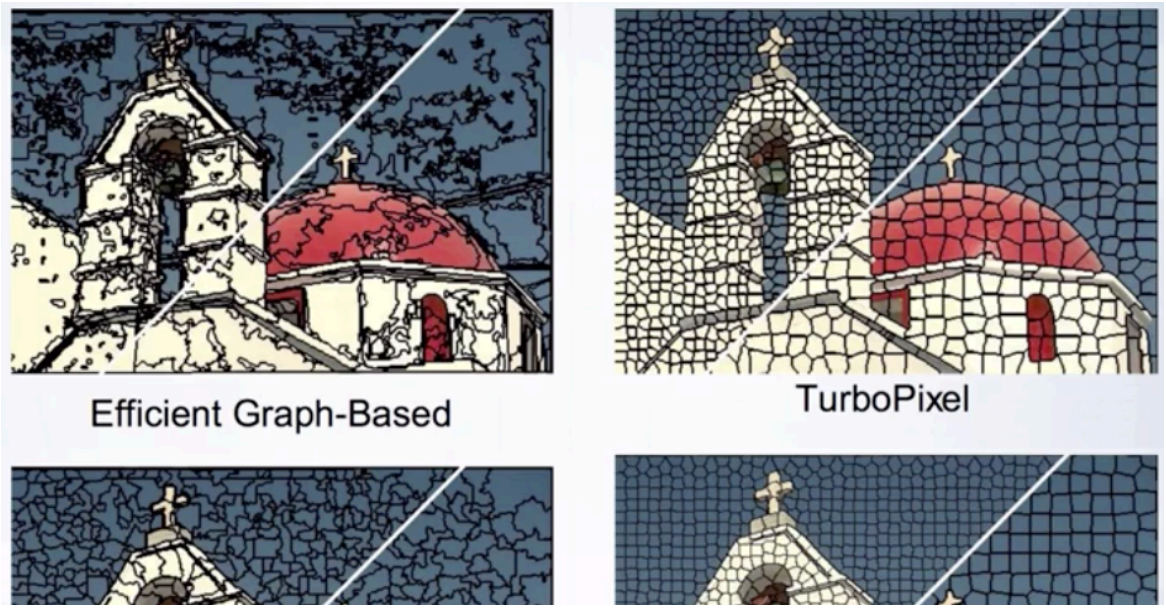
Image Segmentation and Synthesis

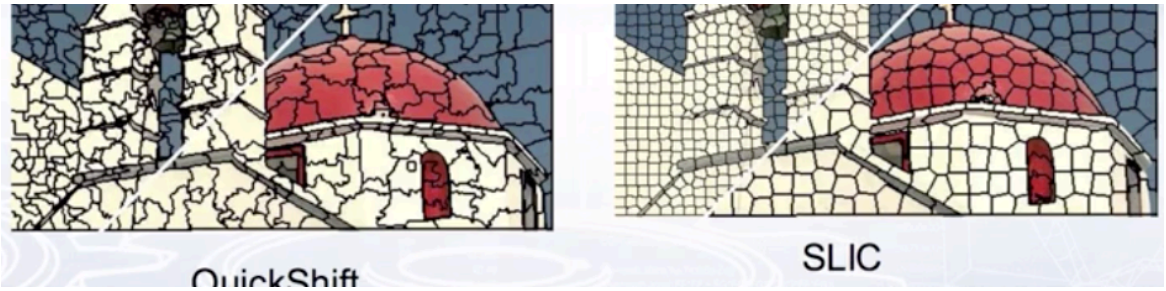
Saturday, March 23, 2019 8:21 PM

★ Image Segmentation

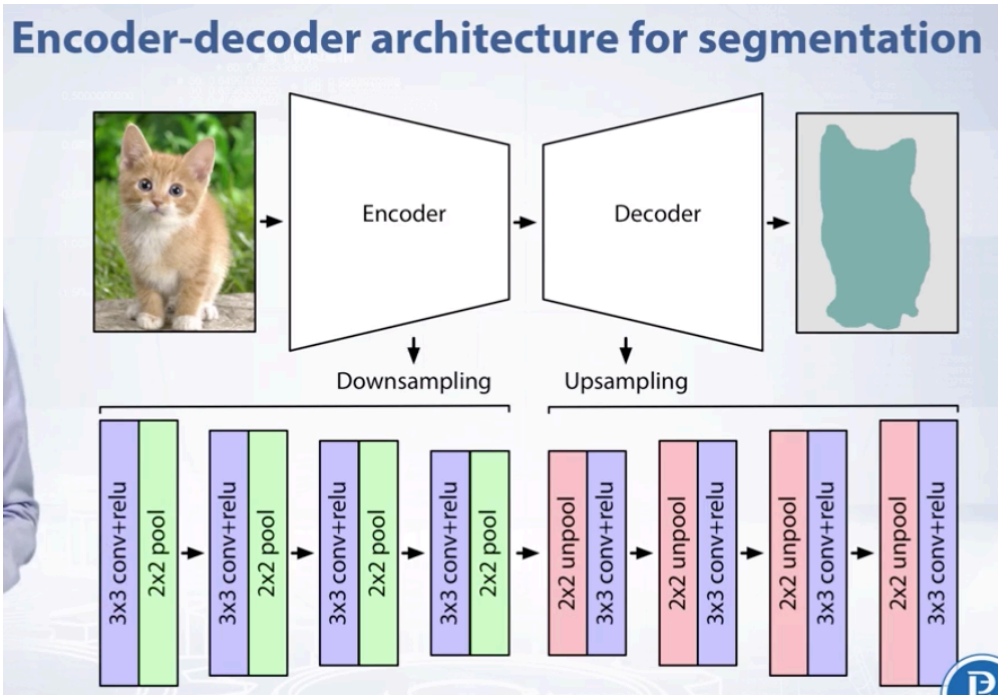
Instance segmentation : label all individual objects of a given instances
Semantic segmentation : per-pixel markup of image where label corresponds to a specific object
Object extraction : selecting a given object
Co-segmentation: Exclude the same object from different images
Unsupervised segmentation: Grouping pixels

1. Over segmentation
   Segment the object into fragments/subcomponents.
   a. Heuristic methods (region growing, split and merge regions)
   b. Graph-based methods
   c. Energy-based methods (snakes, TurboPixels)
   d. Clustering-based approaches (Mean shift, QuickShift, SLIC)
      Mean Shift: For each point, find the nearest mode of the distribution density, Split the entire sample into clusters according to the modes closet to them.
