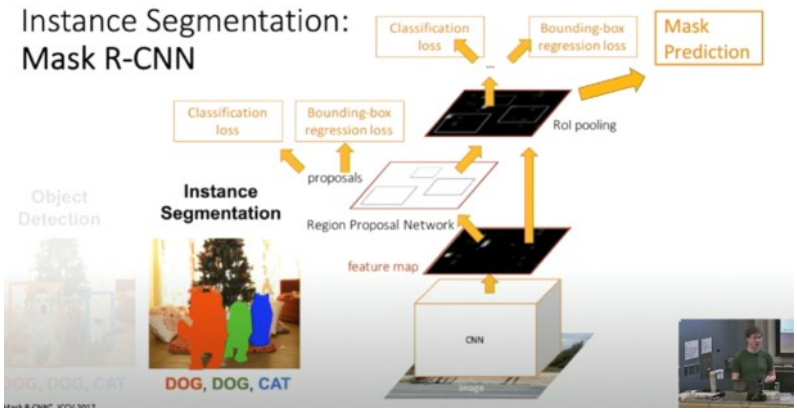
6 Segmentation

2024年12月18日 19:58

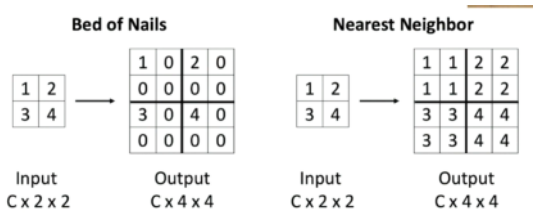
label each pixel in the image with a label. 不知道哪几个pixel属于cow, 有多少只cow

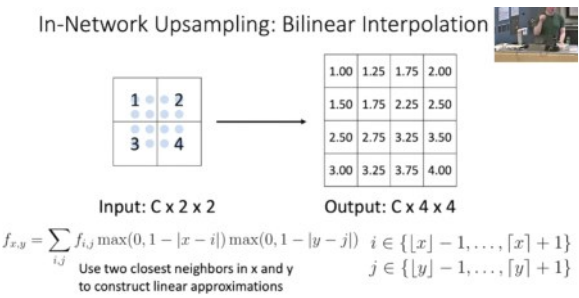
1. patching window. inefficient, not shared features between patches
2. fully conv: bunch of conv o make predictions for pixels all at once.
   a. loss: per-pixel cross-entropy
   b. all conv
      i. effective receptive field size is linear.
      ii. expensive computation
   c. conv downsampling then ***upsampling*** ⟶
3. Things (individual obj instances), stuff (cannot separa)
   a. obj detection: only things
   b. segmentation: semantic:both; instance segmentation (both, obj detec then predict a seg)
4. Mask R-CNN



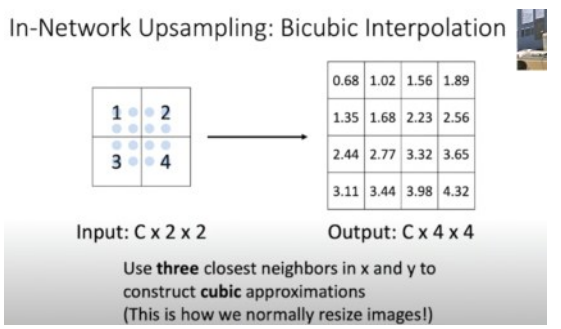
fix functions:

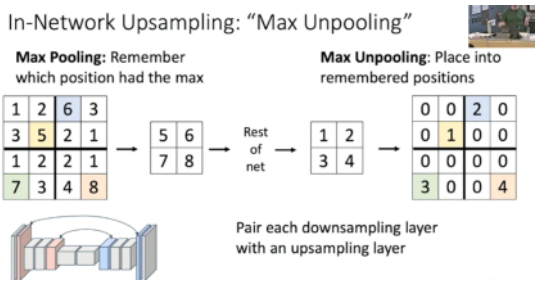
Unpooling





$$f_{x,y} = \sum_{i,j} f_{i,j} \max(0, 1 - |x - i|) \max(0, 1 - |y - j|) \quad i \in \{\lfloor x \rfloor - 1, \ldots, \lceil x \rceil + 1\} \quad j \in \{\lfloor y \rfloor - 1, \ldots, \lceil y \rceil + 1\}$$





和前一步的上采样结合

Learnable upsampling: transposed convolution