

深度學習 Pytorch手把手實作 影像分割

黃志勝 (Tommy Huang)

義隆電子 人工智慧研發部

國立陽明交通大學 AI學院 合聘助理教授

國立台北科技大學 電資學院合聘助理教授



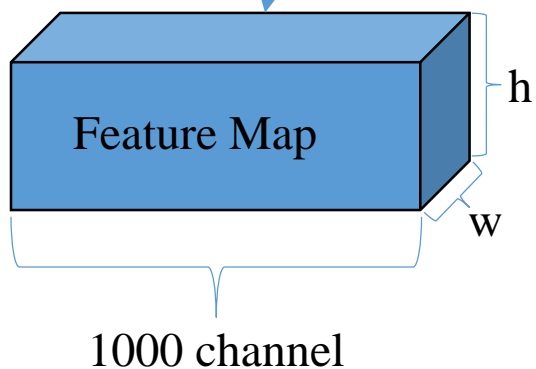
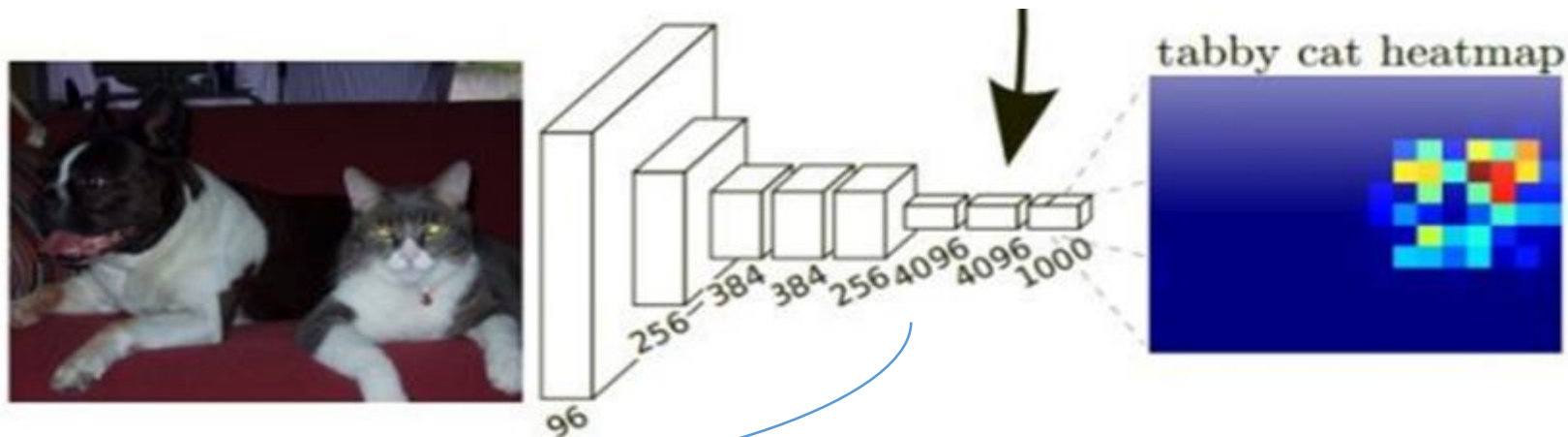
Introduction

- Brief Introduction
- What is image segmentation?
- Image segmentation architectures.
- Hand-by-hand pytorch implementation.



Deep Learning: Feature Extractor

- Feature map可以做什麼？



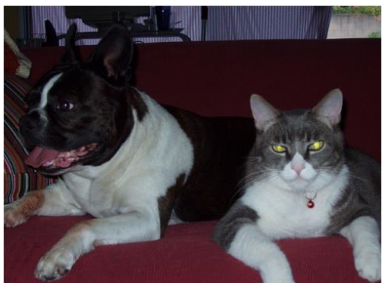
- 這張圖是貓還是狗 (Classification)
- 框出圖片內有貓和狗的位置 (Object detection)
- 從框出的物件(貓和狗)中的描繪出實際物件的輪廓 (Instance Segmentation)
- 把圖片描繪出物件(貓和狗)的輪廓 (Semantic Segmentation)



What is image segmentation?

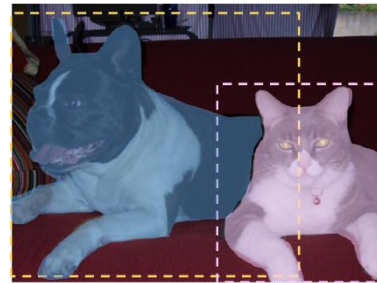
- This is the process of dividing an image into multiple segments.
- Every pixel in the image is associated with an object type.

