Object Classification with ImageNet

Rami Saleh, Lance Barcelona
Department of Computer Engineering, Northeastern University

Problem

- We live in a world of images that would take us years to sort through.
- We need a system that can classify the millions of images we have at our disposal
- That is where computer vision comes in; more specifically, deep learning neural networks.

Deep Learning Neural Networks

- Using deep learning, we are capable of teaching a neural network the ability to sort through and label images for us.
- Intelligent neural networks already surpass the accuracy of a human but we want them to be even more accurate.

Methods

- In tackling this problem, we turned our attention to the ImageNet Large Scale Visual Recognition Challenge (ILSVRC).
- The ImageNet challenge is currently one of the largest competitions in computer vision where participants work to increase the accuracy of their network architectures.
- For our research, we used the validation dataset from ImageNet to test current neural architectures.
- Rather than adding more layers to the current architecture, we focused our research on modifying the amount of crops the network uses when identifying an image in order to increase its accuracy.

Figure 1: Image Classification on Real World

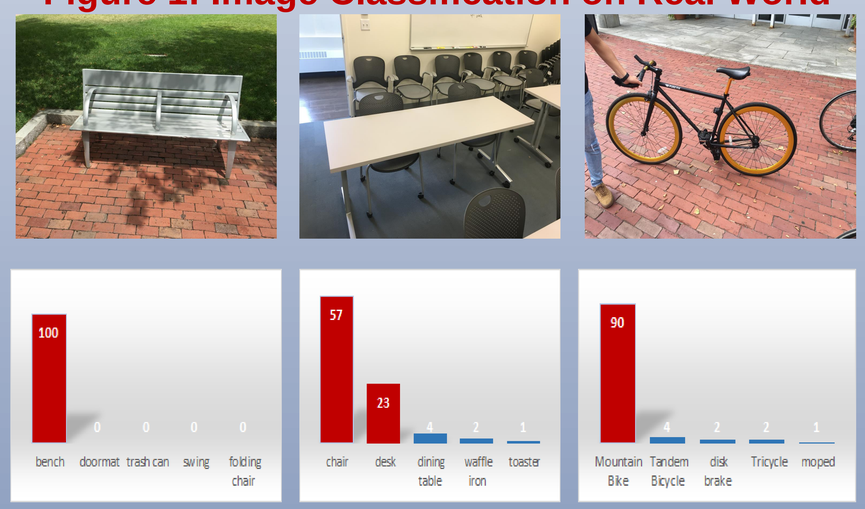

Figure 1: These three pictures were taken in Northeastern's campus. We ran our classifying code on the pictures to test the efficiency of our code. For the images of a bench and a bike, we got excellent results as shown. The image of the chairs and the desk was taken in our workplace. The mixed results that the classifier predicted were expected due to the presence of several chairs and only one desk.

Figure 2: Confusion Matrix for Classified Images

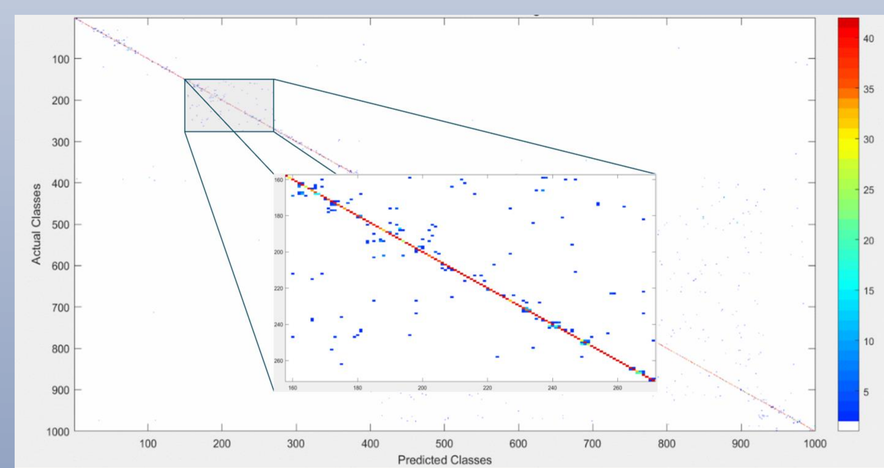

Figure 2: A confusion matrix that graphically displays how many images are classified correctly or incorrectly. In a confusion matrix, the predicted classes are on the x-axis while their actual classes are on the y-axis; images that fall on the diagonal are the images correctly classified while those that do not are misclassified

Figure 3: Increasing the Crops

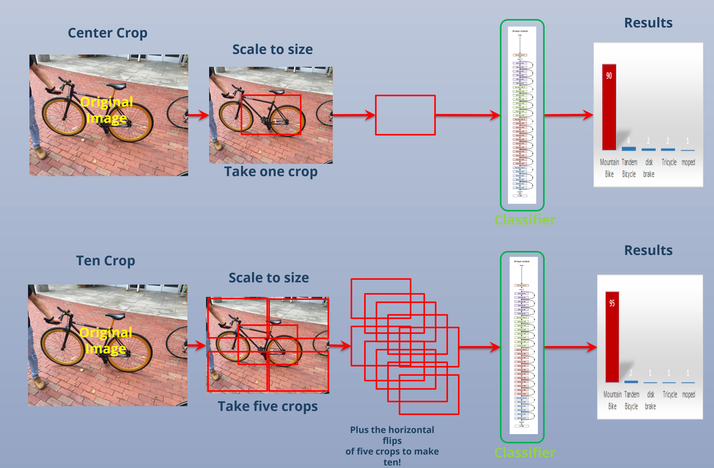

Figure 3: A depiction of the algorithm an image goes through when being classified. First, the original image is scaled to a size. In our classifier, it is scaled to 256. Then, parts of the image are cropped. Those crops are then fed through the convolutional neural network that finally, outputs the results.

Results

- The accuracy of a classifier is measured in top-1 and top-5 error rates.
- An error occurs if the first predicted class (top-1) or one of the five predicted classes (top-5) does not match the ground truth label.
- There was a decrease in the classifier's error rate as we increased the number of crops used.

Neural Architecture	Top-1 Error	Top-5 Error
ResNet-200 (Center Crop)	21.66	5.79
Resnet-200 (10-Crop)	19.89	4.81
ResNet-200 (144-Crop)	19.15	4.37

Conclusions

Our research shows that we are able to achieve a better error rate using our classifier just by increasing the number of crops on each image being analyzed. However, the rate at which the classifier analyzes the dataset will decrease (longer runtime).

Acknowledgments

Dr. Yun (Raymond) Fu, Associate Professor.
Dr. Yu Kong, ECE Department. Yue Wu, PhD Candidate.
National Science Foundation.

References

[1] Karpathy (2014) What I learned from competing against a ConvNet on ImageNet, http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet/. Accessed July 13, 2016.
[2] Eckel (2015) Microsoft Researchers' Algorithm Sets ImageNet Challenge Milestone, https://www.microsoft.com/en-us/research/microsoft-researchers-algorithm-sets-imagenet-challenge-milestone/. Accessed July 14, 2016.