# CS231n Lecture 11

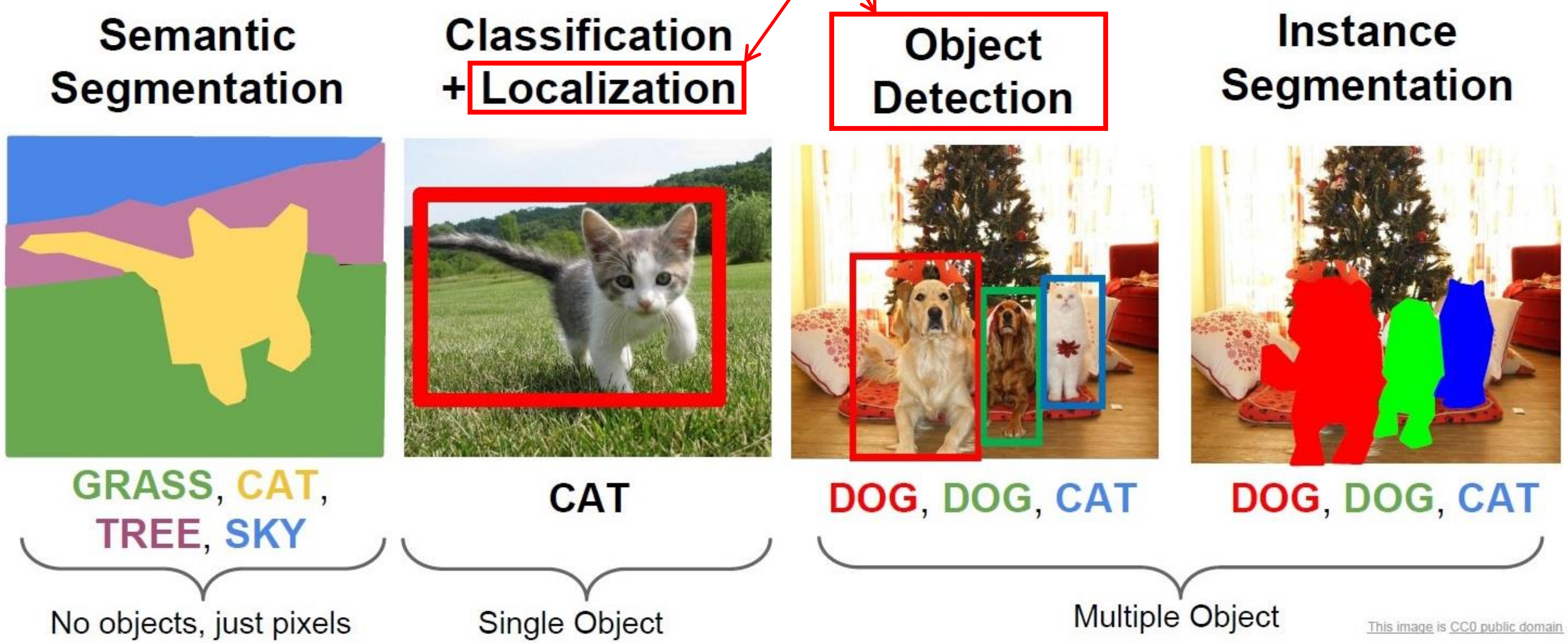

- BOAZ 10기 박성현
- BOAZ 11기 김태희
- BOAZ 11기 홍지민
- BOAZ 10기 김용규


BOAZ

01

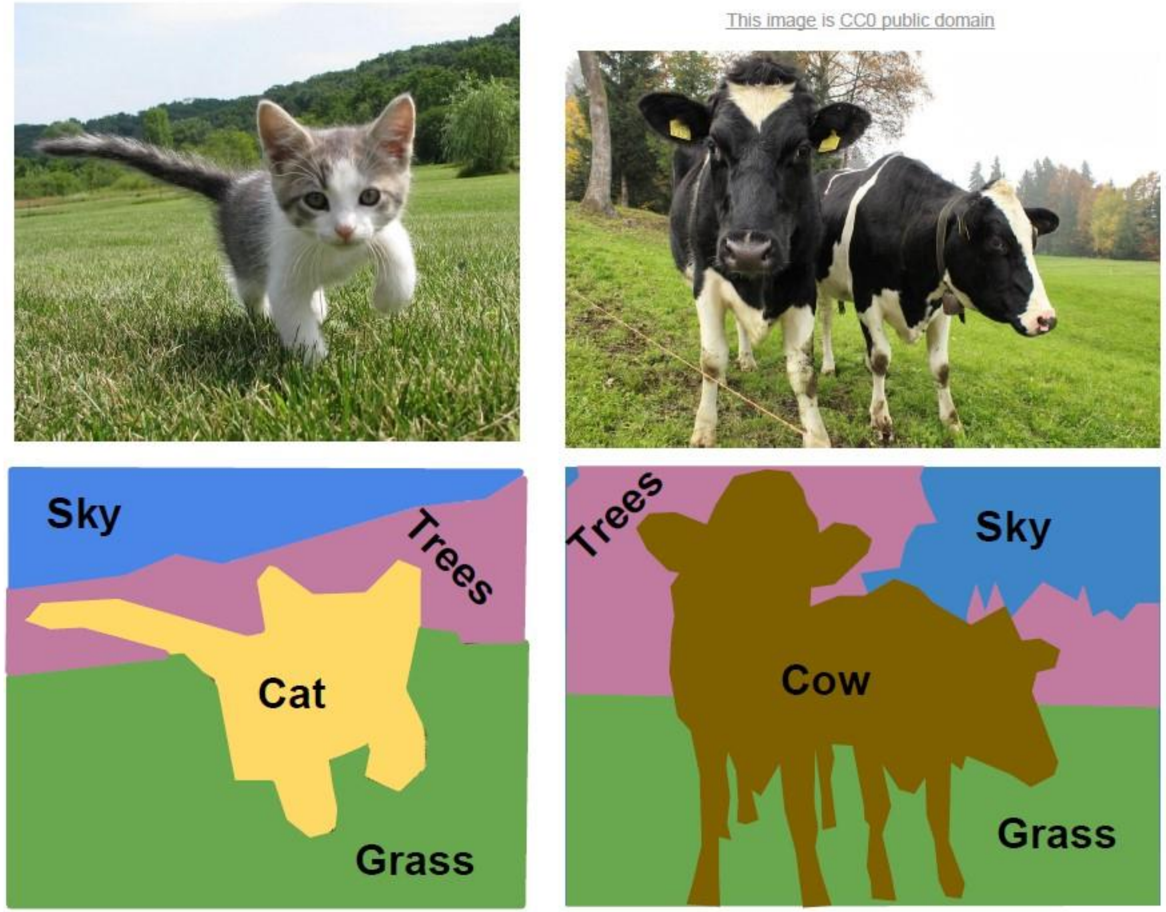