Semantic Segmentation



pixel-by-pixel
藍色Mask是狗
紅色Mask是貓
橘色Mask是沙發
綠色是Background

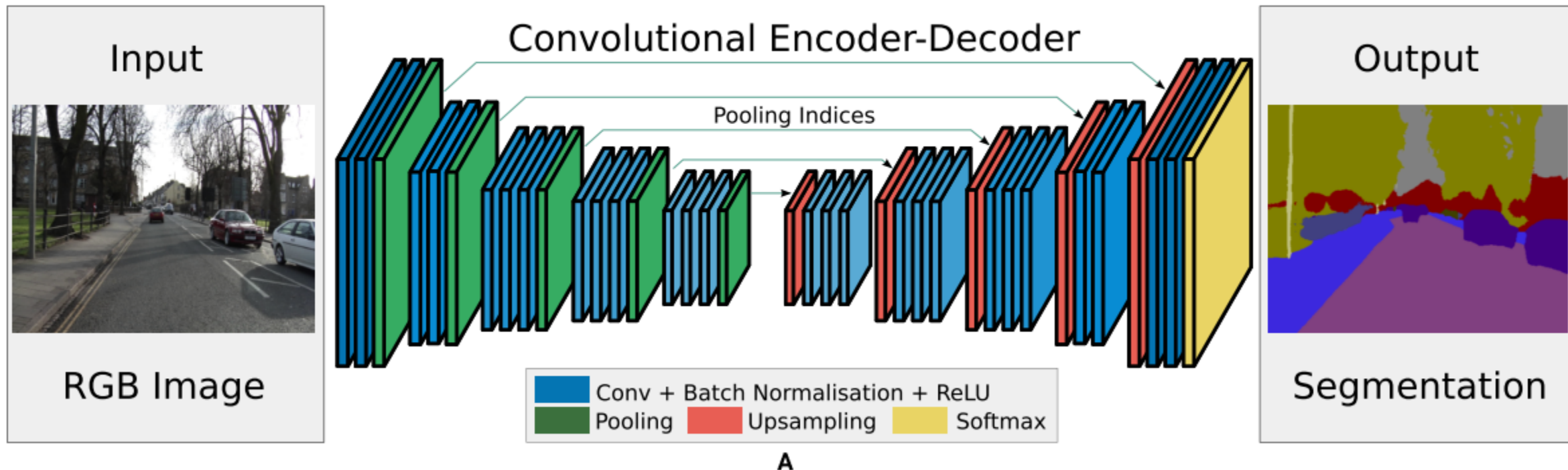
Instance Segmentation



Object detection後的框框內那些pixel是實際的物件。



Image segmentation architectures

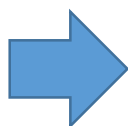
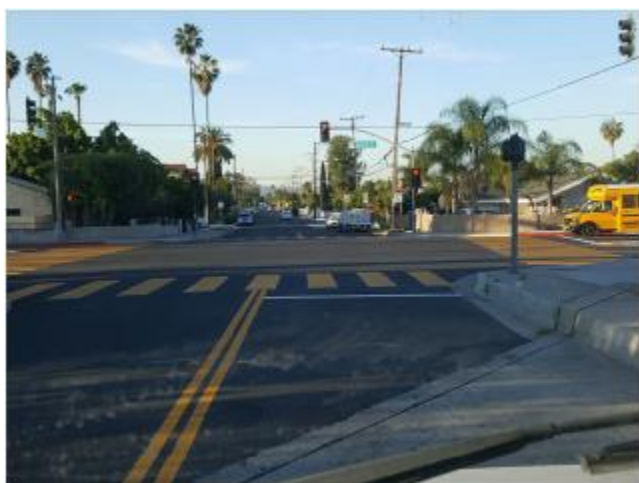
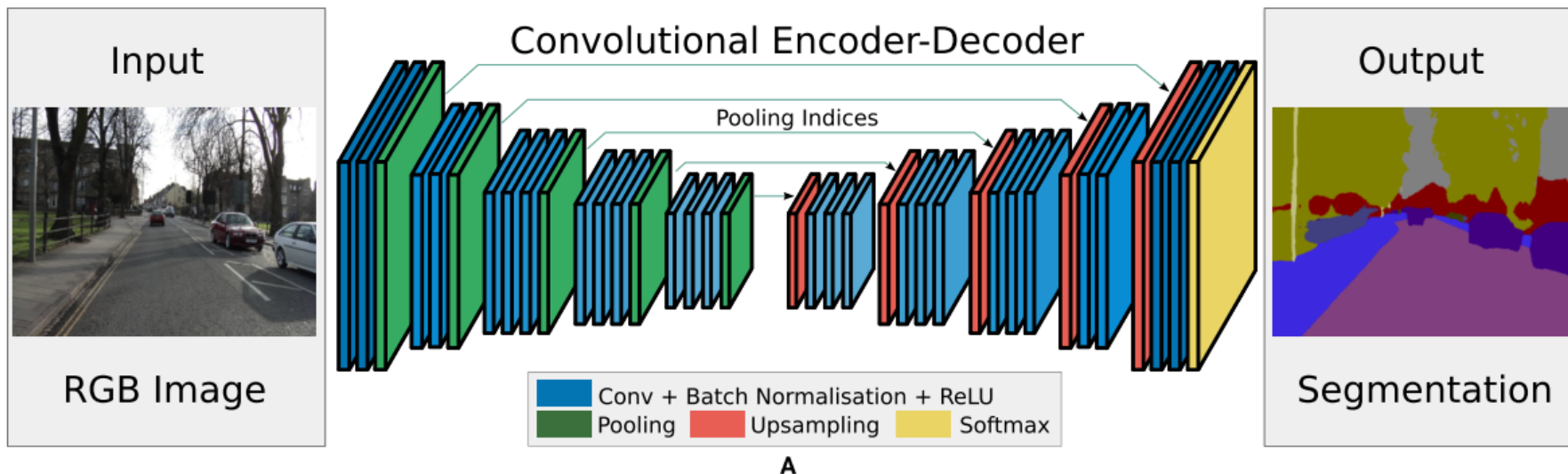


可原圖大小去訓練和inference:
缺點:計算資源需求大, 內存要很夠。
優點: 結果較為準確。

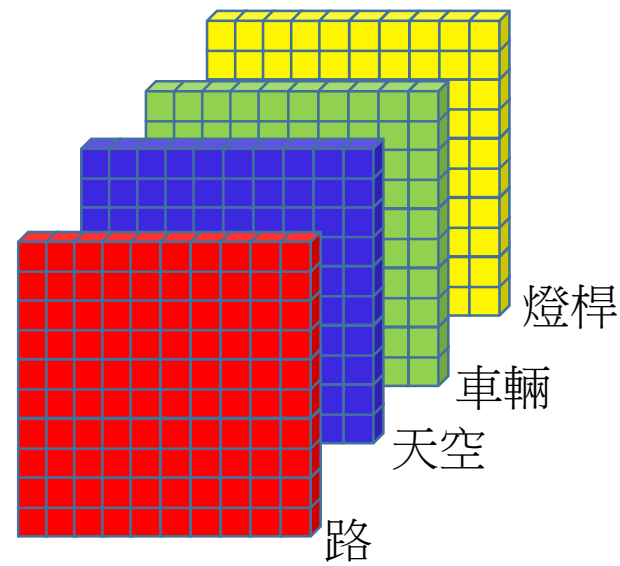
可原圖大小(或resize到固定大小)去訓練, 但結果只從1/8大小
的結果輸出:
缺點:結果較差。
優點: 快速。



