VIR 2019 **Name:** _____

Test 2-A

Time Limit: **Points:** _____

1. In all questions, assume stride denotes length of convolutional stride, pad denotes symmetric zero-padding, rate is dilatation rate of atrous convolution. conv stands for convolution or atrous convolution layer and max is max-pooling layer.

   You are given input feature map (image) $\mathbf{x}$ and kernel $\mathbf{w}$:

   $$\mathbf{x} = \begin{array}{|c|c|c|}\hline 1 & 0 & 2 \\ \hline 2 & 1 & -1 \\ \hline 0 & 0 & 2 \\ \hline \end{array} \qquad \mathbf{w} = \begin{array}{|c|c|}\hline 1 & -1 \\ \hline 0 & 2 \\ \hline \end{array}$$

   Compute outputs of the following layers:

   - $\text{conv}(\mathbf{x},\ \mathbf{w},\ \text{stride} = 1,\ \text{pad} = 0,\ \text{rate} = 2) =$

   - $\text{conv}(\mathbf{x},\ \mathbf{w},\ \text{stride} = 3,\ \text{pad} = 1,\ \text{rate} = 1) =$

   - $\max(\mathbf{x},\ 2 \times 2) =$

2. You are given network (without loss layer) which consists of the convolutional layer and the max-pooling layer. The structure is defined as follows:

$$f(\mathbf{x}, \mathbf{w}) = \max\Big(\text{conv}(\mathbf{x},\ \mathbf{w}),\ 1 \times 2\Big)$$

- Draw computational graph and compute the feed-forward pass for input feature map (image) $\mathbf{x}$ and convolutional kernel $\mathbf{w}$:

$$\mathbf{x} = \begin{array}{|c|c|c|}\hline 2 & 1 & 2 \\ \hline \end{array} \quad \mathbf{w} = \begin{array}{|c|c|}\hline 1 & 0 \\ \hline \end{array}$$

- Estimate gradient wrt kernel $\mathbf{w}$ (i.e. compute local gradients for the max-pooling layer and the convolutional layer and substitute edge-values computed in the feed-forward pass).

$$\frac{\partial f(\mathbf{x}, \mathbf{w})}{\partial \mathbf{w}} =$$

- Update weights using pure SGD ($\alpha = 0.5$).

- Update weights using SGD+momentum ($\alpha = 0.5, \rho = 0.9$). Assume that accumulated speed from previous iterations is $\mathbf{v}_t = [1, 2]$.

3. You want to design a convolutional layer which maps $100 \times 100 \times 3$ input (height$\times$width$\times$depth) on $40 \times 40 \times 6$ output feature map. Propose suitable number of kernels, kernel size, stride and padding.

   number_of_kernels =

   kernel_size_h =

   kernel_size_w =

   kernel_size_d =

   stride =

   pad =

4. You are given function $\mathbf{y} = f(\mathbf{x}, \mathbf{w})$, which maps N-dimensional input $\mathbf{x}$ to N-dimensional output $\mathbf{y}$ (preserves dimensionality) and it is parameterized using weight vector $\mathbf{w}$.

   - Draw computational graph of two layer neural network, where each layer consists of the function $f(\mathbf{x}, \mathbf{w})$ with additional the skip-connection.

   - Derive gradient of the output from the second layer with respect to weights in the first layer.

5. Possible new question - You are navigating a mobile robot through rough environment without outlined path. Design an architecture of neural network model, that takes information from visual sensors (RGB/RGB+Depth) and tells the robot where is the path without obstacles. Neural network should consist of three convolutional layers. However, due to the other processes running on the robot, three layers with minimal number of channels to perserve necessary informations are not computationally possible.

   - Draw computational graph of your neural network. Write how would output of the network look like and why? How would you solve in architecture the problem of computational capacity and still keep necessary information from first layers throughout the network?

- Which loss function would you use and how you calculate it from output?