# Other Computer Vision Tasks

차이점 : 몇 개의 object를 잡아낼 수 있는지



# Semantic Segmentation

This image is CC0 public domain

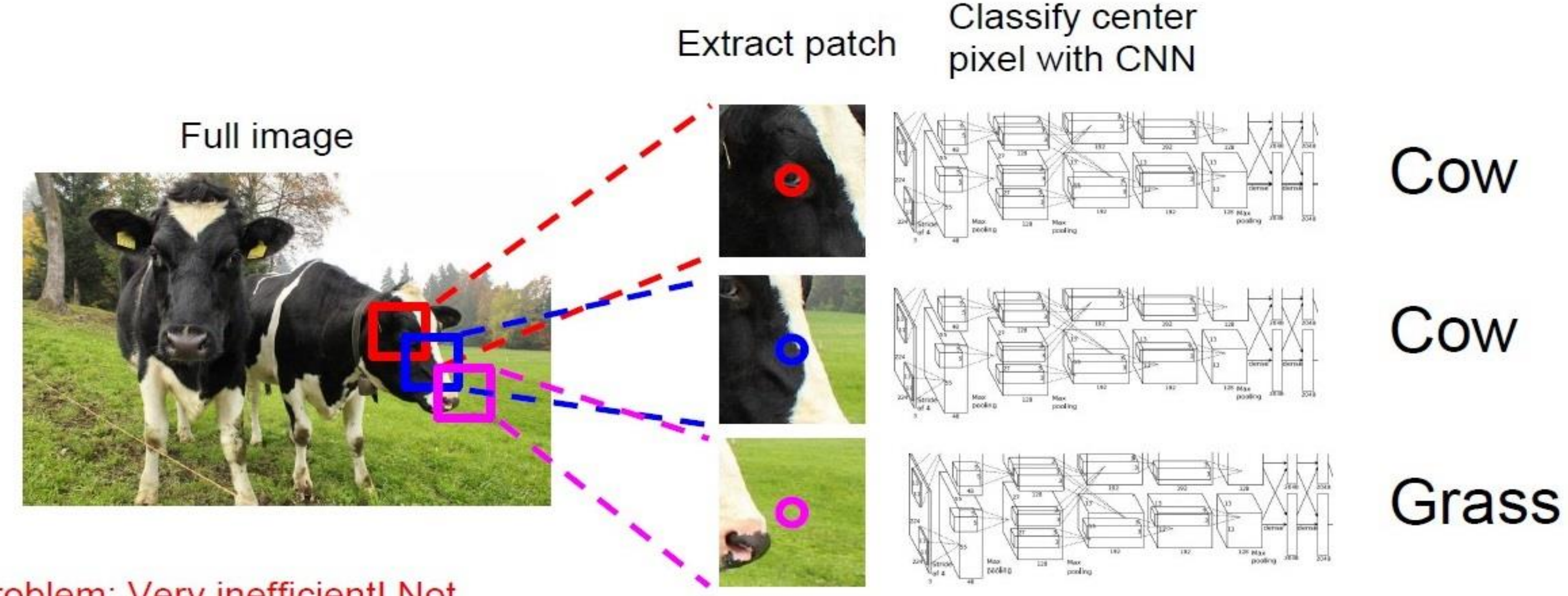



Classification : 하나의 사진 단위

Segmentation : 하나의 pixel 단위
