East China Normal University

2020-2021 学年暑期学期 实验项目报告

课程名称：人工智能前沿

课程类别：专业选修课

选题三 ：基于 Autoencoder 的 Cifar10 数据集异常值检测

院　　系：计算机科学与技术

专　　业：计算机科学与技术

姓名：俞辰杰　学号：10192100571

小组成员：刘诺纬 10191900446

林　霁 10194810417

周佳仪 10205102504

2021 年 7 月 8 日

一、问题描述

Cifar10 数据集包含飞机、鸟、猫、鹿等 10 类图片，训练集 60000 张，测试集 10000 张，32*32 大小。我们设置图片飞机为正常值，其他 9 种图片为正常值。也即，训练集正样本为【飞机】; 测试集正样本为【飞机】 负样本为【汽车，鸟，猫，鹿，狗，青蛙，马，船】，共 10000 张。该问题是异常检测问题，使用 Autoencoder 可进行无监督的异常检测，即异常样本是无需进行学习的。

二、方法描述

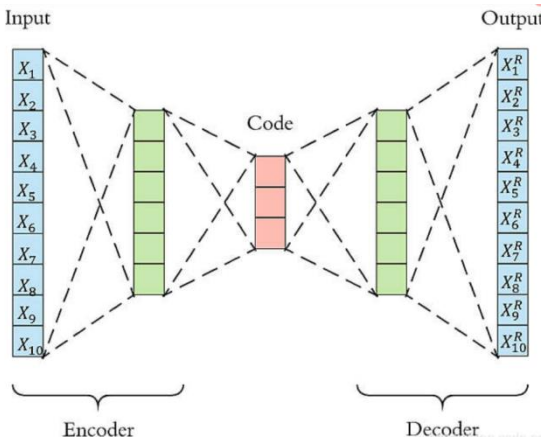


图 1 autoencoder 架构图

Autoencoder 自编码器，是用于无监督学习中的一种常用模型，它主要包括编码器 encoder 和解码器 decoder。其本质就是一个神经网络来产生一个高维输入的低维表示，在本项目中就是对输入图片的一个特征抽取。Autoencoder 的训练目标就是希望抽取出的图片特征能够尽可能的使用解码器进行还原，从而使得 decoder 在训练时，能够抽取出最有信息量的特征。

由于 Autoencoder 能够很好的对训练样本进行重构的特性，使得该方法在无监督异常检测中变得可行。在无监督异常检测中，异常样本是无标签的，并且假设其分布和正常样本是不同的。在训练时，使用 80%的正常样本进行训练，模型能够很好将正常样本进行还原，即 MSE 值相差较小。也可从模型对正常样本过拟来理解 autoencoder 的作用。因此可以通过重构差值来对正常样本和异常样本进行区别。

在 autoencoder 模型实现时，encoder 包含三个下采样部分，每层卷积后都连

接一层 relu 非线性层，decoder 部分包含对应的相反的三个上采样部分，最后层转置卷积层后连接 sigmoid 非线性函数。模型具体代码见 Model.py。模型的训练使用重构 MSE 损失来使得 autoencoder 能够对正常样本进行特征提取并复原，通过该方法能成功的将正常样本和异常样本进行区分。

三、绘制 ROC 曲线

为了刻画 ROC 曲线，我们首先将按照属于“正样本（positive）”的概率将所有样本进行排序，可计算出相应的混淆矩阵，从而可以进一步计算其 TPR 值，在图标上绘制一个点。

以此类推，分类阈值设为每个样例的预测值，即依次将每个样例划分为正确。则可获得所需要的 ROC 曲线，采样样本越多，曲线越平滑。

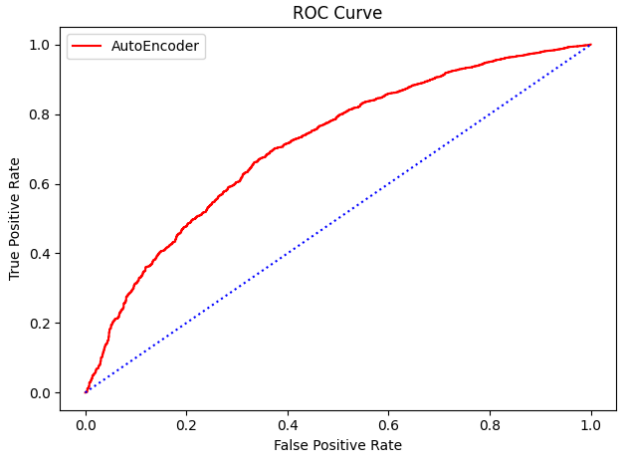
四、结果展示

我们使用了 ROC 曲线来和 AUC 值来显示我们模型的性能：

ROC 曲线指受试者工作特征曲线 / 接收器操作特性曲线, 是反映敏感性和特异性连续变量的综合指标,是用构图法揭示敏感性和特异性的相互关系，它通过将连续变量设定出多个不同的临界值，从而计算出一系列敏感性和特异性，再以敏感性为纵坐标、（1-特异性）为横坐标绘制成曲线，曲线下面积越大，诊断准确性越高。在 ROC 曲线上，最靠近坐标图左上方的点为敏感性和特异性均较高的临界值。AUC 值为该曲线下的面积。ROC 曲线和 AUC 值是评价异常检测分类器很好的性能指标。