      Simple Linear Iterative Clustering(SLIC): initialize the clusters over the grid at a distance s, use k-means procedure, stop when the change is less than the threshold.





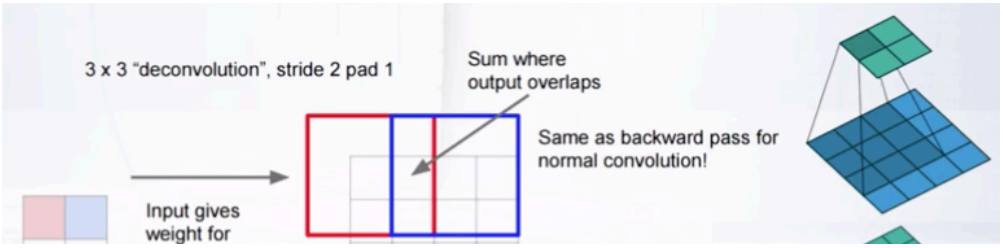
2. Deep learning for image segmentation
   a. Naïve approach: fully convolutional network
      Take a pre-trained CNN, convolutionalize fully- connected layers
      Use multi-class cross entropy loss function
      Fine-tune the fully-convolutional network
   b. Encoder-decoder architecture for segmentation



E.g. SegNet Model:

topologically identical to the 13 convolutional layers in the VGG16 network, transposed convolution in the decoder network, unpooling mechanism with Max Location swithces.

   i. The transposed convolution layer
      Kernel's and stride's sizes remain the same, but now we should use zero padding with appropriate size.



ii. The unpooling layer
   1) Bilinear interpolation
   2) Bed of nails
   3) Max location switches

E.g. U-net

Typical downsampling-upsampling architecture, skip connections with concatenation.