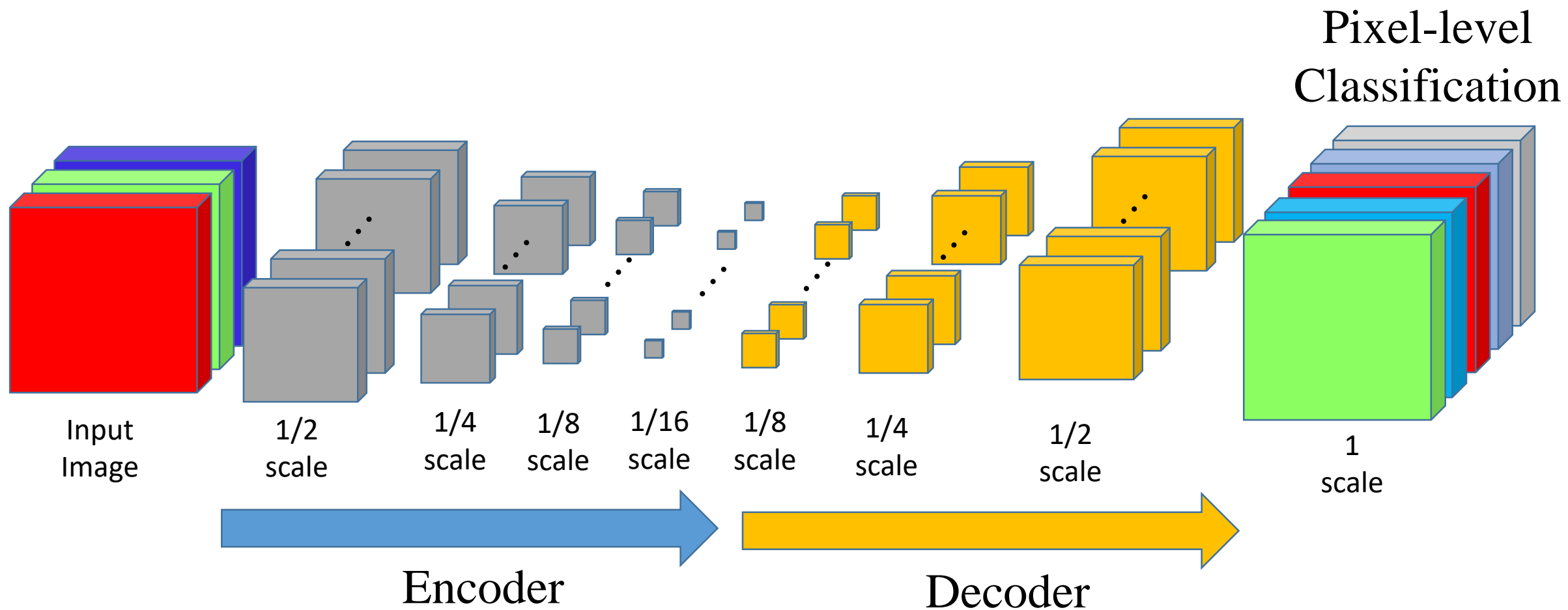
Image segmentation architectures



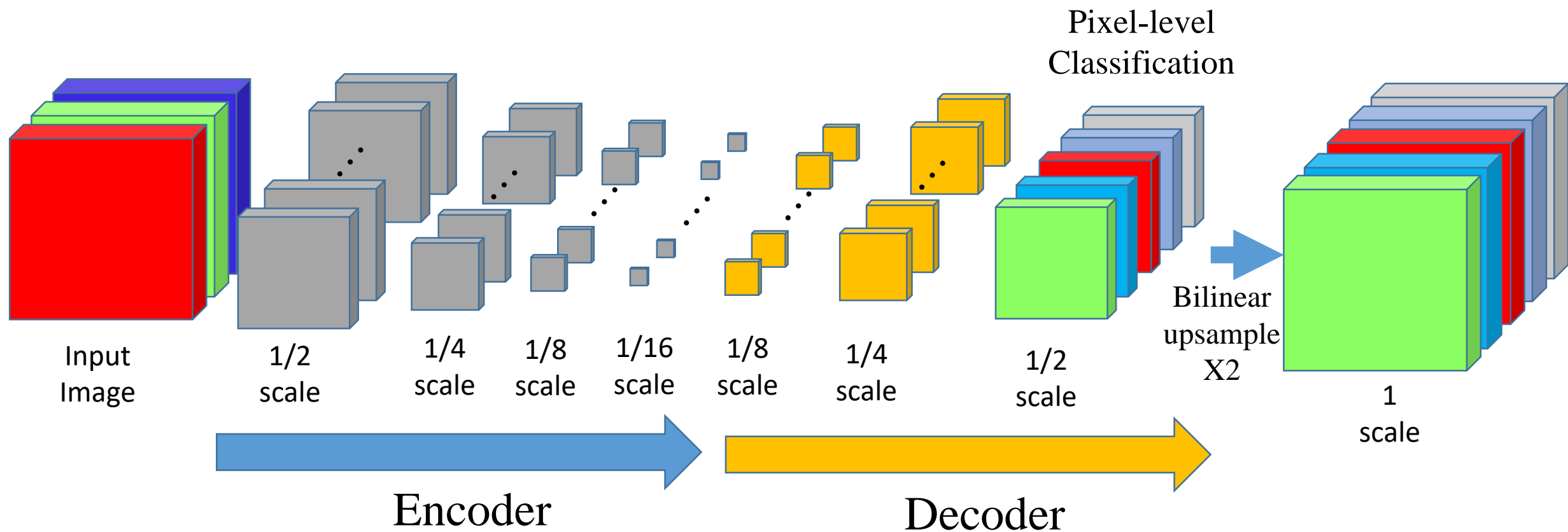
CNN



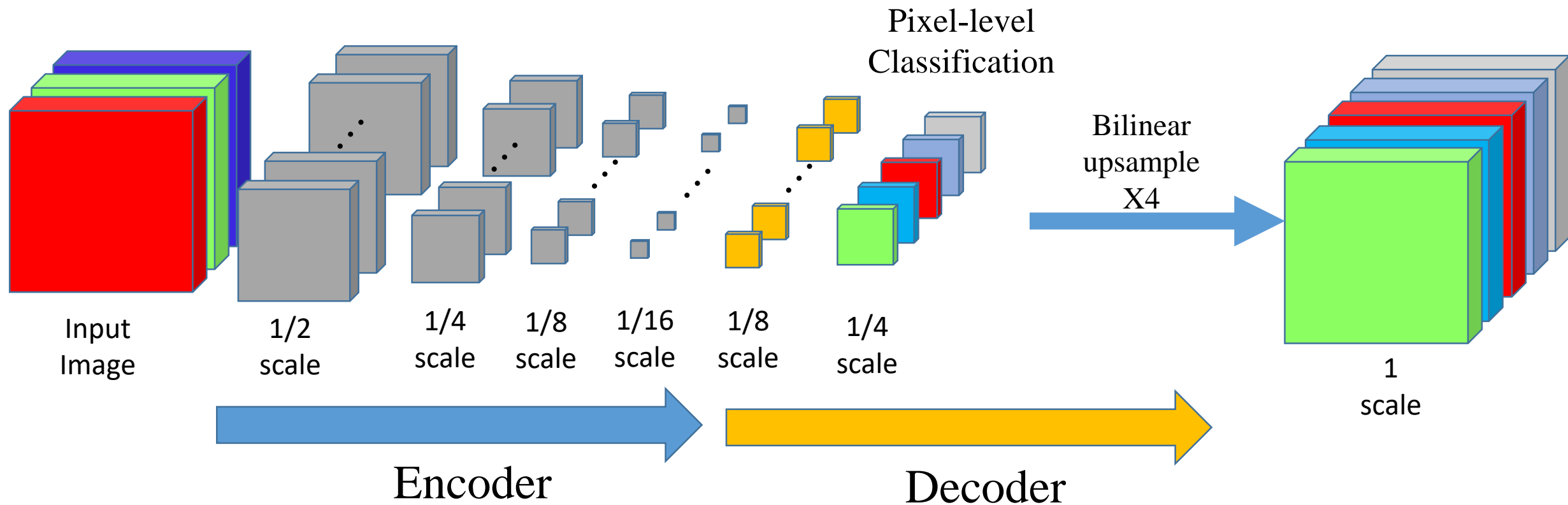
