SimCenter NHERI
Center for Computational Modeling and Simulation

2021 AI Bootcamp

Image Segmentation

Barbaros Cetiner

NHERI SimCenter
University of California, Los Angeles

NSF Award: CMMI 1612843

Outline

- Why do we use image segmentation?
- Image segmentation algorithms
- Segmentation datasets
- A simple framework for model development

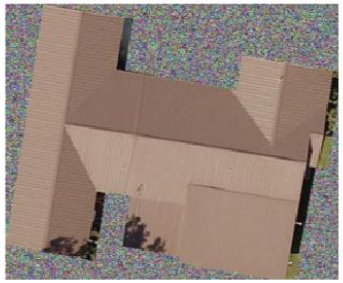
Why Do We Use Image Segmentation?

What is Image Segmentation?

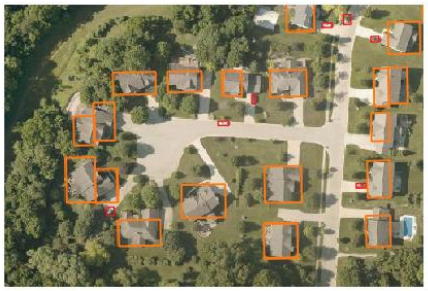
Image classification: Classifies an image into a certain category

Object detection: Identifies (rough) locations of objects in an image

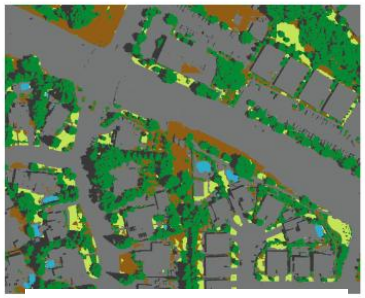
Image segmentation: Identifies the exact pixel regions occupied by objects



Classification



Detection



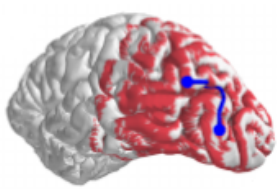
Segmentation

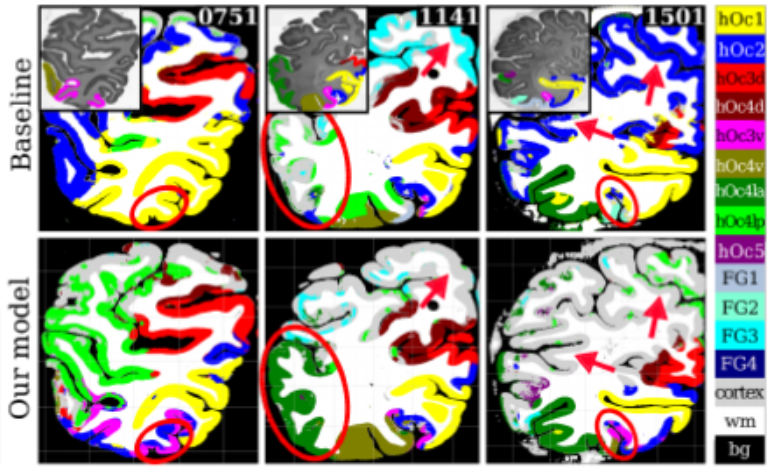
Why Use Image Segmentation?

Image segmentation enables more detailed analysis of the attributes of objects in an image

Using segmentation, **one can better attain object information such as**

- **Height/Width**
- **Area**
- **Texture**





Why do we use semantic segmentation?

We use semantic segmentation to **extract detailed geometric/textural properties of objects from an image**

