

AlexNet深度卷积神经网络

2022/12/17 海南海口

《ImageNet Classification with Deep Convolutional Neural Networks》

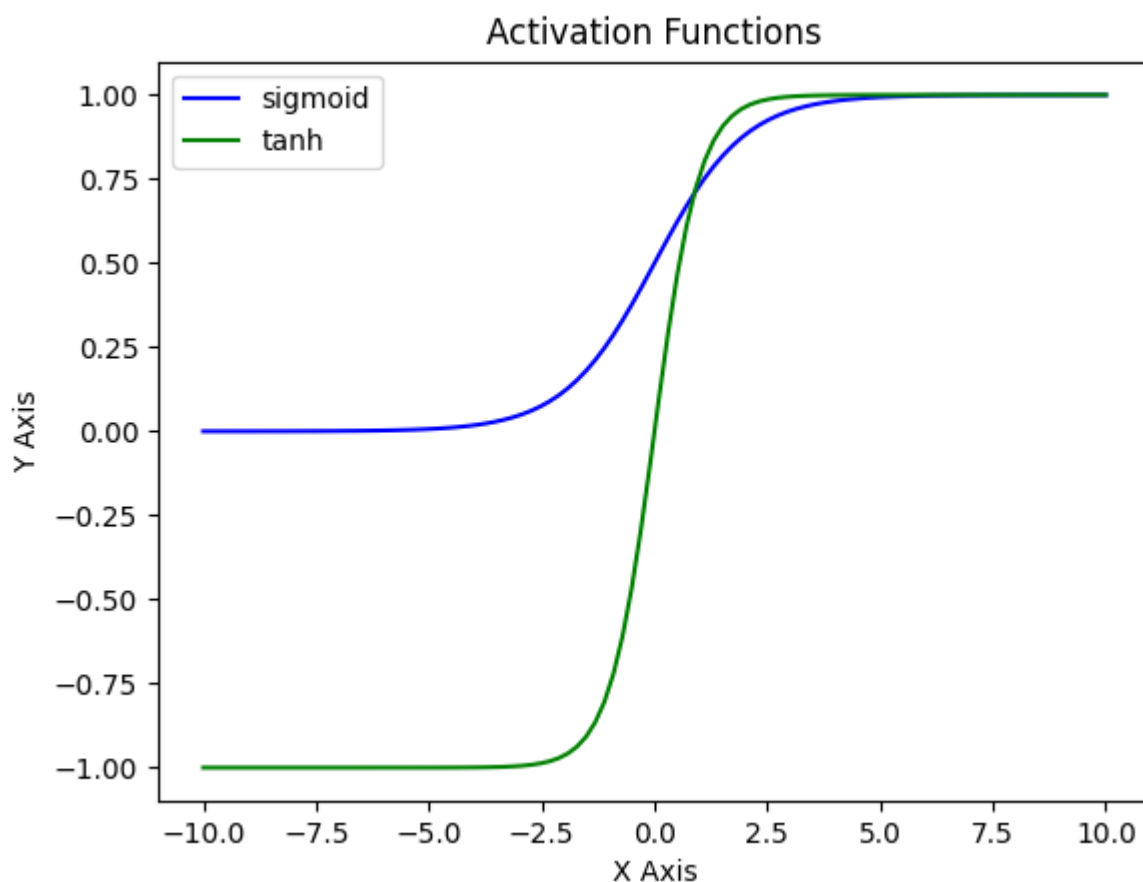
-----By Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton 2012

