VIR 2019 **Name:** _____

Test 2-B

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 $\mathbf{w}$:

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

   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 $\mathbf{v}_t = [2, 1]$.

3. You have a neural network which consists of a single convolution layer and single fully connected layer. The input image $8 \times 8 \times 3$. The convolution layer consists of a 3×3 kernel with stride 1 and creates 6 feature maps, preserving the dimensions of the image. The fully connected layer takes in the vectorized input from the feature maps and outputs a vector of dimension 2. Assume that there are no activation functions and that the vectorization is free. Calculate the following (Label each step of your calculation, it will be easier to give points):

   - Number of parameters (in megabytes) of the a)convolutional and b)fully connected layers. Weights are stored as single precision floats.

- Approximately how many operations does a feed forward pass take? Assume every addition and multiplication counts as a single operation.

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.