Instance를 구분하는 게 아닌 pixel에 대해서 집중한다.

03

# Semantic Segmentation Idea : Sliding Window



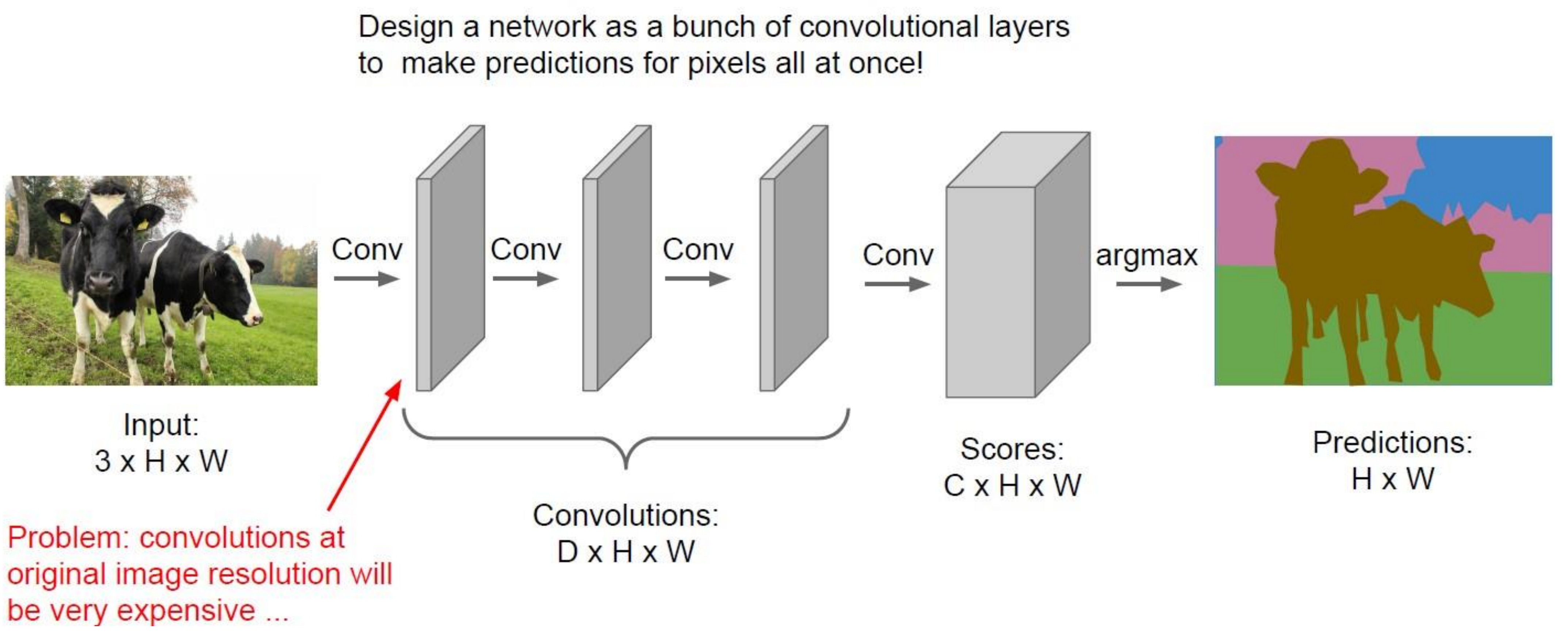

Problem: Very inefficient! Not reusing shared features between overlapping patches