Image Segmentation Algorithms

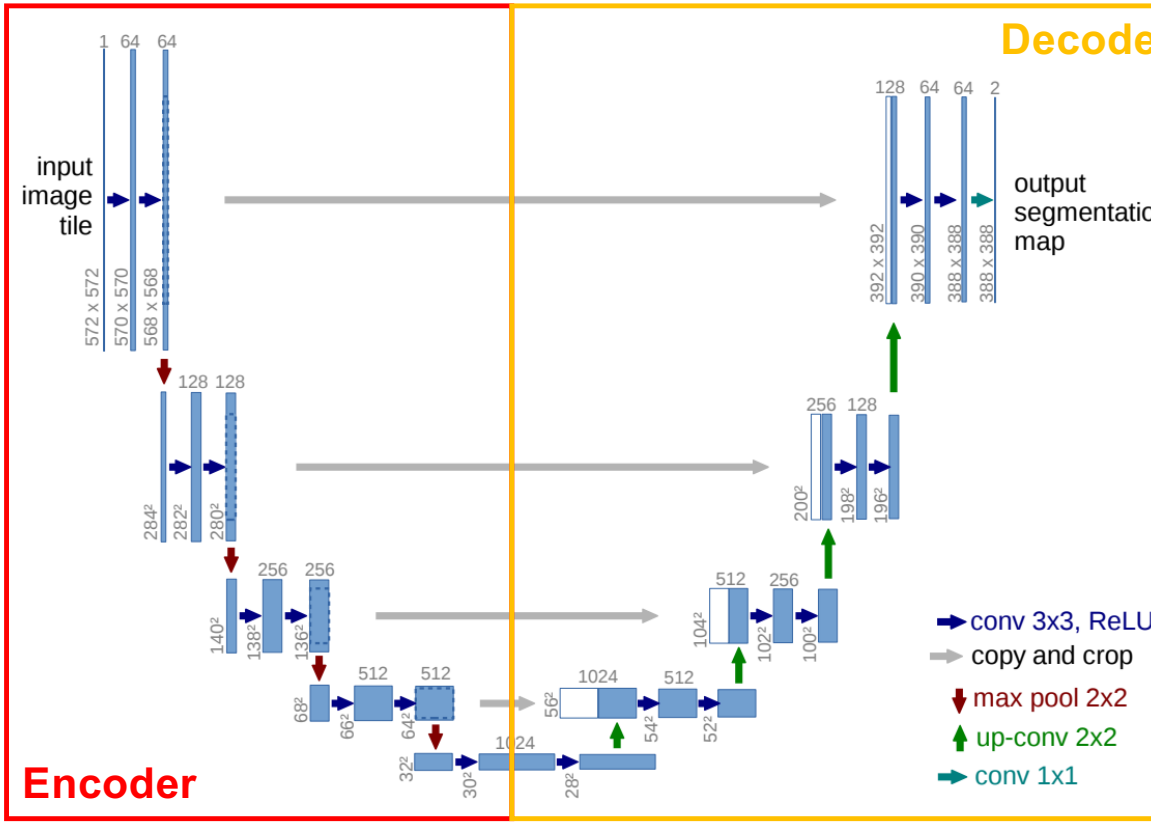
U-Net

Encoder: a pre-trained classification network that extracts features at multiple levels

Decoder: an upsampling network that projects features learnt by the encoder onto the pixel

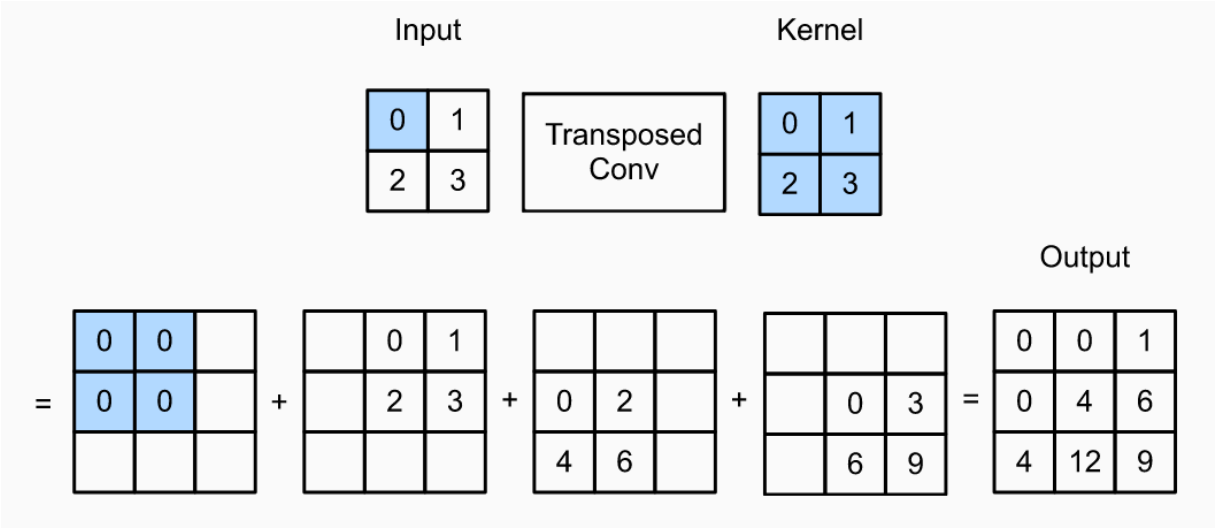
Feature space is
- **semantically rich**
- **has low spatial resolution**

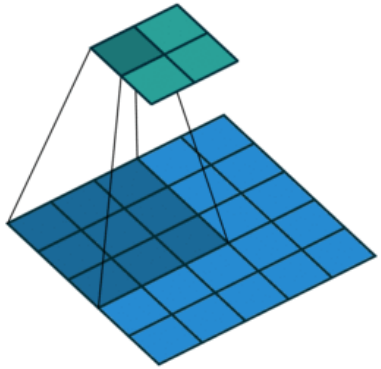
Pixel space
- **has high spatial resolution**
- **lacks the semantics of feature space**



