Mahidol University

# ITCS498 Special Topics in Computer Science

## Lecture 11 - Semantic Segmentation

Instructor: Asst. Prof. Dr. Akara Supratak

Contact: akara.sup@mahidol.edu

# Agenda

- Semantic Segmentation
- Lab Exercise

ref:

# Semantic Segmentation

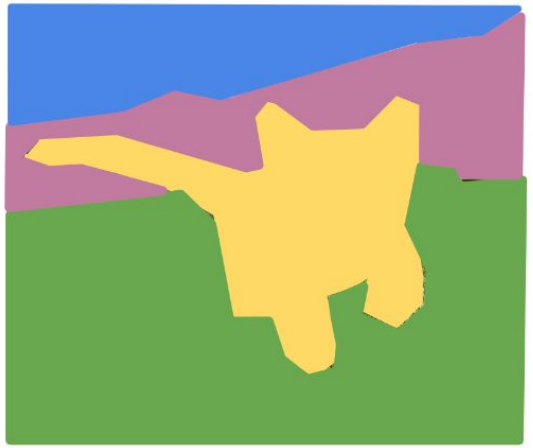
# Computer Vision Tasks

**Classification**

**CAT**

No location information

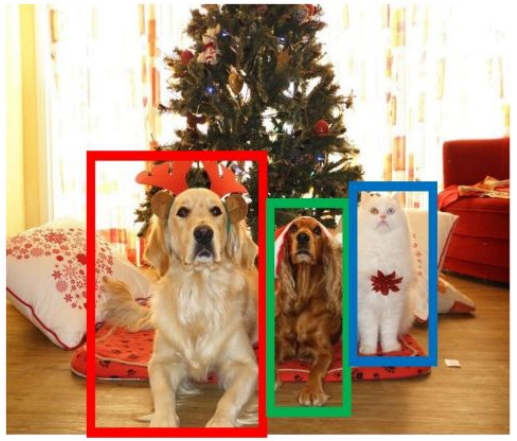
**Semantic Segmentation**



**GRASS, CAT, TREE, SKY**

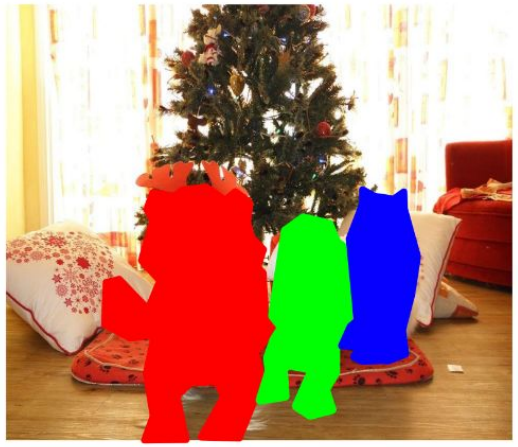
No object, just pixels

**Object Detection**



**DOG, DOG, CAT**

**Instance Segmentation**

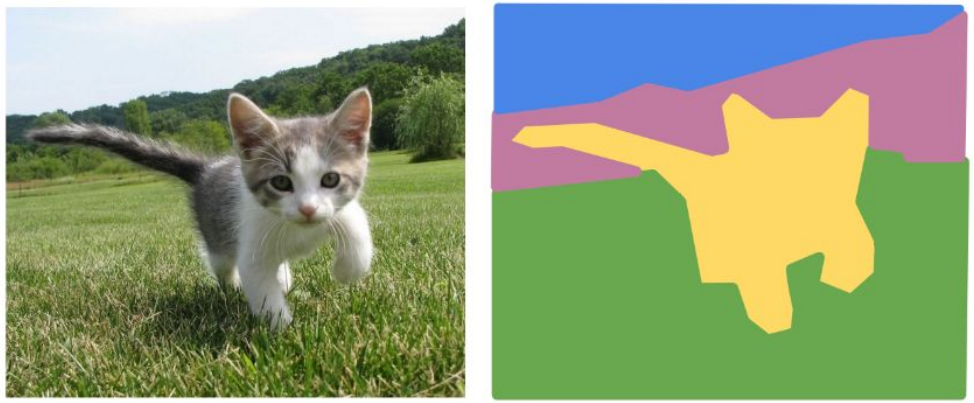


**DOG, DOG, CAT**

Multiple objects

ref: [35]