Semantic Segmentation



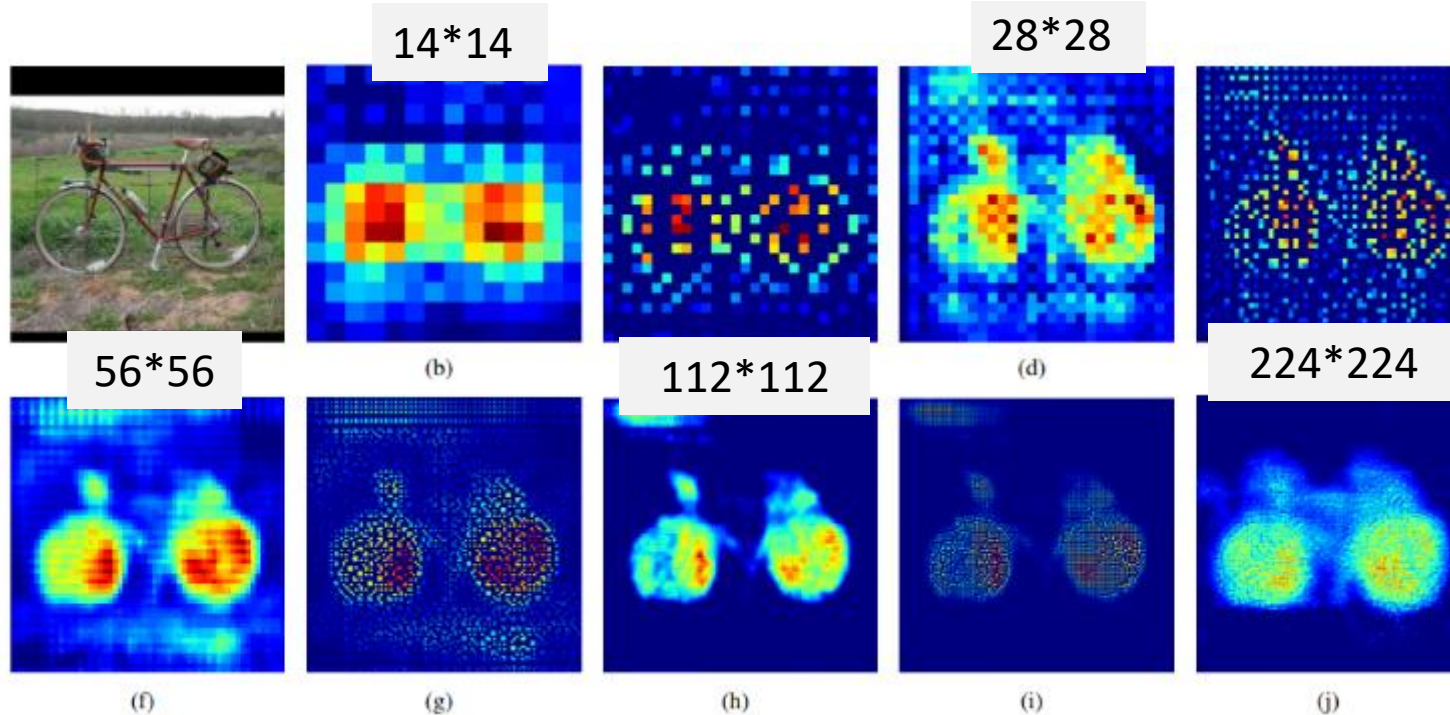
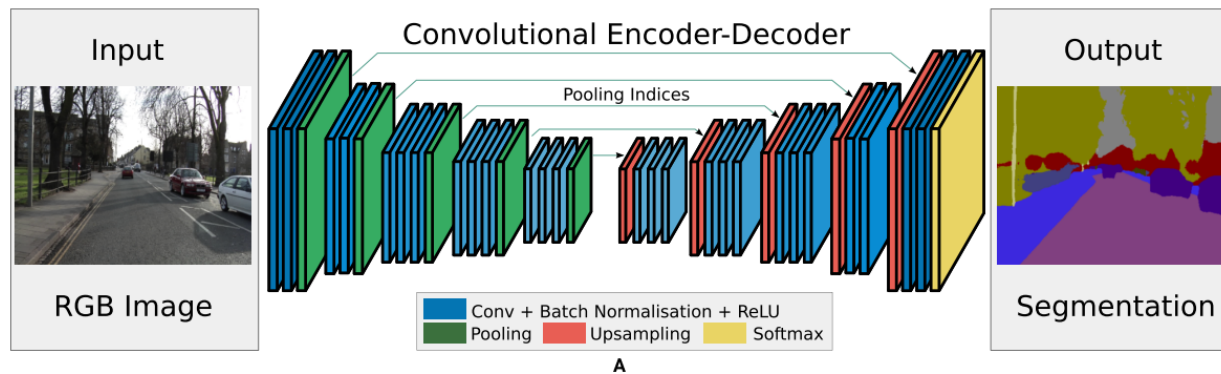
Semantic Segmentation



