Sidenote: ImageNet Dataset

Recitation 4

Spring 2022, CMU 10-403

ImageNet

Dataset for visual object recognition

- > 14 Million images, > 20,000 categories
- Provides ontology over image categories

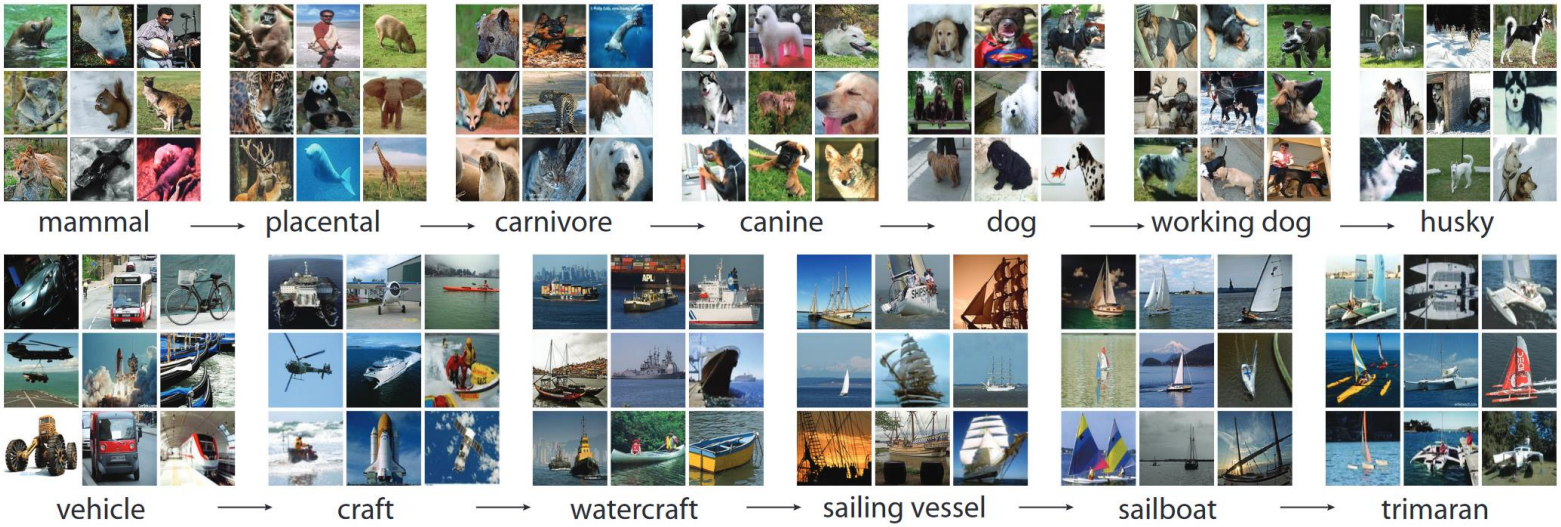


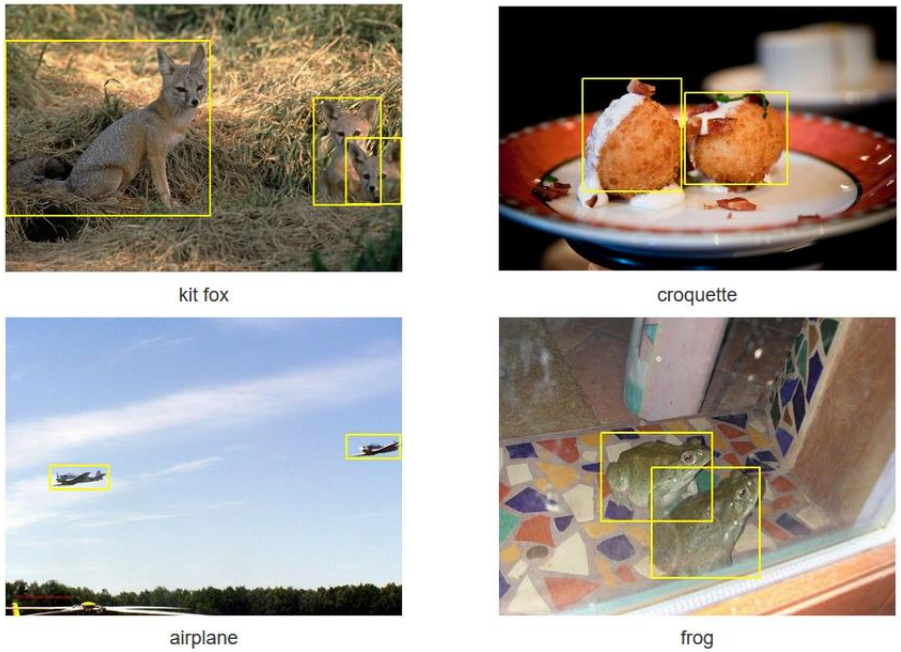
Figure 1: A snapshot of two root-to-leaf branches of ImageNet: the **top** row is from the mammal subtree; the **bottom** row is from the vehicle subtree. For each synset, 9 randomly sampled images are presented.

