Semi-supervised training of CNNs

Student: Kevin Duarte
Mentors: Dr. Boqing Gong
Yang Zhang

Simple Network results

- Autoencoder structure
  - Input(3, 32, 32) - (128)3c - 2p - (64)3c - (3)3c - (64)3c - (128)3c - 2up - (3)3c=Output(3, 32,32)
  - (128)3c means a convolutional layer with 128 feature maps and a kernel size of 3, 2p means a 2x2 max pooling layer, and 2up means an unpooling layer.
- Trained on 50,000 cifar10 images, using L2 loss

Autoencoder	Image reconstruction loss
Basic max-pooling/upscaling with 1 final loss	1.54628
What-where pooling/unpooling with intermediate losses	1.19630

Simple Network results

- Classification CNN structure
  - Input(3,32,32) - (128)3c - 2p - (64)3c - (3)3c - 128fc - 10fc
  - 10fc is a fully connected layer with 10 hidden units
- Results:

CNN	Accuracy
Randomly initialized, trained on 10,000 images	40.25%
Randomly initialized, trained on 50,000 images	48.64%
Initialized using basic autoencoder, trained on 10,000 images	40.53%
Initialized using ww autoencoder, trained on 10,000 images	42.91%

Simple Network results

- The semi-supervised learning did not show much of an improvement in this network.
- What-where pooling/unpooling only increased the accuracy by 2%, probably since the network was so shallow

Coming Weeks

- I will create a deeper CNN/Autoencoder for semi-supervised learning
- We will use the VGG16 CNN as the encoder