Upsampling using Transpose Convolutions

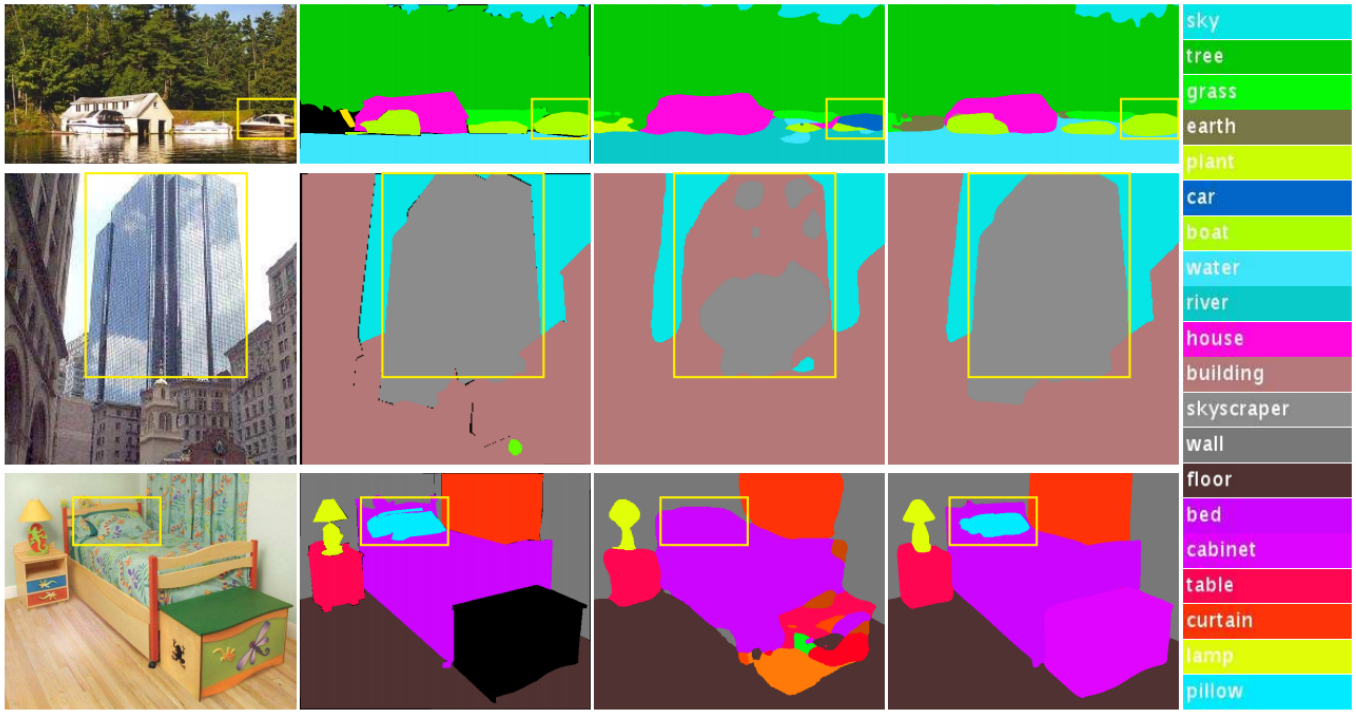
Transpose convolutions: upsample the input feature map by trying to remove the effect of convolutions





PSPNet: Pyramid Scene Parsing Network

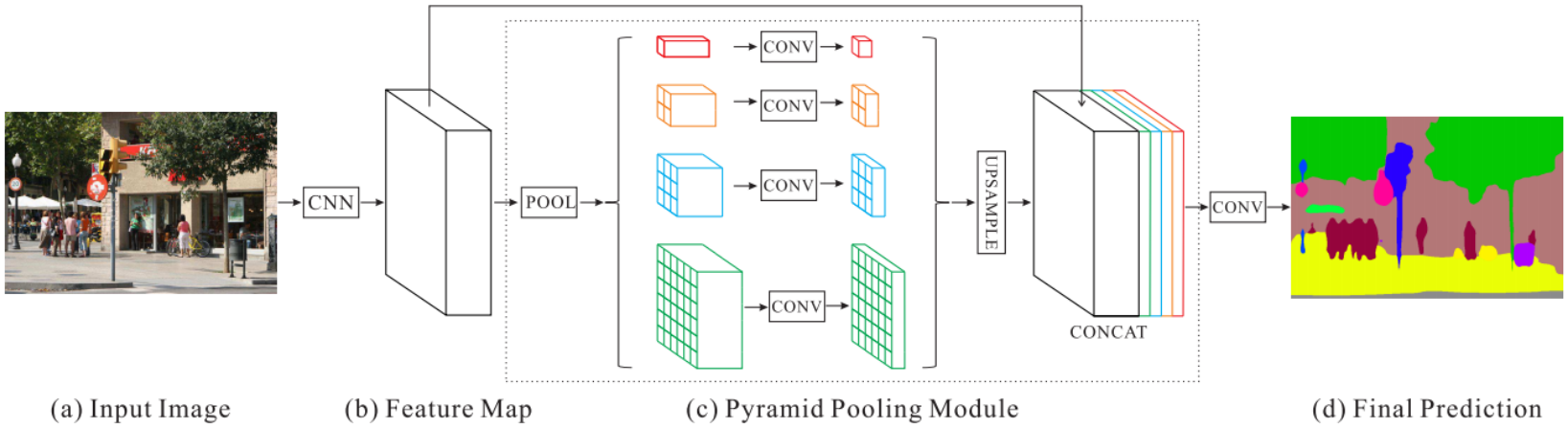
Accurate semantic segmentation requires capturing context information



