**Homework 2**

**ELEC 677 Deep Learning**

**Dr. Ankit Patel**

**11/17/16**

**Joshua Michalenko**

Zeiler and Fergus present a non-parametric visualization of convolutional neural network hidden units that show which features in the data set activate the feature map the most. Their approach allows visualization of layers in the neural network that are more intermediate than the first layer or two of the network like previous works have show. Their approach uses LeCunNet and Krizhevsky’s nueral network from 2012 as model organisms to view hidden unit activations and is a way to map network activations from intermediate layers to the input pixel space with a Deconvolutional Network (deconvnet). To look at each layer, a deconvnet is attached to each of the layers in the network using unpooling, rectification and filtering to learn the input (feature) from the previous layer to that layer that maximizes it’s activation. This allows the deconvnet to learn the previous layer’s input that maximizes the response in the image space in the following layer.

The authors show the top 9 feature patterns reconstructed from the validation set that maximize the layers feature map output as well as the corresponding images for several layers on the different networks. They show that layer 1 looks for edges, layer two looks for simple patterns, layer 3 looks for larger objects such as wheels, pumpkins, and bird beaks, layer 4 is most excited looks at bigger objects such as parts of animals, and lastly, layer five are complete scenes. Each feature maps shows a strong grouping of activations with greater invariances (such as occlusion) at higher layers in the network.

By visualizing the layers in Krizhevsky’s network, they saw that the first and second layers capture high and lower frequencies respectively. To remedy this, they change the filter sizes and create a network that performs better than the original by focusing on intermediate frequencies as well. They continued this process and came up with a network that surpasses all networks and becomes state of the art in object classification. This new visualization technique shows how they can debug the network with some intuition. They show many experiments from the Imagenet dataset, Caltech datasets, and PASCAL VOC 2012 dataset. They achieve a top 5 misclassification result of 14.8% which beats Krizhevsky’s network by 1.7%. They also show some impressive results that quantify occlusion in images and how well the network can cope with a covered part of the image, signaling that the network generalizes well. They show this by using the same network but training a different softmax regression layer on the Caltech datasets and performing state of the art. They also show how prediction accuracy changes as a function of the number of training images per class. Another analysis they do is training a softmax and SVM classifier on several different layers on the network and show that the classification accuracy increases with depth. Overall they show how visualizing the network can lead to tremendous gains in prediction performance and how CNNs and generalize well to a variety of tasks.