思考：

1 如何理解这里的Attention？

2 什么是high-dimensional feature representation

3 Fisher Vector

4 B-CNN

网络基本架构：

特征提取网络:

VGG-19 [27] (pre-trained on ImageNet) for bird and car datasets, and VGG-16 for dogs