Farabet et al, "Learning Hierarchical Features for Scene Labeling," TPAMI 2013
Pinheiro and Collobert, "Recurrent Convolutional Neural Networks for Scene Labeling", ICML 2014

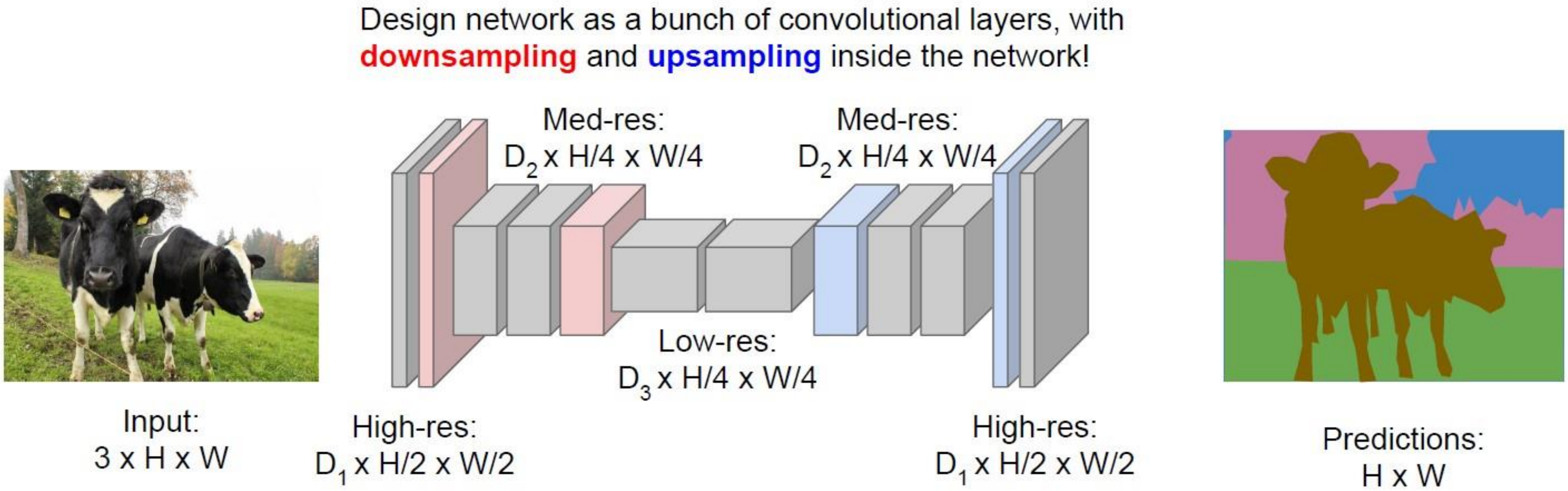
1. 작은 patch들을 일일이 CNN의 input으로 사용하는 비효율성
2. 겹치는 patch들 사이에 shared feature를 재사용하지 않음

