EE6310 Image and Video Processing, Fall 2021

Indian Institute of Technology Hyderabad
Homework 5, Assigned 08.11.2021, Due **11:59 pm on 15.11.2021**

The only source of knowledge is experience – Albert Einstein

1. In this assignment, implement a Convolutional Neural Network (CNN) in the *feed forward mode*. The CNN architecture must is as follows: (30)

   - Conv layer with 4 kernels of size 3×3 and Rectified Linear Unit (ReLU) activation
   - Max pooling layer of size 2×2 with a stride of 2 along each dimension
   - Conv layer with 8 kernels of size 3×3 and Rectified Linear Unit (ReLU) activation
   - Max pooling layer of size 2×2 with a stride of 2 along each dimension
   - Flattening layer
   - An MLP with one hidden layer that accepts as input the flattening layer output and maps it to 10 output nodes. Use ReLU activation for the MLP as well.

   Instructions:

   - Use color images from the CIFAR-10 dataset – `https://www.cs.toronto.edu/~kriz/cifar.html`. In other words, the input to the CNN is of size $32 \times 32 \times 3$ i.e., spatial dimension of 32×32 and three channels - Red, Green and Blue.
   - As discussed in class, ensure that the kernels have the same depth (or number of channels) as the input at each layer.
   - Write a function that performs channel-wise convolution. Refer to the class notes. The function must accept the image and the kernel as input and return the convolved image as the output. Ensure that the convolution output has the same spatial dimensions as the input. To do so, use appropriate indexing into the regular convolution output that picks out the "central" region.
   - Randomly initialize the weights of all the network parameters.
   - Pick 5 images from each of the 10 classes. For each input image do the following:
     - Visualize the activations from every kernel in every layer before and after pooling.
     - Instead of randomly initializing the CNN kernels, use band pass filters (difference of Gaussian). Experiment with different variances of the Gaussian to achieve different band pass characteristics. Refer to the class notes for Gaussian filters. Repeat the previous step and output the resulting activations.
     - For each of the cases (random and band pass kernels), visualize the output of the flattening layer using t-SNE plots (using open source library). Note down your observations. With the band pass kernels, do you notice any class separation in the t-SNE plots?