(a) Image (b) Ground Truth (c) FCN (d) PSPNet (e) ColorMap

PSPNet: Pyramid Scene Parsing Network

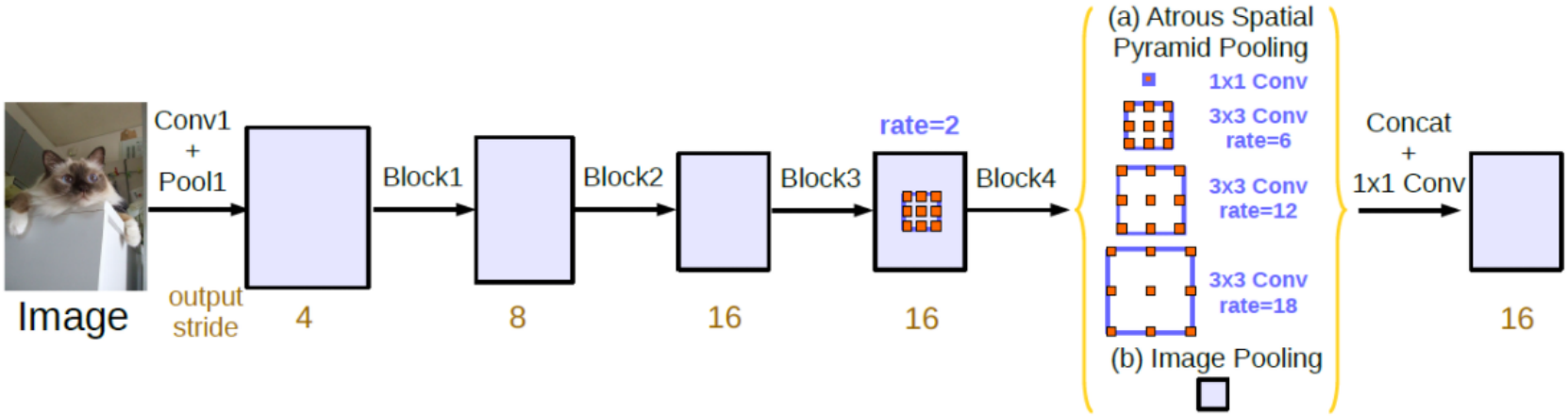
- Feature map size is 1/8 of the input image
- Performs sub-region average pooling over each feature map from coarse to fine level
- Upsampling is performed using bilinear interpolation