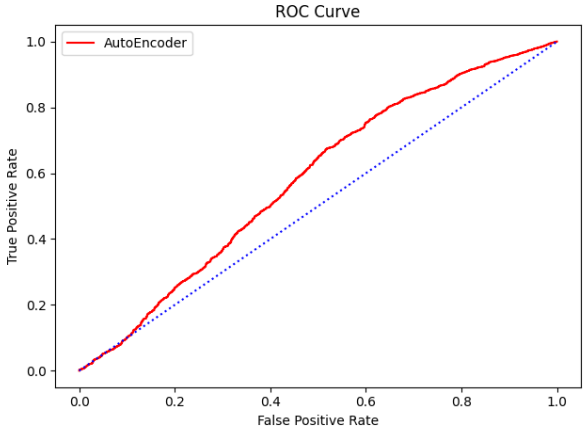
我们对每个类别都进行了如下的实验，将某一类别样本当作正常样本对异常检测模型进行训练，而将其他类别样本当作异常样本，仅在测试时使用。

我们首先将飞机作为正常类，其 ROC 曲线如下所示：



图一 飞机异常检测模型 ROC 曲线

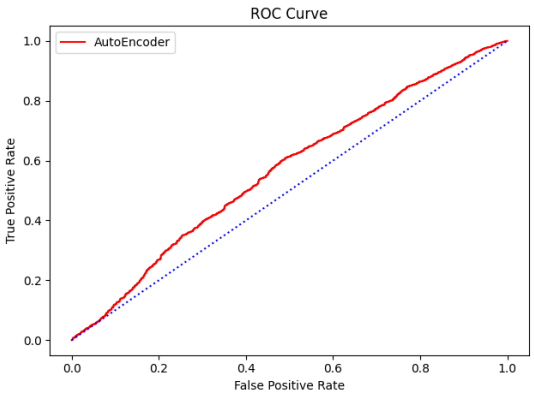
从 ROC 曲线上来看，该飞机异常检测模型已经具备了一定异常检测的能力，该图对应的 AUC 值为 0.7125。由于选择 Autoencoder 模型较为简单，因此在后续模型的精度提升上，可选用更复杂的模型，例如 Unet 架构增加跳跃连接等。



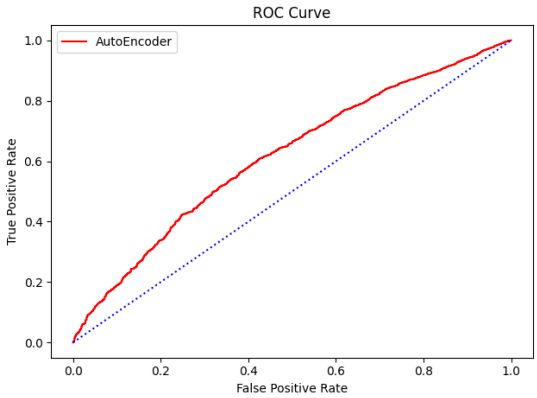
图二 鸟异常检测模型 ROC 曲线

该鸟异常检测模型对应的 AUC 值为 0.5814，相较于飞机的异常检测模型其值略低，分析其可能的原因是鸟的数据较为复杂，模型对于复杂数据的特征提取能力仍然受限。

正常类	飞机	鸟	鹿	青蛙	船
AUC 值	0.7125	0.5811	0.5655	0.5504	0.6169



图三 鹿异常检测 ROC 曲线



图四 船异常检测 ROC 曲线

五、文件组织结构

选题 3_：

1. 选题 3_工程代码（存放代码）
    a. AUCCalculator.py（计算 AUC 代码）
    b. DatasetSplit.py（数据预处理代码）
    c. ImageDisplay.py （图像展示代码）
    d. Model.py （autoencoder 模型代码）
    e. TrainTool.py （模型训练代码）
2. 选题 3_实验结果（存放训练模型的 ROC 曲线和 AUC 结果）
3. 选题 3_PPT.pptx （大作业 PPT）
4. 选题 3_课程报告.pdf（大作业报告）

六、代码运行指南

打开 TainTool.py，对模型进行训练。