논문들로 알아보는

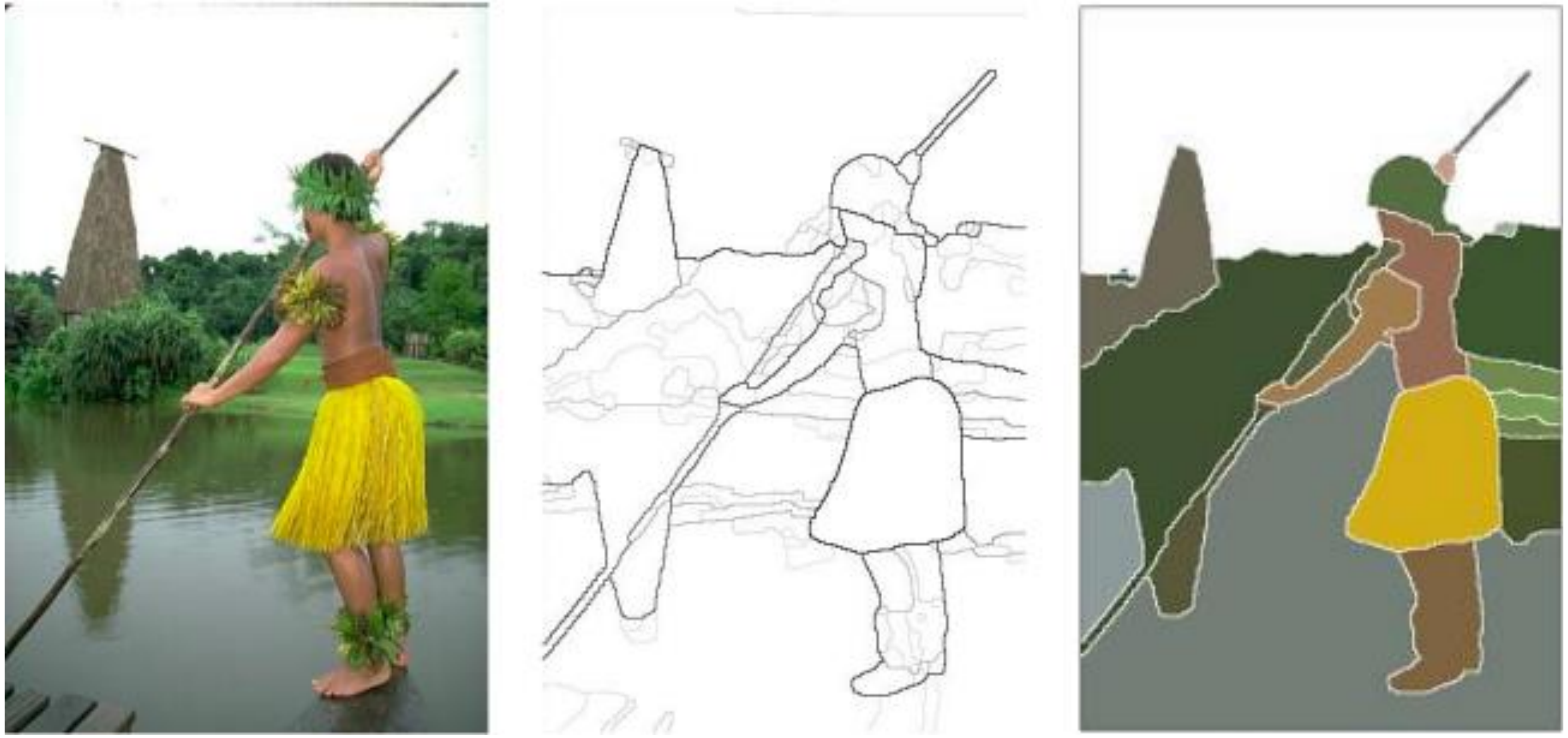
# Basic Algorithms of Semantic Segmentation

rayjang111@gmail.com
HyunSukJang

# 01 Introduction

# 01. Introduction

What is Semantic Segmentation??



# 01. Introduction

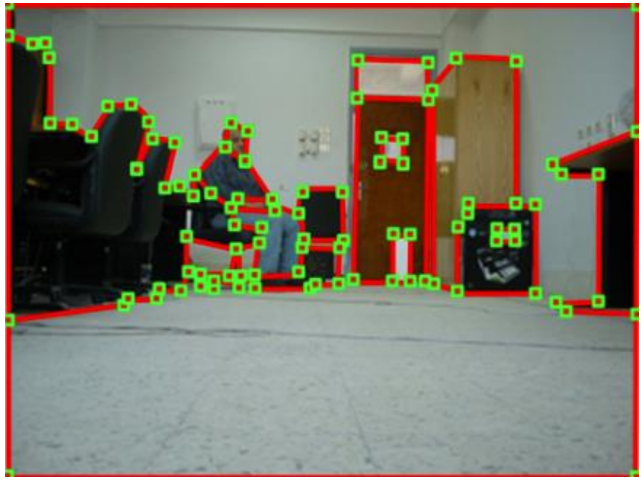

1.recognizing
2.understanding
what's in the image in pixel level.

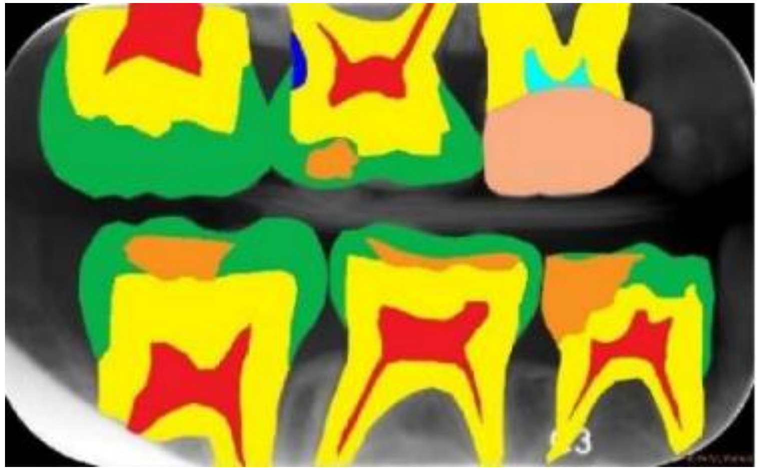
# Applications of Semantic segmentation

robot vision and understanding

autonomous driving

medical purposes (IS BI Challenge)





## 01. Introduction

### Introducing 4 Main Papers of Semantic Segmentation

1. Fully Convolutional Networks for Semantic Segmentation

2. Semantic Image Segmentation with deep convolutional nets and fully connected CRFs (DeepLab V1)

3. Learning Deconvolution Network for Semantic Segmentation

4. Semantic Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs

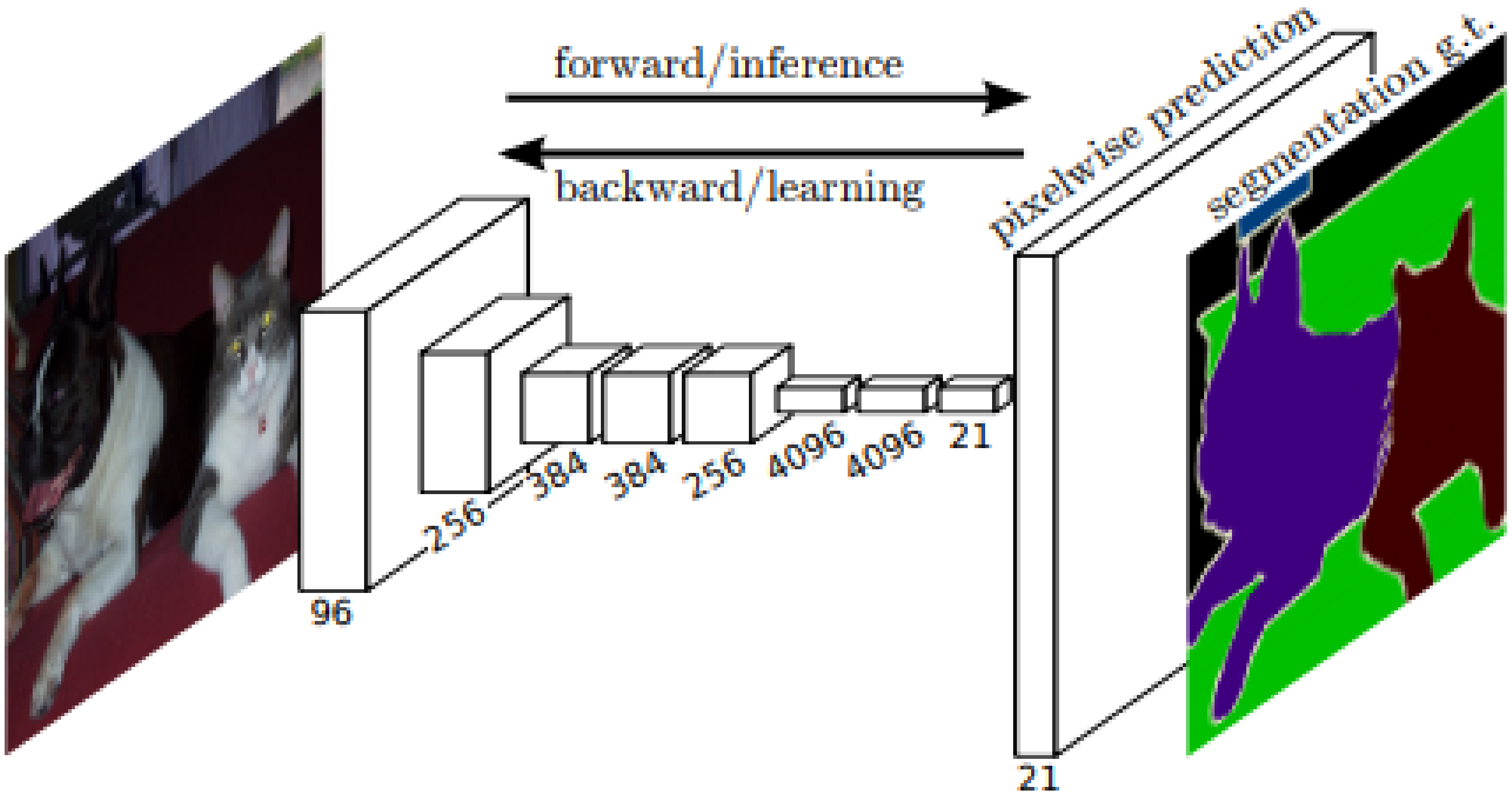
# 02FCN(Fully Convolutional Network)