Bilinear interpolation results in blobby segmentation masks

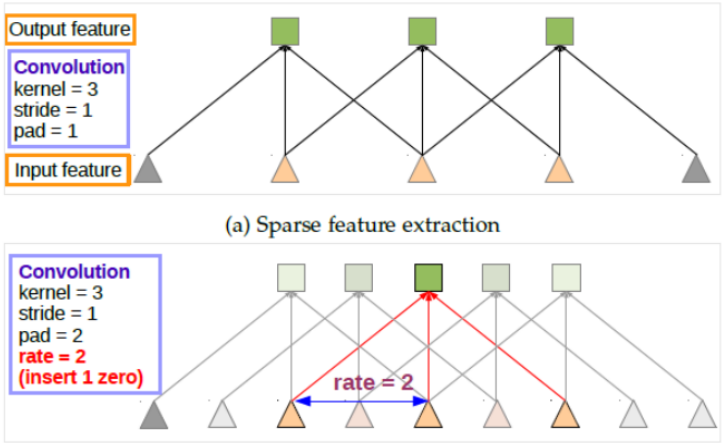
DeepLabV3

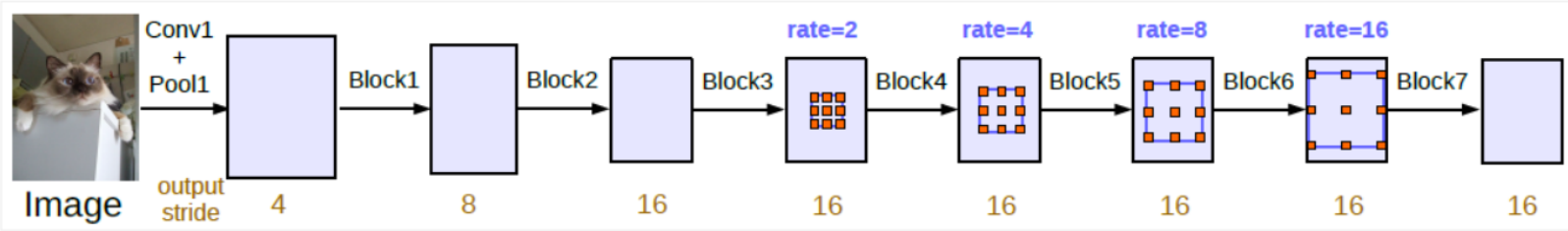
- Attains better segmentation accuracy than PSPNet
- Sharper masks



Atrous Convolution

- Results in a feature map (preserves spatial information)
- Increasing the rate increases the field of view
- **Pyramid pooling is required:**
  sampling rate ↑, # valid filter weights↓





What is one key architectural difference between segmentation and classification algorithms?

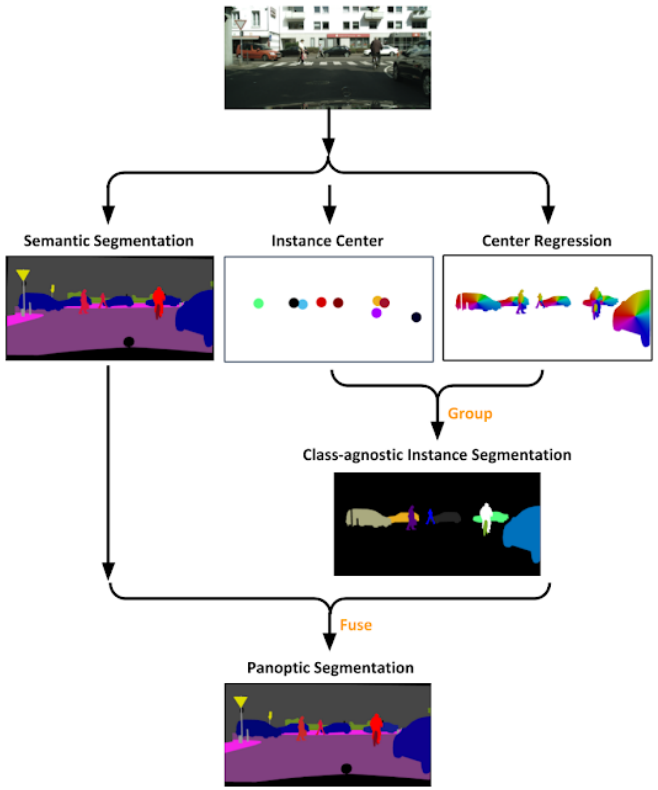
Segmentation algorithms ***fuse feature maps with spatial information for accurate mask prediction***

In classification algorithms, ***final output is based on feature maps only***

Why does Atrous Convolution (DeepLabV3) perform better than Bilinear Interpolation (PSPNet)?

