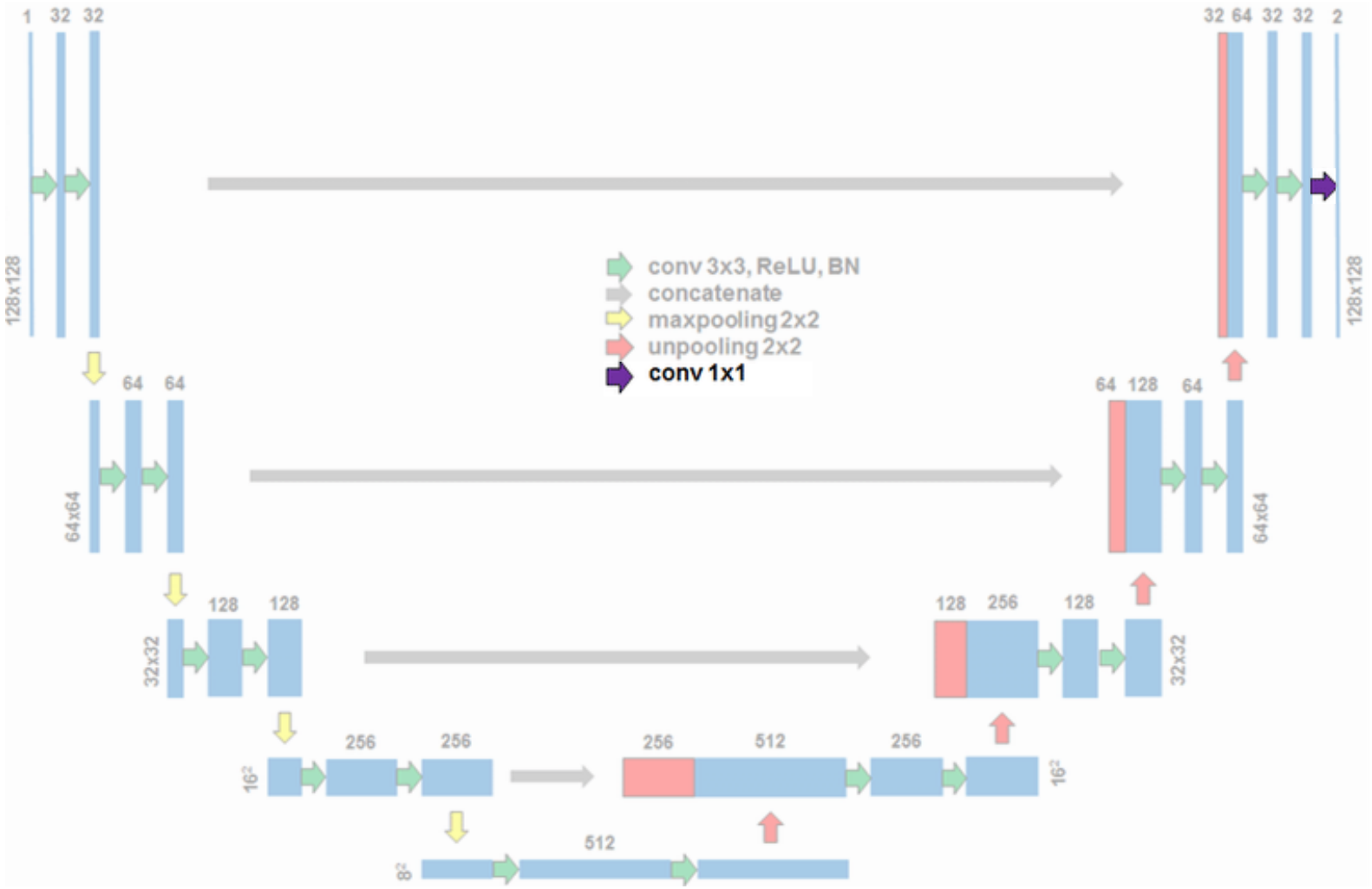
Skip connections

- Propagate multi-scale information
- Improved spatial information



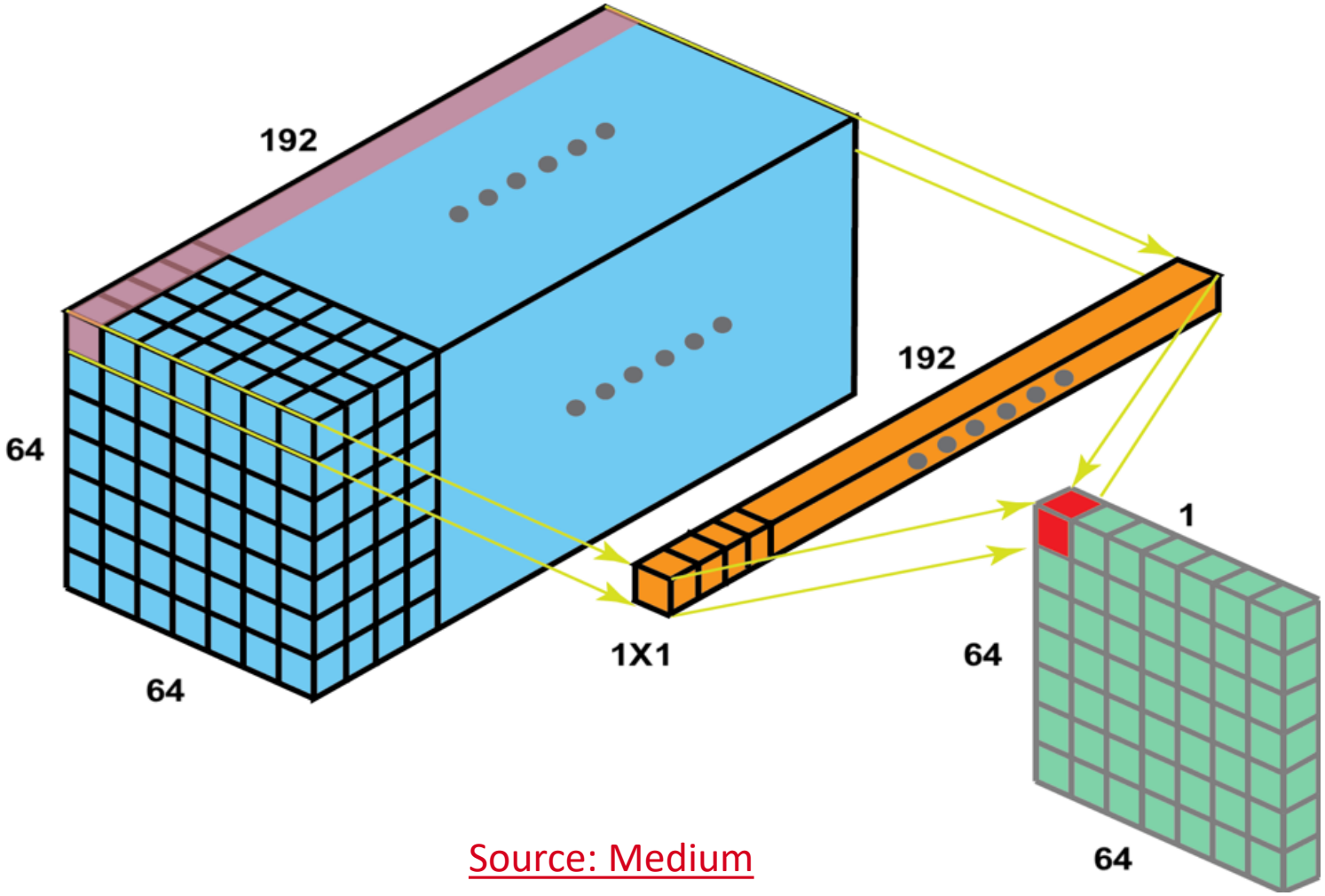
UNIVERSITY OF CALGARY

