

Semantic Segmentation

D.A. Forsyth

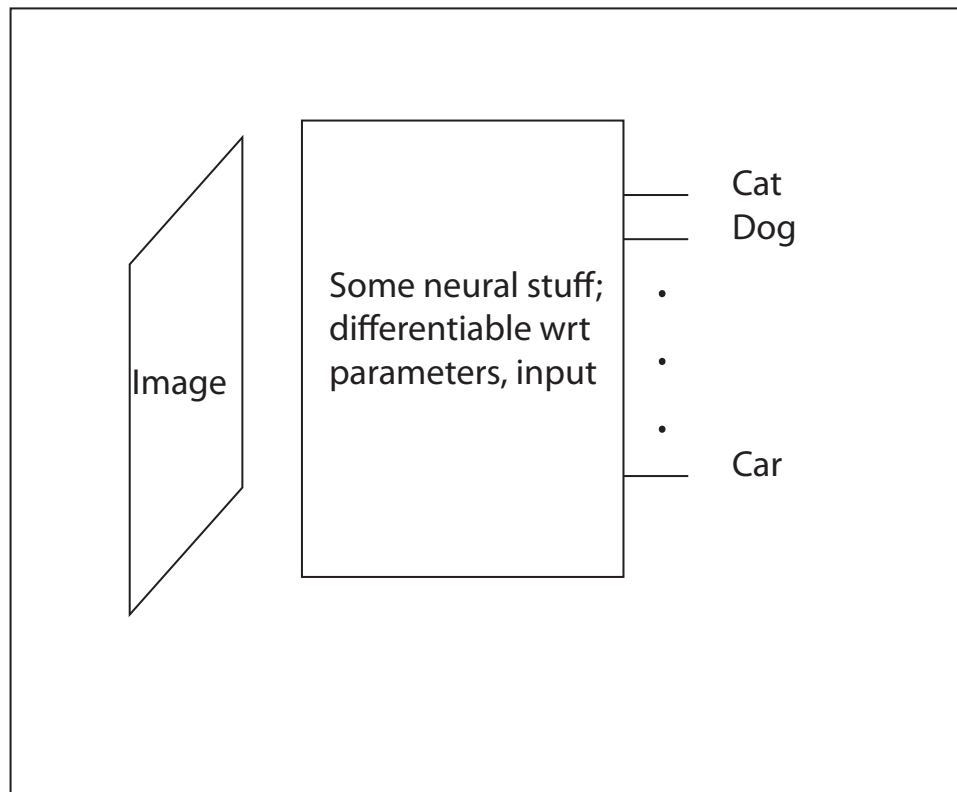
University of Illinois at Urbana Champaign

Ideas:

- Simple Image Classification:
 - Classify an image into one of two classes by
 - Stack some pooling, FC layers on an encoder
 - Adjust result into a vector
 - Pass to linear classifier
 - Learn all parameters with SGD and various losses
 - get training examples right
- Simple image regression:
 - Build U-Net
 - input is image
 - output is image-like thing (depth, normal, etc)
 - train to produce the right answer
- **Merge these to produce class scores at each pixel**
 - rather than a depth...

Some more detail...

- Construct decoder with k-dim vector
 - at each location



Semantic segmentation

- Label pixels in images with labels taken from a vocabulary
- Notice there are two kinds of label
 - “things” == count nouns == what you can count
 - car, pedestrian, bike, dog, ...
 - “stuff” == mass nouns == others
 - grass, sky, water, cloud, ...

Semantic segmentation

- A number of variants
 - ? - thing labels only
 - ? - stuff labels only
- Instance segmentation – distinguish between individual objects, so car 1, car 2, pedestrian 1, pedestrian 2
- Panoptic segmentation – every pixel gets a label

Evaluation

At a high level, semantic segmentation is evaluated by scoring whether pixel labels are correctly predicted. Obtain a test set consisting of pairs (images, label maps). Choose a class c . For that class, there are two interesting sets: \mathcal{G}_c , the pixels that are labelled with c and \mathcal{P}_c , the pixels where the label c is predicted.

Definition: 22.4 *IoU or intersection over union for a class*

The IoU is given by

$$\text{IoU}_c = \text{IoU}_c(\mathcal{G}_c, \mathcal{P}_c) = \frac{\#(\mathcal{G}_c \cap \mathcal{P}_c)}{\#(\mathcal{G}_c \cup \mathcal{P}_c)}.$$

Definition: 22.6 *IoU of a semantic segmenter*

The IoU is

$$\frac{1}{C} \sum_{c \in \mathcal{C}} \text{IoU}_c$$

Variants

- Some labels are class labels ('car', 'bicycle'),
 - some are category labels ('vehicle')
- IoU per class
 - 0.76 on Kitti leaderboard
- IoU per category
 - 0.9 on Kitti leaderboard

Semantic segmentation on Kitti (labelled data)



MS-CoCo dataset

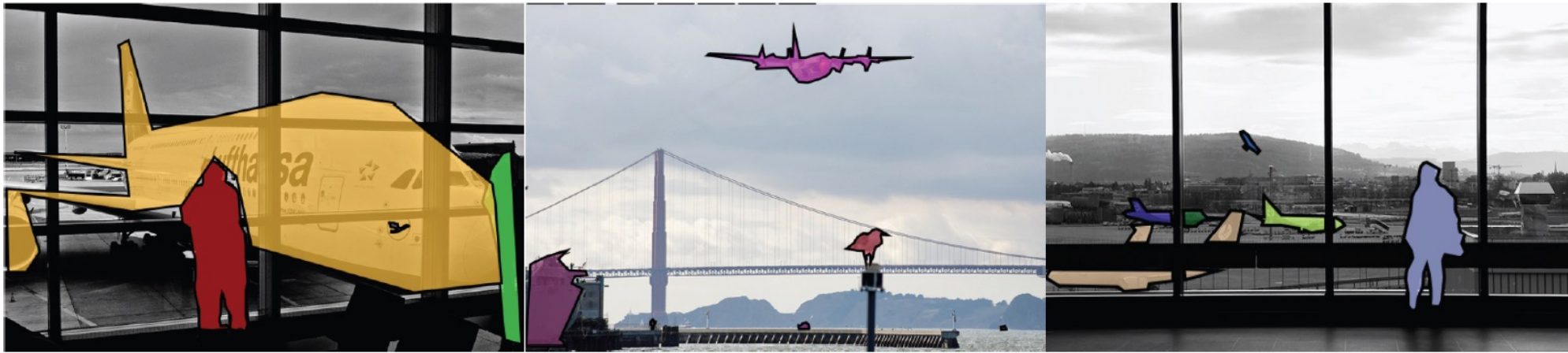


FIGURE 22.8: *The three images in the MS-CoCo dataset that contain an aeroplane, a bird and a person (the bird in the image on the **left** is the airline's logo). Note how objects are delineated with polygons and the relatively rich context in which the objects occur.*

Resources

- Datasets in text
- Numerous segmenters online

Think about this...

- 22.3.** Does stuff have a shape?
- 22.4.** What would be a good procedure to choose the taxonomy of labels for a semantic segmentation problem?
- 22.5.** Give an example of a situation where important things in images occupy few pixels (think, for example, about what a junction looks like to a driver). Use this example to explain why IoU scores may not be the only evaluation you need for a semantic segmenter.
- 22.6.** What is the iIoU score of Section 22.2.1 good for?