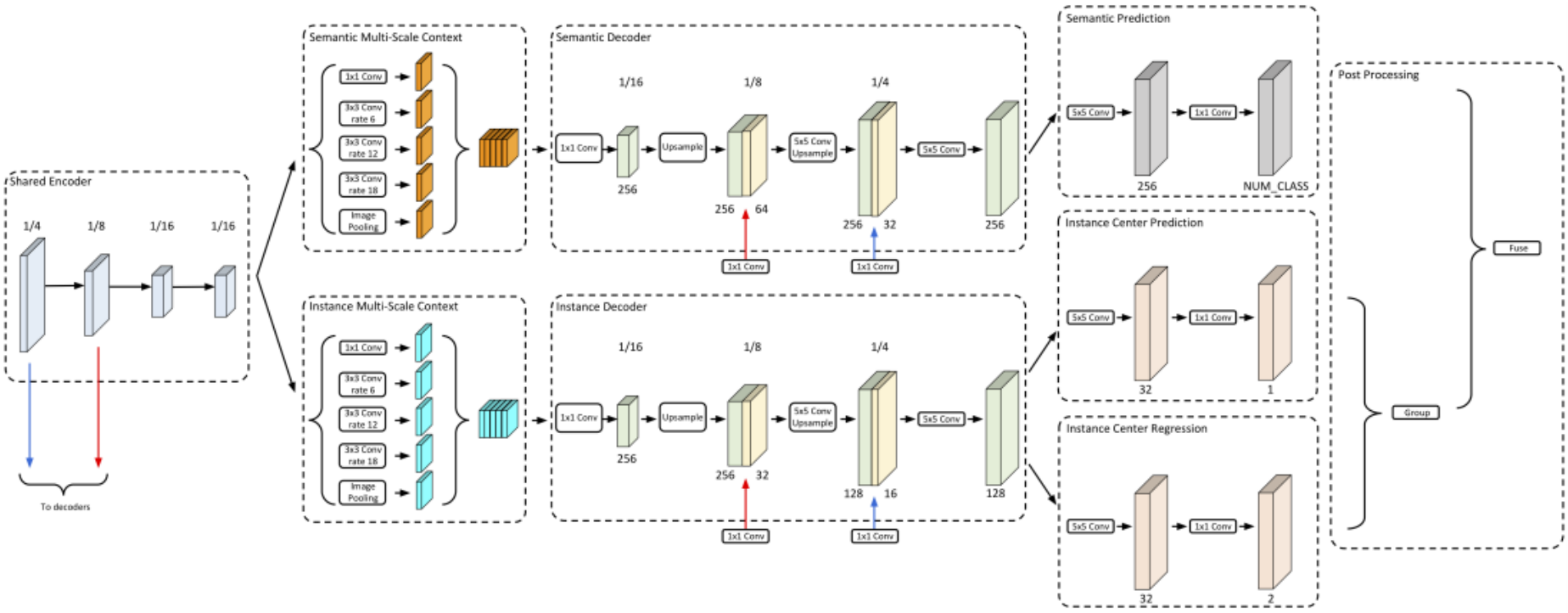
Atrous Convolution ***better preserves spatial resolution in extracting features***

SOTA Segmentation: Panoptic-DeepLab



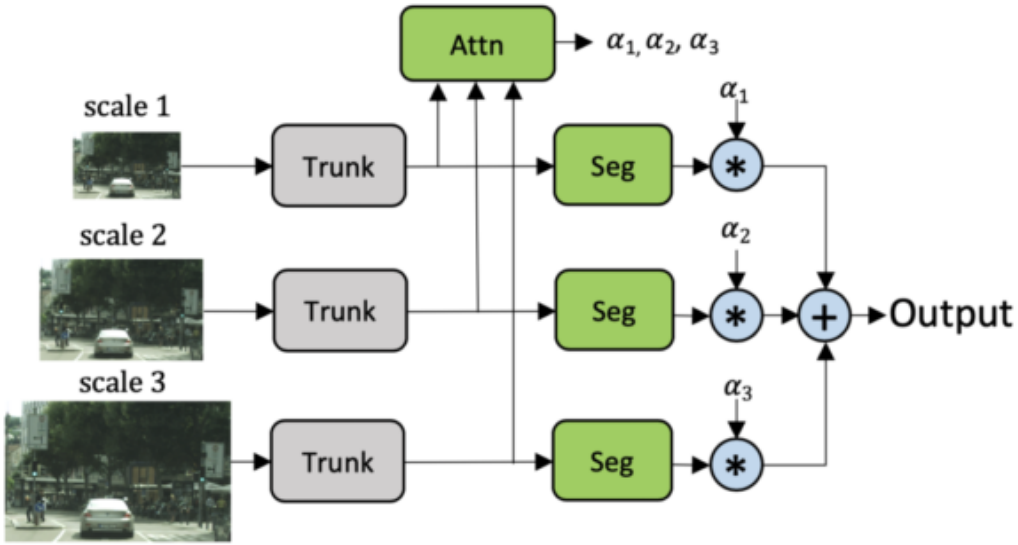
SOTA Segmentation: Panoptic-DeepLab

Dual-context and dual-decoder modules for semantic segmentation and instance segmentation predictions