1x1 Convolution



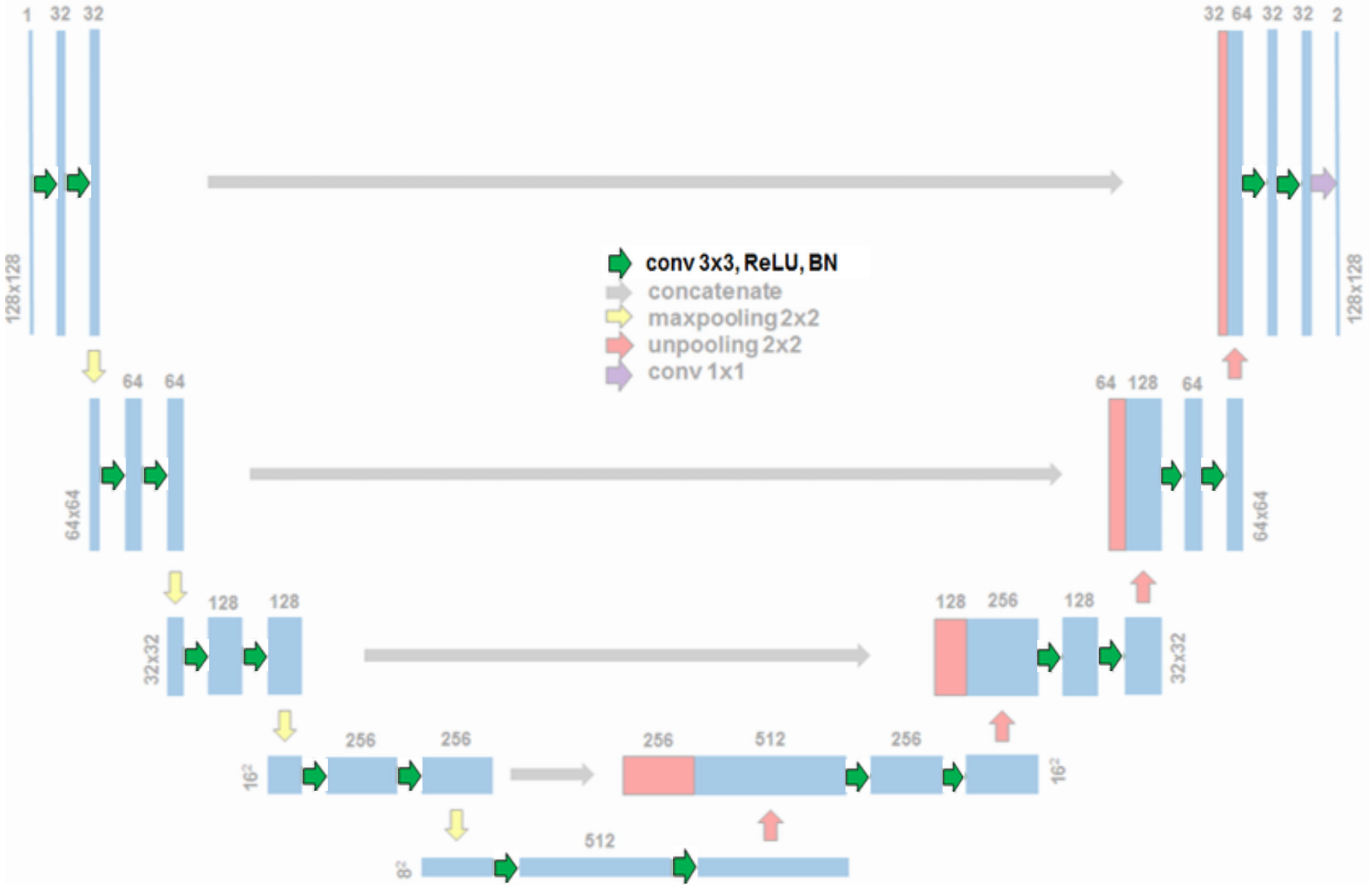
UNIVERSITY OF CALGARY

1x1 Convolution



Source: Medium

Batch normalization



UNIVERSITY OF CALGARY

