

# 面向磁浮轨道异常检测的大数据分析框架研究

## 小组汇报

2024-9-3    By 刘震

# Part I: 卷积核沿空间维度移动，实现降维

## 基本思路：

Step I:

矩阵输入形式：1127\*28 and 1127\*4\*7，即分别设置输入通道in\_channels为1、4，输出通道out\_channels均设置为1，利用卷积核滑动的方式得到1127\*1的一维向量

Step II:

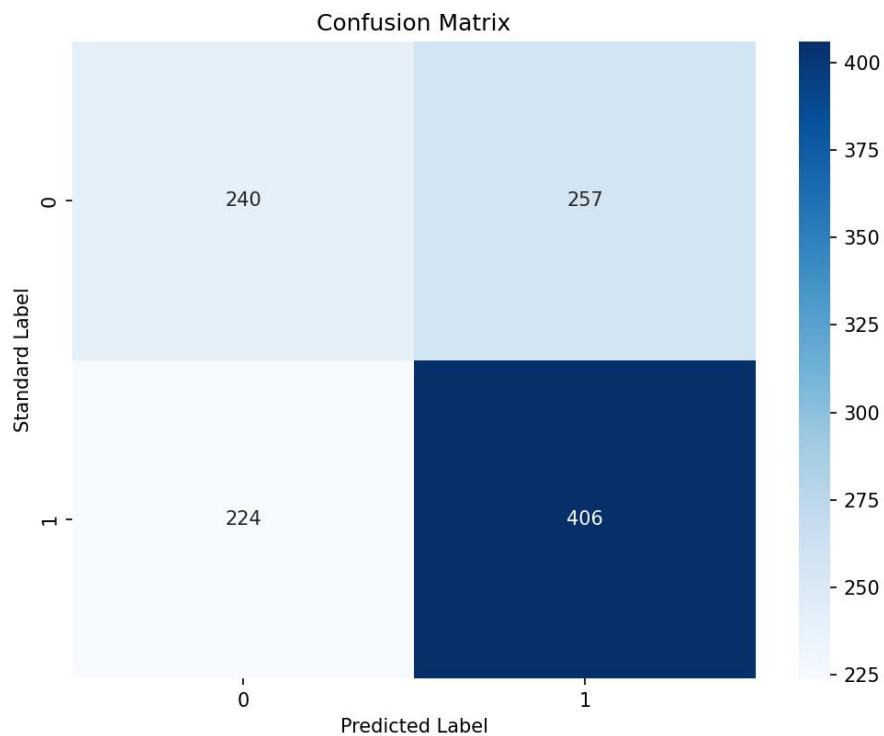
利用Isolation Forest方法检测出异常点（contamination的值分别设定为0.4117(464/1127)、0.0292(33/1127)）

Step III:

与标签对比，检测confusion matrix, recall, precision, F1-score指标，比较降维结果的好坏

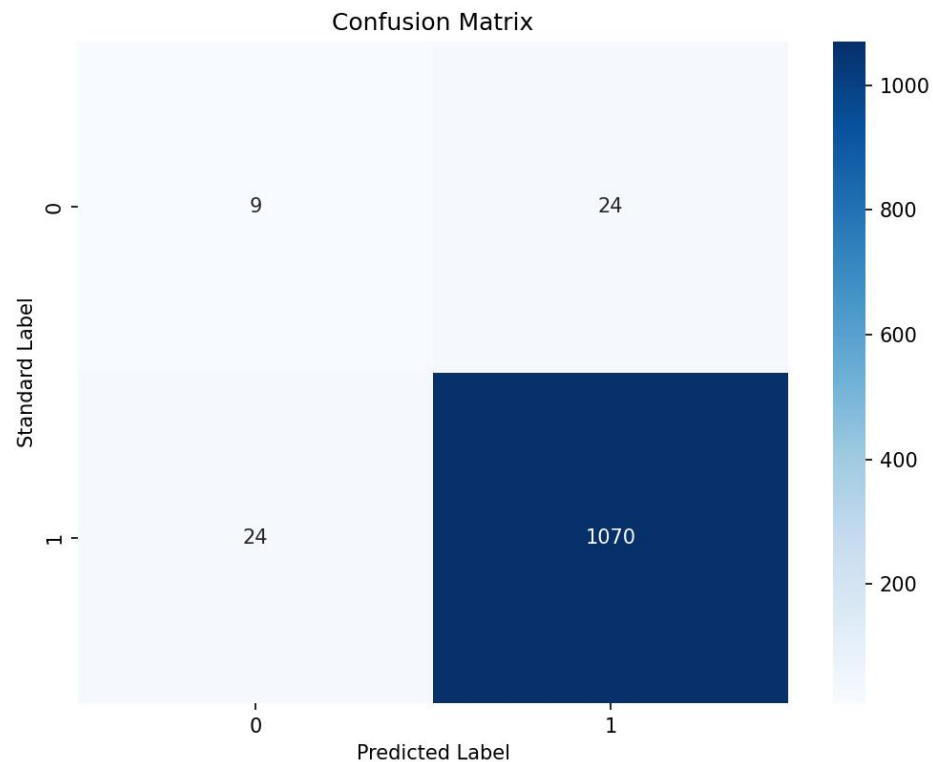
# I. 矩阵以 1127\*4\*7 的形式输入，in\_channels设置为4