04

# Semantic Segmentation Idea : Fully Convolutional

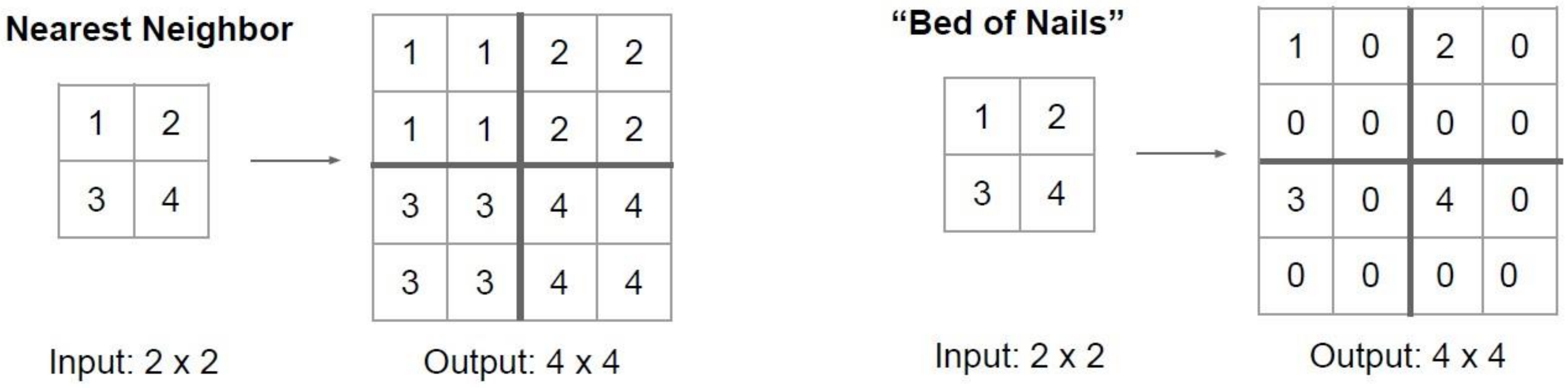


# 05 Semantic Segmentation Idea : Fully Convolutional



06

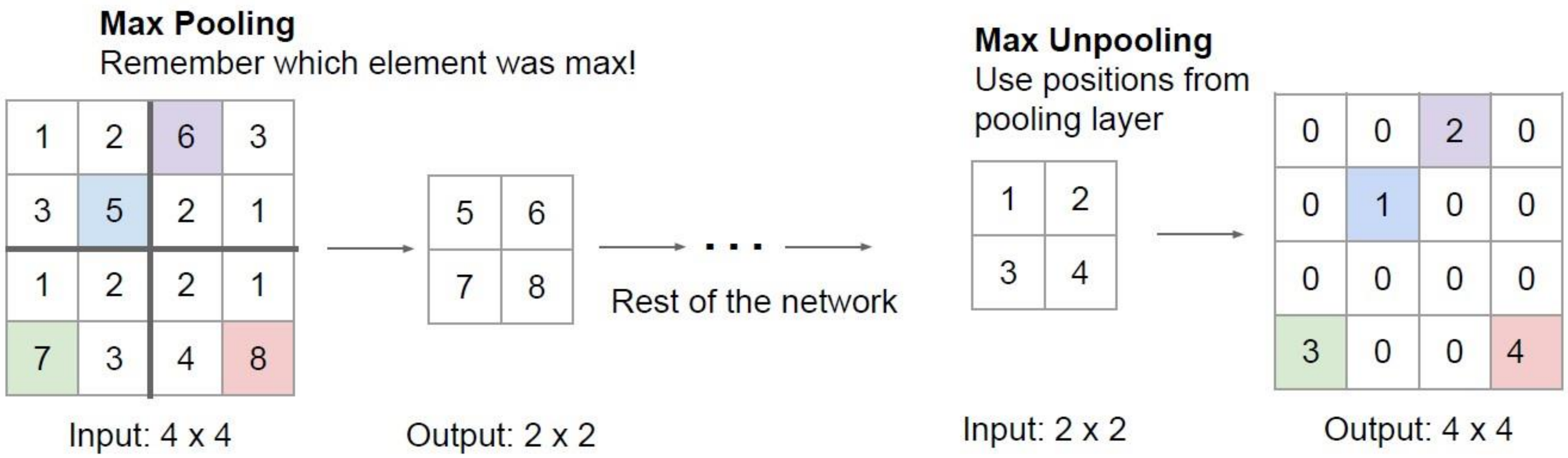
# In-Network upsampling : “Unpooling”