# Semantic Segmentation

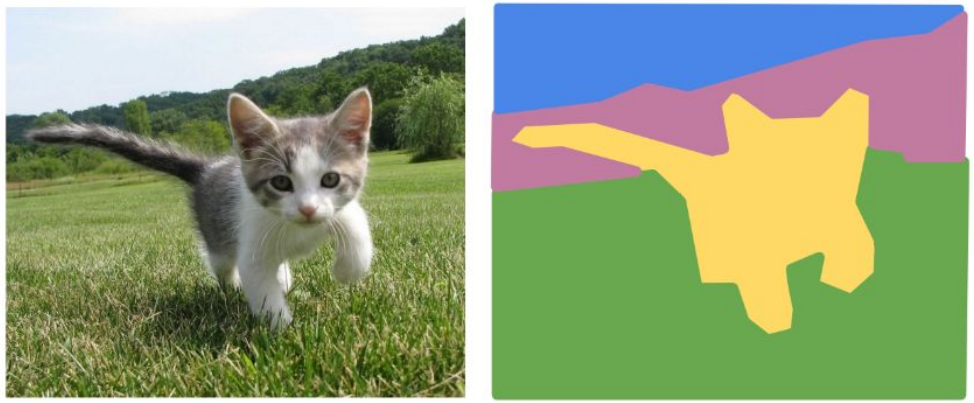


GRASS, CAT,
TREE, SKY, ...

Paired training data: for each training image,
each pixel is labeled with a semantic category.

ref: [35]

# Semantic Segmentation



GRASS, CAT,
TREE, SKY, ...

Paired training data: for each training image, each pixel is labeled with a semantic category.

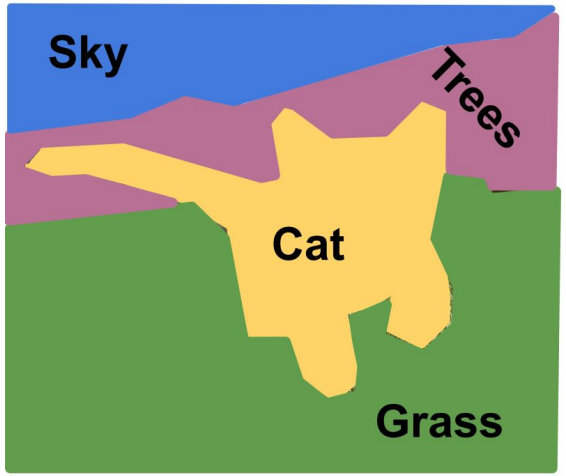

At test time, classify each pixel of a new image.

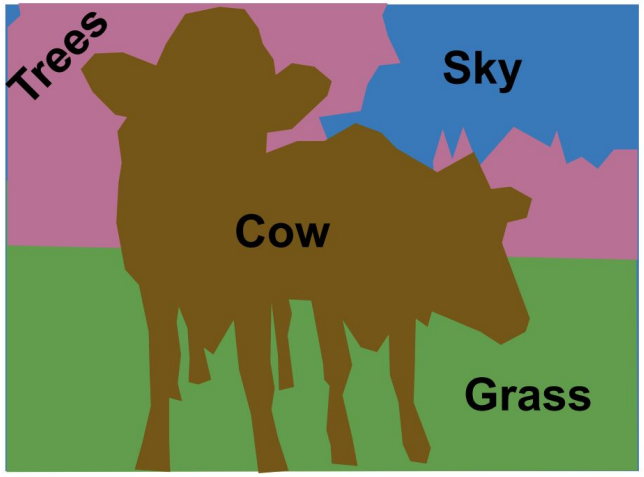
ref: [35]