对普通异常点的检测效果：（ Confusion Matrix ）



```
Precision: 0.6123680241327301  
Recall: 0.6444444444444445  
F1 Score: 0.6279969064191802
```

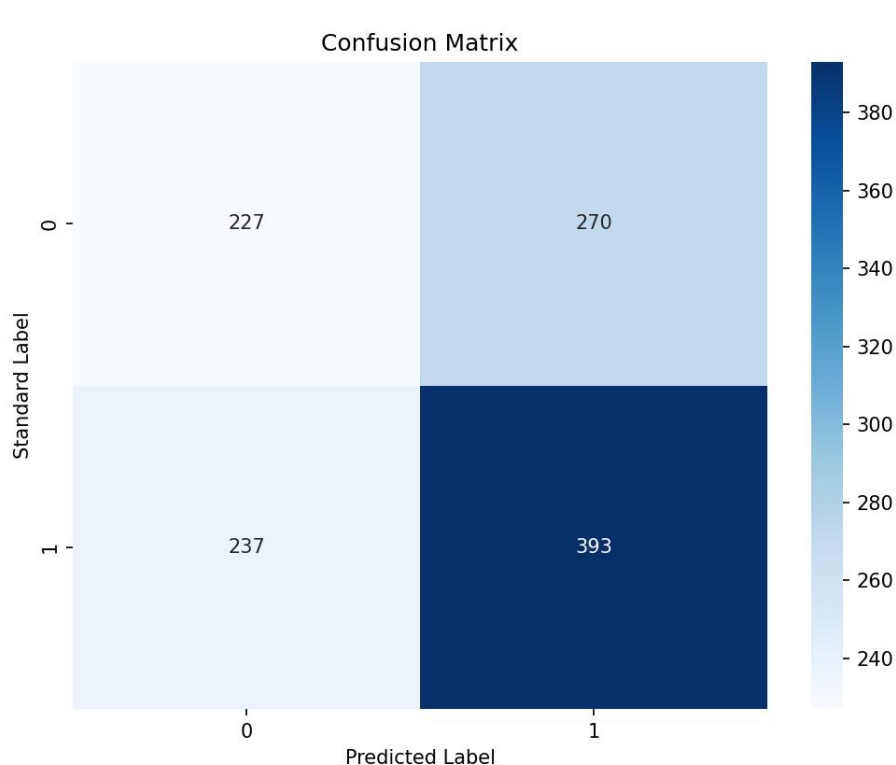
对明显异常点的检测效果：（ Confusion Matrix ）



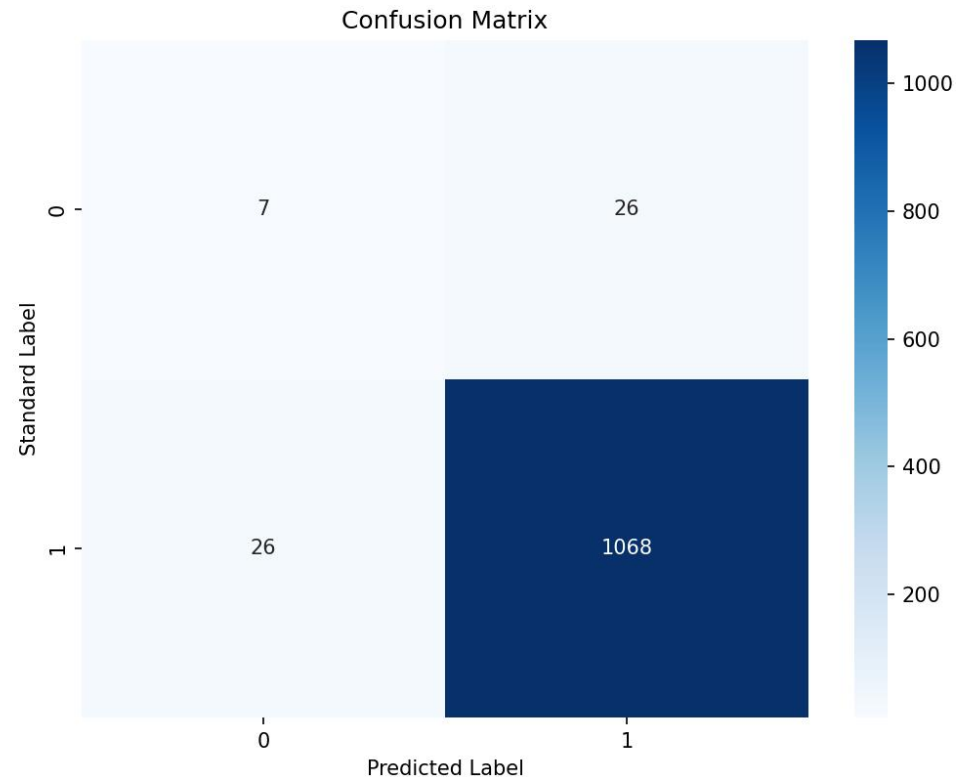
```
Precision: 0.9780621572212066  
Recall: 0.9780621572212066  
F1 Score: 0.9780621572212066
```

## II. 矩阵以 1127\*28 的形式输入，in\_channels设置为1

对普通异常点的检测效果：（ Confusion Matrix ）      对明显异常点的检测效果：（ Confusion Matrix ）



```
Precision: 0.5927601809954751  
Recall: 0.6238095238095238  
F1 Score: 0.6078886310904872
```



```
Precision: 0.9762340036563071  
Recall: 0.9762340036563071  