# Fully Convolutional Networks for Semantic Segmentation

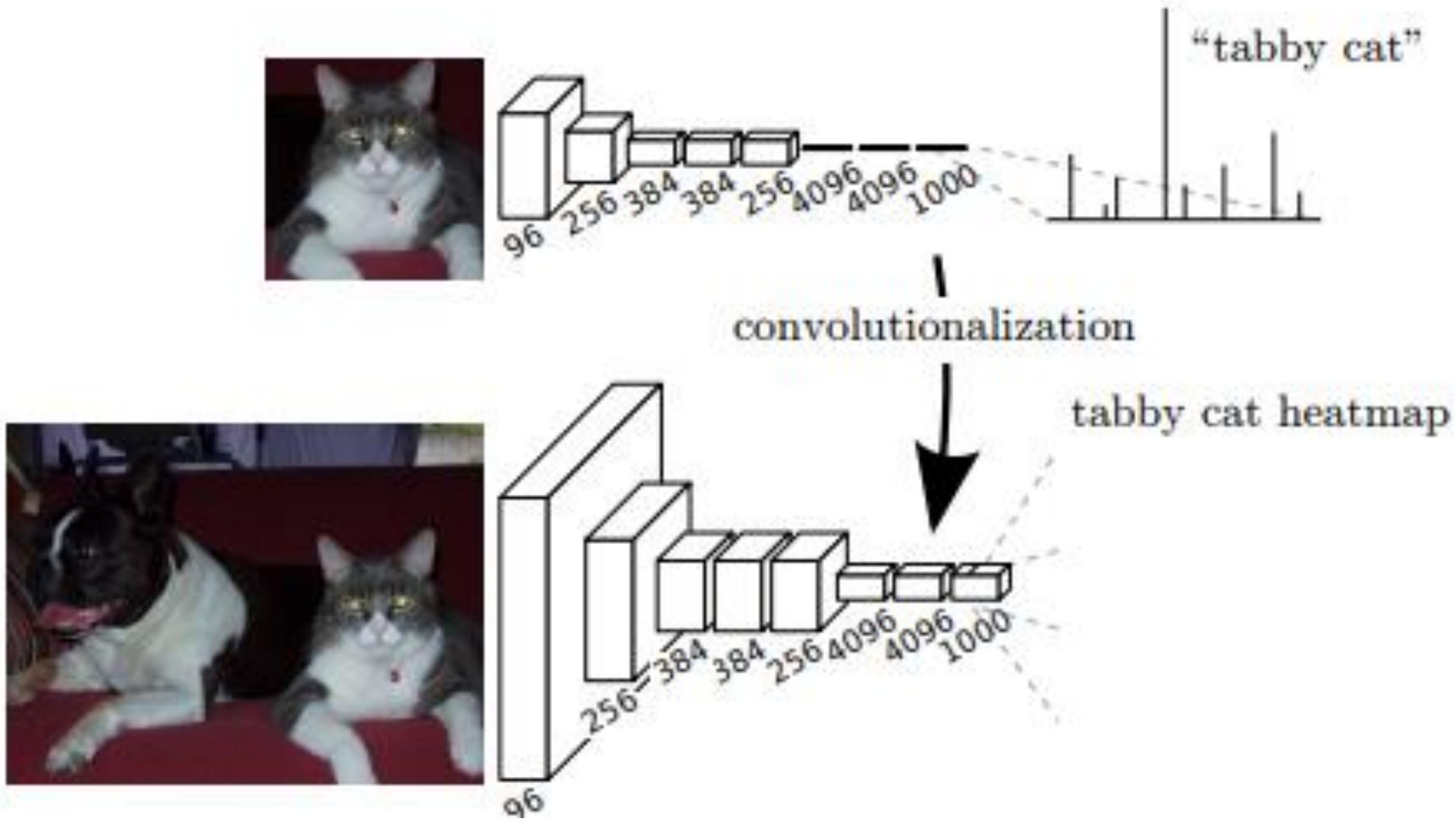
Jonathan Long* Evan Shelhamer* Trevor Darrell
UC Berkeley
{jonlong,shelhamer,trevor}@cs.berkeley.edu

