Disclaimer: This presentation is a work of fiction written from the perspective of a 2020 researcher traveling back in time to mid 2013 to share some 2020 xNN based application ideas; references to credit the actual inventors of the various ideas is provided at the end

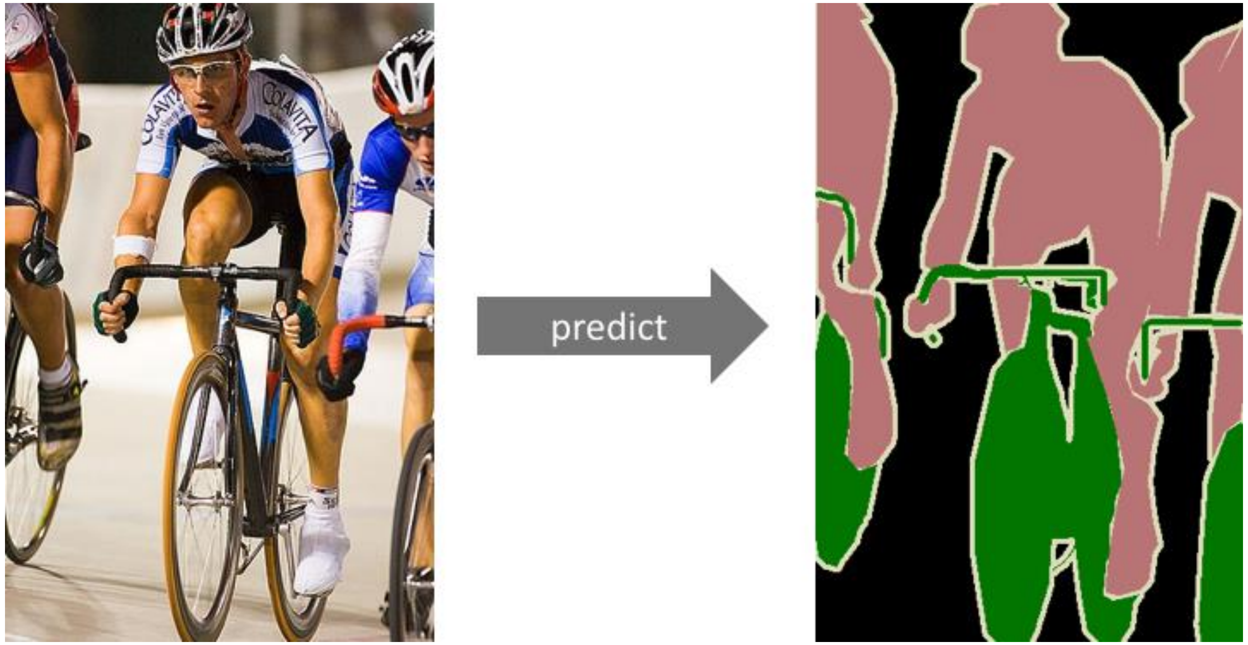
The New Semantic Segmentation

Harrison Jansma
hsj180000@utdallas.edu
May 1, 2020

Disclaimer: This presentation is a work of fiction written from the perspective of a 2020 researcher traveling back in time to mid 2013 to share some 2020 xNN based application ideas; references to credit the actual inventors of the various ideas is provided at the end

Motivation

- Involves realistic understanding of image contents. (what and where)
- Medical imaging, Autonomous cars, …
- Sliding window prediction
  - Slow implementation
  - Loss of localization in large windows



Disclaimer: This presentation is a work of fiction written from the perspective of a 2020 researcher traveling back in time to mid 2013 to share some 2020 xNN based application ideas; references to credit the actual inventors of the various ideas is provided at the end