F1 Score: 0.9762340036563071
```

# 结果评估:

平均来看，两种卷积方式的降维结果中，输入通道为4时，降维效果相对好，误检率和漏检率相对较低；

输入通道为4

```
Precision: 0.6123680241327301
Recall: 0.6444444444444445
F1 Score: 0.6279969064191802
```

```
Precision: 0.9780621572212066
Recall: 0.9780621572212066
F1 Score: 0.9780621572212066
```

输入通道为1

```
Precision: 0.5927601809954751
Recall: 0.6238095238095238
F1 Score: 0.6078886310904872
```

```
Precision: 0.9762340036563071
Recall: 0.9762340036563071
F1 Score: 0.9762340036563071
```

检测普通异常点

检测明显异常点

## Part II: 训练GAN深度学习模型，检测异常点

### 基本思路：

Step I:

分别训练生成器和鉴别器，将1127组数据输入训练

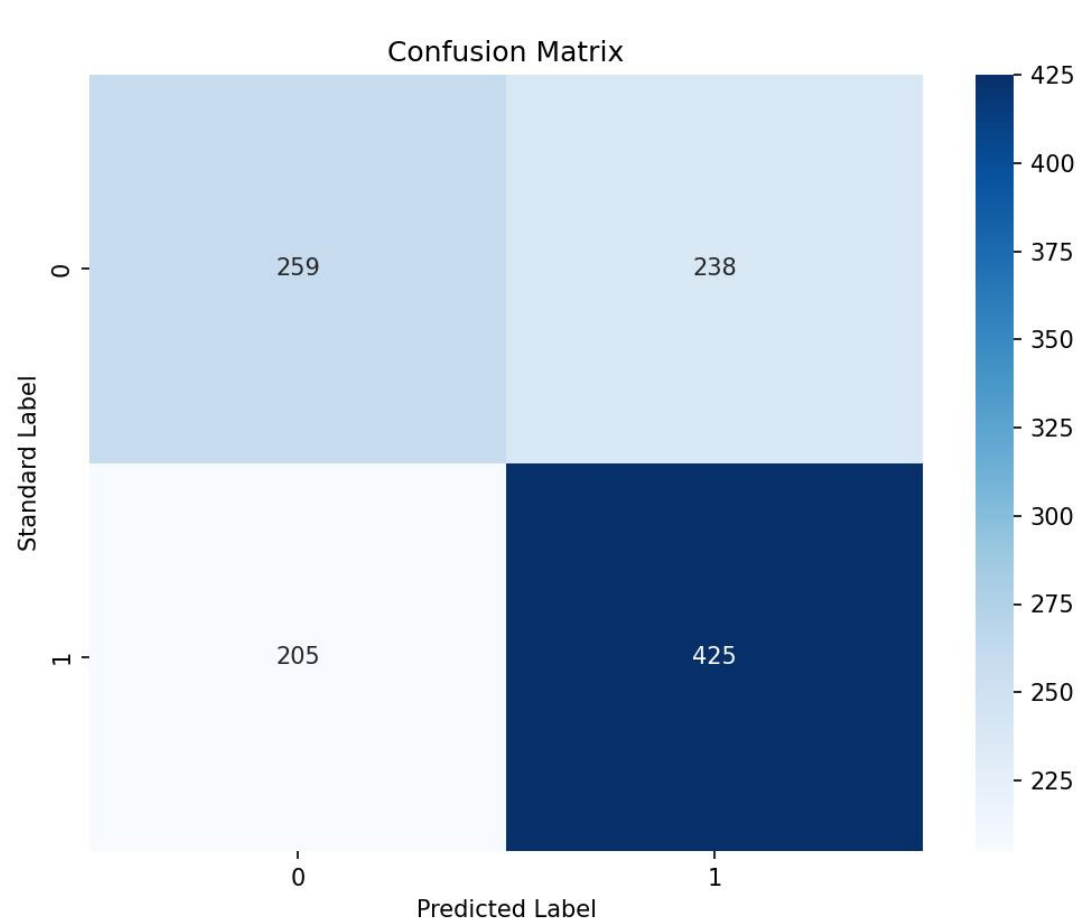
Step II:

计算重构误差，以标签数据中普通异常点、明显异常点的比例设置阈值；分别检测出异常点

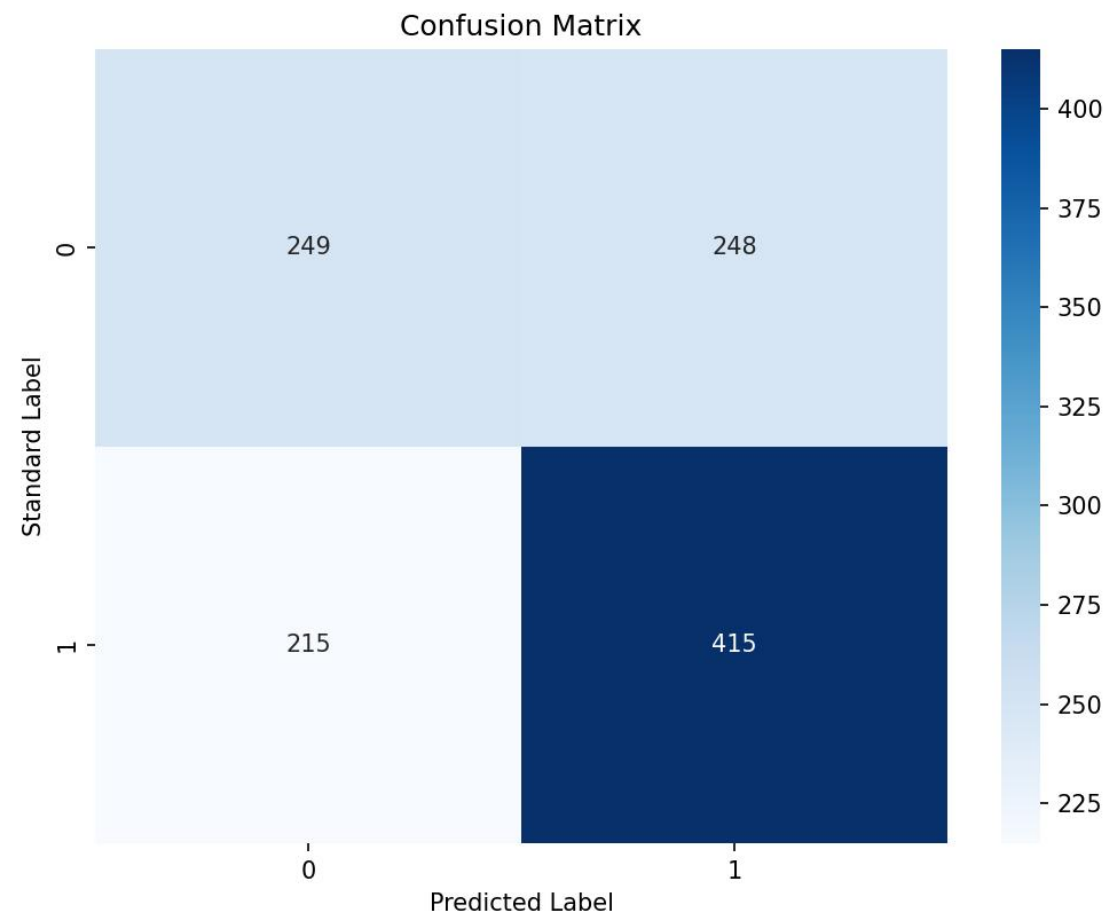
Step III:

与标签对比,检测confusion matrix, recall, precision, F1-score指标，比较降维结果的好坏

对于普通异常点：（训练轮数100）

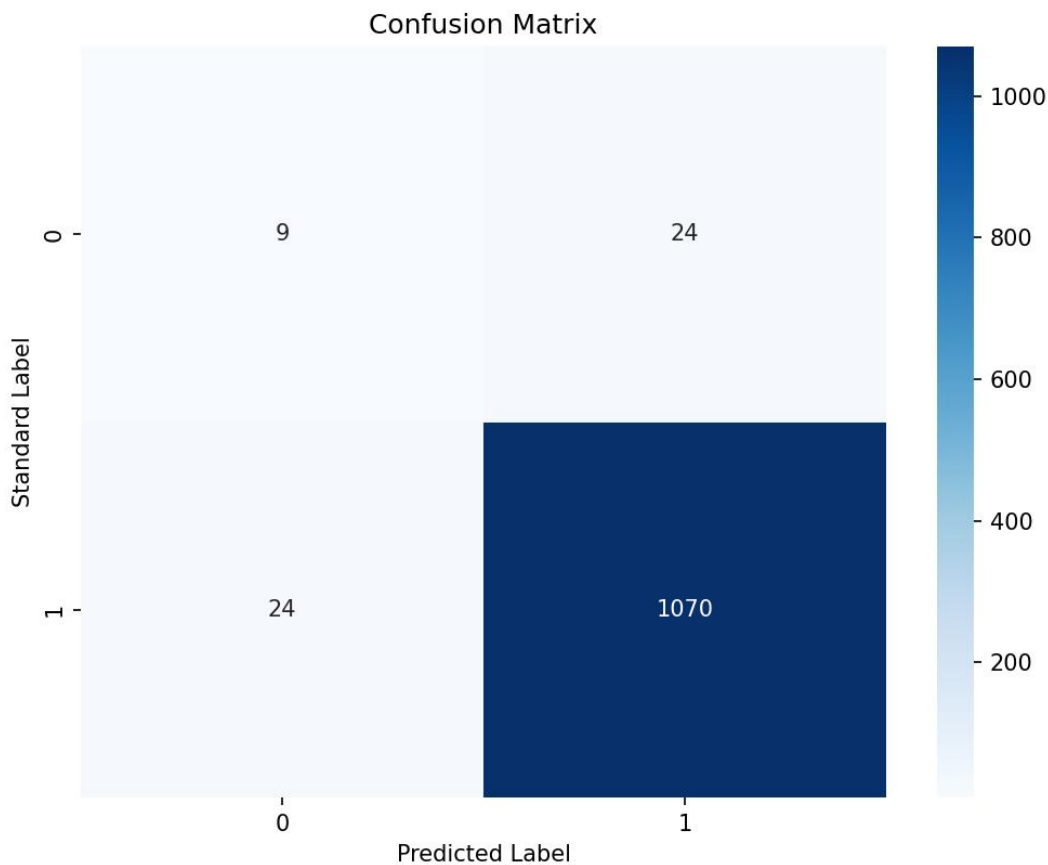


Precision: 0.6259426847662142  
Recall: 0.6587301587301587  
F1 Score: 0.6419180201082754