# Semantic Segmentation: Labeled Dataset

- Label each pixel in the image with a category label

- No differentiate between instances, only care about the class of each pixel

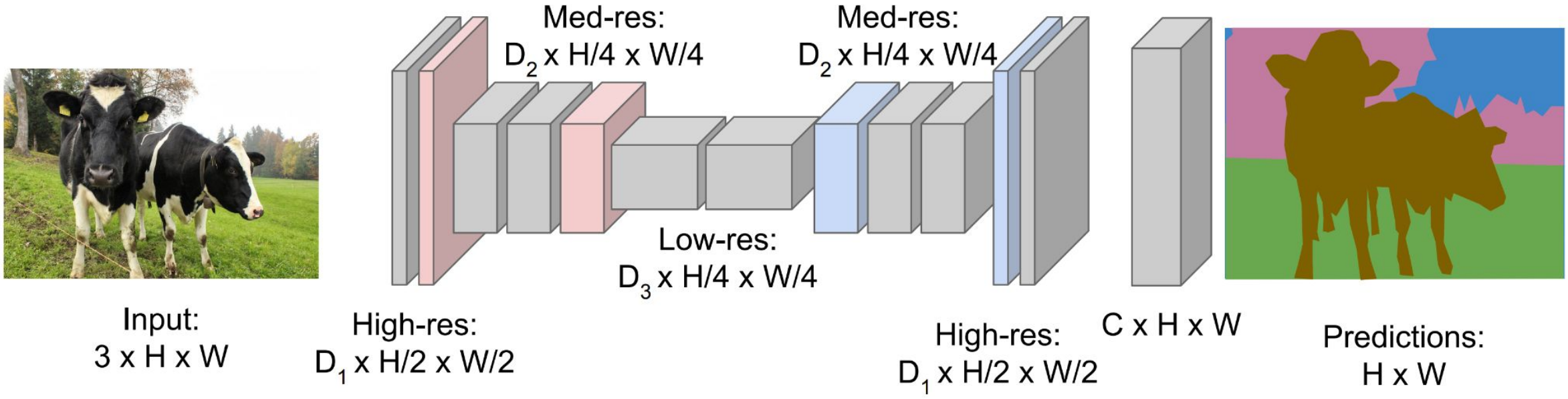




ref: [35]

# Semantic Segmentation: Fully-Convolutional

Design network as a bunch of convolutional layers, with **downsampling** and **upsampling** inside the network!




Long, Shelhamer, and Darrell, "Fully Convolutional Networks for Semantic Segmentation", CVPR 2015
Noh et al, "Learning Deconvolution Network for Semantic Segmentation", ICCV 2015

ref: [35]

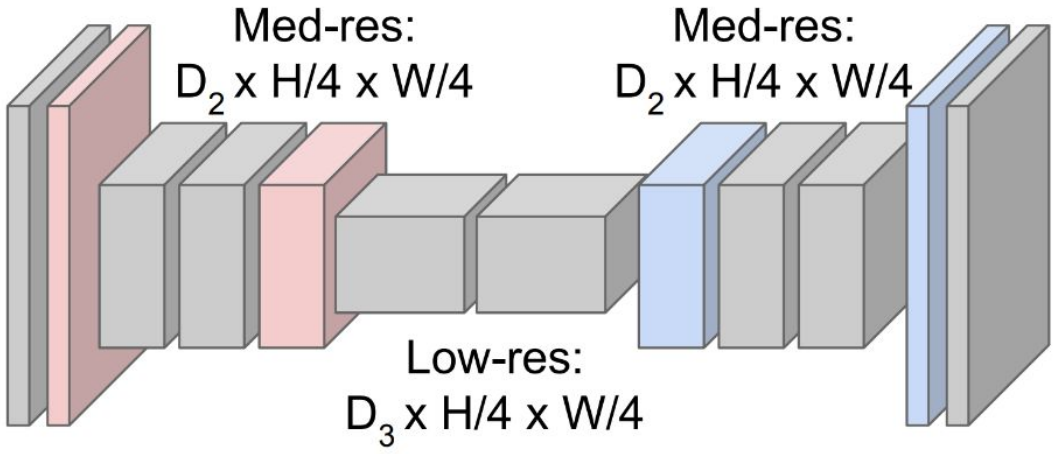
# Semantic Segmentation: Fully-Convolutional

**Downsampling**:
Strided convolution, pooling

Design network as a bunch of convolutional layers, with **downsampling** and **upsampling** inside the network!