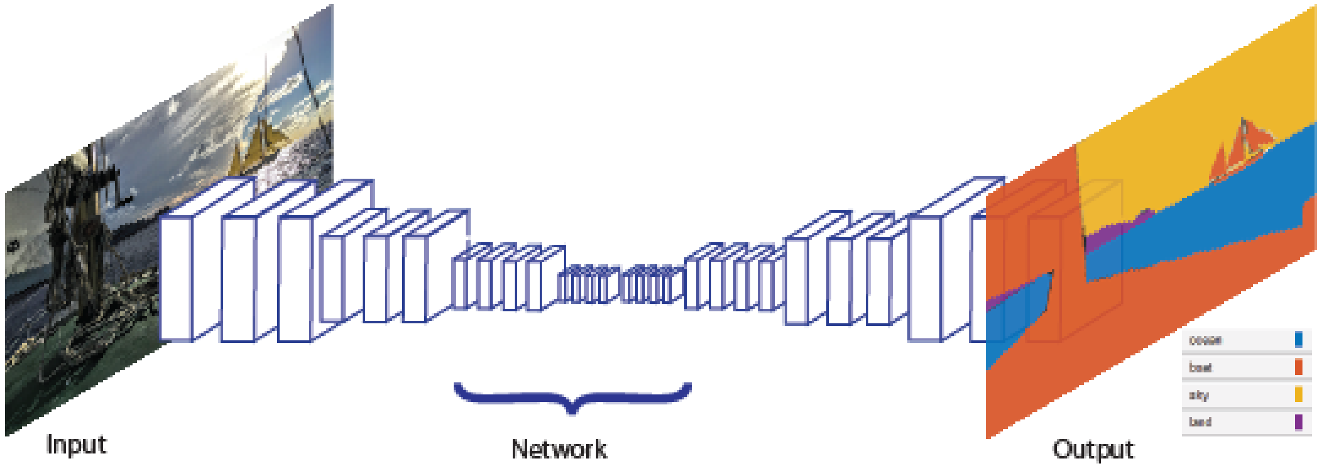
Key Insight

- Map from image -> pixel predictions

- Solution needs to support:
  - Varied in-picture object sizes
  - Fast inference for applications

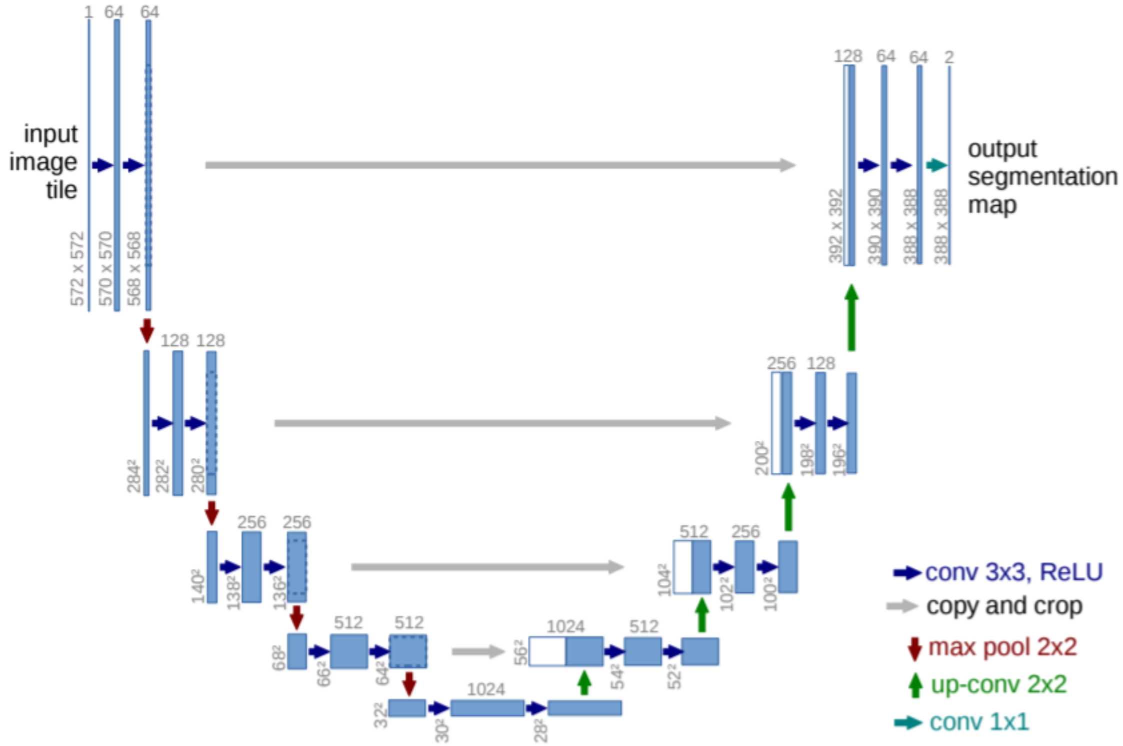
- Fully Convolutional Network with:
  - Image to Image -> Arbitrary image size
  - Depthwise Sep Conv -> Efficient