# 02. FCN

## Pixelwise prediction with end to end learning



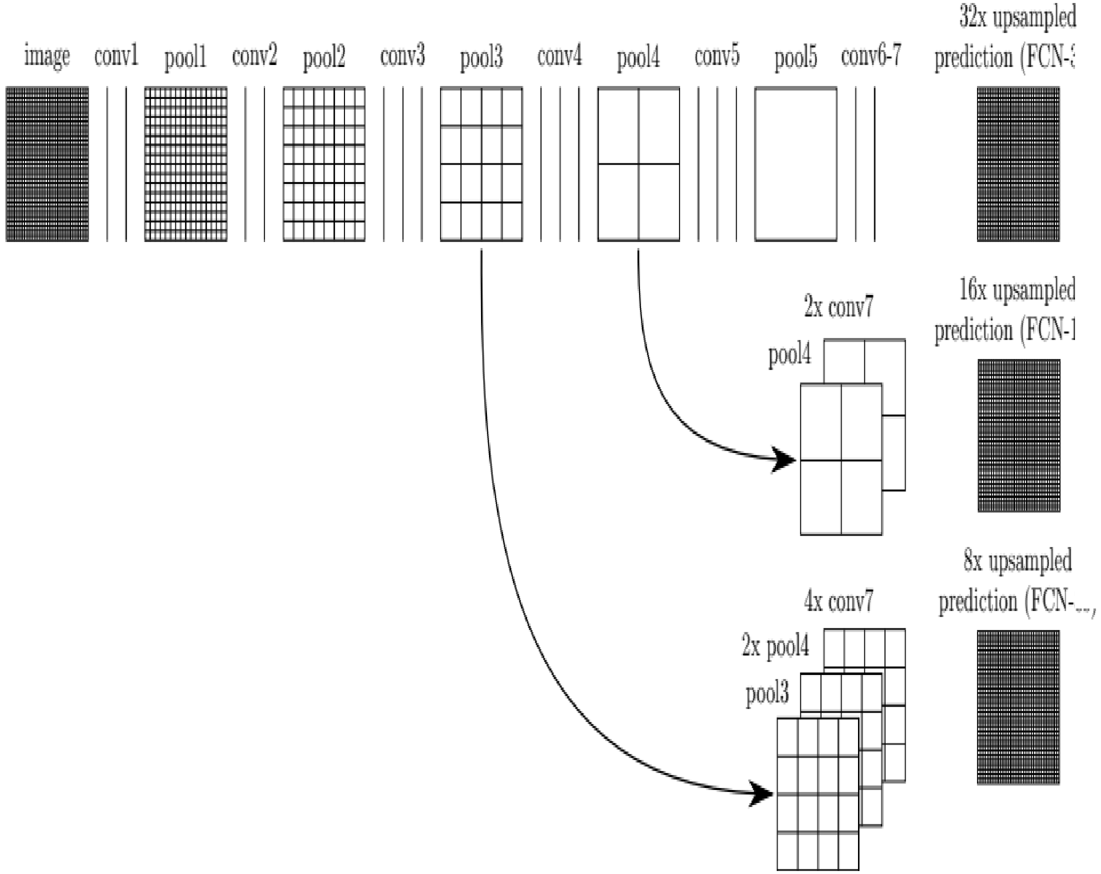
# Convolutionize

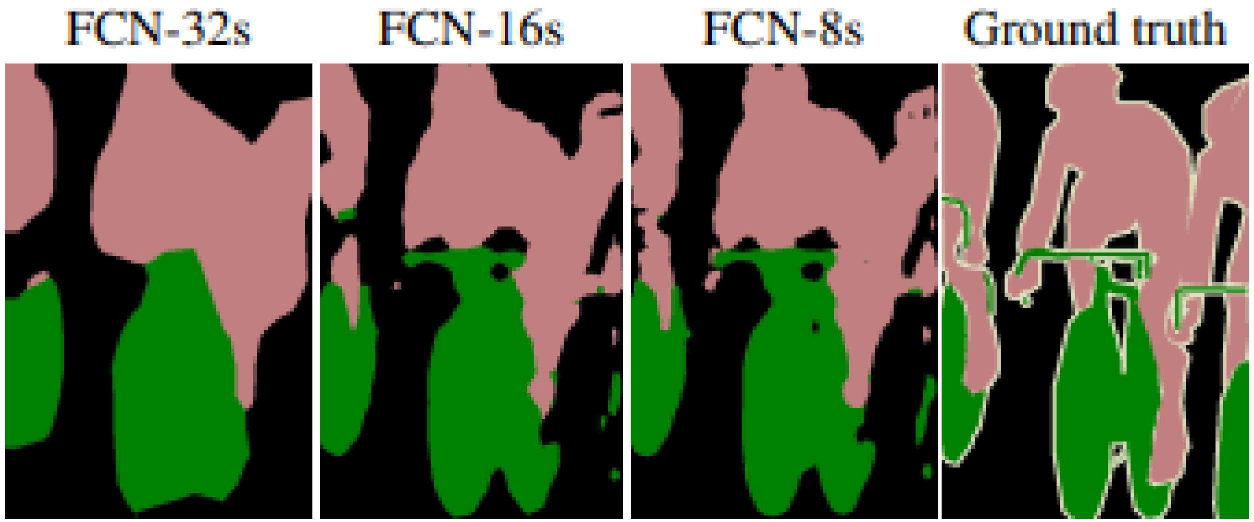



Fully convolutional versions of existing networks predict dense outputs from arbitray-sized inputs

# 02. FCN

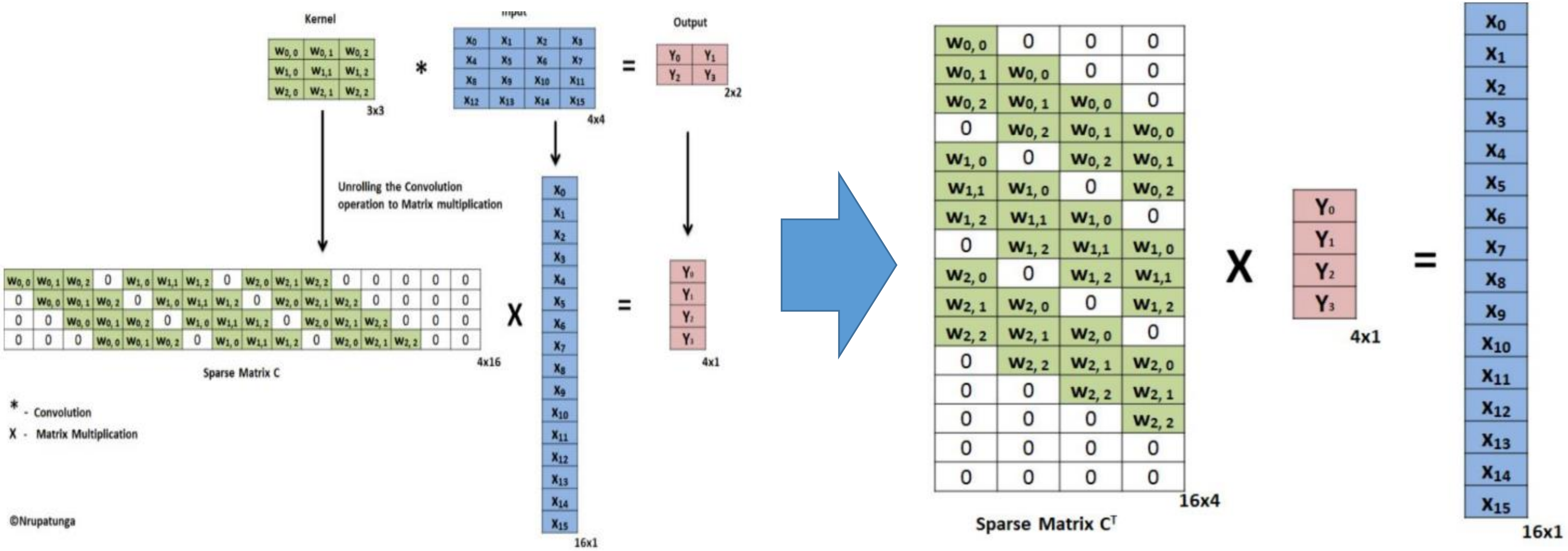
## Skip Connection



# 02. FCN

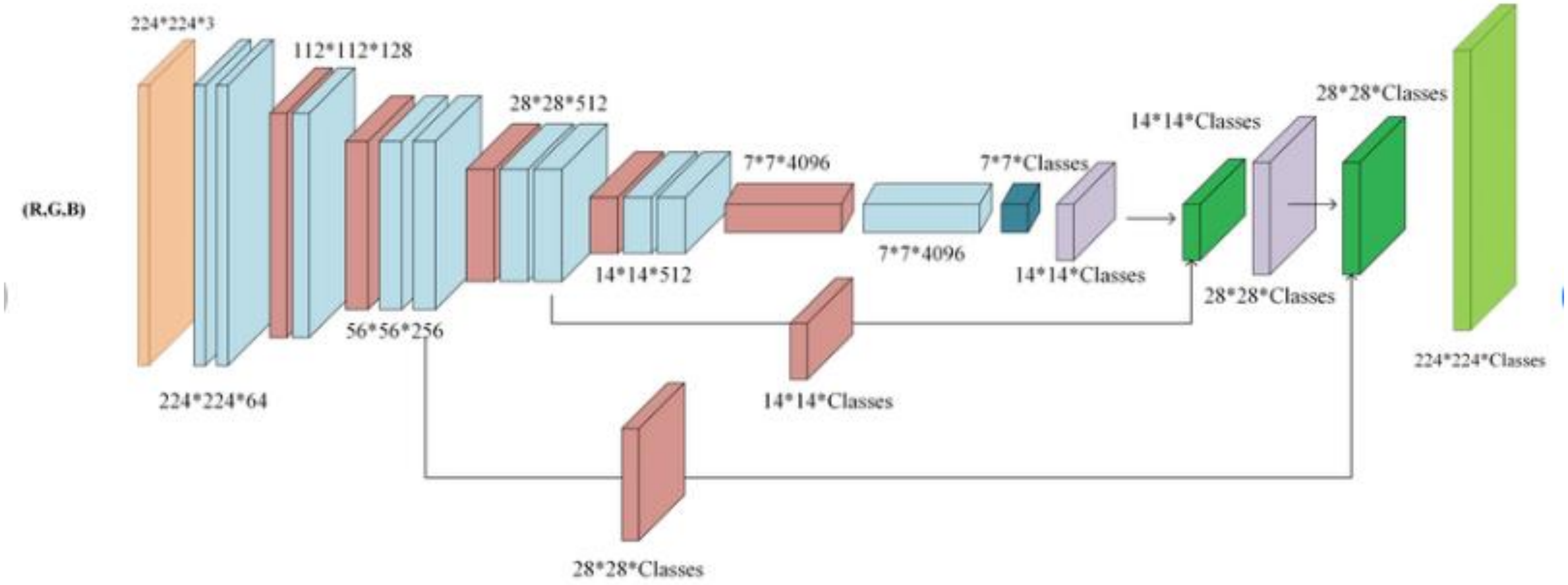
## Transposed Convolution(Deconvolution)



