

# 开题报告 - 机器学习纳米学位（进阶）

Sam 优达学城

2017年12月09

## 猫狗大战

### 项目背景

*(approx. 1-2 paragraph)*

在这个项目我使用的是卷积神经网络解决猫和狗的分类问题，[卷积神经网络](#)是89年Yann LeCun提出的一个新的模型，简称为CNN，当时能够对字符进行分类，准确率达到99%，打破了以前所有记录，但CNN是自行学习特征，在2012年，另外一名研究者Alex Krizhevsky在一年一度的ImageNet比赛上使用了它，准确率达到新高85%，自此之后被Google用于识别搜索图片，Facebook用它来自动添加标签，想了解更多可以点击[这里](#)。

卷积神经网络是由三种主要类型的层来构建ConvNet体系结构：卷积层，池化层和完全连接层，我们将堆叠这些层来形成一个完整的ConvNet架构。当然每层的作用都不一样，例如：CNN的第一层很清楚的选择了非常简单的形状、图案，例如直线和色块。第二层认出了圈，条纹或者长方形。第三层捕捉了第二层特征的复杂组合等等。当然你可以自己增加和组合不同的层来解决你的面对的问题，此次我选择卷积神经网络中一个VGG神经网络，简称：[VGGNet](#)，来解决猫狗大战的分类问题。

### 问题描述

*(approx. 1 paragraph)*

猫狗大战是个二分类问题，可以用很多类型的卷积网络来解决问题，你可以重新建立一个的模型，但效率很低，为了提高效率，我决定使用[权重迁移](#)，权重迁移主要作用是使用有经验的模型，加以修改交给数据训练后成为新的模型。

### 数据或输入

*(approx. 2-3 paragraphs)*

数据集来源于Kaggle的[Dogs vs. Cats](#)，训练集有2.5万照片，一半是猫，一半是狗，数据的分布是均衡的，两个类各有1.25万张图片，这里我把总训练集切分成80%训练集，10%验证集，10%测试集。从训练集随机抽取猫和狗的照片：



每张图片的大小不同，而且拥有对应的背景，其中狗的背景还带有猫或人，对应的第一张照片中的猫是模糊的，这些图片不同的特征都会增加模型识别的难度。

### 解决方法描述

*(approx. 1 paragraph)*

猫狗大战为二分类问题，刚好ImageNet竞赛有对自然动物进行过训练，应该保留很多动物重要的特征，所以我们可以使用现有训练好的模型来解决该问题，可这里有三个解决方案：

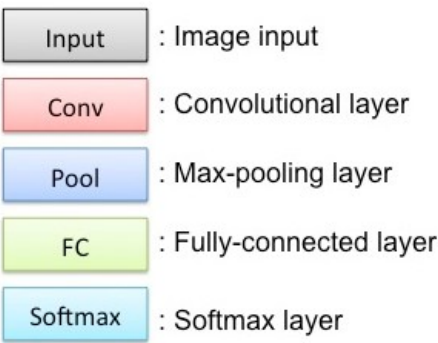
- 1.Build Model：重新建立一个模型，在训练过程中，根据准确率或者其他指标，不断调试模型的参数和网络结构，达到最佳的识别准确率。但这个方法比较耗时间和资源。
- 2.Bottleneck Features：把训练好的VGGNet,GoogLeNet,ResNet模型和特征文件下载下来，然后建立新的全连接层，对其进行训练，如果这些特征符合本项目的数据特征，效果是不错的，因为训练过的模型，已经有重要的特征对其识别。该方法计算量最少并节省时间。（该结论已经过实践验证）。
- 3.Feature extraction：选择训练好的VGGNet模型，将倒数第二全连接层之前所有卷积作为提取特征器，经过训练后，将这些特征文件保存为离线形式，修改最后连接层，再对其进行训练，就能形成一个新的模型。该方法计算量较多但能节省时间。

第二个计算量比第一个方案要多。但我认为准确率比第一种方案要高，因为第二种方案的特征方向不一定全符合本项目的数据特征，也许偏离较远。如果对训练好的网络再一次训练数据，特征方向会靠拢本项目图像的特征，模型准确率会更高。为验证我的想法，我选择第3个方案来解决该分类问题。

## 基准模型

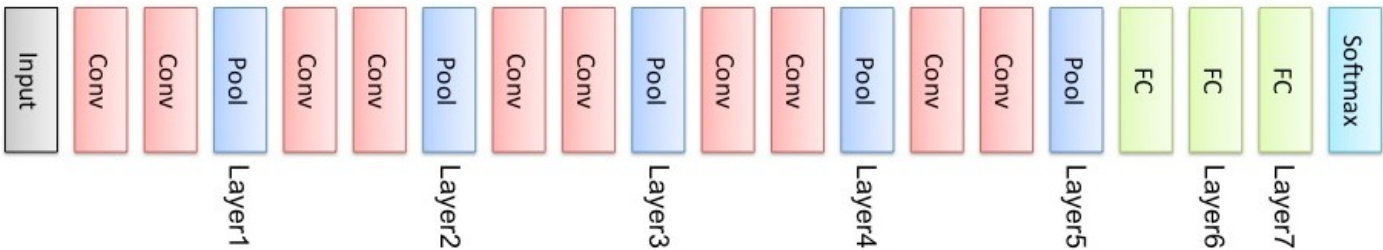
(approximately 1-2 paragraphs)

如下包括VGGNet模型的节点描述和基本模型：



（图1：节点描述）

### VGGNet



（图2：VGGNet模型架构）

这是本次使用的模型架构VGGNet，里面两个卷积层接最大池化+两个卷积层接最大池化...五个这种类型卷积层叠在一起，最后有三个连接层+Softmax函数，重点是Layer6，提取Layer6前所有特征作为特征提前器（也能称为：卷积码），然后修改Layer7全连接层，使用卷积码结合训练集训练Layer7，新的模型就完成训练了。

## 评估标准

(approx. 1-2 paragraphs)

评估标准选择为准确度（Accuracy），是计算模型正确的预测结果在所有被预测结果的占比。选择准确度是因为两种类型的数据分布是均衡的，而不是偏斜的，所以准确度是能够衡量模型预测能力。跟kaggle要求的logloss相比，其实本质是相同的，logloss越低，准确度越高，由于准确度更加直观，所以选择它。

## 项目设计

(approx. 1 page)

项目设计初步计划如下：

### 一、图像路径处理

- （1）将训练集中的猫和狗分成两个文件夹；

### 二、训练VGGNet；

- （1）使用训练集训练VGGNet，得到特征提取器（截止到layer6）；
- （2）提取特征文件，保存文件在本地；

### 三、预处理训练集

- （1）训练集切分成80%训练数据，10%验证数据，10%测试数据；
- （2）训练集标签同样独热处理后进行微调；

### 四、构建模型

- （1）修改最后的layer7，让其输出为256，甚至可以更多，这里我选择256单元；
- （2）利用layer7和标签输出的维度，计算出logits，进行softmax，得出预测结果；
- （3）使用交叉熵计算出代价，然后使用Adam优化算法进行反向传播，求出最小的损失；（这里Adam跟SGD是相似的，但它的计算比SGD复杂一些，具体请看[这里](#)）

### 五、训练模型

- （1）提取特征文件；
- （2）这里使用了批次函数来对训练集进行分批训练；
- （3）使用验证集，查看每次训练后的准确度；

### 六、测试模型

- （1）使用测试集进行验证准确度；
- （2）在自带的测试集中，随机抽取一张图片，加以验证模型准确度；