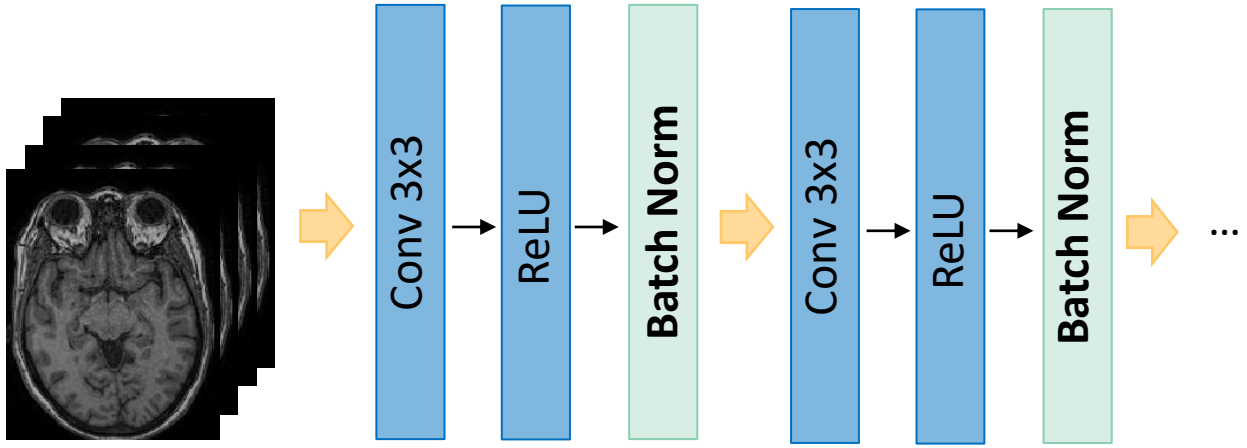
Batch normalization

- Normalizes your batch of data in between layers

$$y = \frac{x - \mathrm{E}[x]}{\sqrt{\mathrm{Var}[x] + \epsilon}} * \gamma + \beta$$

