Gen. idea

- take two images, produce a new image that reflects the content of one but the artistic style of the other.
- loss function that matches the style & content of each resp. image in the feature space of the deep network, & then performing gradient descent on the pixels of the image itself.
- VGG19, AlexNet → feature extractor
  - trained on ImageNet

Loss

- weighted sum of 3 terms: content loss + style loss + total variation loss.
- penalize deviations from the content of the content image & deviations from the style of the style image.
- perform grad. descent not on the parameters of the model, but on the pixel values of our original image.
- Content loss
  - feature map of gene. image differs from feature map of source image.
  - Content repr / feature map of one layer : A $(1 \times C \times H \times W)$. (reshape)
  - F : feat. map of current image

    P : feat. map of content source image.
  - $L_c = w_c \sum_{ij} (F_{ij} - P_{ij})^2$
- Style loss
  - Compute gram matrix G : correlations betn values in each channel of the feature map.
  - We want activation statistics of our generated image to match the activation stats of our style image.
  - $G_{ij} = \sum_K F_{iK} F_{jK}$
    - Gram matrix from the feature map of the current image.
  - A : Gram matrix from the feature map of the source style image.
  - $L_s^{\ell} = w_s \sum_{ij} (G_{ij} - A_{ij})^2$
  - $L_s = \sum_{\ell \in \text{set of layers}} L_s^{\ell}$
- Total variation regul' (used in perceptual losses for real-time style transf.)

  spatial smoothness
  - that penalises the wiggles / total variation in pixel values.
  - sum of squares of differenc in pixel values for all pairs of pixels that are next to each other
  - $L_{VR} = w_{VR} \left(\sum_{c=1}^{3} \sum_{i=1}^{H-1} \sum_{j=1}^{W} (x_{i+1,c} - x_{i,j,c})^2 + \sum_{c=1}^{3} \sum_{i=1}^{H} \sum_{j=1}^{W-1} (x_{i,j+1,c} - x_{i,j,c})^2 \right)$

Models:

ResNeXt

- VGG-16,19 pretrained
- Our own architecture (ResNet, GoogleNet, good stem & head)
- Model trained on Assignment 2

Dataset:

- ImageNet

Dash

- fast-dash by Kedar.