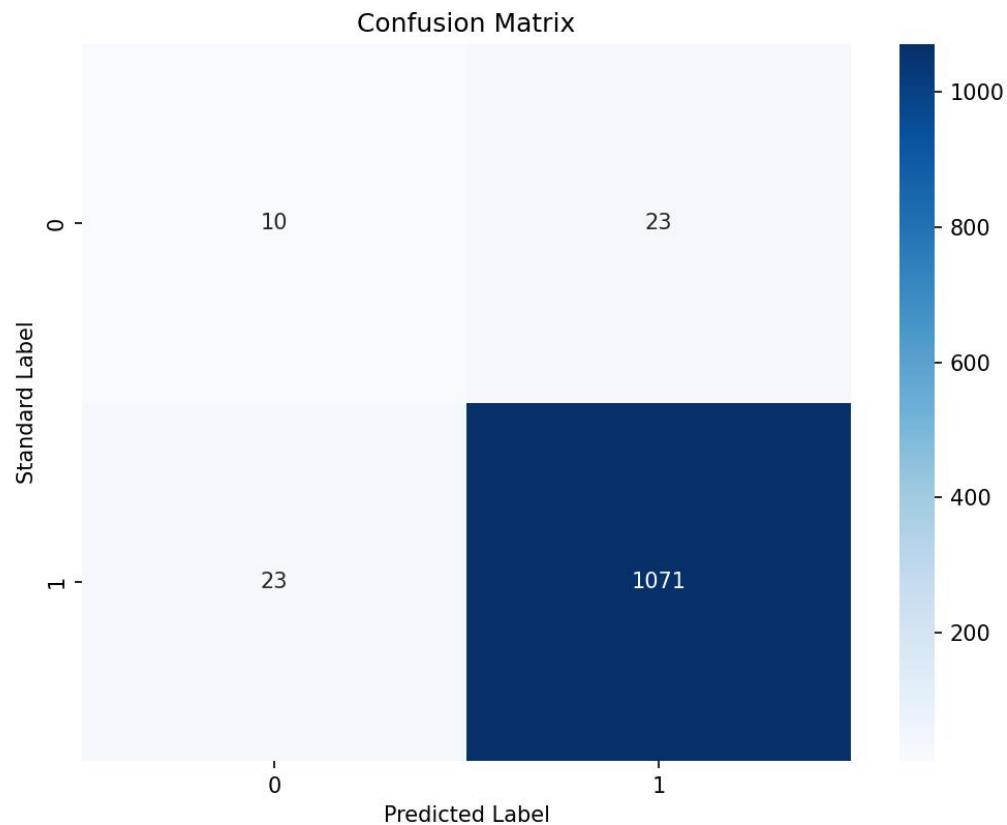


Precision: 0.6259426847662142  
Recall: 0.6587301587301587  
F1 Score: 0.6419180201082754

对于明显异常点：（训练轮数100）



Precision: 0.9780621572212066  
Recall: 0.9780621572212066  
F1 Score: 0.9780621572212066



Precision: 0.9789762340036563  
Recall: 0.9789762340036563  
F1 Score: 0.9789762340036563



## 结论：

与卷积核滑动的方式相比，利用GAN深度学习模型检测普通异常点的准确率要略胜一筹，而检测明显异常点的准确率差异不大；

## However：

结合上一次任务：利用PCA以及其他降维方法检测普通异常点的效果比本次任务中的方法均要优秀

PCA的最佳结果：(取前7个主成分再IF)

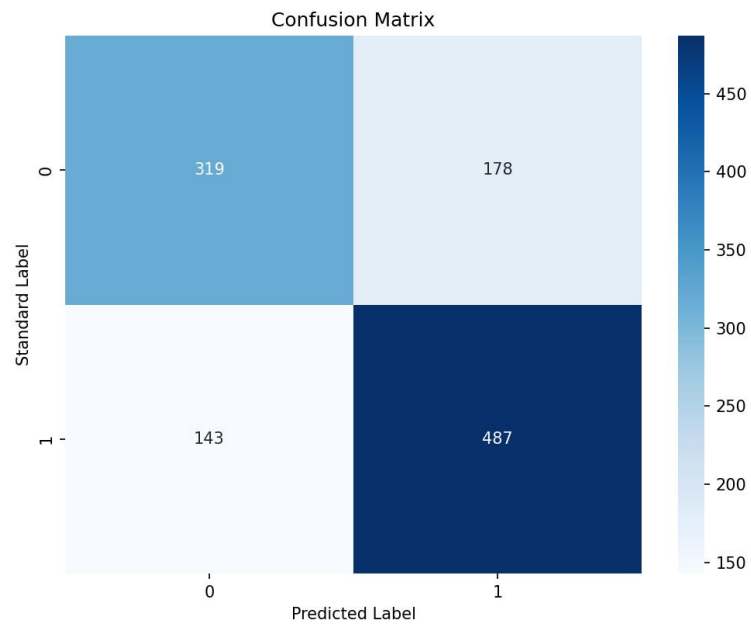
```
Precision: 0.7443609022556391  
Recall: 0.7857142857142857  
F1 Score: 0.7644787644787645
```