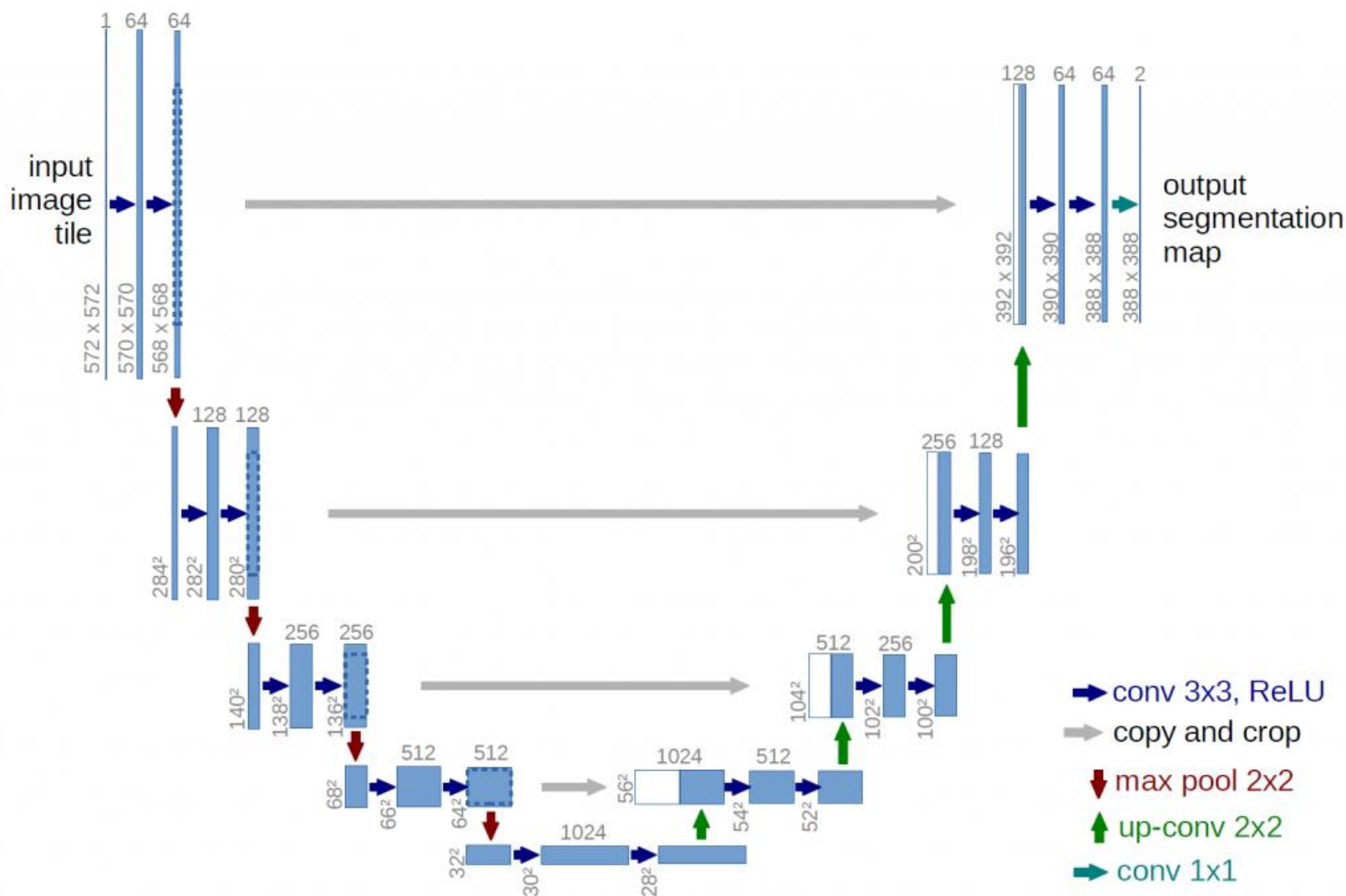
Semantic Segmentation



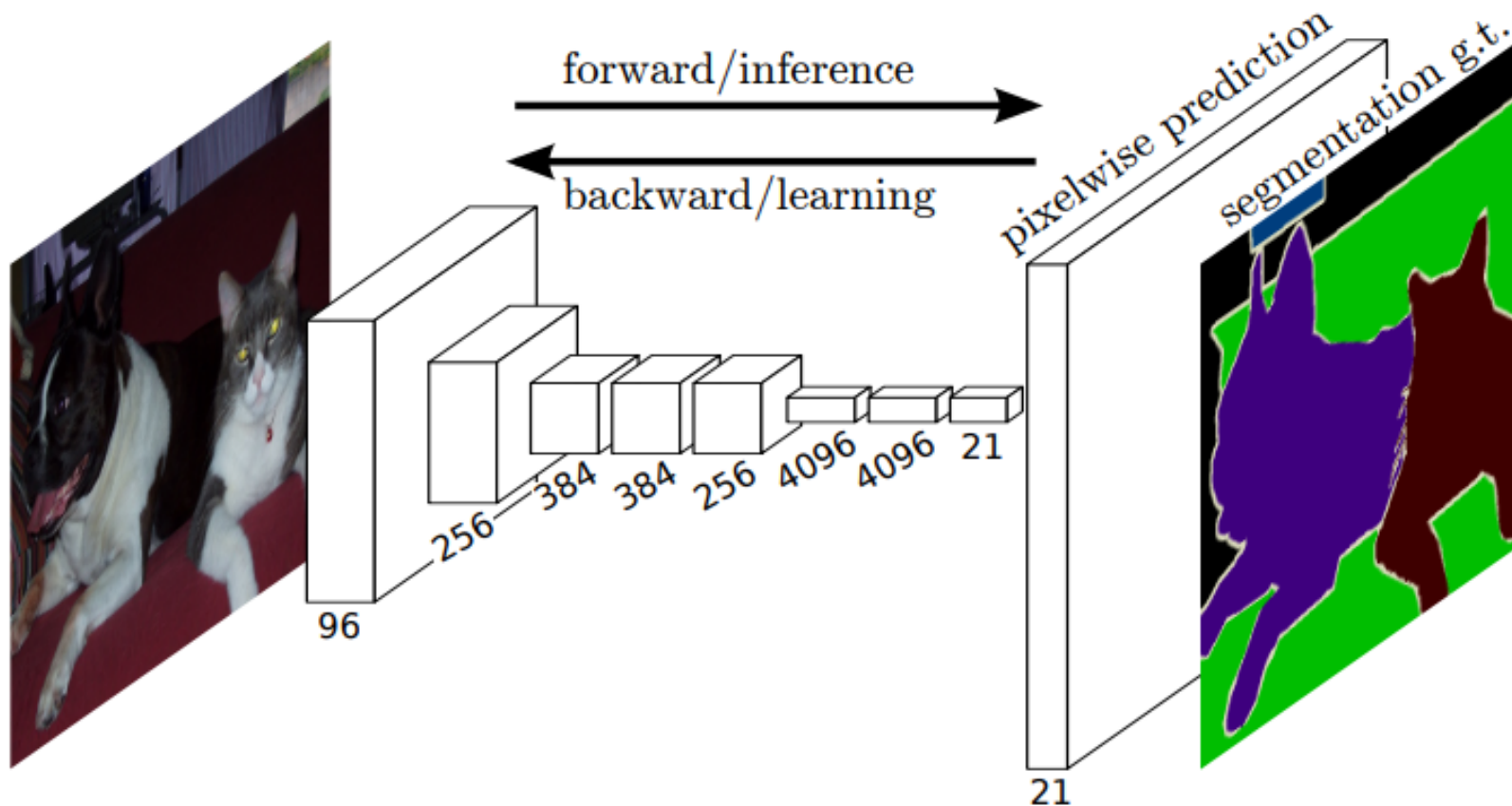