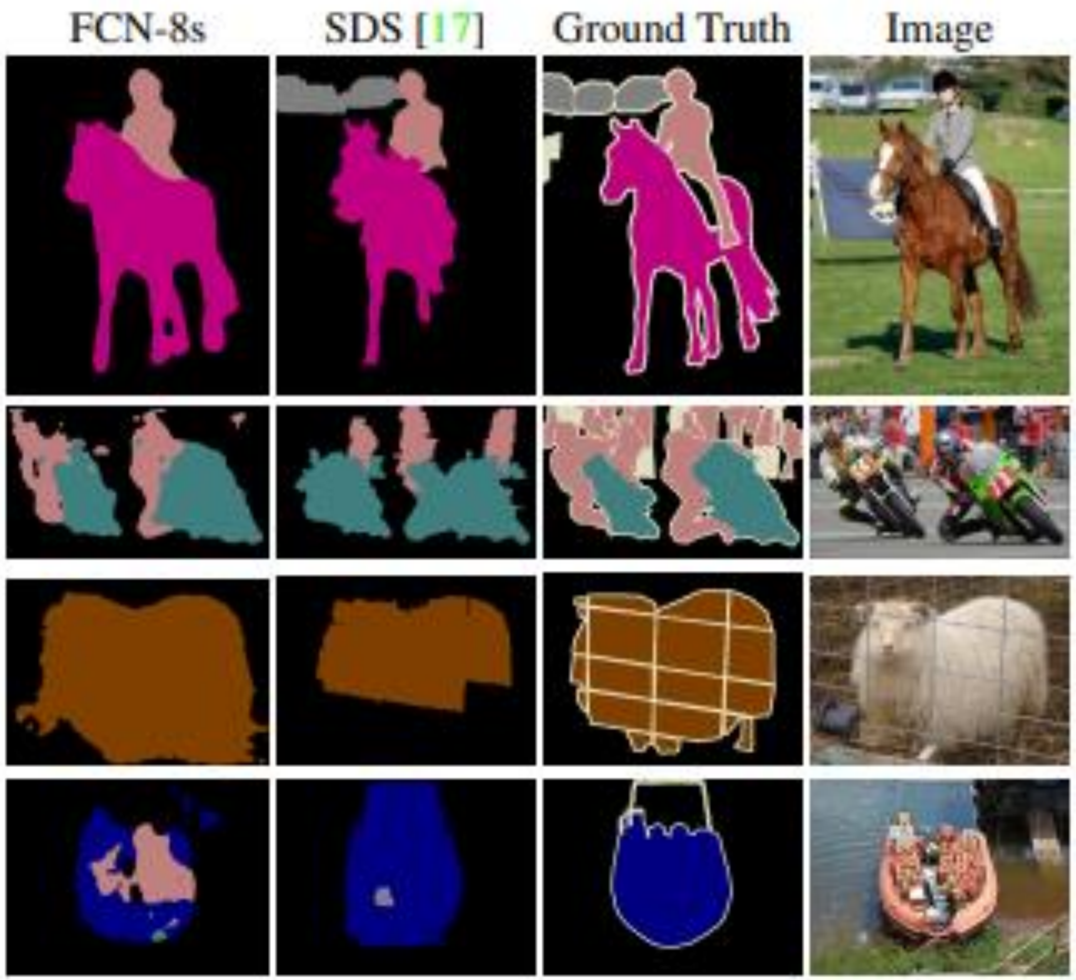

수학적 역관계는 아니지만 연결된 위치를 활용할 수 있다는 점에서 유용한 일종의 interpolation 기법

# 02. FCN

## FCN Architecture



# 02. FCN




Works well, but needs to catch more details

# 03. DeepLabV1

SEMANTIC IMAGE SEGMENTATION WITH DEEP CONVOLUTIONAL NETS AND FULLY CONNECTED CRFS

**Liang-Chieh Chen**
Univ. of California, Los Angeles
lcchen@cs.ucla.edu

**George Papandreou** *
Google Inc.
gpapan@google.com

**Iasonas Kokkinos**
CentraleSupélec and INRIA
iasonas.kokkinos@ecp.fr

**Kevin Murphy**
Google Inc.
kpmurphy@google.com

**Alan L. Yuille**
Univ. of California, Los Angeles
yuille@stat.ucla.edu

ABSTRACT

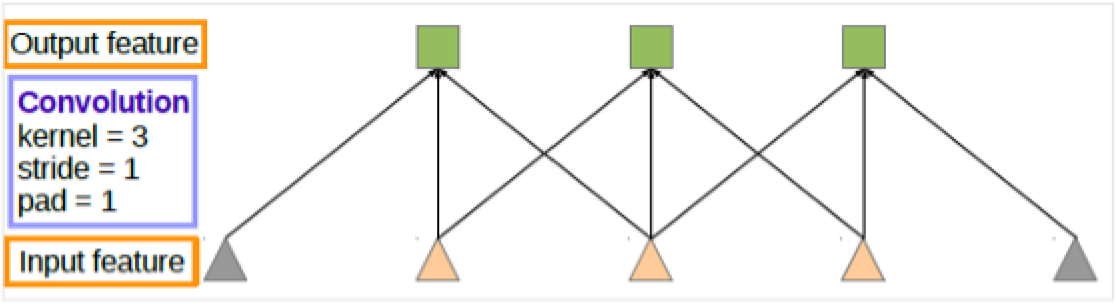
Deep Convolutional Neural Networks (DCNNs) have recently shown state of the art performance in high level vision tasks, such as image classification and object detection. This work brings together methods from DCNNs and probabilistic graphical models for addressing the task of pixel-level classification (also called "semantic image segmentation"). We show that responses at the final layer of DCNNs are not sufficiently localized for accurate object segmentation. This is due to the very invariance properties that make DCNNs good for high level tasks. We overcome this poor localization property of deep networks by combining the responses at the final DCNN layer with a fully connected Conditional Random Field (CRF). Qualitatively, our "DeepLab" system is able to localize segment boundaries at a level of accuracy which is beyond previous methods. Quantitatively, our method sets the new state-of-art at the PASCAL VOC-2012 semantic image segmentation task, reaching 71.6% IOU accuracy in the test set. We show how these results can be obtained efficiently: Careful network re-purposing and a novel application of the 'hole' algorithm from the wavelet community allow dense computation of neural net responses at 8 frames per second on a modern GPU.

# Two contributions

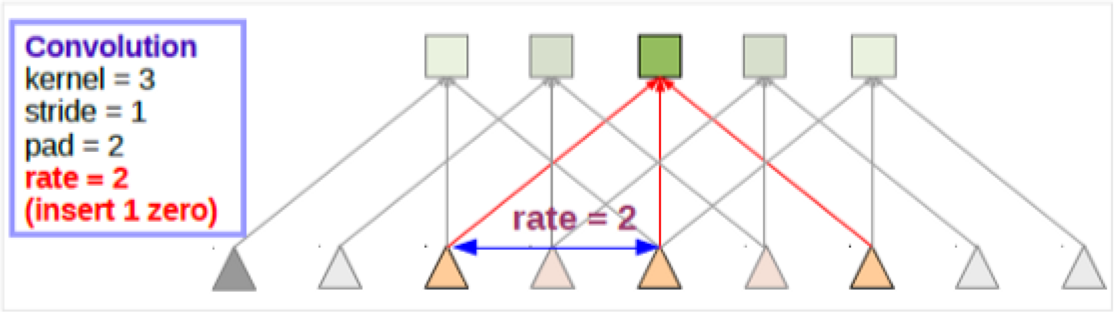
1.Signal downsmapling:
=>Atrous Algorithm