- HairNet 2: The Thinnest Prediction



Proposed Approach

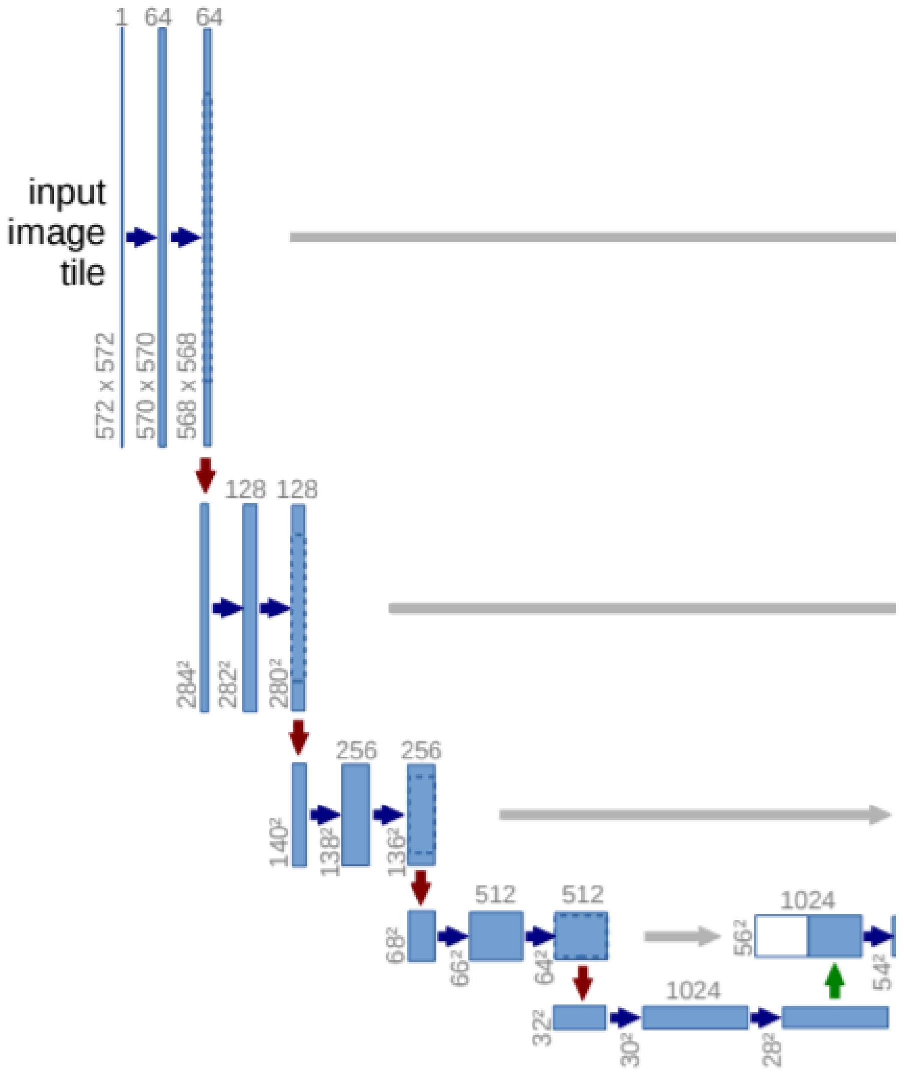
- Combine rich deep features with highly localized early network feature maps
- Many features allow context to propagate during upsampling
- Apply cropping and concatenation to combine early/late feature maps