GAN模型的结果：

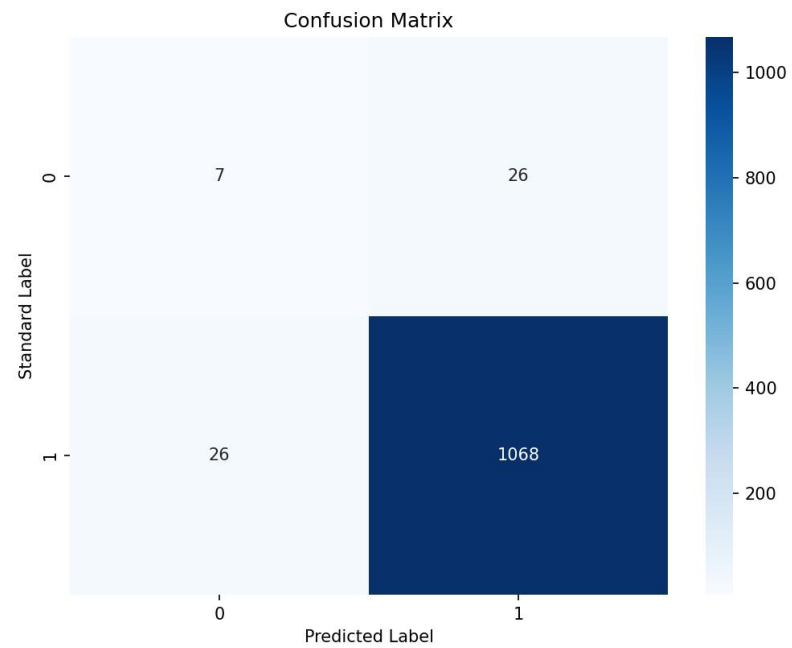
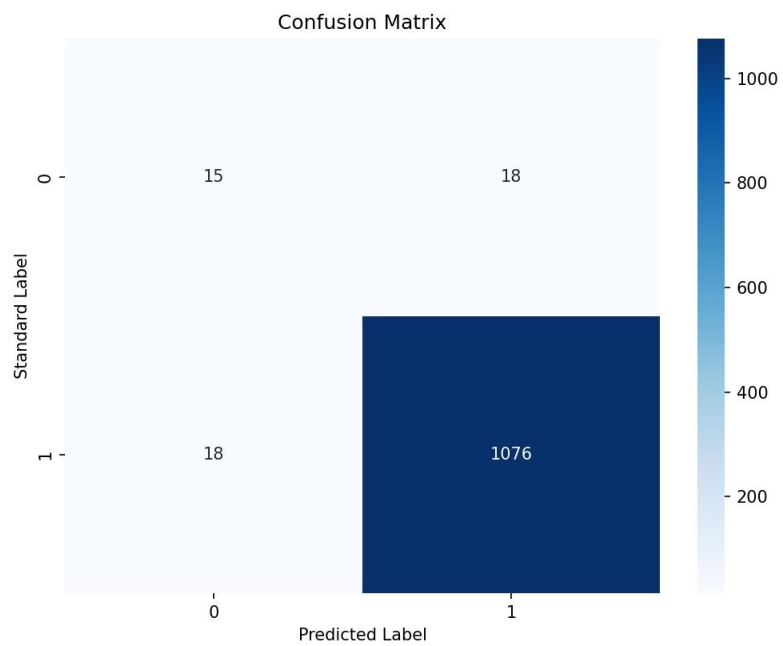
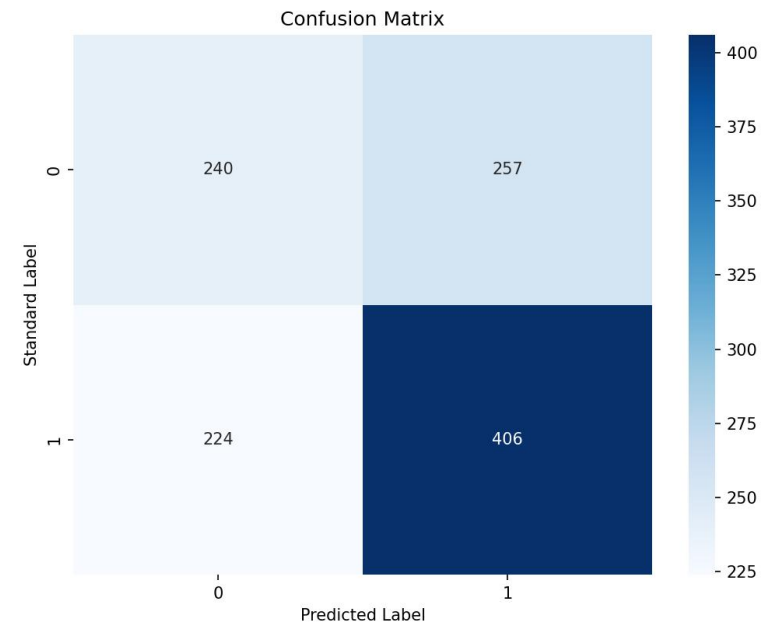
```
Precision: 0.6259426847662142  
Recall: 0.6587301587301587  
F1 Score: 0.6419180201082754
```