2.Spatial intensity:
=>Conditional random field

# Atrous Convolution(Dilated Convolution)




(a) Sparse feature extraction



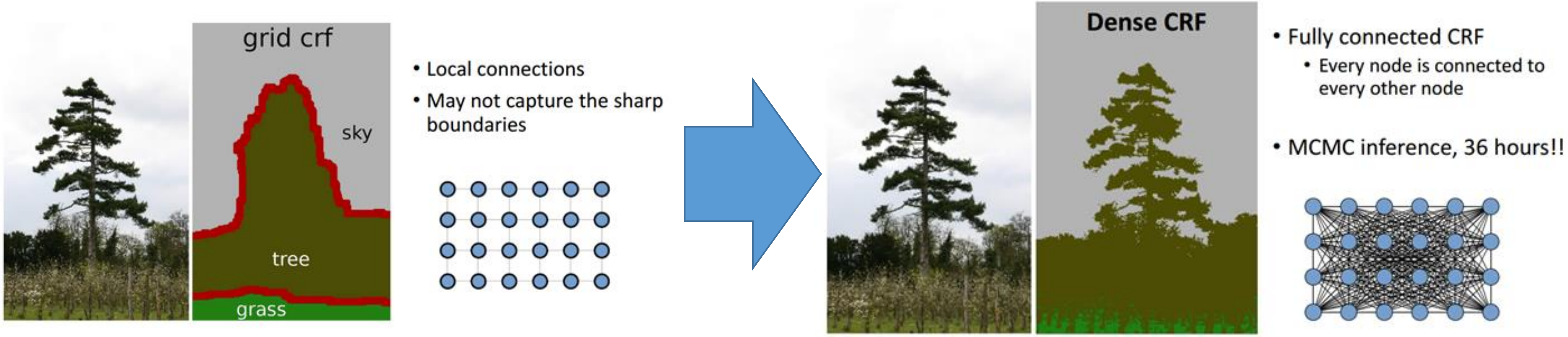

(b) Dense feature extraction

1. makes the output feature map larger

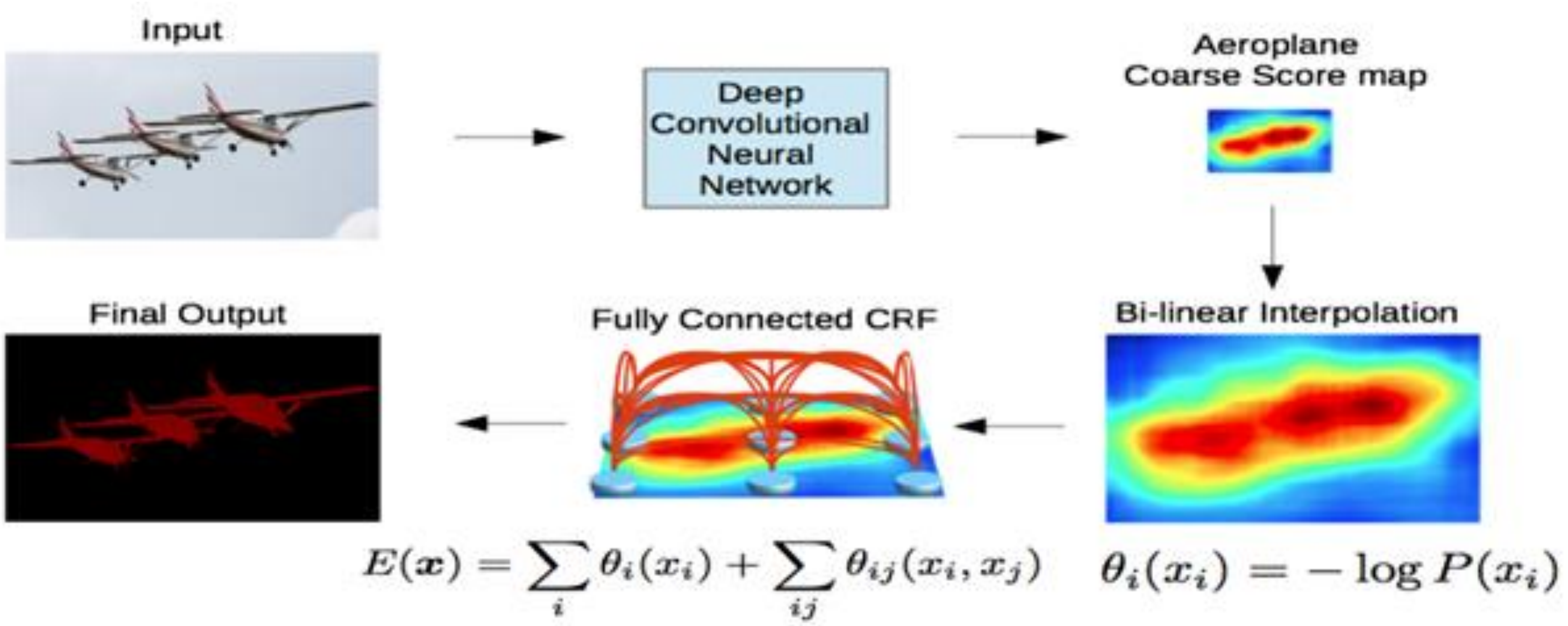
2. allows us to enlarge the field o f view of filters to incorporate l arger context

03. DeepLabV1

# Conditional Random Field



# Architecture



# 03. DeepLabV1

(a) FCN-8s vs. DeepLab-CRF

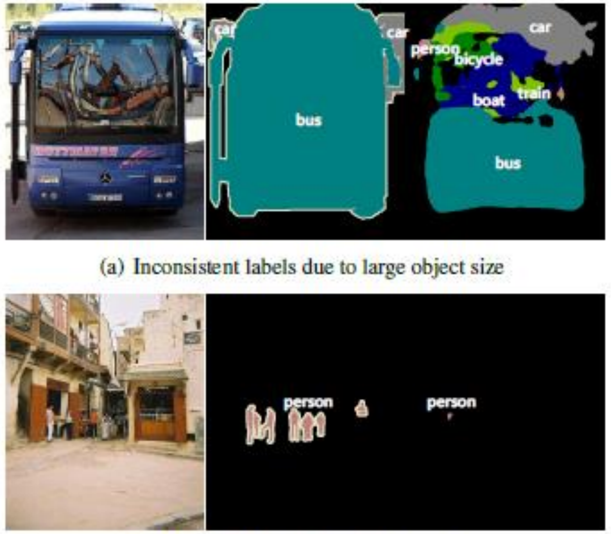
# 04. Deconvolutional Network

## Learning Deconvolution Network for Semantic Segmentation

Hyeonwoo Noh  Seunghoon Hong  Bohyung Han
Department of Computer Science and Engineering, POSTECH, Korea
{hyeonwoonoh_,maga33,bhhan}@postech.ac.kr

### Abstract

*We propose a novel semantic segmentation algorithm by learning a deep deconvolution network. We learn the network on top of the convolutional layers adopted from VGG 16-layer net. The deconvolution network is composed of deconvolution and unpooling layers, which identify pixel-wise class labels and predict segmentation masks. We apply the trained network to each proposal in an input image, and construct the final semantic segmentation map by combining the results from all proposals in a simple manner. The proposed algorithm mitigates the limitations of the existing methods based on fully convolutional networks by integrating deep deconvolution network and proposal-wise prediction; our segmentation method typically identifies detailed structures and handles objects in multiple scales naturally. Our network demonstrates outstanding performance in PASCAL VOC 2012 dataset, and we achieve the best accuracy (72.5%) among the methods trained without using Microsoft COCO dataset through ensemble with the fully convolutional network.*




(a) Inconsistent labels due to large object size

(b) Missing labels due to small object size

Figure 1. Limitations of semantic segmentation algorithms based on fully convolutional network. (Left) original image. (Center) ground-truth annotation. (Right) segmentations by [19]

# Two Problems

1.Network has a predifined fixed size receptive filed
=>Too small or too big objects neglected

2. Details are often lost

# 04. Deconvolutional Network

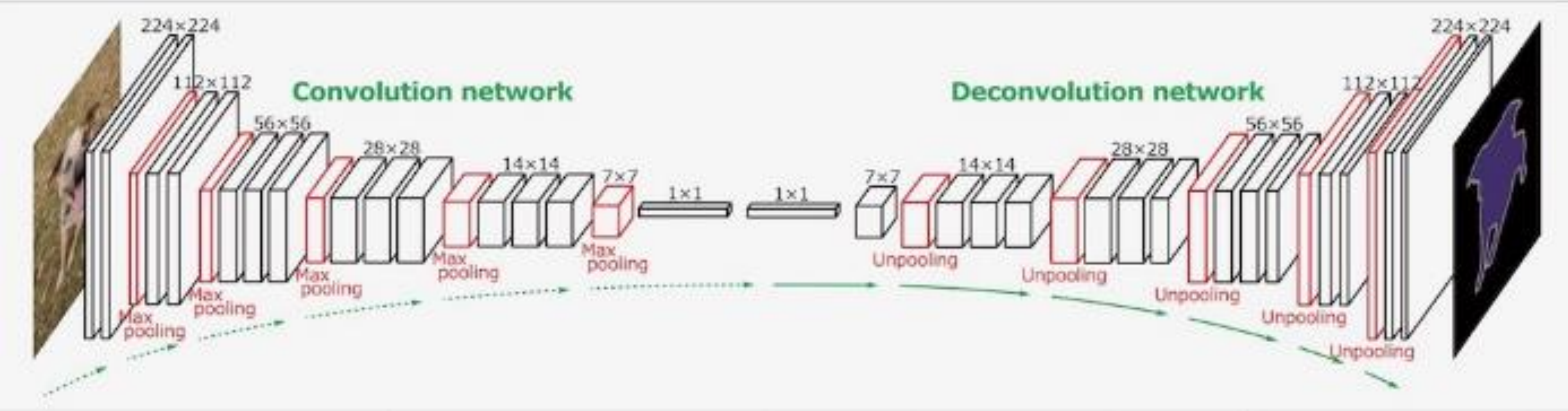
## Too small or too big objects neglected