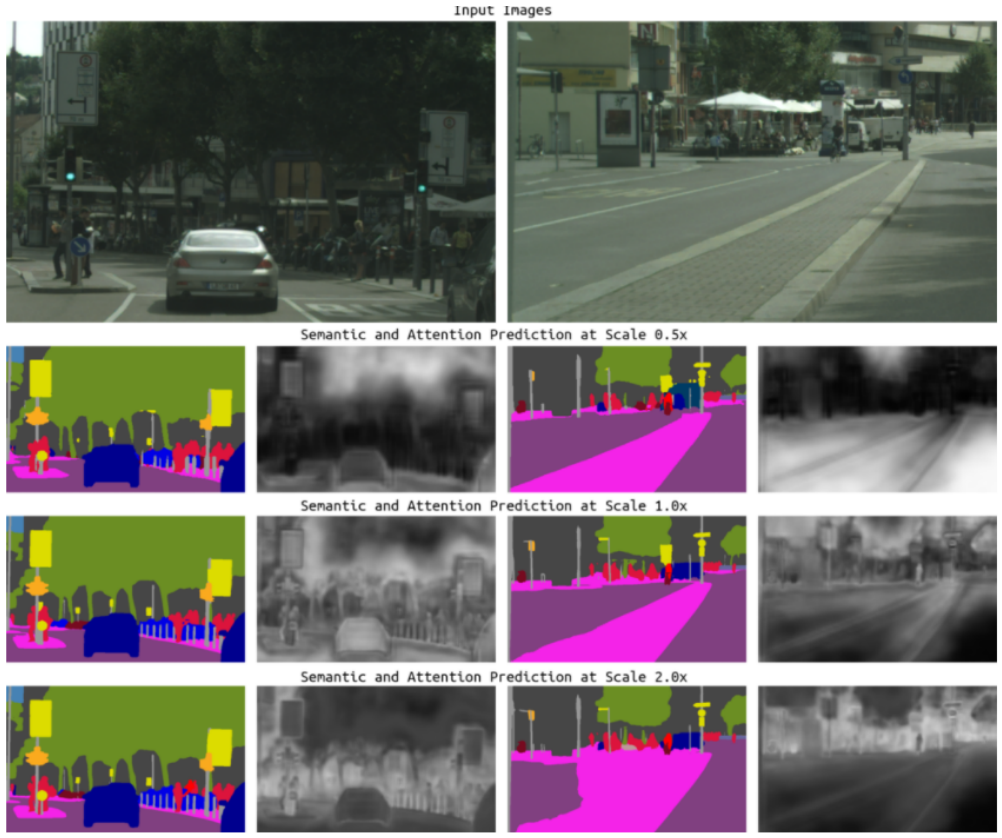
SOTA Segmentation: HRNet-OCR

Unlike previous methods, this approach uses Attention networks to combine multi-scale predictions



SOTA Segmentation: HRNet-OCR

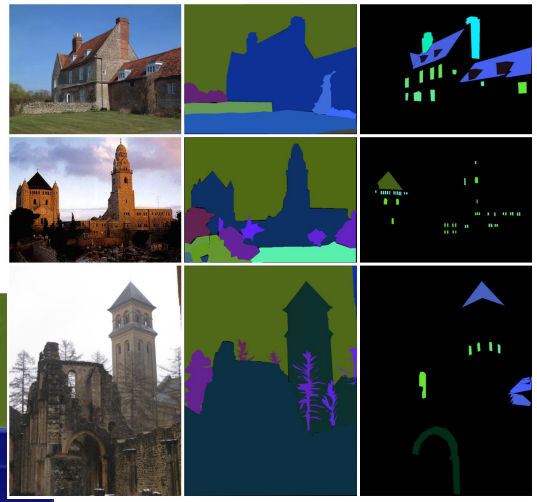
White regions indicate the areas that draw higher levels of attention

