ECE C147/247, Winter 2023 **Homework #5**
Neural Networks & Deep Learning Prof. J.C. Kao
UCLA, ECE TAs: T.M, P.L, R.G, K.K, N.N, S.R, S.P, M.E

Due Friday, 3 March 2023, by 11:59pm to Gradescope.
100 points total.

1. (10 points) **Understanding receptive fields**
In a fully connected network, the value/output of every unit (i.e. neuron) depends on the entire input to the network. However, in convolutional neural networks, we typically use filters smaller in size than the input and hence the value of a neuron depends on a part of the input. This region of input space that affects the value of a neuron is called receptive field of the neuron.
In other words, we say that the receptive field of a neuron is n if it's output depends on an $n \times n$ patch of the input. More rigorously, receptive field is a region of space and must be represented as $n \times n$. However, when we consider square kernels, we can denote it using a single variable, since both dimensions are equal. Receptive fields are key to diagnosing potential issue with CNNs. They also help us identify any blind spots in the input and ensure visibility of all input pixels at the output so that no relevant information is missed out while making a prediction.

Consider a toy CNN with 3 convolution layers (CL), powered by ReLU activations, a fully connected layer followed by the output, as shown below:

$$(Input) \rightarrow (CL_1) \rightarrow (ReLU_1) \rightarrow (CL_2) \rightarrow (ReLU_2) \rightarrow (CL_3) \rightarrow (ReLU_3) \rightarrow (FCL) \rightarrow (Output) \rightarrow (Softmax) \rightarrow Loss$$

The kernel size of these layers are $CL_1 : m_1 \times m_1$, $CL_2 : m_2 \times m_2$, $CL_3 : m_3 \times m_3$ and stride is set to 1 for all layers.
(**Note**: For the following questions it maybe be helpful to draw a couple of layers of a simple network to visualize the neurons' fields of view)

(a) (1 point) What is the receptive field of each neuron in (CL_1)?

Solution: Each neuron in (CL_1) sees an $m_1 \times m_1$ patch of the input image. Therefore the receptive field of every neuron in $(CL_1) = m_1$

(b) (2 points) What is the receptive field of each neuron in (CL_2) ?

Solution: A neuron in (CL_2), say n_2, sees an $m_2 \times m_2$ patch of the feature map from (CL_1). The first neuron in this patch (from CL_1) sees m_1 inputs. Now, n_2 sees

$m_2 - 1$ more neurons from CL_1, each of which add one additional input (because stride is 1) to n_2's field of view. Therefore, the resultant receptive field of $n_2 = m_1 + (m_2 - 1) * 1$

(c) (3 points) Suppose we introduced a stride of s_1, s_2, s_3 on conv layers CL_1, CL_2 and CL_3 respectively, how would this affect the receptive fields of CL_1 and CL_2 ?

Solution: Stride of a given layer would increase the receptive field of neurons in the following layers. The receptive field of CL_1 is unaffected by it's stride and would remain m_1.
The stride s_1 of CL_1 amplifies the receptive field of CL_2 i.e., each of the $m_2 - 1$ neurons that n_2 sees add s_1 additional inputs to it's field of view. Thus, the new receptive field becomes $m_1 + (m_2 - 1) * s_1$.

(d) (2 points) Based on the patterns you observed in the previous question, write down a generalized expression for the receptive field of a neuron in the k^{th} layer of a CNN with n convolution layers (take m_i and s_i as the kernel size and stride of each layer).

Solution: To get a better sense of the generic pattern, let us solve for the receptive field of a neuron n_3 in CL_3. Following the analysis in part (c), we can say that n_3 sees an $(m_3 \times m_3)$ patch of CL_2's feature map. The first neuron in this patch sees $m_1 + (m_2 - 1) * s_1$ inputs (note that this is just receptive field of n_2). Rest of the $m_3 - 1$ neurons see s_2 additional neurons in CL_1, each of which in-turn adds s_1 new inputs to n_3's field of view.
Therefore, receptive field of $n_3 = m_1 + (m_2 - 1) * s_1 + (m_3 - 1) * s_1 * s_2$. Extending this logic, we can write the generic expression for the receptive field of a neuron in the n^{th} layer as,
$1 + \sum_{j=1}^{k} (m_j - 1) \Pi_{i=1}^{j-1} s_i$

(e) (2 points) Mention two ways to increase the receptive field of neurons in a CNN.

Solution: Some ways to increase the receptive field include:
1. Stack up more layers to make the network deep
2. Increase stride
3. Add pooling layers
4. Increase the kernel size

Coding : You should complete the notebooks in order, i.e., CNN-Layers, followed by CNN-BatchNorm, followed by CNN. This is due to potential dependencies. Note however, that CNN can be completed without CNN-Layers, since we provide the fast implementation of the CNN layers to be used in question 4.

2. (40 points) **Implement convolutional neural network layers.** Complete the CNN-Layers.ipynb Jupyter notebook. Print out the entire workbook and relevant code and submit it as a pdf to gradescope. Download the CIFAR-10 dataset, as you did in earlier homework.
Solution: As a policy, we do not release solutions to coding questions.

3. (20 points) **Implement spatial normalization for CNNs.** Complete the CNN-BatchNorm.ipynb Jupyter notebook. Print out the entire workbook and relevant code and submit it as a pdf to gradescope.

   Solution: As a policy, we do not release solutions to coding questions.

4. (30 points) **Optimize your CNN for CIFAR-10.** Complete the CNN.ipynb Jupyter notebook. Print out the entire workbook and relevant code and submit it as a pdf to gradescope.

   Solution: As a policy, we do not release solutions to coding questions.