3. Human pose estimation as image segmentation

   Top-down approach: person detector -> single-person pose estimation for each detection
   a. Naïve regression task

   Bottom-up approach:
   a. DeepCut: initial detections/part candidates and pairwise terms -> cluster parts belong to one person

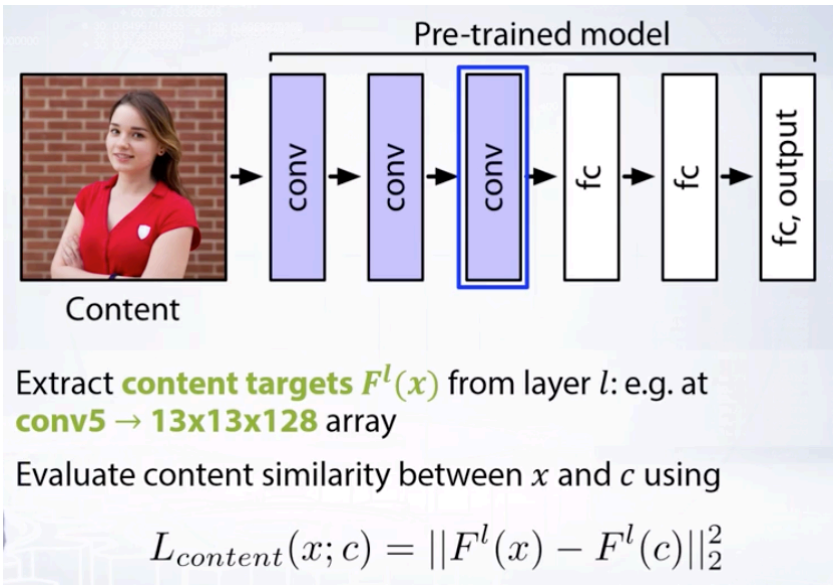
★ Image Segmentation

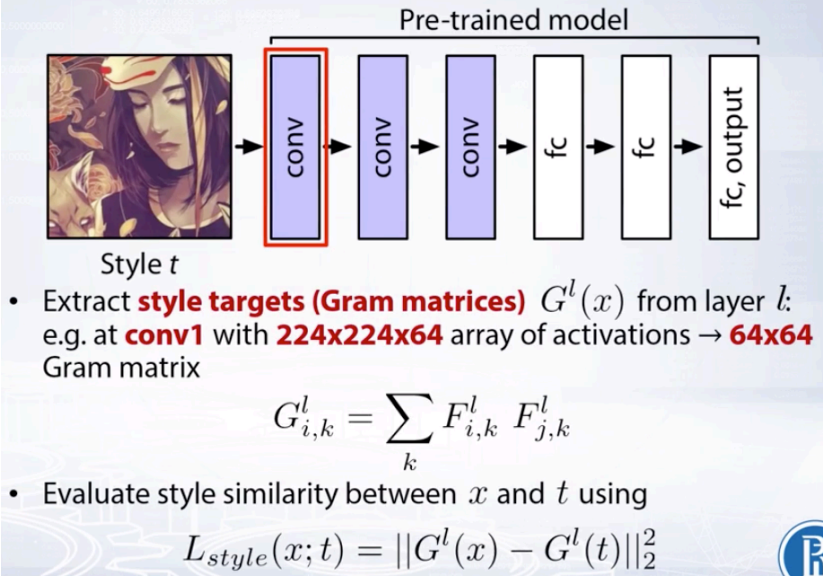
1. Style transfer
   a. Formulation:

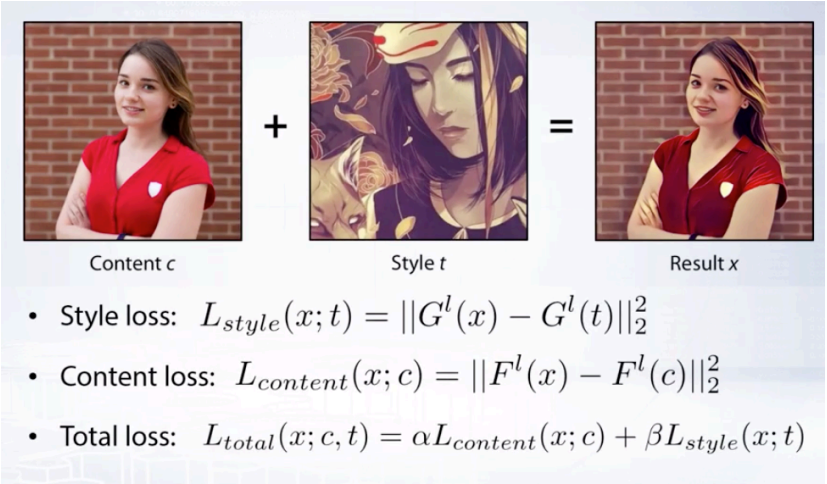
      Content **c** and Style ***t*** are presented ->

      Goal: produce output ***x*** matching the content and the style ->

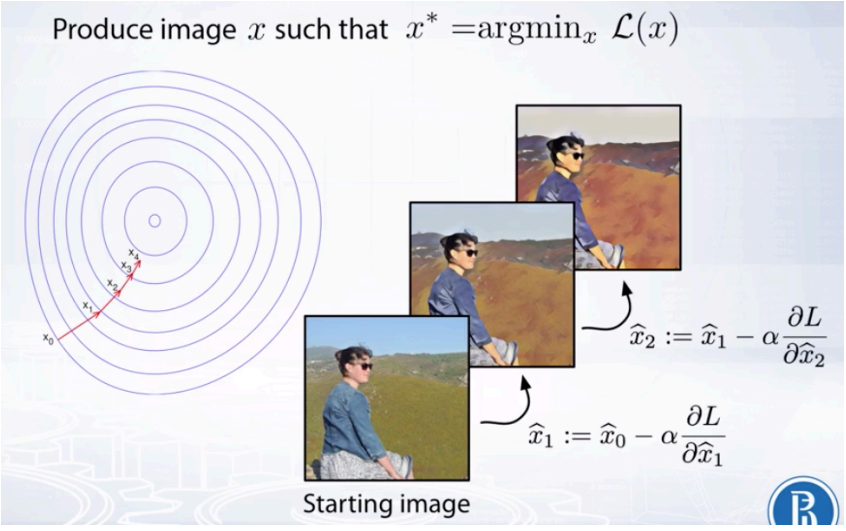
      Using feature activations in a convolutional network as a similarities metric