Semantic Segmentation



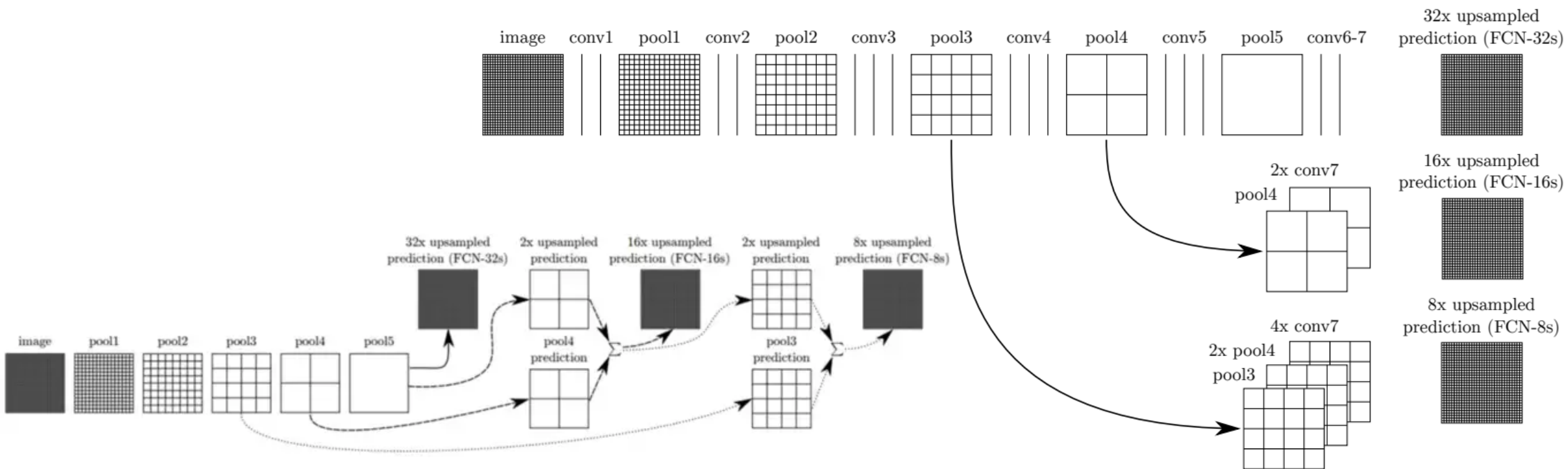
U-NET(2015)