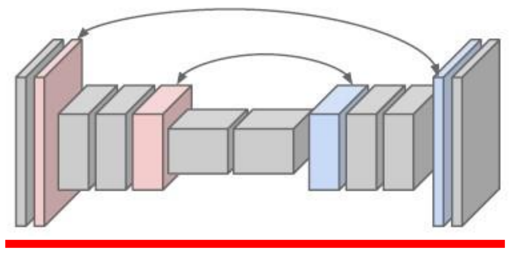

Nearest Neighbor
: 주변 값들을 모두 같은 수로 변경

Bed of Nails
: 맨 왼쪽, 맨 위의 값만을 채움

# In-Network upsampling : “Max Unpooling”



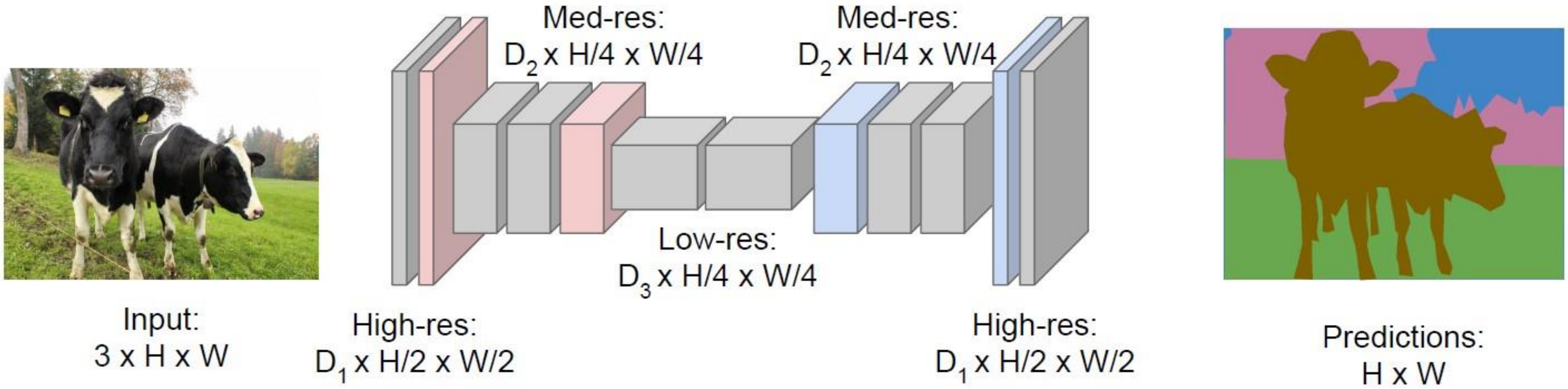

Corresponding pairs of downsampling and upsampling layers



이전에 Max Pooling 자리를 기억
이를 이용해서 Max Unpooling

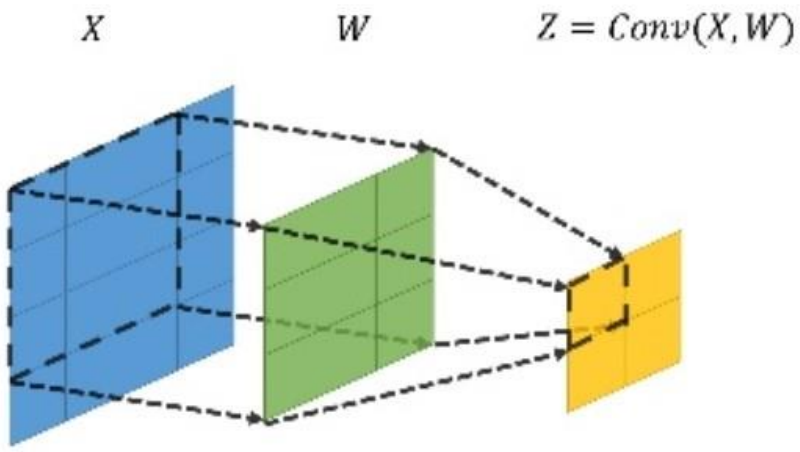
# Semantic Segmentation Idea : Fully Convolutional

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



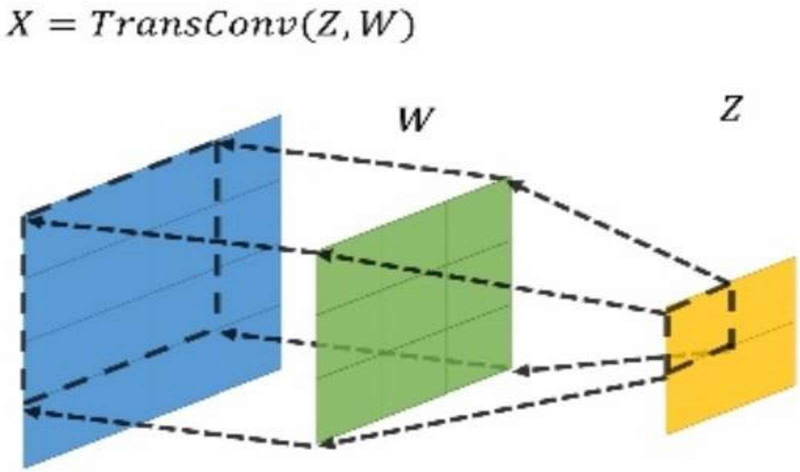
# Convolution vs Transpose Convolution

## Convolution Network




X : image
W : filter
Z : feature

## Transpose Convolution Network




X : feature
W : filter
Z : input

Transpose convolution filter 또한 convolution filter의 특징을 가지고 있다.