**Upsampling**: ?

Input:
3 x H x W




High-res:
D1 x H/2 x W/2

High-res:
D1 x H/2 x W/2

C x H x W

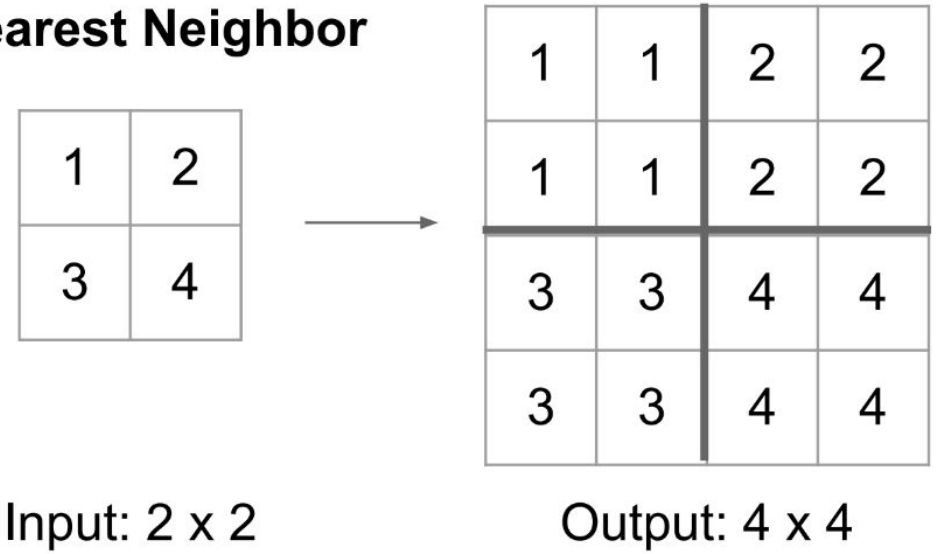
Predictions:
H x W

Long, Shelhamer, and Darrell, "Fully Convolutional Networks for Semantic Segmentation", CVPR 2015
Noh et al, "Learning Deconvolution Network for Semantic Segmentation", ICCV 2015

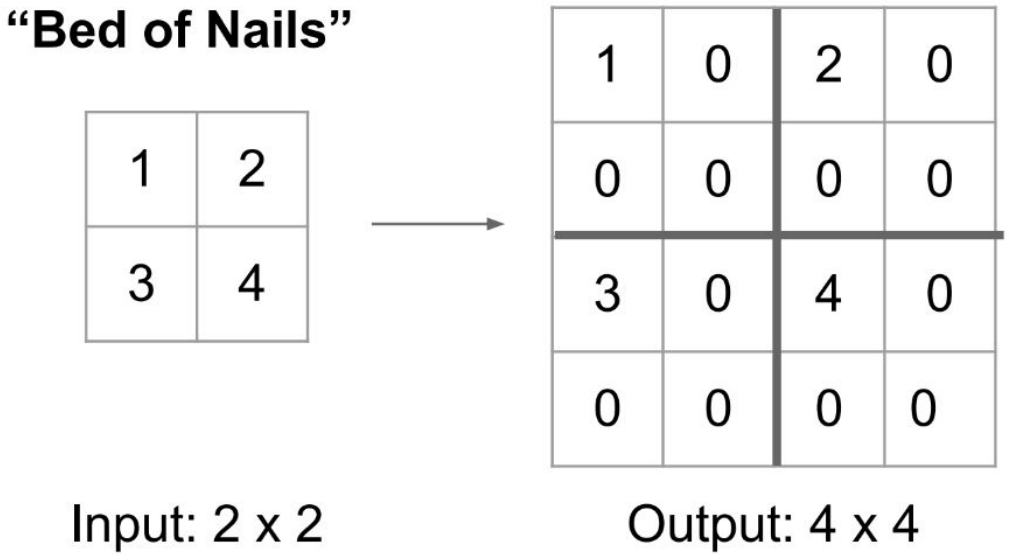
ref: [35]