(a) Inconsistent labels due to large object size


(b) Missing labels due to small object size

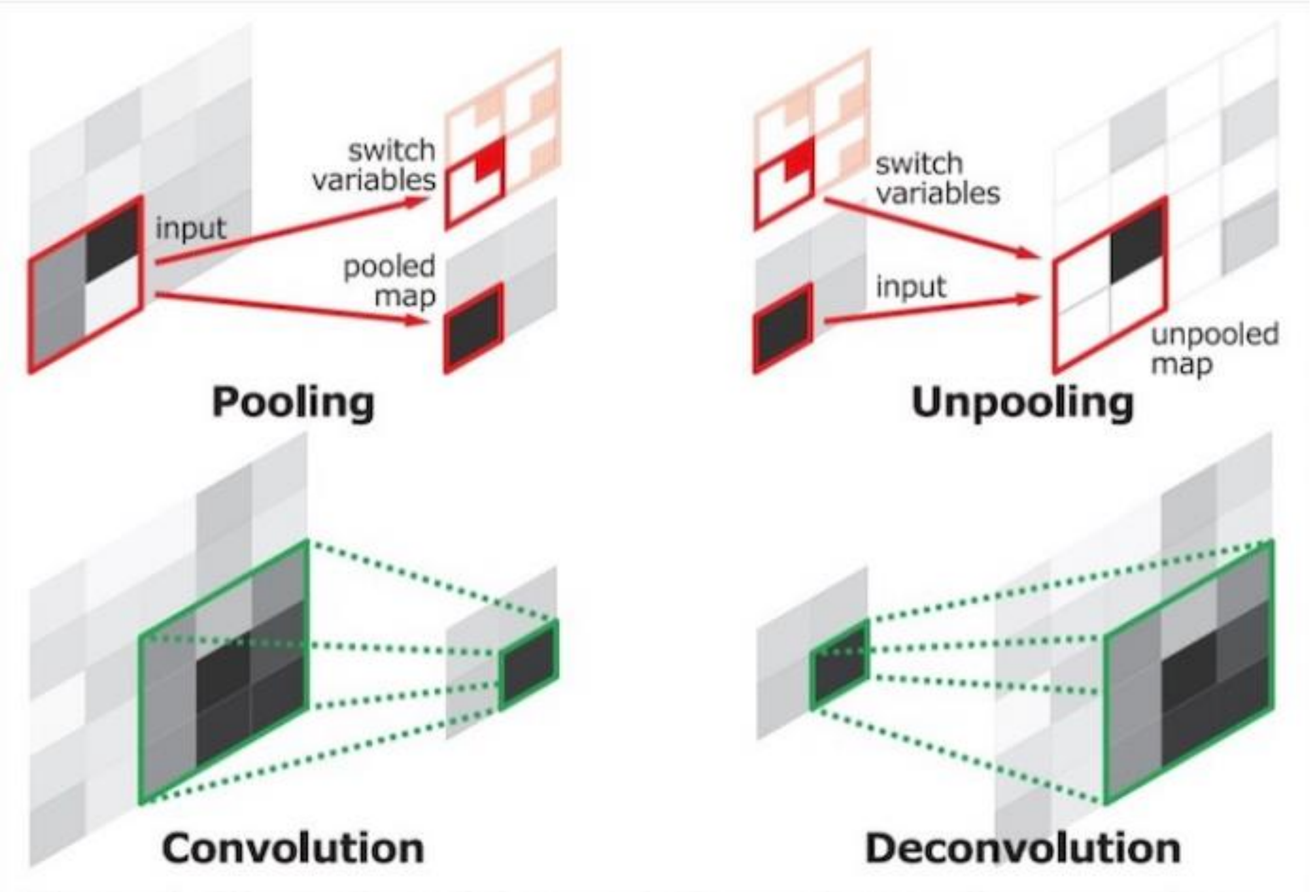
# 04. Deconvolutional Network

## Deconvolutional layers

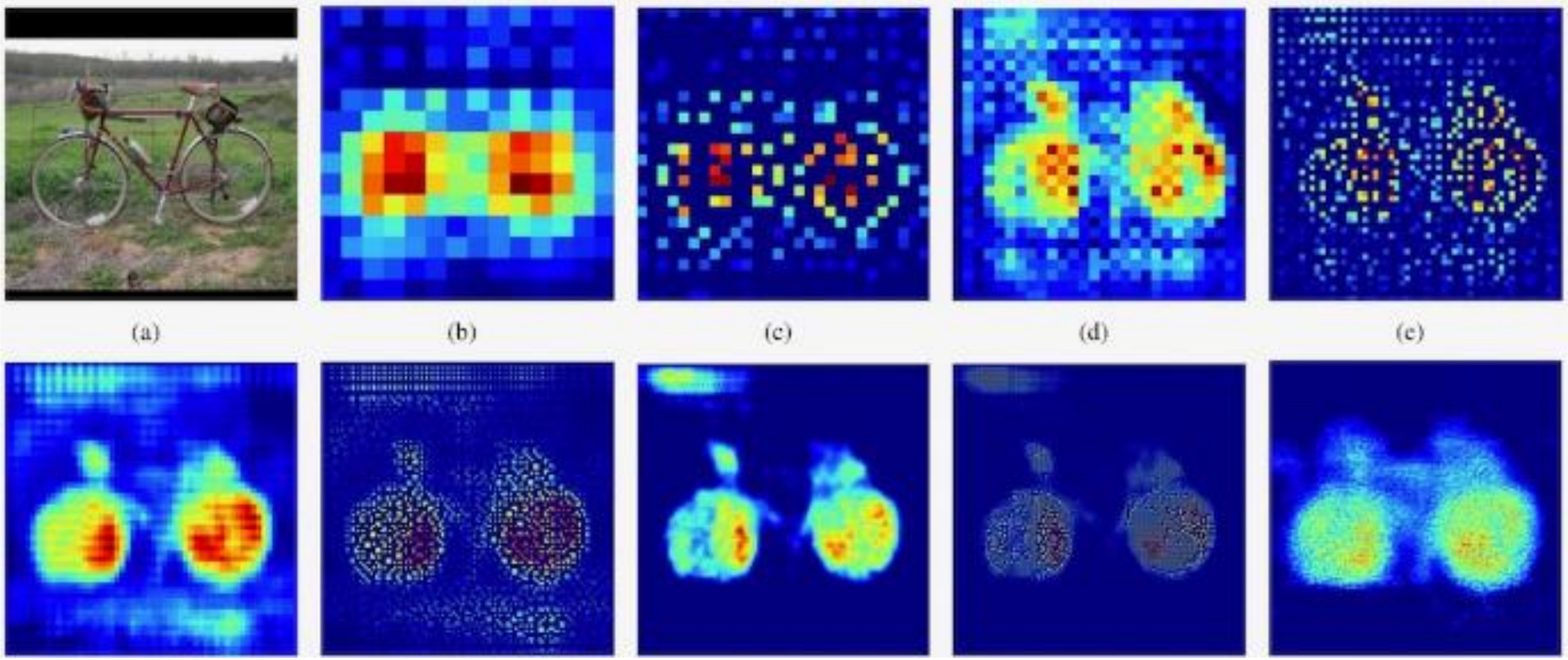


# 04. Deconvolutional Network

## Unpooling layer



# 04. Deconvolutional Network



# 05. DeepLab V2

# DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs

Liang-Chieh Chen, George Papandreou, *Senior Member, IEEE*, Iasonas Kokkinos, *Member, IEEE*, Kevin Murphy, and Alan L. Yuille, *Fellow, IEEE*

**Abstract**—In this work we address the task of semantic image segmentation with Deep Learning and make three main contributions that are experimentally shown to have substantial practical merit. *First*, we highlight convolution with upsampled filters, or 'atrous convolution', as a powerful tool in dense prediction tasks. Atrous convolution allows us to explicitly control the resolution at which feature responses are computed within Deep Convolutional Neural Networks. It also allows us to effectively enlarge the field of view of filters to incorporate larger context without increasing the number of parameters or the amount of computation. *Second*, we propose atrous spatial pyramid pooling (ASPP) to robustly segment objects at multiple scales. ASPP probes an incoming convolutional feature layer with filters at multiple sampling rates and effective fields-of-views, thus capturing objects as well as image context at multiple scales. *Third*, we improve the localization of object boundaries by combining methods from DCNNs and probabilistic graphical models. The commonly deployed combination of max-pooling and downsampling in DCNNs achieves invariance but has a toll on localization accuracy. We overcome this by combining the responses at the final DCNN layer with a fully connected Conditional Random Field (CRF), which is shown both qualitatively and quantitatively to improve localization performance. Our proposed "DeepLab" system sets the new state-of-art at the PASCAL VOC-2012 semantic image segmentation task, reaching 79.7% mIOU in the test set, and advances the results on three other datasets: PASCAL-Context, PASCAL-Person-Part, and Cityscapes. All of our code is made publicly available online.