Disclaimer: This presentation is a work of fiction written from the perspective of a 2020 researcher traveling back in time to mid 2013 to share some 2020 xNN based application ideas; references to credit the actual inventors of the various ideas is provided at the end

Encoder (Spatial Compression)

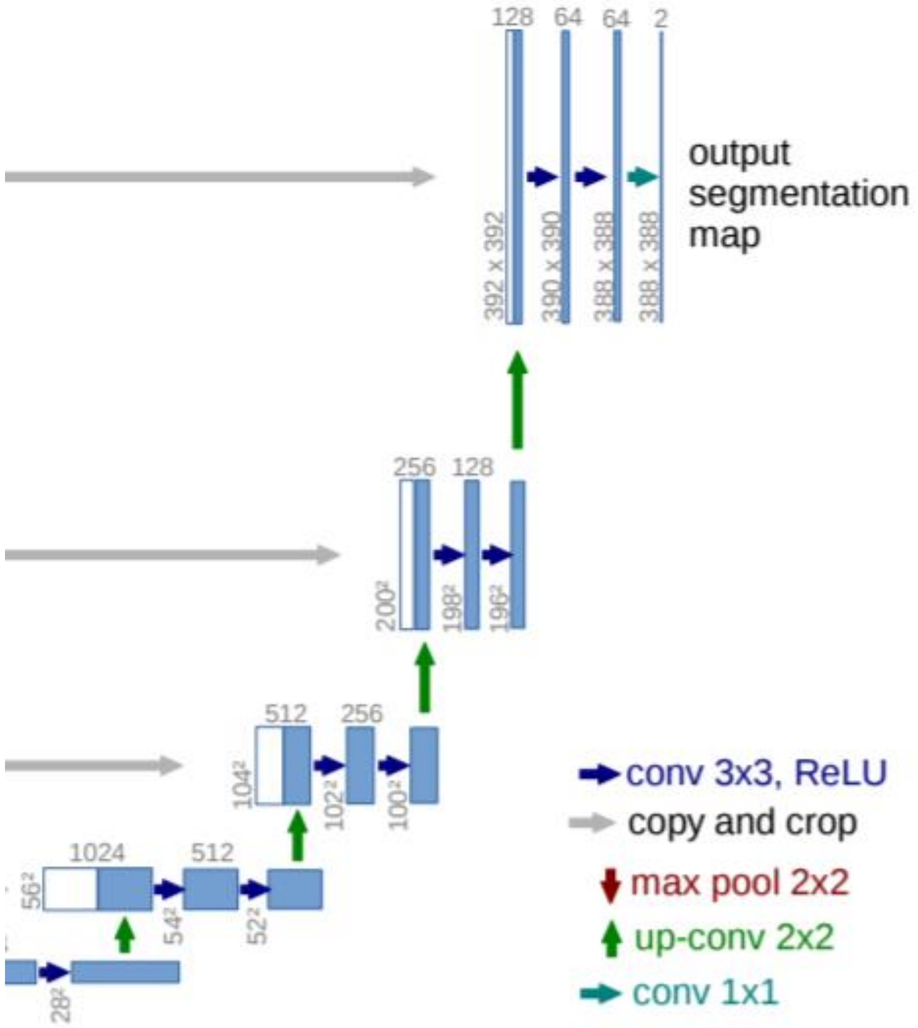
- Standard convolutional architecture
- Series of 3x3 convolutions (or depthwise separable convolutions) followed by maxpool
- Deep feature representations learn WHAT is in a segment of an image.
  - Lose info on WHERE object is in segment



Disclaimer: This presentation is a work of fiction written from the perspective of a 2020 researcher traveling back in time to mid 2013 to share some 2020 xNN based application ideas; references to credit the actual inventors of the various ideas is provided at the end