# Upsampling: Unpooling

**Nearest Neighbor**

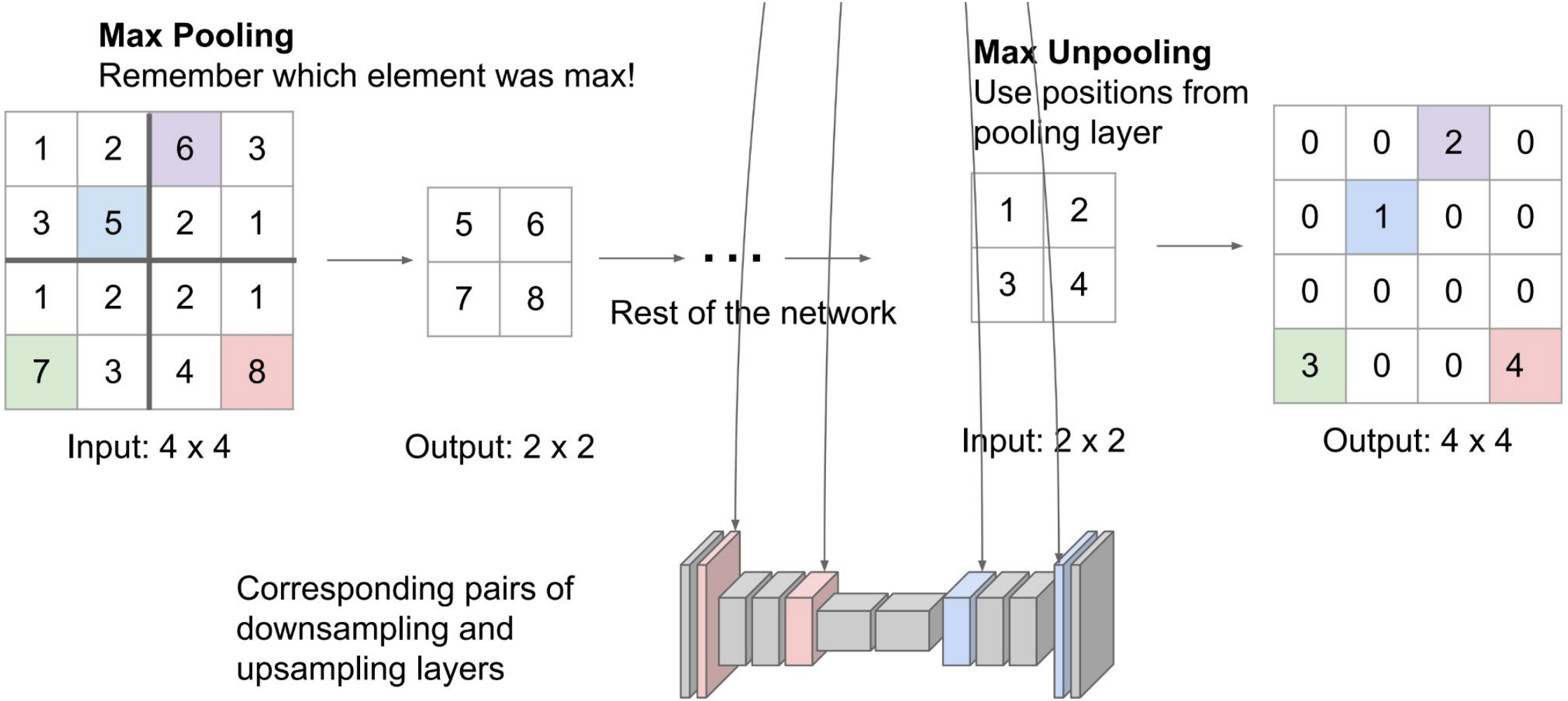



**“Bed of Nails”**




ref: [35]

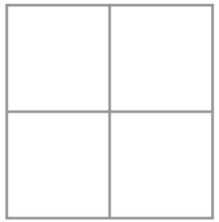
# Upsampling: Max-Unpooling



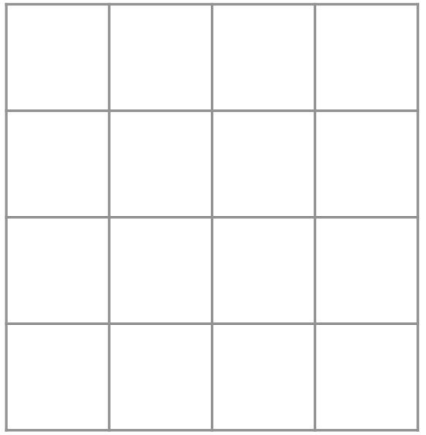

ref: [35]