**Index Terms**—Convolutional Neural Networks, Semantic Segmentation, Atrous Convolution, Conditional Random Fields.

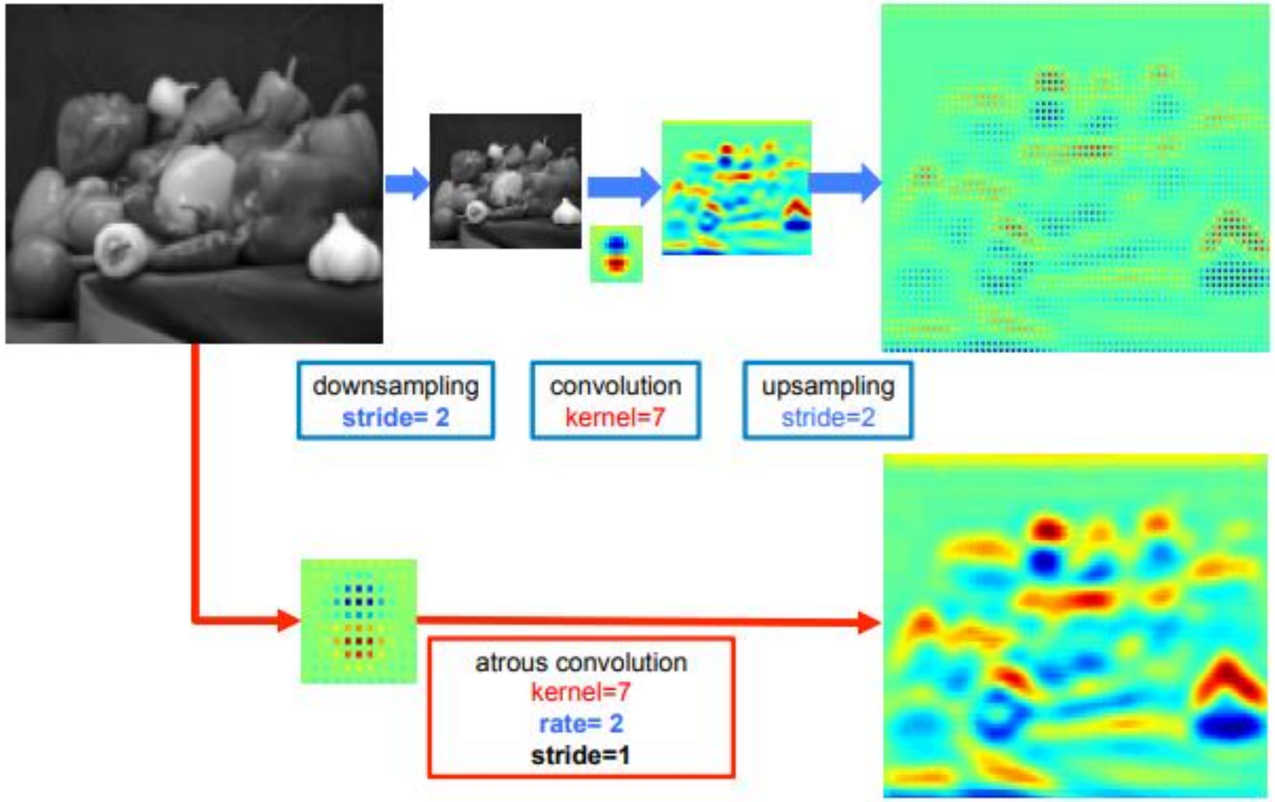
# Challenges

1. Reduced feature resolution

2. Existence of objects at multiple scales

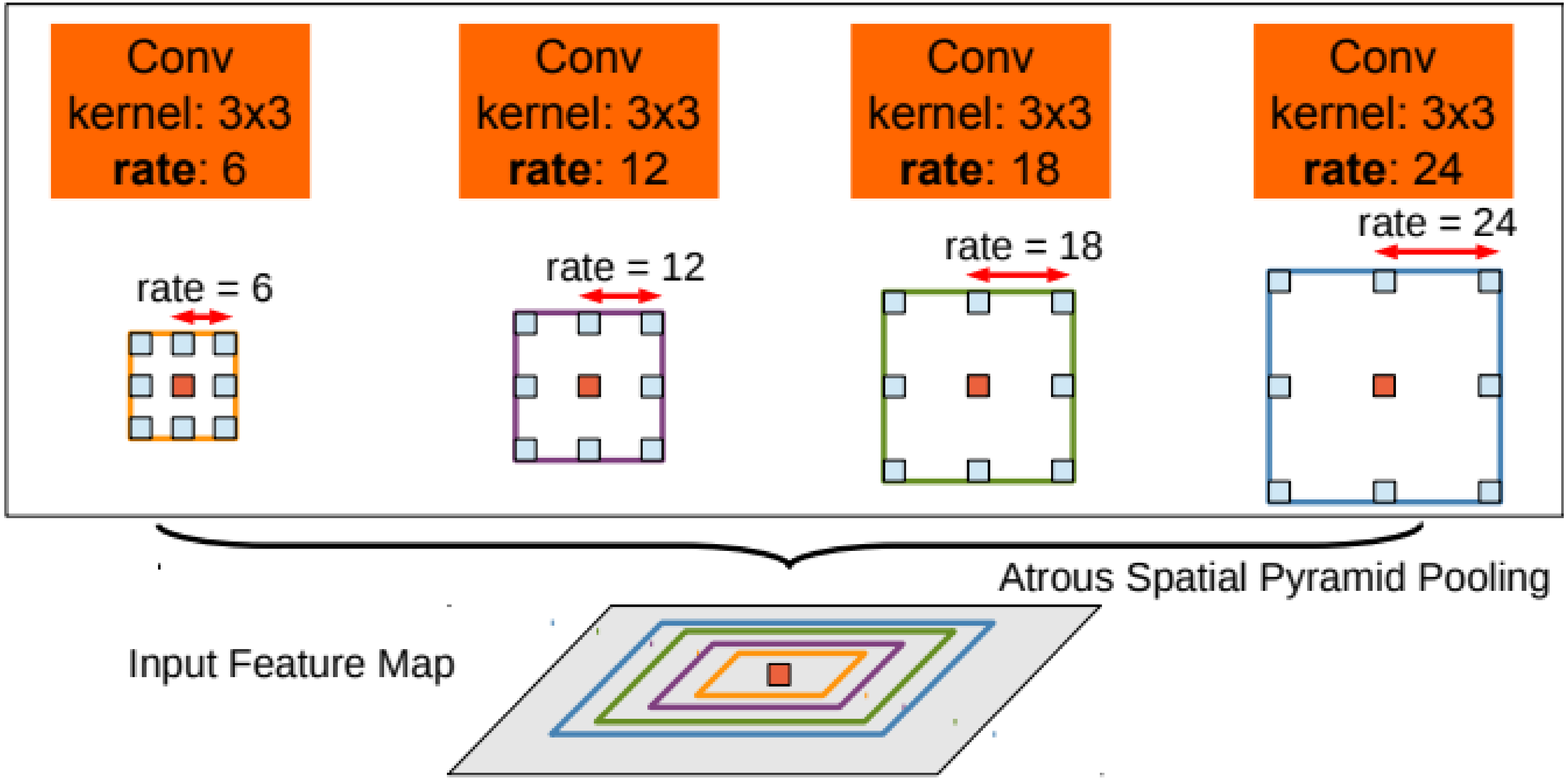
05. DeepLab V2

# Pooling+ CNN vs Atrous convolution

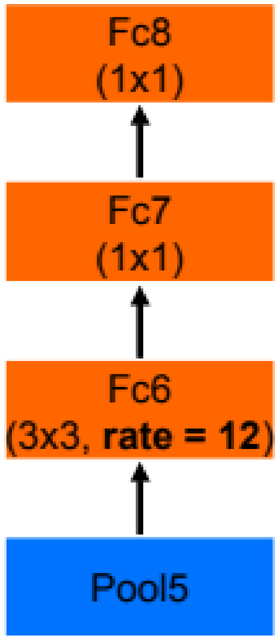