Computed parameters

Learnable parameters



Batch normalization

- Each feature/feature map is normalized!
- During training, running mean and variance computed for use during inference

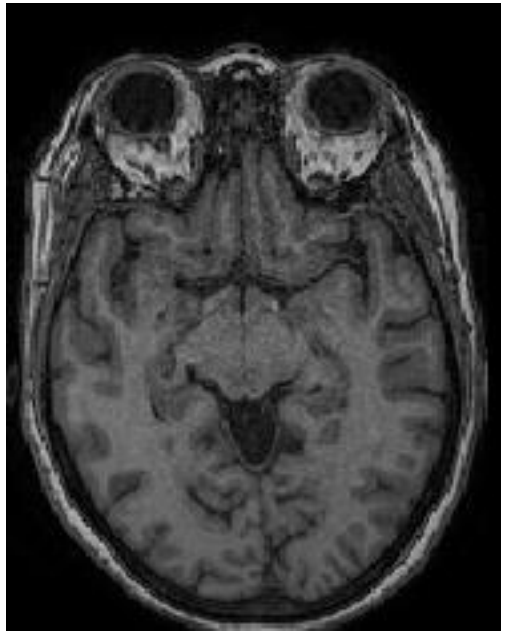
$$\mathrm{E}[x]_{run} = \alpha \mathrm{E}[x]_{run} + (1-\alpha)\mathrm{E}[x]$$

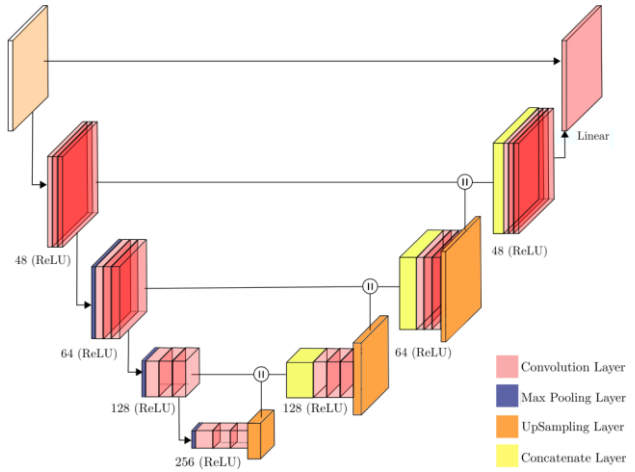
$$\mathrm{Var}[x]_{run} = \alpha \mathrm{Var}[x]_{run} + (1-\alpha)\mathrm{Var}[x]$$