# Upsampling: Transpose Convolution

3 x 3 **transposed** convolution, stride 2 pad 1



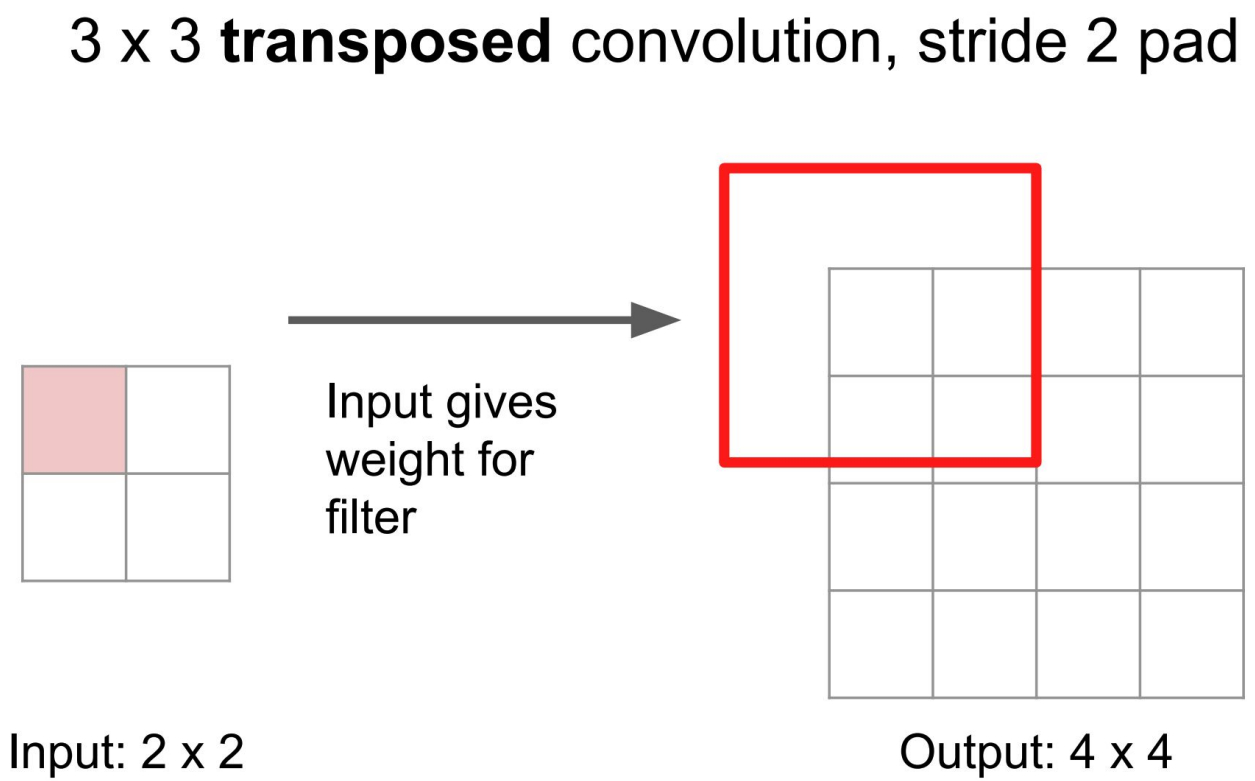
Input: 2 x 2



Output: 4 x 4

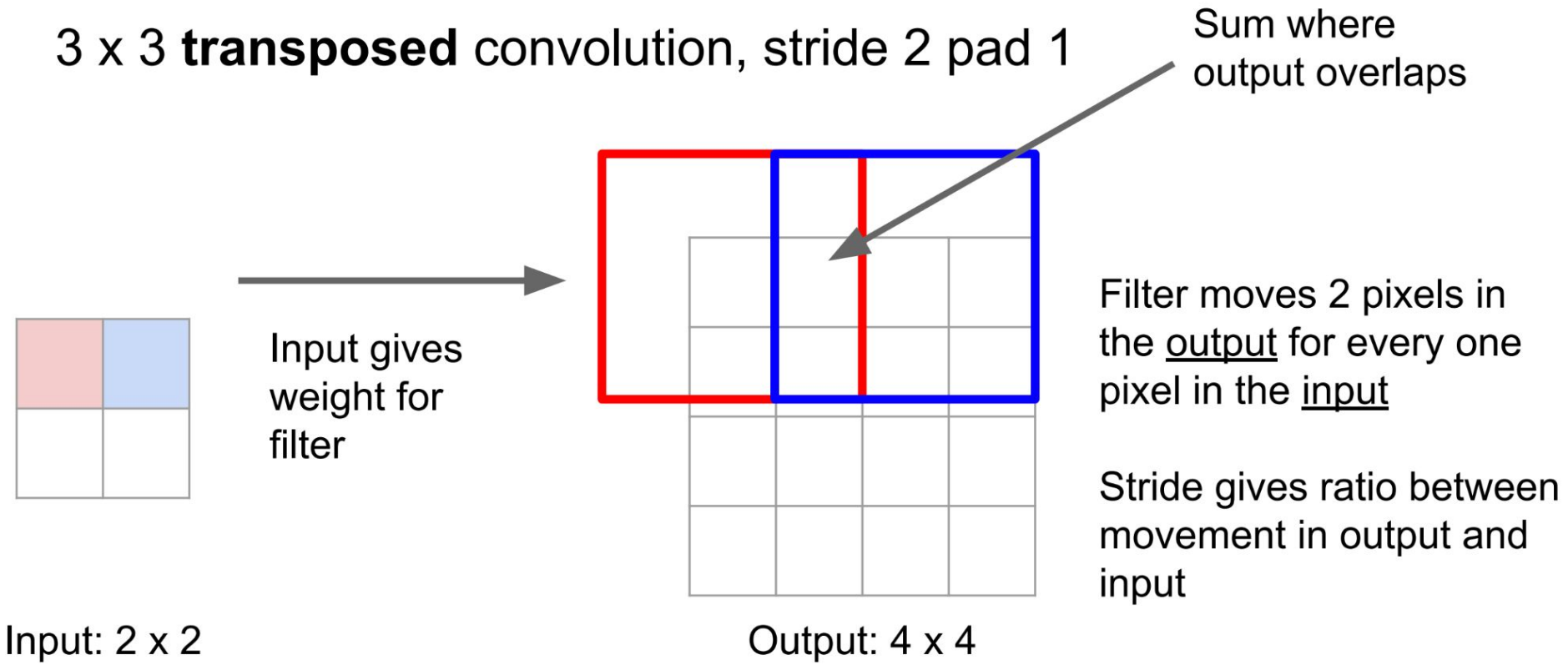
ref: [35]

