# 机器学习大作业实验报告

## 1.数据

### （1）问题

任务二（样本不均衡问题）：

数据集中各个危险品类别的数据量是不均衡的。如何解决样本不均衡条件下模型训练的类别偏好问题是一个热点。安检机返回的x光图像为RGB彩色图像。

训练集中的危险品包括带电芯充电宝和不带电芯充电宝两个类别。比例为1:10（带电芯充电宝：不带电芯充电宝，500：5000）

### （2）数据格式处理

转换为VOC标准格式，便于训练。如图1所示：

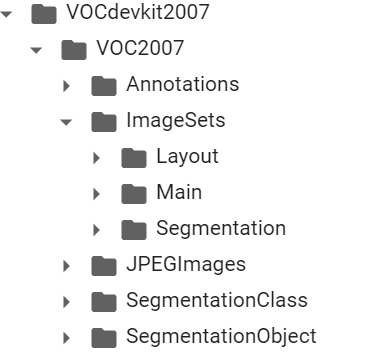


图1 voc基本格式

Annotations：保存的是xml格式的标签文件，每个xml对应JPEGImage中的一张图片

ImageSets：包含main文件，main文件包含以下文件：

train.txt、core\_battery\_train.txt、coreless\_battery\_train.txt、

test.txt、core\_battery\_test.txt、coreless\_battery\_test.txt、

trainval.txt、val.txt、core\_battery\_val.txt、coreless\_battery\_val.txt、

里面都是存放的图片编号

JPEGImages：包含所有的图片，包括训练图片以及测试图片，验证图片

SegmentationClass和SegmentationObject：保存的是物体分割后的数据

## 2.模型

### （1）Faster R-CNN

目标检测算法大致分为两类：一步法和二步法。我们选择两步法中的Faster R-CNN。

Faster R-CNN由四个部分组成：

1）卷积层(conv layers)，包含了conv，pooling，relu三种层，用于提取图片的特征，CNN网络的输入为整张图片，输出为提取出的特征称为feature maps，feature maps被共享用于后续RPN层和全连接层。

2）RPN网络(Region Proposal Network)，用于推荐候选区域，这个网络是用来代替之前的search selective的。输入为图片featrue maps，输出为多个候选区域。经过CNN网络前向传播至最后共享的卷积层，一方面得到供RPN网络输入的特征图，另一方面继续前向传播至特有卷积层，产生更高维特征图。

3）Roi pooling，将不同大小的输入转换为固定长度的输出。供RPN网络输入的特征图经过RPN网络得到区域建议和区域得分，并对区域得分采用非极大值抑制【阈值为0.5】，输出其Top-N得分的区域建议给RoI池化层。

第2步得到的高维特征图和第3步输出的区域建议同时输入RoI池化层，提取对应区域建议的特征。

4）分类和回归，这一层的输出是最终目的，输出候选区域所属的类，和候选区域在图像中的精确位置。得到的区域建议特征通过全连接层后，输出该区域的分类得分以及回归后的bounding-box。