# Semantic Segmentation Idea : Fully Convolutional

**Downsampling**:
Pooling, strided convolution

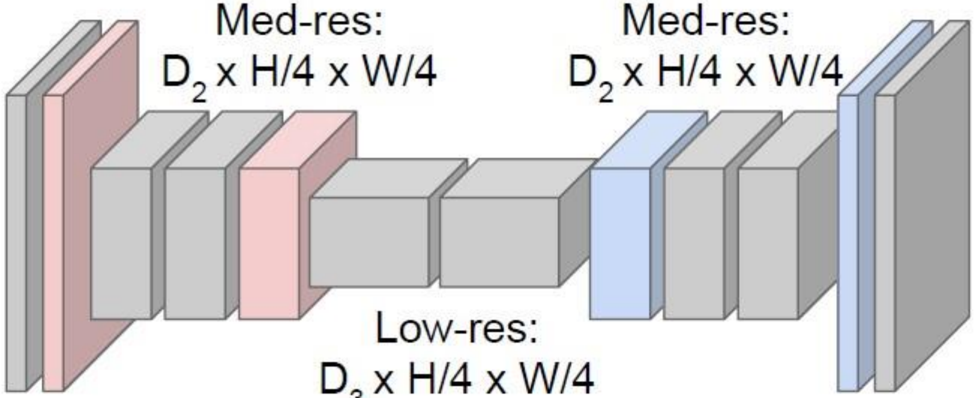
Design network as a bunch of convolutional layers, with **downsampling** and **upsampling** inside the network!