# Upsampling: Transpose Convolution



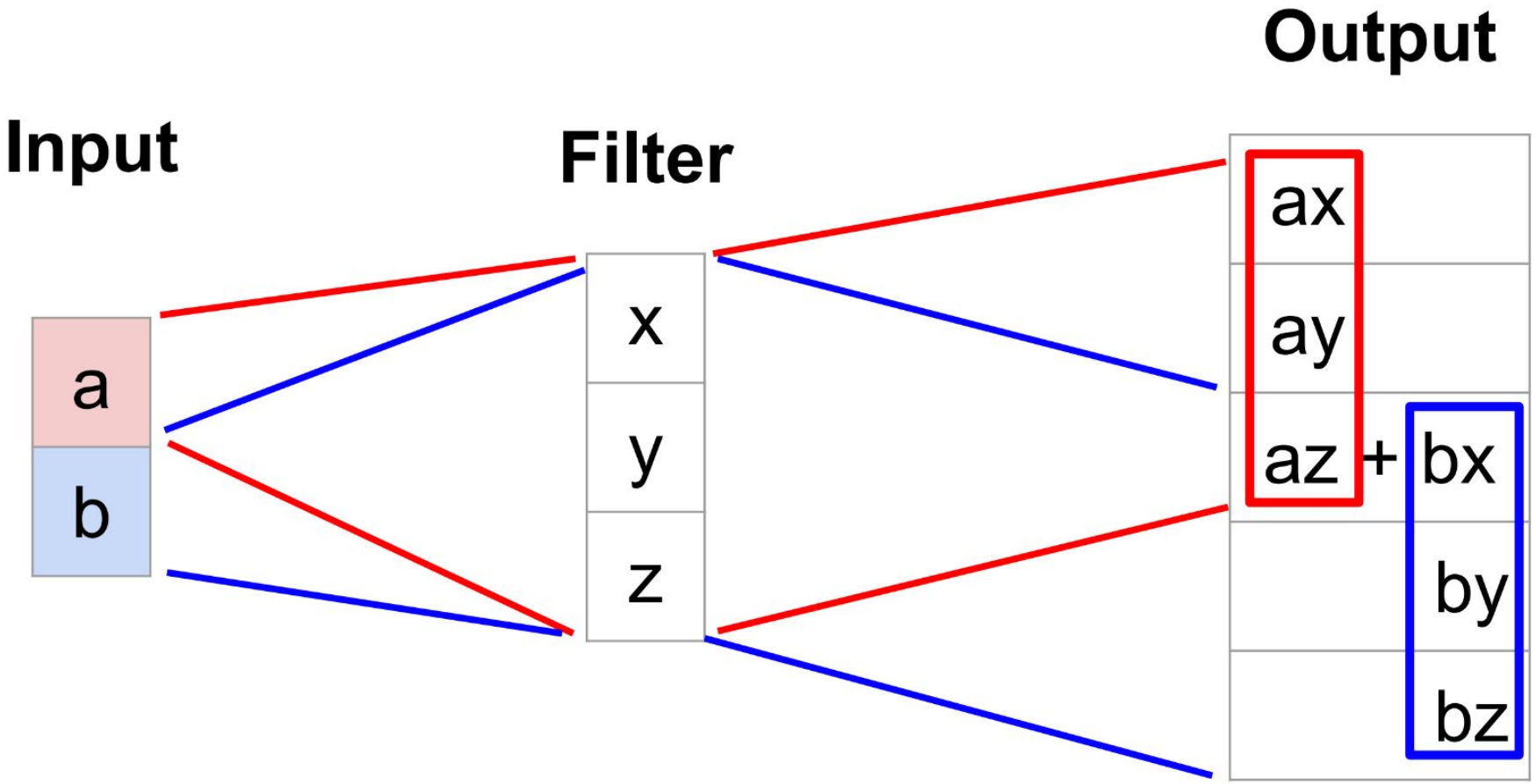

ref: [35]

# Upsampling: Transpose Convolution




ref: [35]

# Upsampling: Transpose Convolution




Output contains copies of the filter weighted by the input, summing at where at overlaps in the output

ref: [35]

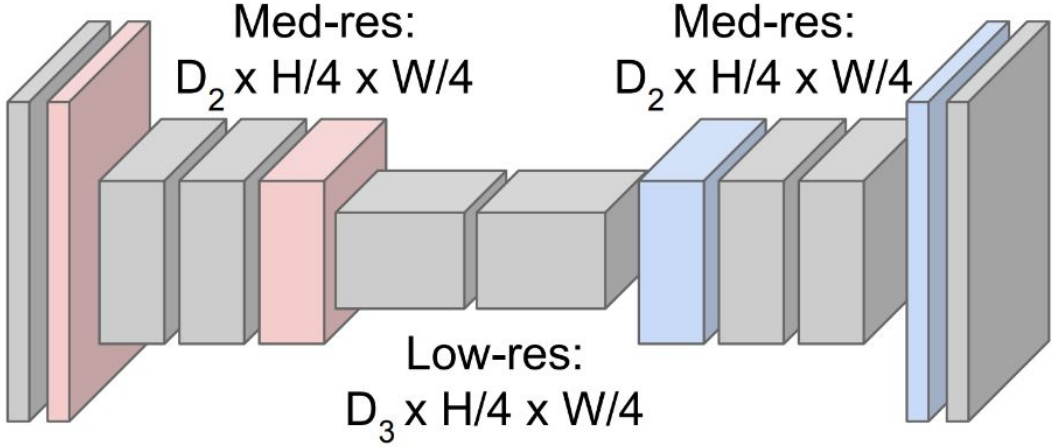
# Semantic Segmentation: Fully-Convolutional

**Downsampling**:
strided convolution, pooling

Design network as a bunch of convolutional layers, with **downsampling** and **upsampling** inside the network!

**Upsampling**:
strided transposed conv, unpooling

Input:
3 x H x W




High-res:
$D_1$ x H/2 x W/2

High-res:
$D_1$ x H/2 x W/2

C x H x W

Predictions:
H x W

Long, Shelhamer, and Darrell, "Fully Convolutional Networks for Semantic Segmentation", CVPR 2015
Noh et al, "Learning Deconvolution Network for Semantic Segmentation", ICCV 2015

ref: [35]

# Lab Exercises