

卷积

李沐 · AWS



分类猫和狗的图片



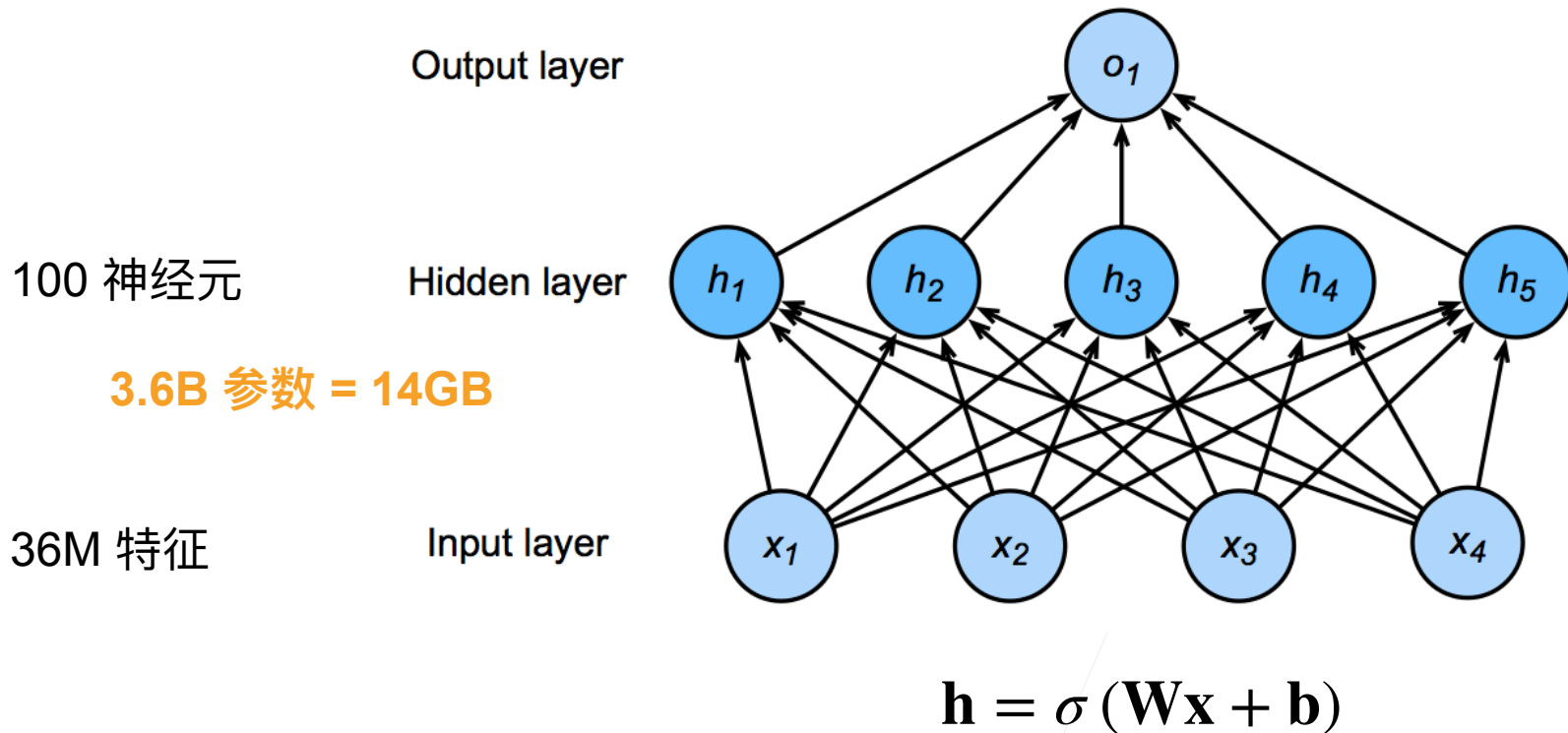
- 使用一个还不错的相机采集图片（12M像素）
- RGB 图片有 36M 元素
- 使用100大小的单隐藏层 MLP，模型有 3.6B 元素
 - 远多于世界上所有猫和狗总数（900M 狗，600M 猫）



VS



回顾：单隐藏层 MLP



Waldo 在
哪里？



两个原则

- 平移不变性
- 局部性





重新考察全连接层

- 将输入和输出变形为矩阵（宽度，高度）
- 将权重变形为4-D张量 (h, w) 到 (h', w')

$$h_{i,j} = \sum_{k,l} w_{i,j,k,l} x_{k,l} = \sum_{a,b} v_{i,j,a,b} x_{i+a,j+b}$$

- V 是 W 的重新索引 $v_{i,j,a,b} = w_{i,j,i+a,j+b}$



原则 #1 - 平移不变性

- x 的平移导致 h 的平移 $h_{i,j} = \sum_{a,b} v_{i,j,a,b} x_{i+a,j+b}$
- v 不应该依赖于 (i, j)
- 解决方案: $v_{i,j,a,b} = v_{a,b}$

$$h_{i,j} = \sum_{a,b} v_{a,b} x_{i+a,j+b}$$

这就是 2 维~~卷积~~ 交叉相关



原则 #2 - 局部性

$$h_{i,j} = \sum_{a,b} v_{a,b} x_{i+a,j+b}$$

- 当评估 $h_{i,j}$ 时，我们不应该用远离 $x_{i,j}$ 的参数
- 解决方案：当 $|a|, |b| > \Delta$ 时，使得 $v_{a,b} = 0$

$$h_{i,j} = \sum_{a=-\Delta}^{\Delta} \sum_{b=-\Delta}^{\Delta} v_{a,b} x_{i+a,j+b}$$

总结



- 对全连接层使用平移不变性和局部性得到卷积层

$$h_{i,j} = \sum_{a,b} v_{i,j,a,b} x_{i+a,j+b}$$



$$h_{i,j} = \sum_{a=-\Delta}^{\Delta} \sum_{b=-\Delta}^{\Delta} v_{a,b} x_{i+a,j+b}$$