### 神经网络模型

选取了预训练模型vgg16。图2是vgg的结构，其中D就是vgg16的结构，vgg16包括13个卷积层，5个池化层，3个全连接层。

vgg16模型中输入图像的尺寸是224\*224\*3，经过64个通道3\*3\*3的卷积核，卷积两次，输出图像尺寸为224\*224\*64；

经过滤波器为2\*2的池化，图像尺寸变为112\*112\*64；

经过128个通道3\*3\*3的卷积核，卷积两次，输出图像尺寸为112\*112\*128；

经过滤波器为2\*2的池化，图像尺寸变为56\*56\*128；

经过256个通道3\*3\*3的卷积核，卷积三次，输出图像尺寸为56\*56\*256；

经过滤波器为2\*2的池化，图像尺寸变为28\*28\*256；

经过512个通道3\*3\*3的卷积核，卷积三次，输出图像尺寸为28\*28\*512；

经过滤波器为2\*2的池化，图像尺寸变为14\*14\*512；

经过512个通道3\*3\*3的卷积核，卷积三次，输出图像尺寸为14\*14\*512；

经过滤波器为2\*2的池化，图像尺寸变为7\*7\*512；

经过2层1\*1\*4096，一层1\*1\*1000的全连接层；

最后通过softmax输出预测结果。

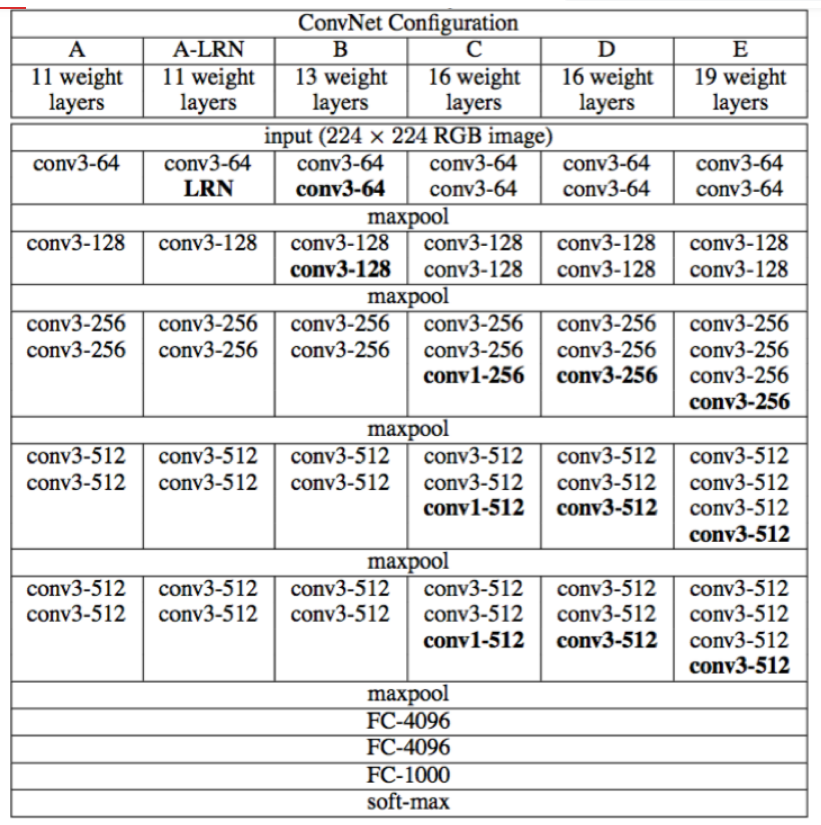


图2(a) vgg结构

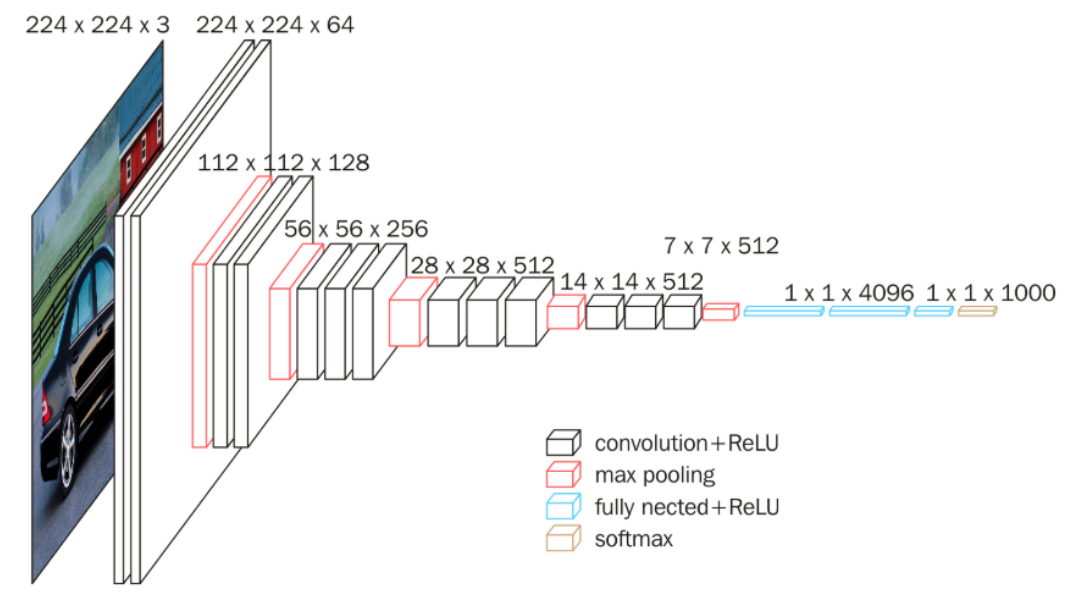


图2(b) vgg16结构

## 样本不平衡解决方法

解决样本不平衡问题可以使用过采样，欠采样，设置权重这些方法，本次实验采取的过采样方法，通过对core类图像操作生成新的图片，使2个类别的数量大致平衡。

生成新样本的方法及命名方式如下：

镜像（左右）:core\_battery0001xxxx

旋转45度 :core\_battery0002xxxx

旋转90度 :core\_battery0003xxxx

平移:core\_battery0004xxxx

缩放:core\_battery0005xxxx

镜像（上下）:core\_battery0006xxxx

高斯扰动:core\_battery0007xxxx

颜色扰动:core\_battery0008xxxx

锐化:core\_battery0009xxxx

其中xxxx为所变换的图片的名称中的值。剔除不合格的样本，最终得到了约4200个新样本，此时带芯与不带芯充电宝数量大概为1：1。

## 4.优化目标

### （1）IoU

一个给定的边界框的正确性的度量标准是“交并比”（Intersection over Union, IoU）。

利用 IoU ，我们现在要分辨检测结果是否正确。最常用的阈值是0.5：如果 IoU > 0.5，那么认为这是一个正确检测，否则，认为这是一个错误检测。

### （2）mAP

使得预测的正确率最高，mAP是平均精度均值

1）对于某类别C，在一张图片上：



即P=一张图片上类别C识别正确的个数 / 这张图片上类别C的总个数

2）对于类别C，在多张图片上



即AP=每张图片上的Precision求和  /  含有类别C的图片数目

3）对于整个数据集而言，存在多个类别C1、C2、C3...



即mAP=所有类别的AP求和  /  总的类别数目