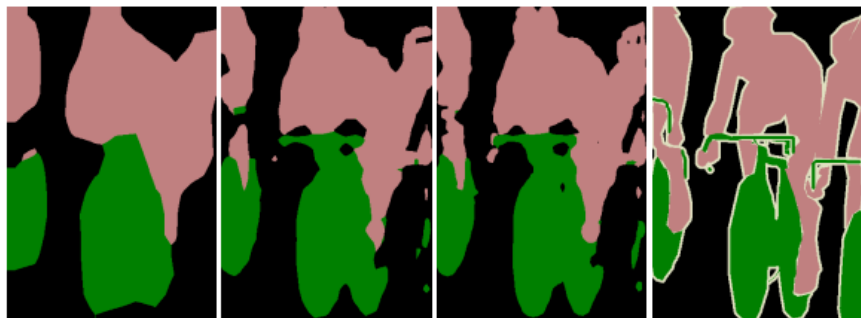
FCN — Fully Convolutional Network (2015)



FCN — Fully Convolutional Network (2015)



FCN-32s FCN-16s FCN-8s Ground truth



HRNET(2019)

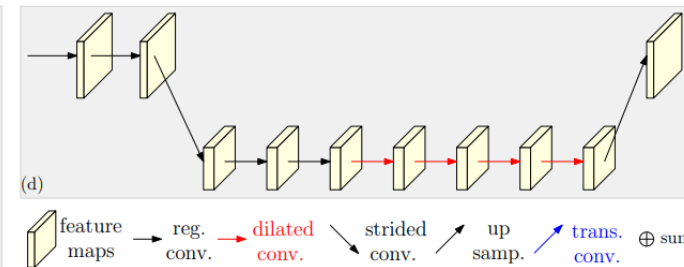
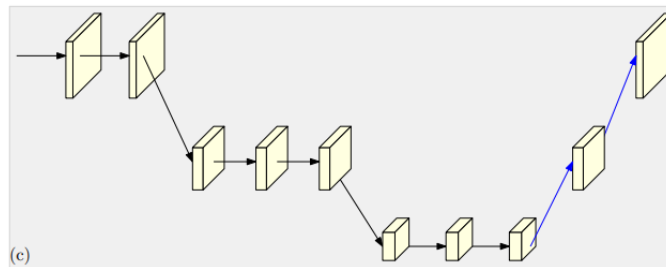
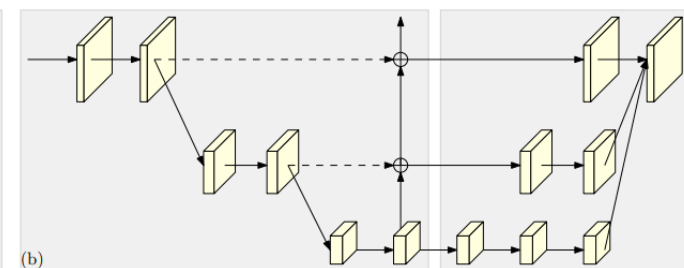
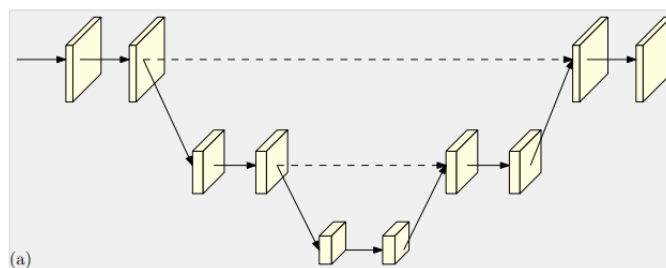
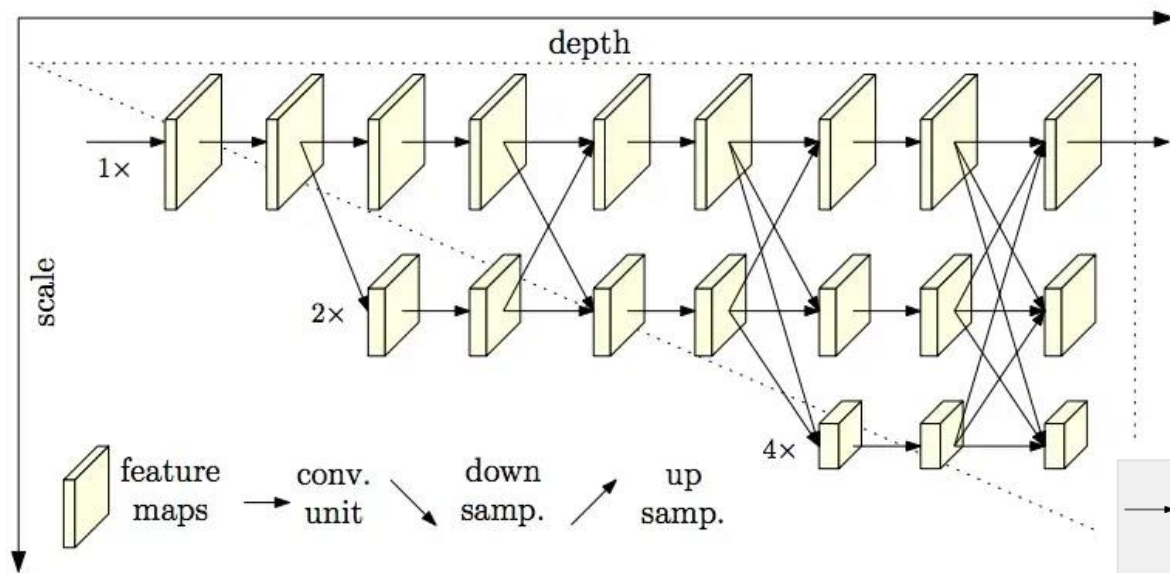