**Upsampling**:
Unpooling or strided transpose convolution

Input:
3 x H x W

Med-res:
$D_2$ x H/4 x W/4

Med-res:
$D_2$ x H/4 x W/4



Low-res:
$D_3$ x H/4 x W/4

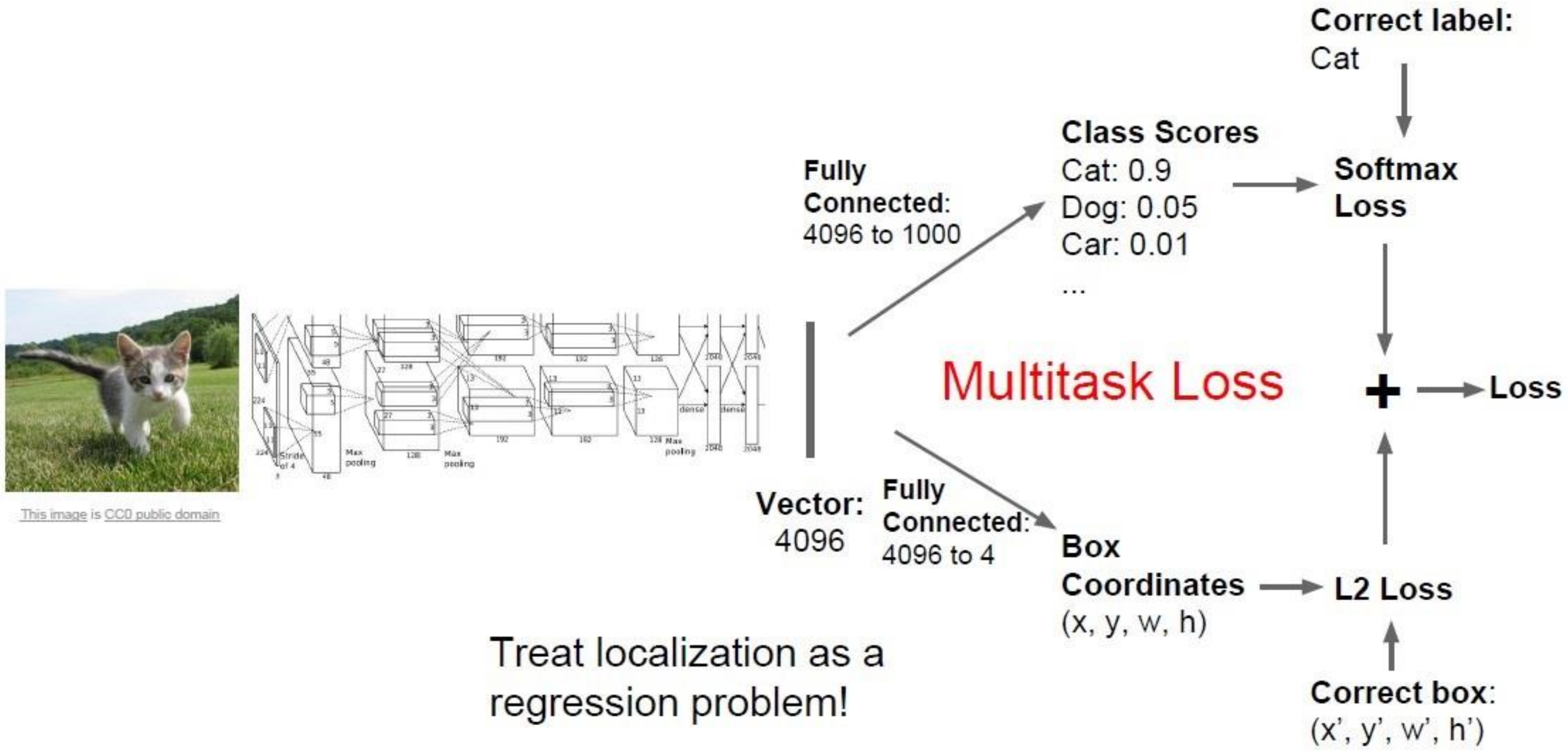
High-res:
$D_1$ x H/2 x W/2

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

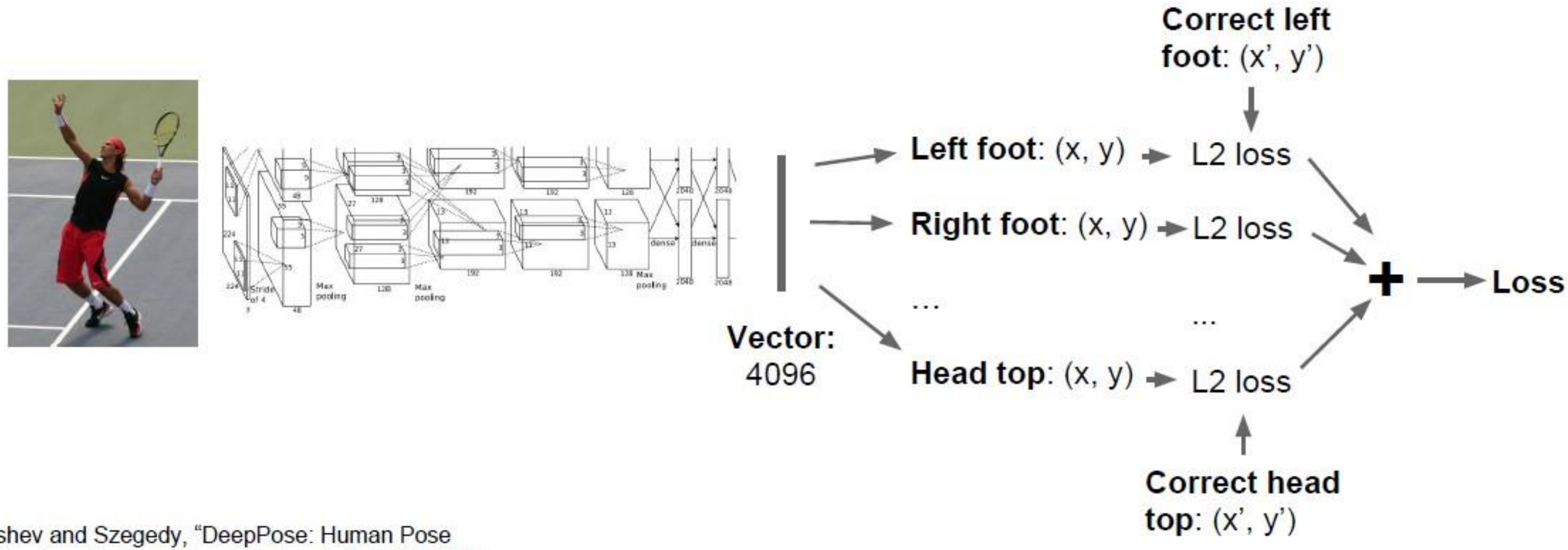
Predictions:
H x W

11

# Classification + Localization



# Human Pose Estimation




Toshev and Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks", CVPR 2014

End

# 참고 자료


CS231n : http://cs231n.stanford.edu/syllabus.html

website : https://www.slideshare.net/ssuserb208cc1/transposed-convolution


BOAZ