Source: He

ImageNet

Dataset for visual object recognition

- Provides bounding boxes for over 3000 classes
- On average 150 images per class



Source: He

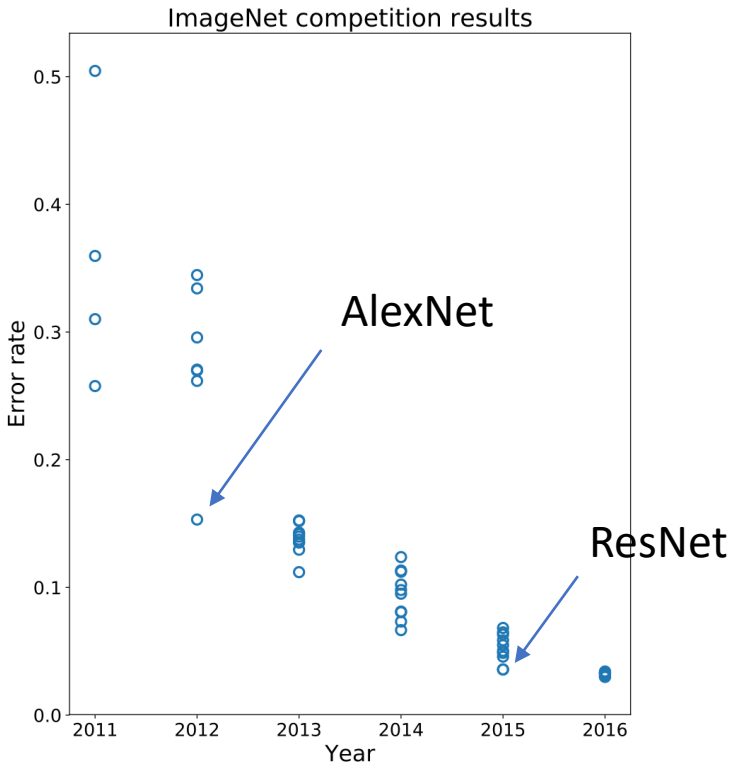
ImageNet

Popular benchmark dataset which has been used for annual competitions and many research papers

ImageNet 2012 Challenge attracted major public attention

-> Deep learning approach reduced top-5 error by 10.8% compared to other machine learning approaches