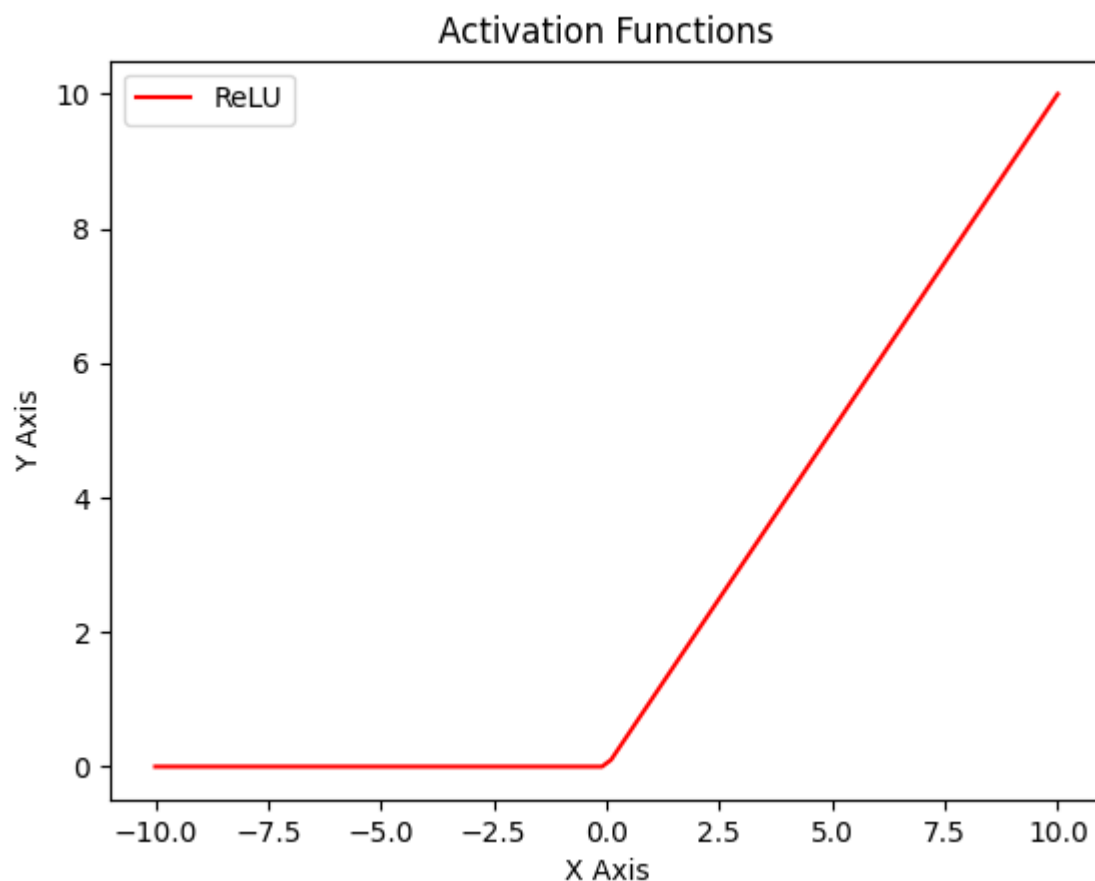
创新点

1. 首先提出了端到端的训练；
2. 使用了 ReLU 激活函数；
3. 使用了 dropout 随机失活；

非线性激活单元：ReLU

相比于 sigmoid 和 tanh 激活函数，这两个函数在 x 值趋向于无穷时的导数接近于 0，也就造成了梯度消失的问题，ReLU 作为不饱和的激活函数就可以解决梯度消失的问题。





Dropout随机失活

主要分析一下 dropout 依据：

1. 模型集成。
2. 记忆随机抹去。
3. dropout减少了神经元之间的联合依赖性。

SGD随机梯度下降

使用动量值为 0.9，权重衰退为 0.0005 作为参数。

$$v_{i+1} := 0.9 \cdot v_i - 0.0005 \cdot \epsilon \cdot w_i - \epsilon \cdot \left\langle \text{angle} \left(\frac{\partial L}{\partial w} \right), w_i \right\rangle_{D_i}$$

$$w_{i+1} := w_i + v_{i+1}$$

实验：定量分析

在论文中做了三个实验，画了两张图。

实验一

实验一展示了模型在两块gpu中，相应的第一层卷积核的卷积结果可视化。实验结论是 gpu1 更多的检测到了图像的边缘信息；gpu2更多的检测到了图像的颜色信息。

Figure 3. Ninety-six 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 7.1 for details).



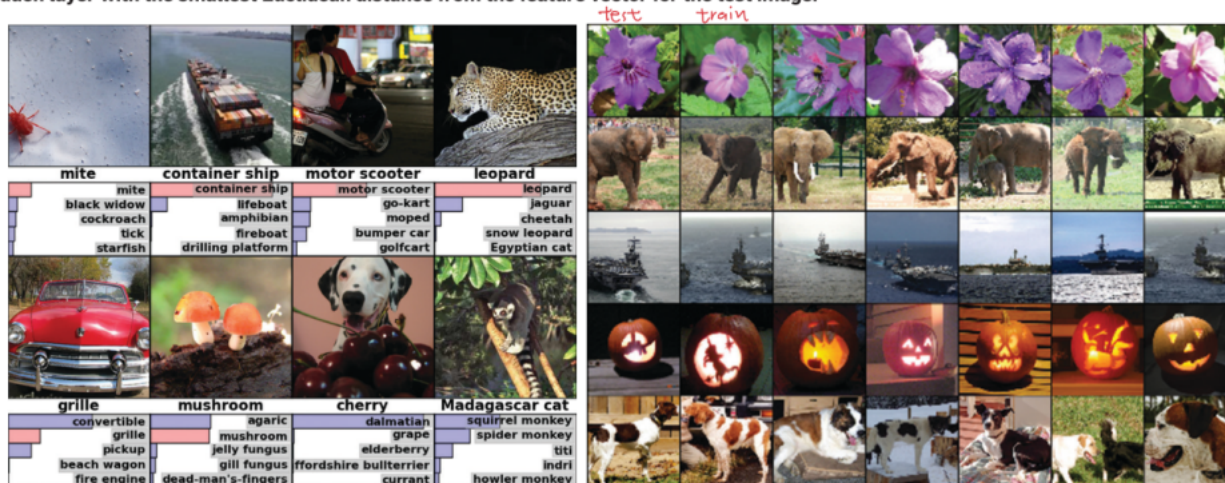
李沐老师提到过，现在许多的模型网络都是在块gpu上计算的，模型并行的方式适用于大规模模型的训练，例如 BERT 的预训练模型。如果想多 gpu 训练非大模型参数的模型，通常使用数据并行等等方式，所以实验一的过程可以了解一下。

实验二

实验二（左图）展示了模型在测试集上 8 张图片的 Top-5 预测种类。实验结论如下：

1. 非处于中心物体同样可以被正确识别。
2. Top-5 的预测种类名，大多都是和正确标签类别相似的名称。

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.



实验三

在上图的右侧为实验三的图示。实验把测试集的 5 张图像作为查询图像，在训练集上搜索与测试图像最相似的图像。具体做法是，把测试图像经过模型计算得到最后一层（输出层）的 4096 维向量，与训练集每张图像的 4096 维向量做欧式距离，度量结果最小的训练集图像被挑选出来。

换句话说，网络提取出的是图像的语义级信息 semantic information，并不是像素级信息 pixel information，并且可以看到在训练集中选择出的花朵有着不同的光照、大象有着不同的姿态。

结论

神经网络需要做的更深。