Figure 2. Illustration of representative pose estimation networks that rely on the high-to-low and low-to-high framework. (a) Hourglass [40]. (b) Cascaded pyramid networks [11]. (c) SimpleBaseline [72]: transposed convolutions for low-to-high processing. (d) Combination with dilated convolutions [27]. Bottom-right legend: reg. = regular convolution, dilated = dilated convolution, trans. = transposed convolution, \oplus sum.

Hand-by-hand pytorch implementation.

- Dataset: <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/>

Development Kit

The development kit provided for the VOC challenge 2007 is available. You can:

- Download the [training/validation data](#) (450MB tar file)
- Download the [development kit code and documentation](#) (250KB tar file)
- Download the [PDF documentation](#) (120KB PDF)
- Browse the [HTML documentation](#)
- View the [guidelines](#) used for annotating the database

Download link: http://host.robots.ox.ac.uk/pascal/VOC/voc2007/VOCtrainval_06-Nov-2007.tar

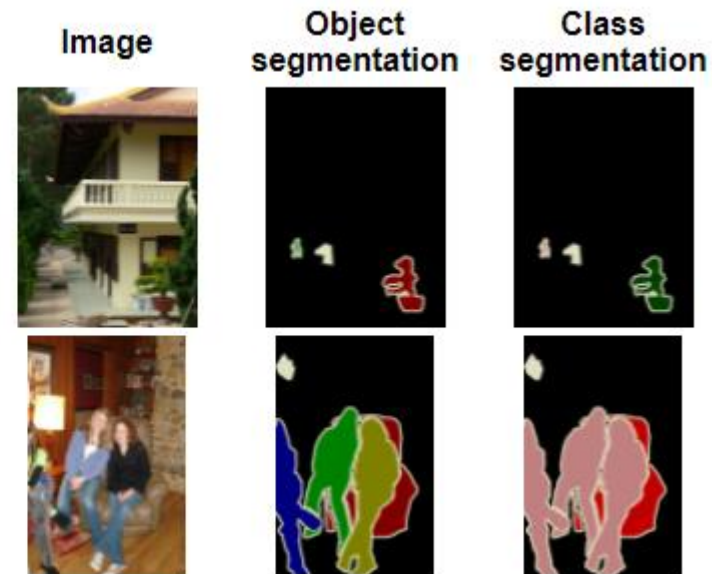
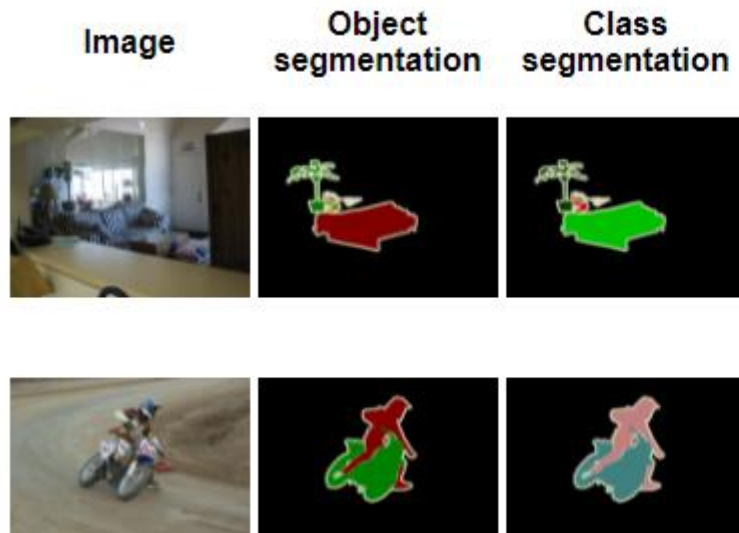


Dataset (VOC 2007)

- **the class segmentation**

(1=aeroplane, 2=bicycle, 3=bird, 4=boat, 5=bottle, 6=bus, 7=car, 8=cat, 9=chair, 10=cow, 11=diningtable, 12=dog, 13=horse, 14=motorbike, 15=person, 16=potted plant, 17=sheep, 18=sofa, 19=train, 20=tv/monitor)

- Index 0 corresponds to background and index 255 corresponds to 'void' or unlabeled.



Dataset (VOC 2007)

tion

檢視

ser > deep learning course 2021 > dataset > VOCdevkit > VOC2007 > ImageSets > Segmentation

名稱	修改日期	類型	大小
 train.txt	2007/4/8 上午 12:36	文字文件	2 KB
 trainval.txt	2007/4/8 上午 12:36	文字文件	3 KB
 val.txt	2007/4/8 上午 12:36	文字文件	2 KB



UNet Structure

- In my example code, I build a Unet-like model for training/evaluating.

