基于**一维卷积神经网络**对分布式光纤振动传感器事件进行**定位和分类**

团队编号： T02053124

目录

[**1.** **说明** 2](#_Toc119712192)

[**2.** **创作背景** 2](#_Toc119712193)

[**3.** **运行环境搭建** 3](#_Toc119712194)

[a) **所用的库及其用途** 3](#_Toc119712195)

[**4.** **问题分析** 3](#_Toc119712196)

[**a)** **需求分析** 3](#_Toc119712197)

[**b)** **解决思路** 4](#_Toc119712198)

[**i.** **定位问题** 4](#_Toc119712199)

[**ii.** **分类问题** 4](#_Toc119712200)

[**iii.** **模型组合** 4](#_Toc119712201)

[**5.** **程序概要设计** 5](#_Toc119712202)

[**a)** **训练模型-流程** 5](#_Toc119712203)

[**b)** **组装模型-流程** 5](#_Toc119712204)

[**6.** **程序运行指南** 6](#_Toc119712205)

[**7.** **数据集处理以及程序的详细设计** 6](#_Toc119712206)

[**8.** **运行效果** 6](#_Toc119712207)

[**a)** **定位模型训练结果** 6](#_Toc119712208)

[**b)** **分类模型训练结果** 6](#_Toc119712209)

[**c)** **组合模型与预期结果对比** 7](#_Toc119712210)

[**9.** **创新点** 8](#_Toc119712211)

[**10.** **总结** 9](#_Toc119712212)

1. **说明**

本指南用于指导软件说明本次比赛作品的创作背景、作品采用的技术路线、代码运行环境的搭建、对数据集的处理、主要创新点、以及作品所达成的效果 。

1. **创作背景**

**此次项目的选题受指导老师的启发以及团队自身对于人工智能和超算领域的兴趣所确定的。**

**我们组希望通过这次难得的实践机会，提升对于人工智能在实际运用时的认知，利用实践来打磨、提升自己的编程水平、积攒一些较大型项目开发经验，并尽量多使用并行计算以感受超算的强悍之处。**

1. **运行环境搭建**

**该项目在python3.8.10环境下运行。**

**主要通过ssh导入所需的 \*\*\*.whl 文件后，利用命令安装所需的第三方库：**

pip install --no-index *\*\*\*.whl*

* 1. **所用的库及其用途**

**numpy：用于储存数据集，以及对其进行处理。**

**scipy：利用其中的 fft.fft 函数对数据集进行快速傅里叶变换。**

**openpyxl：相对于numpy，更快更精准的读取大型 .xlsx文件。**

**tensorflow：搭建神经网络。**

**sklearn：使用 train\_test\_split 函数分割数据集，用于训练神经网络。**

**collections：使用Counter对分类结果进行计数**

**multiprocessing：用于并行计算**

**datetime：用于计算程序运行时间**

1. **问题分析**
   1. **需求分析**

分布式光纤振动传感采样频率高、检测距离长的优点，但同样存在着事件发生时精准定位难，识别事件准确率低的问题。因此，程序需要分别解决**定位问题**和**分类问题**。然后将之组合，以达到精准定位事件并识别的目的。

* 1. **解决思路**
     1. **定位问题**

对于题目所提供的数据集，定位，即是找到符合特征的列的范围，故每一列只有**“是”和“否”**两种情况，归根结底，是一个**二分类问题**，那么虽然每一列的特征不同，但是通过利用**一维卷积神经网络**，使用是足够多的单列数据对其进行训练，就可以对每一列的数据进行分类，最终找到连续的列，记为事件的发生位置。

* + 1. **分类问题**

这是一个**多分类问题**，对于已经给出的事件发生位置以及事件标签，将其中的数据从对应的文件中提取出来后，对每个**连续**的32行进行**分割**，将每一块数据进行处理后转为同样大小的一维数组，与其对应标签一同加入另一个**一维卷积神经网络**进行训练，

* + 1. **模型组合**

得出两个神经网络后，将之组合，利用从**定位模型**中得到的结果，提取相应位置的数据，放入**分类模型**中，得出对应位置的数据标签，需求完成

1. **程序概要设计**
   1. **训练模型-流程**
   2. **组装模型-流程**
2. **程序运行指南**

**进入文件夹**: **cd /home/T02053124/源代码**

**提取训练集**：/T02053124/源代码# **python save\_dataset\_to\_csv.py**

**建立模型**：/T02053124/源代码# **./build.sh**

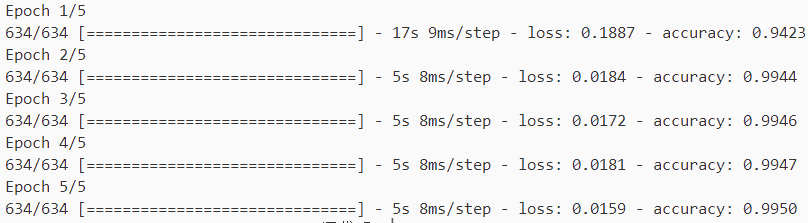
**组合模型测试**：/T02053124/源代码# **python test.py**

**运行主程序**: /T02053124/源代码# **python main.py**

1. **数据集处理以及程序的详细设计**

**见 文档/程序说明.docx**

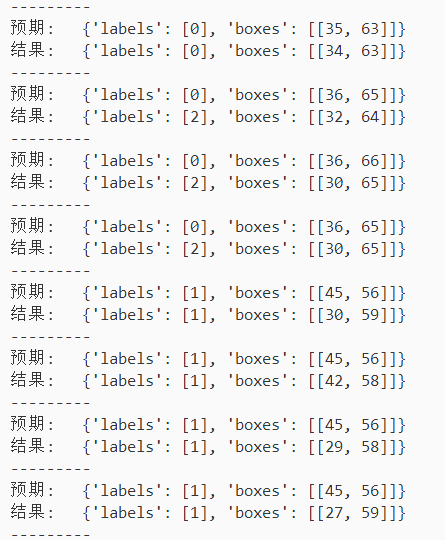
1. **运行效果**
   1. **定位模型训练结果**



* 1. **分类模型训练结果**



* 1. **组合模型与预期结果对比**

****

****



组合模型的运行结果与**"标注.json"**中所给的数据吻合度很高。

1. **创新点**

将**卷积神经网络**用于**分类问题**，把**定位问题**同样视为**分类问题**，运用**快速傅里叶变换**降低信号扰动带来的影响，合理运用**并行计算**减少程序调试时间。

编程尽可能**简洁**，设置更有**巧妙**的函数与类，增加代码的**可复用性**、程序的**可拓展性**，减少运行的**多余操作**，变量名尽可能简单易懂，增加**可读性**

1. **总结**

在这次比赛中，我们组争取充分利用超算平台的算力，力求做到又快又准，收获巨大，收益良多，但发现了许多的不足，这对未来的职业规划有着重要作用。

这次的比赛让我们意识到了团队的力量，越是庞大的项目仅凭借一个人的力量越是无法完成的，因为没有人是全知全能且精力无限的，良好的团队管理和合理的任务分配机制可以让项目开发事半功倍。