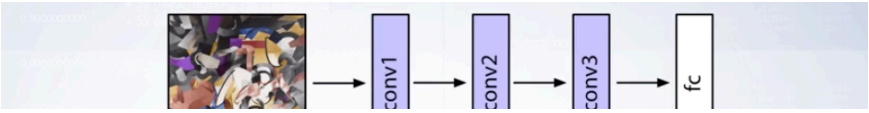


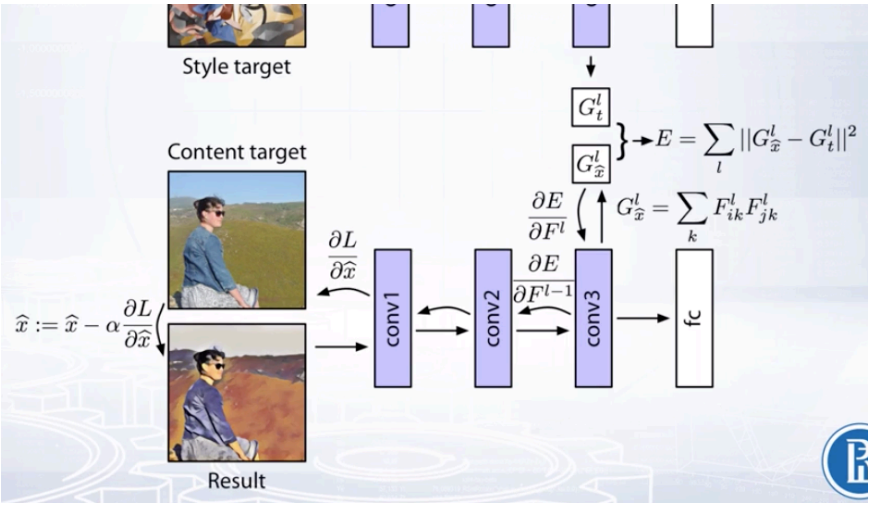




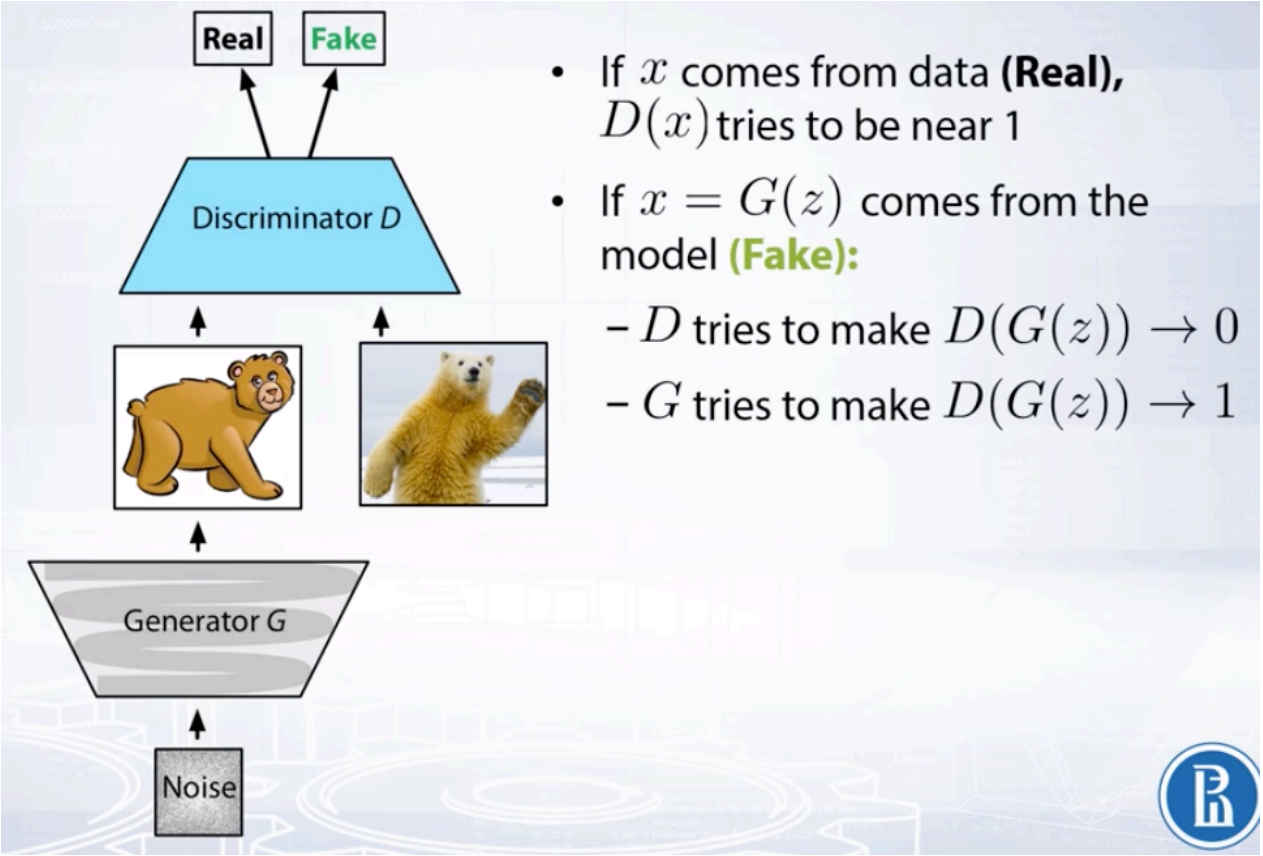
b. Image generation by optimization







2. Generative adversarial network (GAN)



   a. Training objectives for GAN: