ImageNet Classification with Deep Convolutional Neural Networks [Krizhevsky et al. 2012]

1. First Large Scale Application of CNNs.
2. Developed a system called AlexNet
3. Applied on subset of ImageNet dataset
    a) As part of LSVRC (Large Scale Visual Recognition Challenge)
    b) More than 1 million images – 1.2 million for training, 50 k for validation, 150k for testing
    c) More than 1000 object categories
    d) Evaluation on Top 1/Top 5 Accuracy

AlexNet Architecture

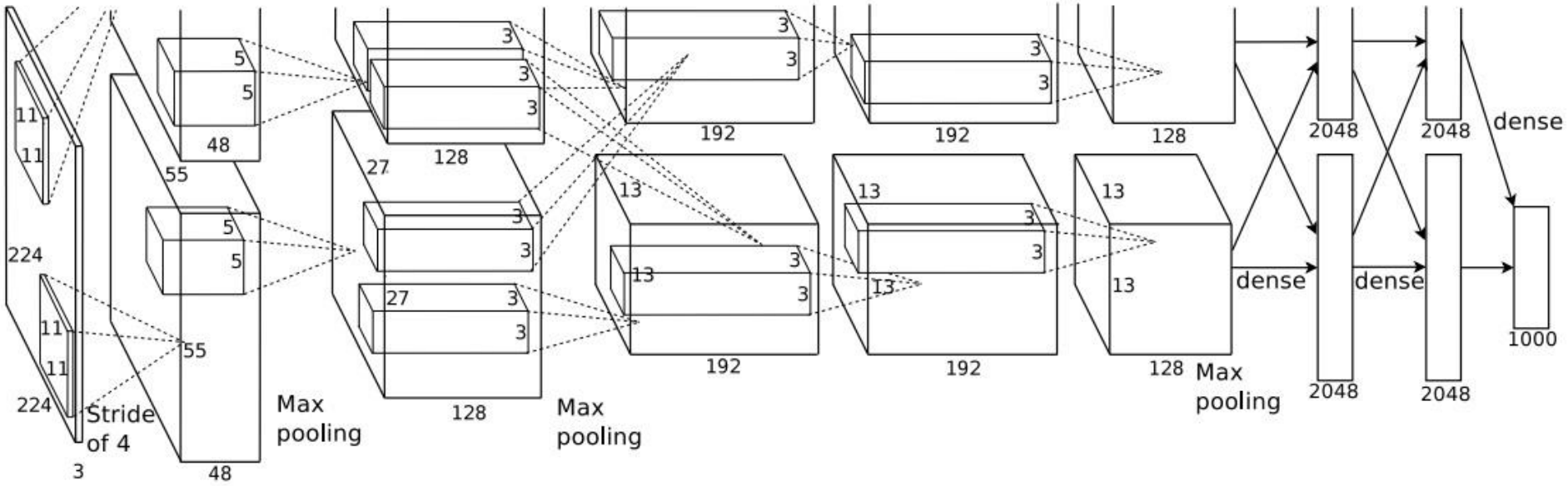
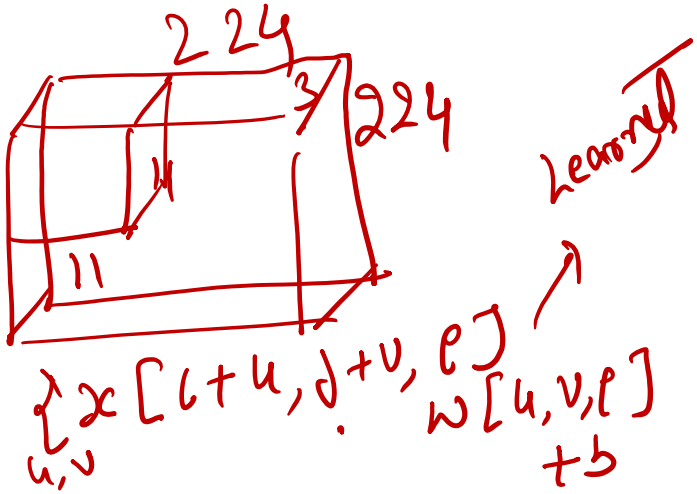


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.

Image & Caption Source: Krizhevsky et al. (2012)

Kernels Learned

1. Kernels can be thought of as feature map
2. Kernels learned in first layer shown on the right side



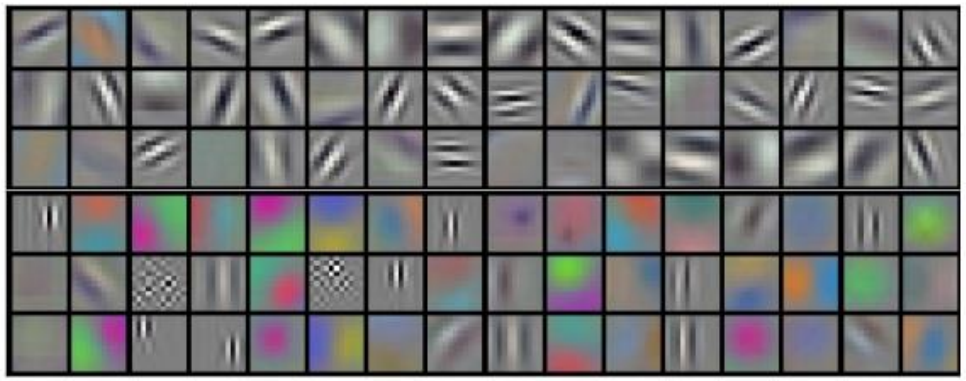


Figure 3: 96 convolutional kernels of size $11 \times 11 \times 3$ learned by the first convolutional layer on the $224 \times 224 \times 3$ input images. The top 48 kernels were learned on GPU 1 while the bottom 48 kernels were learned on GPU 2. See Section 6.1 for details.

Image & Caption Source: Krizhevsky et al. (2012)

Visual Results

Significant Improvement in Results: Up to 8% improvement in results over then state-of-the-art.

Error Rates:
Top 1: 37.5%
Top 5: 17%

(See Table 1 in the Paper)

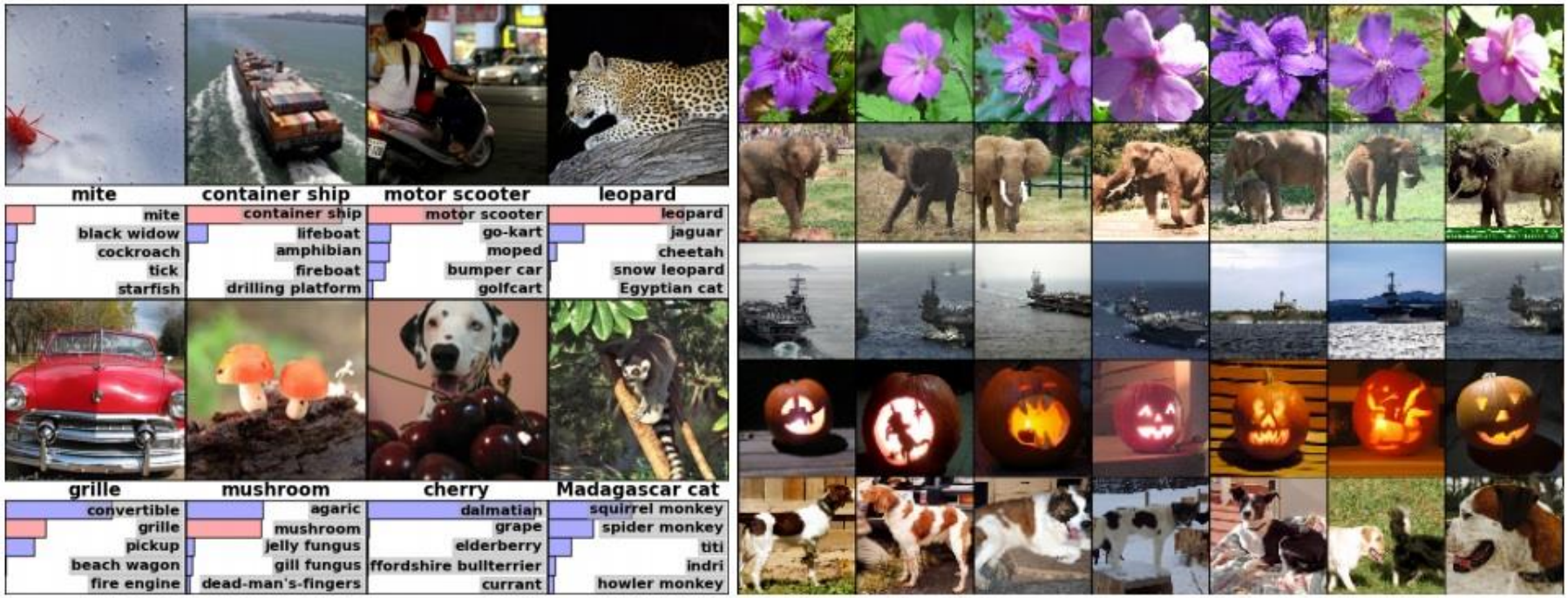


Figure 4: **(Left)** Eight ILSVRC-2010 test images and the five labels considered most probable by our model. The correct label is written under each image, and the probability assigned to the correct label is also shown with a red bar (if it happens to be in the top 5). **(Right)** Five ILSVRC-2010 test images in the first column. The remaining columns show the six training images that produce feature vectors in the last hidden layer with the smallest Euclidean distance from the feature vector for the test image.

Image Source: Krizhevsky et al. (2012)

Some Cool Applications

Following website describes some cool applications of deep learning all the way from colorization of black & white images, to automated text generation.

Check them out:
https://machinelearningmastery.com/inspirational-applications-deep-learning/

Spend about 10 minutes reading through the blog and watching relevant images/videos