Decoder (Spatial Expansion)

- “up-convolution” halves number of feature dims and expands height/width
- Concatenate cropped feature maps from earlier stage in the network
- Follow with a series of 3x3 convolutions
- Final layer uses 1x1 conv to map channel dim to number of classes



Disclaimer: This presentation is a work of fiction written from the perspective of a 2020 researcher traveling back in time to mid 2013 to share some 2020 xNN based application ideas; references to credit the actual inventors of the various ideas is provided at the end

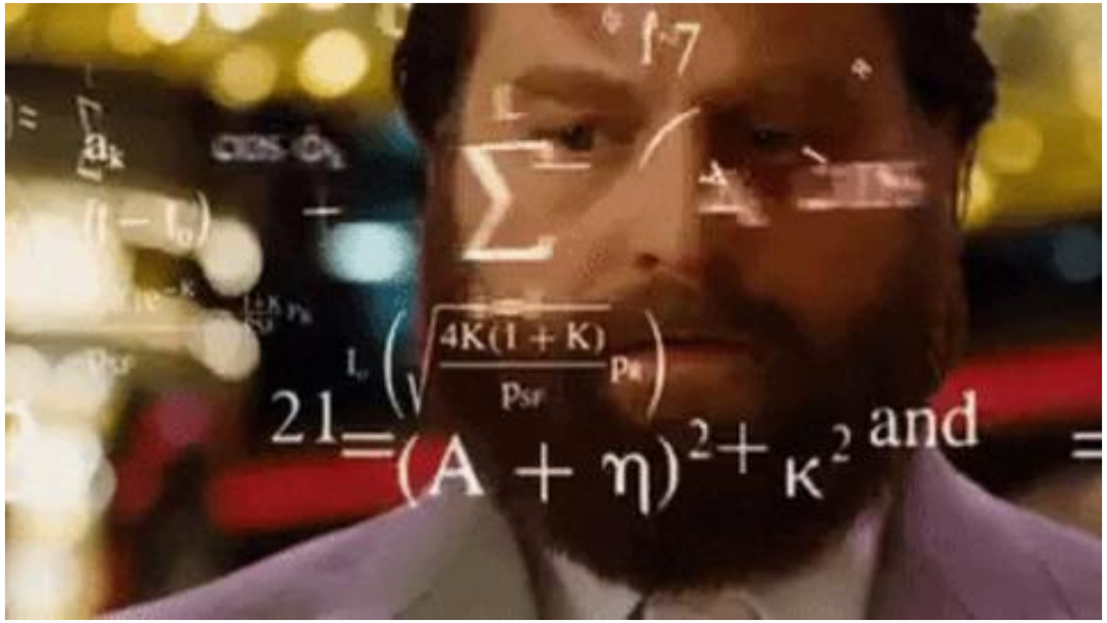
Training Methods

- Adam with high momentum and low batch size.

- Pixel-Wise softmax

$$E = \sum_{\{x \in \Omega\}} w(x) \log(p_l(x))$$

- $l: \Omega \rightarrow [1, \ldots, K]$ is true label
- $w: \Omega \rightarrow R$ a weight map to prioritize more important pixels.



Results

- Image segmentation on Oxford IIIT Dataset
  - https://github.com/harrisonjansma/2020_Notes/tree/master/DL/Courses/CS6301%20CNNs%20UTD/Project/Project%202
- Initialized a MobileNet style encoder
- Trained encoder + decoder to predict pixel classifications

Next Steps

- Some ideas for future experimentation
  - Residual connections with addition instead of concatenations
  - Multiple convolutional filters with different sizes.
- Further experimentation with methods to reduce memory footprint.
  - Bottleneck layers

References

- Note: As stated above, this presentation is a work of fiction; the following are the actual inventors of the ideas described in this presentation

- O. Ronneberger, et. al., “, U-Net: Convolutional Networks for Biomedical Image Segmentation” arXiv:1505.04597, 2015.

- https://github.com/qubvel/segmentation_models

Thank You!