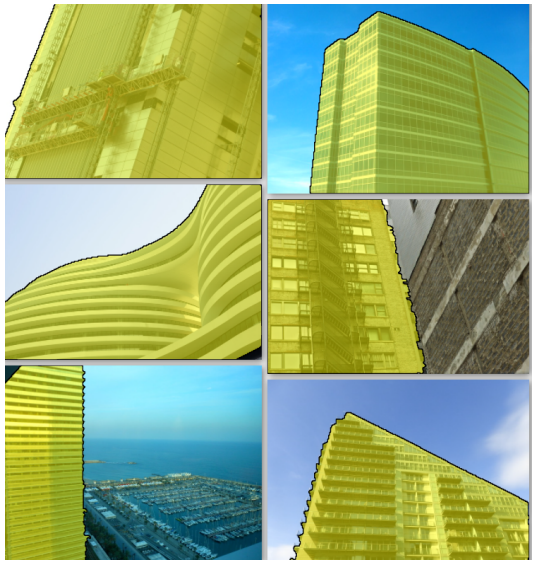


Segmentation Datasets

Popular Object Detection Datasets

1. Cityscapes
2. ADE20K
3. Google Open Images
4. PASCAL VOC
5. PASCAL Context





Custom Datasets

CVAT is a great tool that has automatic labeling support

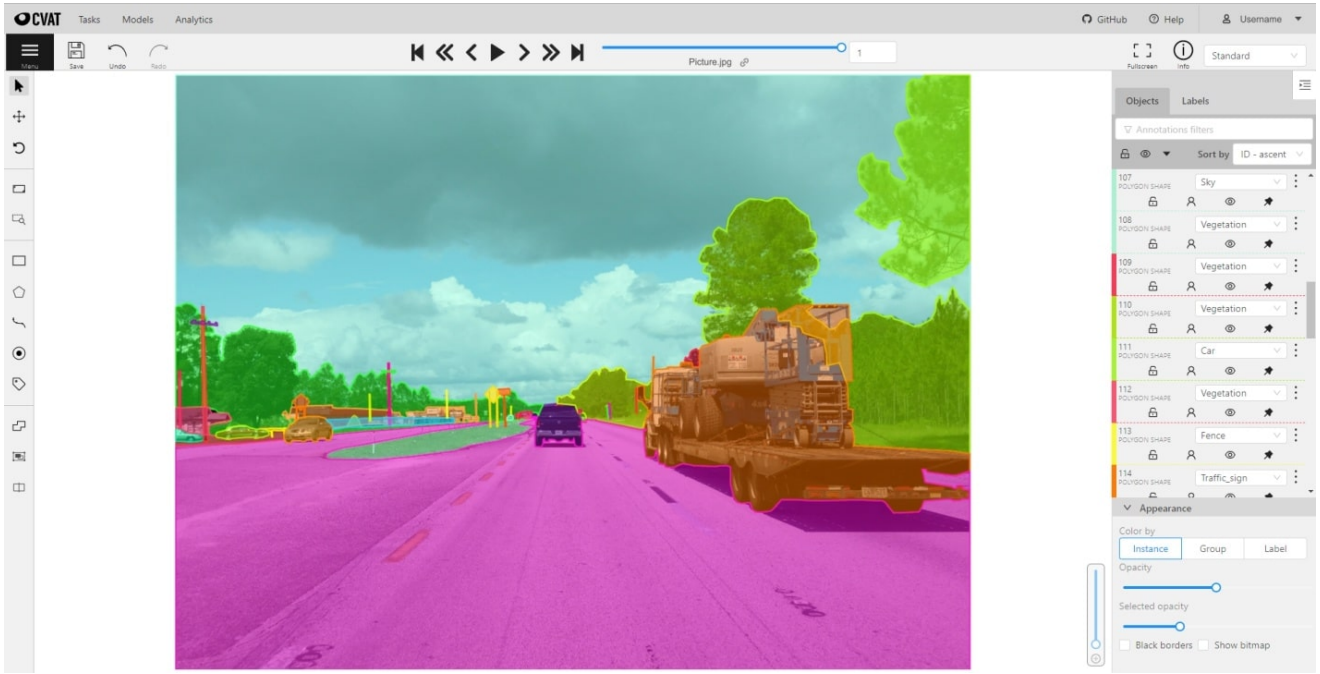


Image Augmentation

1. Augmentation is essential to introducing variability in the dataset, which helps models generalize better.
2. Also, the size of a dataset can be increased with augmentation (important if dataset size is small)

Popular techniques:

- Contrast/brightness/saturation adjustment
- Lighting noise
- Random crop
- Rotate
- Flip

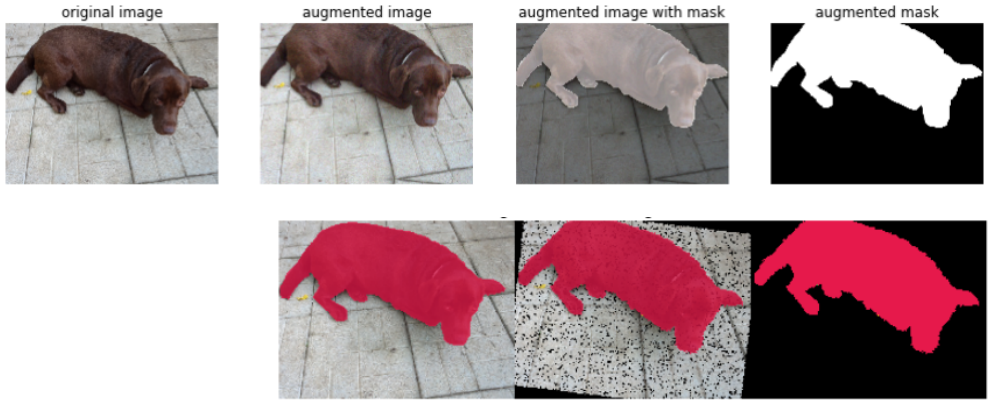


Image Augmentation using Python

Available augmentation libraries:

- Pillow
- Albumentations
- Augmentor
- imgaug

A Simple Framework for Model Development

Simple Framework for Model Development

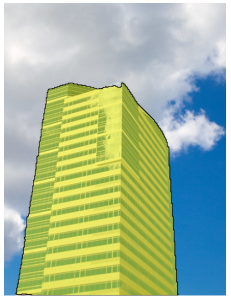


Image labeling → Image Preprocessing → Label preprocessing → Model training → Post-processing