Start of „deep learning revolution“



Source: Wikipedia

AlexNet

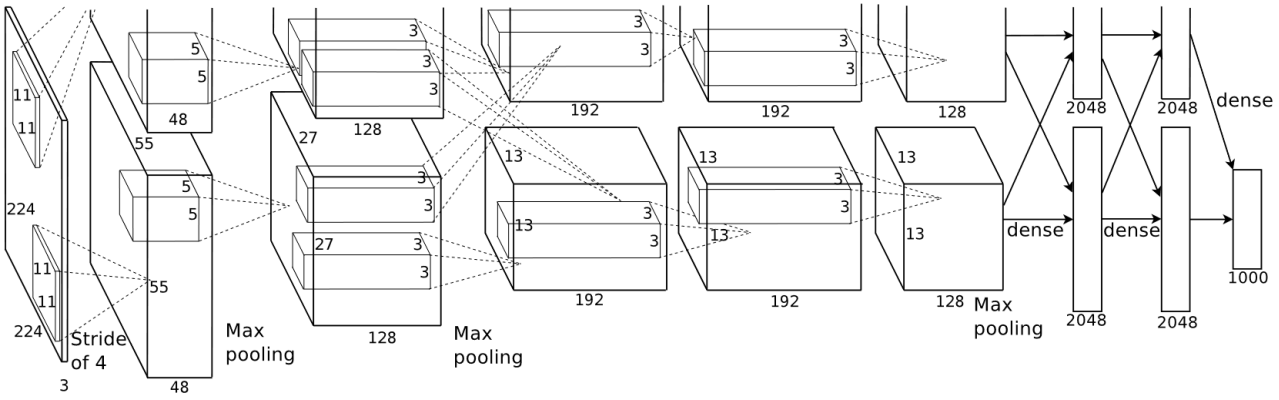
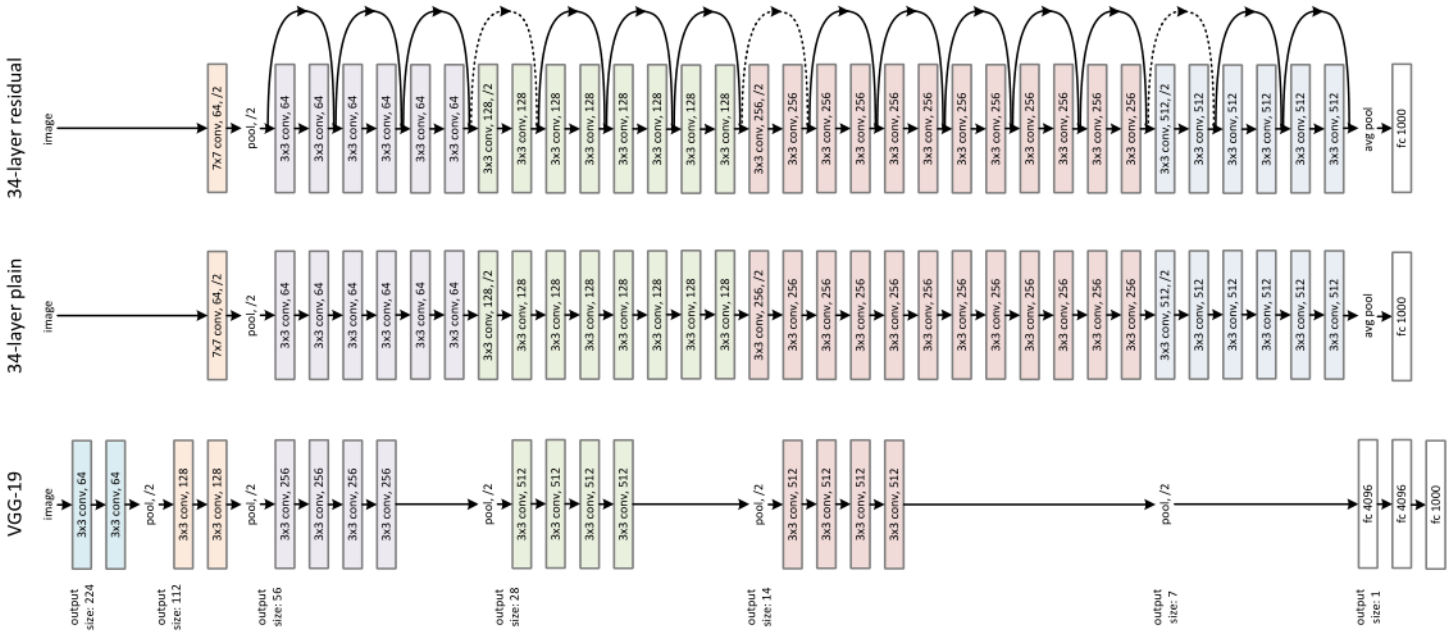


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input 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.

Source: Krizhevsky

- Architectures features two lines of CNN networks which are deployed on separate GPUs
- Winner of ImageNet 2012 Challenge (top-5 error of 15.3%, 10.8% lower than 2nd best)

Residual Network (ResNet)



Source: He

- Idea: Use skip connections to help the input signal to propagate in deep networks.
- Winner of ILSVRC 2015 (top-5 error of 3.57%)