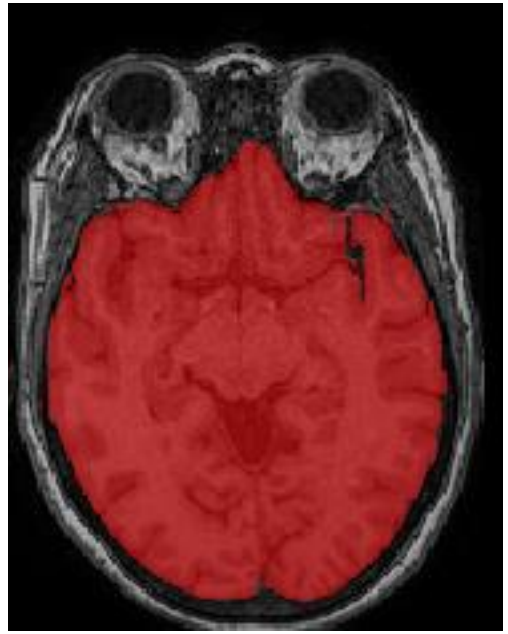
Momentum

U-net - Segmentation

Segmentation = pixel-wise or voxel-wise segmentation

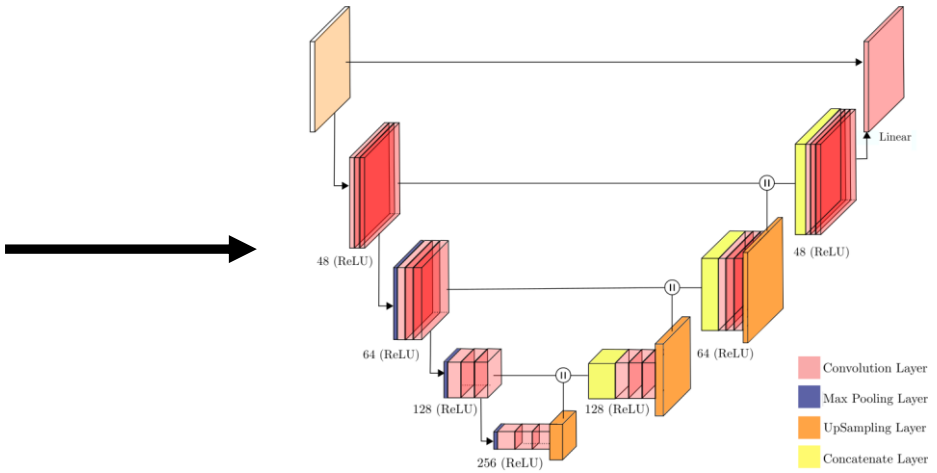






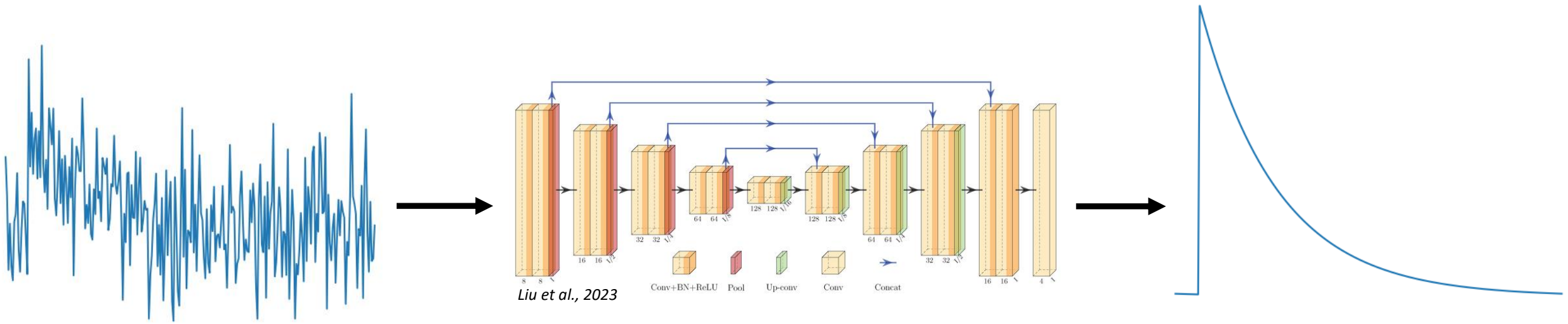
UNIVERSITY OF CALGARY

U-net – 2D Regression



UNIVERSITY OF CALGARY

U-net – 1D Regression



Liu et al., 2023

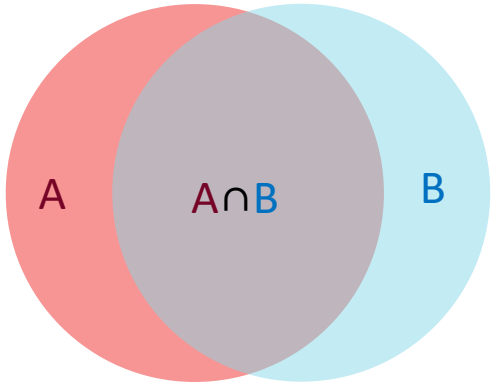
UNIVERSITY OF CALGARY

Metrics

- For regression:
  - Mean squared error
  - Mean absolute error
  - …

y			-	ŷ		
3	4	4		3	3	1
2	3	4		2	3	3
1	2	3		2	2	3

- For segmentation:
  - Dice coefficient
  - Jaccard coefficient
  - …



UNIVERSITY OF CALGARY

Summary

- The U-net is a very powerful deep learning model that maps inputs to outputs of the same size
- The model works across different scales of the input signal/image
- It is a fully convolutional model that is independent of the input size

Thank you!

UNIVERSITY OF CALGARY