优于

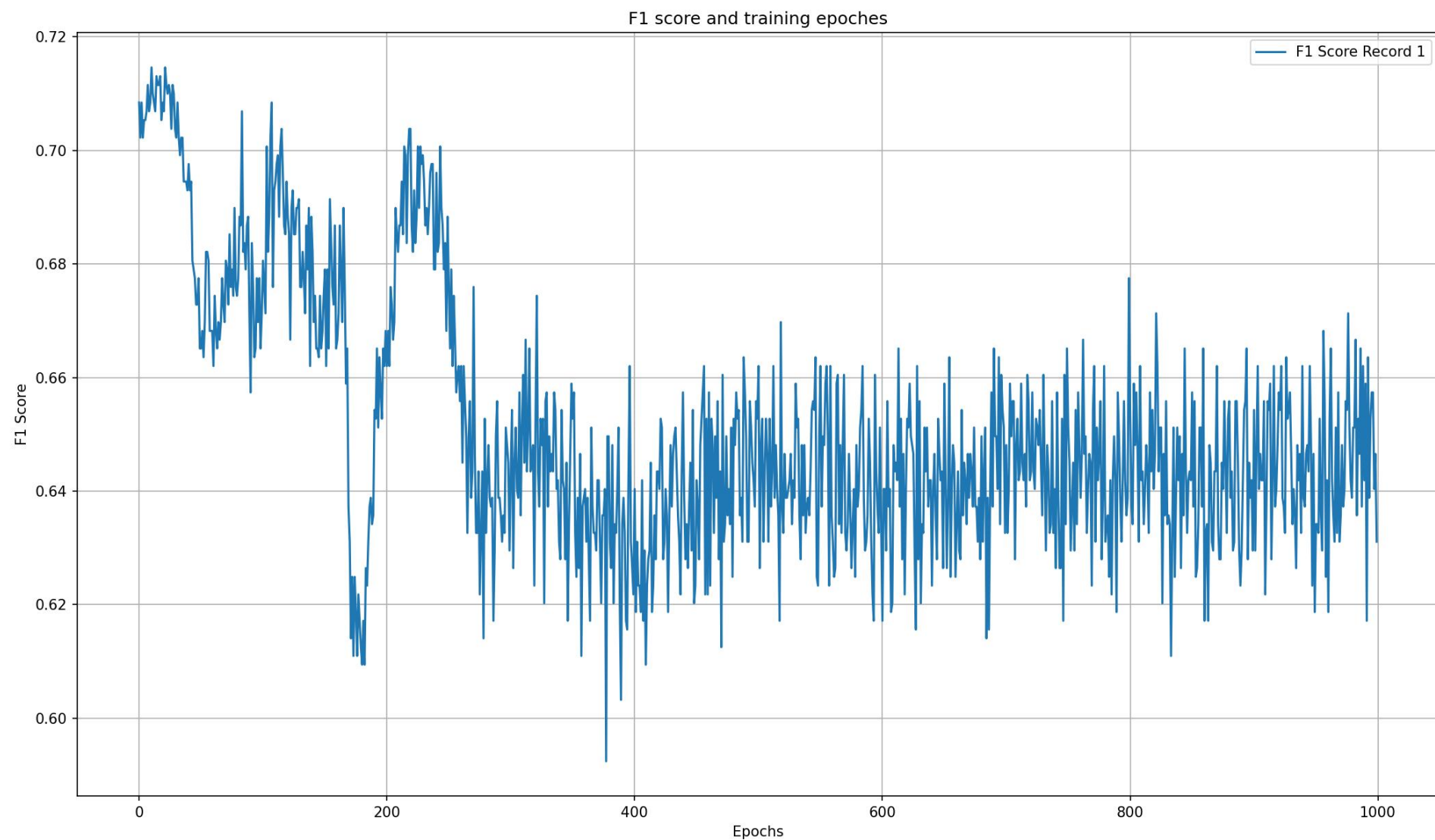
PCA:



卷积核滑动:



## GAN模型准确率较低的原因之一：模型可能存在过拟合



**下一步想法：**

- 1.找出我的GAN模型准确率低的原因，并进行修正**
- 2.学习更多训练模型、调整参数的方法，以提高准确率**

# Thank you for listening!

2024-9-3 By 刘震