COMP9444 Neural Networks and Deep Learning

Quiz 4 (Image Processing)

This is an optional quiz to test your understanding of the material from Week 4.

---

1. List five different Image Processing tasks.

    - image classification
    - object detection
    - object segmentation
    - style transfer
    - generating images
    - generating art
    - image captioning

2. Explain the problem of vanishing and exploding gradients, and how Weight Initialization can help to prevent it.

    The differentials in a deep neural network tend to grow according to this equation

    $\mathrm{Var}[\partial/\partial x] \approx (\Pi_{i=1}^{D} G_1 n_i^{\mathrm{out}} \mathrm{Var}[w^{(i)}])\ \mathrm{Var}[\partial/\partial z]$

    where $w^{(i)}$ are the weights at layer i, n_i^{out} is the number of weights fanning out from each node in layer i, and G_1 estimates the average value of the derivative of the transfer function. If the weights are initialized so that the factor in parentheses corresponding to each layer is approximately 1, then the differentials will remain in a healthy range. Otherwise, they may either grow or vanish exponentially.

3. Describe the Batch Normalization algorithm.

    The mean and variance of the activations $x_k^{(i)}$ at layer i over a batch of training items are estimated or pre-computed, and *normalized* activations are calculated for each node
    $\hat{x}_k^{(i)} = (x_k^{(i)} - \mathrm{Mean}[x_k^{(i)}])/\ \mathrm{sqrt}(\mathrm{Var}[x_k^{(i)}])$
    These activations are then shifted and rescaled by
    $y_k^{(i)} = \beta_k^{(i)} + \gamma_k^{(i)}\ \hat{x}_k^{(i)}$
    where $\beta_k^{(i)}, \gamma_k^{(i)}$ are additional parameters to be learned by backpropagation.

4. Explain the difference between a Residual Network and a Dense Network.

    A Residual Network includes "skip" connections which bypass each pair of consecutive layers. These intermediate layers therefore compute a residual component, which is added to the output from previous layers and corrects their errors, or provides additional details which they were not powerful enough to compute.

    A Dense Network is built from densely connected blocks, separated by convolution and pooling layers. Within a dense block, each layer is connected by shortcut connections to all preceding layers.

5. What is the formula for the $(i, j)^{\mathrm{th}}$ entry G^l_{ij} in the Gram matrix at level l of a convolutional neural network? (remember to define any terms that you use)

    $G^l_{ij} = \Sigma_k F^l_{ik} F^l_{jk}$, where F^l_{ik} is the i^{th} filter at depth l in spatial location k.

6. Explain the difference between Texture Synthesis and Style Transfer (both in their purpose, and their cost function).

Texture Synthesis aims to produce an image which matches the texture of a given (perhaps, smaller) image. Its cost function is

$E_{style} = (1/4) \Sigma_{l=0}^{L} (w_l / N_l^2 M_l^2) \Sigma_{i,j} (G^l_{ij} - A^l_{ij})^2$

where:
w_l is a weighting factor for each layer l
N_l, M_l are the number of features, and size of the feature maps, in layer l
G^l_{ij}, A^l_{ij} are the Gram matrices of the original and synthetic image.

Neural Style Transfer aims to combine the content of one image (x_c) with the style of another, to produce a new image x. Its cost function is

$E = \alpha E_{content} + \beta E_{style}$,
where $E_{content} = (1/2) \Sigma_{i,k} \| F^l_{ik}(x) - F^l_{ik}(x_c) \|^2$