# 05. DeepLab V2

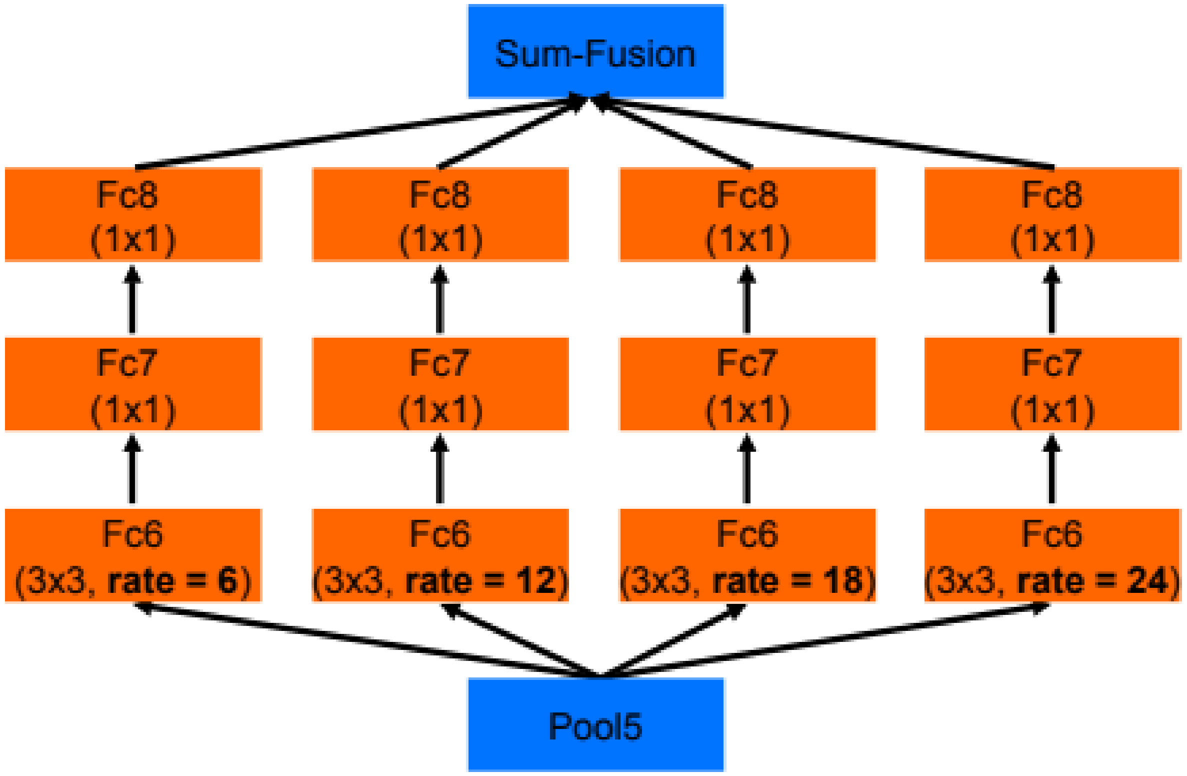
## Multi-scale solution



# 05. DeepLab V2




(a) DeepLab-LargeFOV

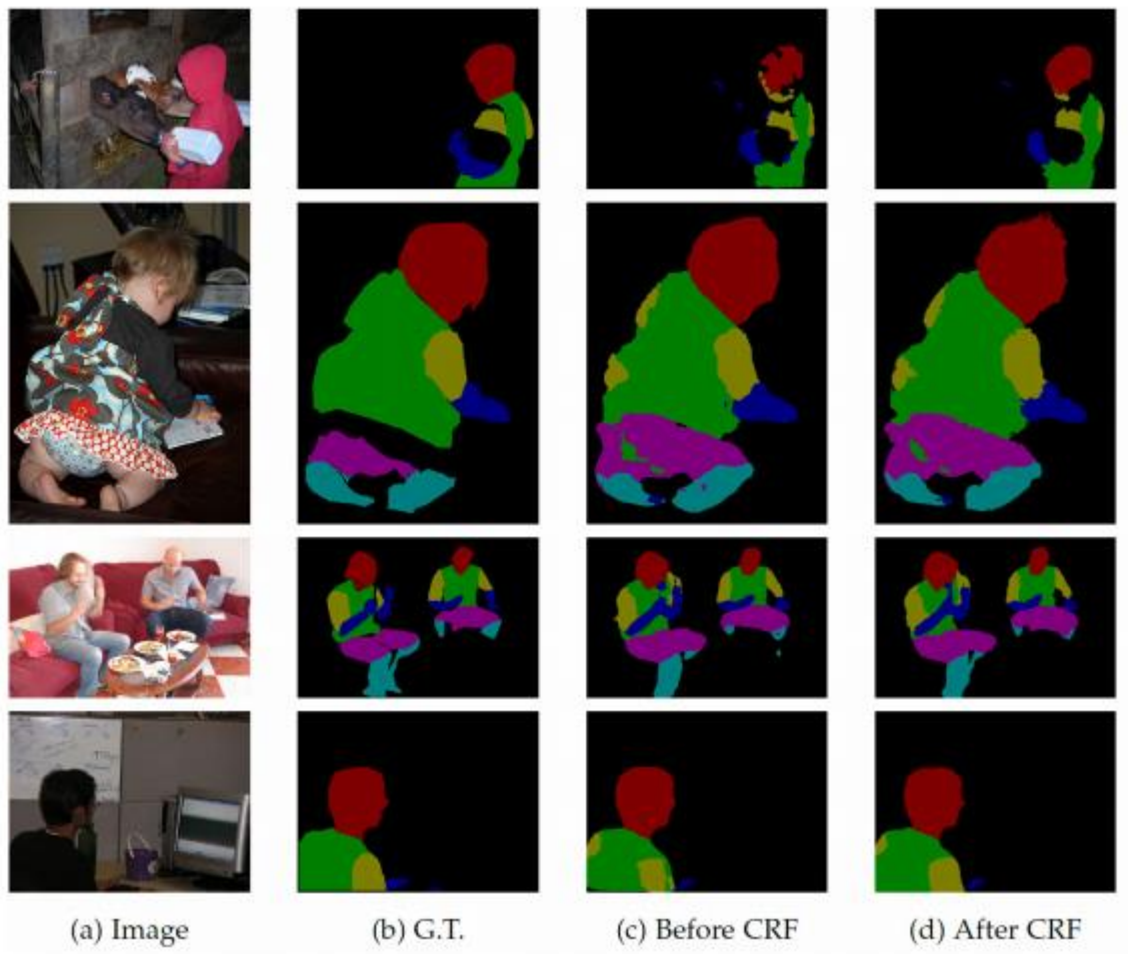


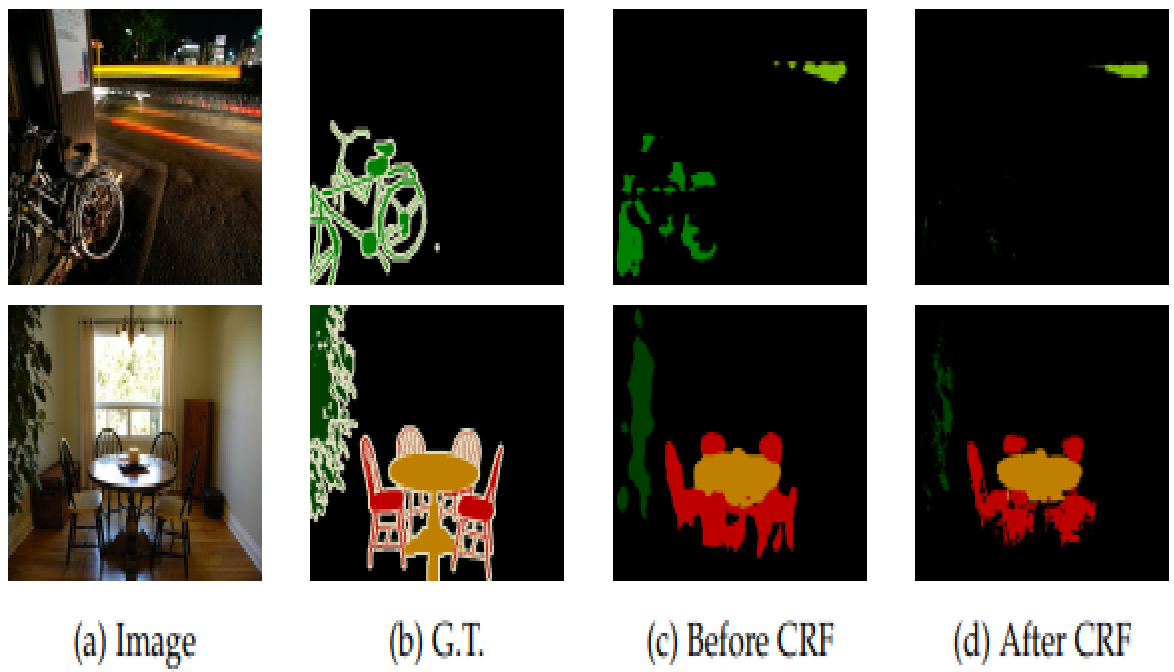

(b) DeepLab-ASPP

| Method | before CRF | after CRF |
|---|---|---|
| LargeFOV | 65.76 | 69.84 |
| ASPP-S | 66.98 | 69.73 |
| ASPP-L | 68.96 | 71.57 |

# 05. DeepLab V2

## CRF



# Thank you