

**2022年秋季学期  
《视听觉信息理解》课程**

**听觉部分实验**

|  |  |
| --- | --- |
| 姓名 | 姚舜宇 |
| 学号 | 1190202107 |
| 班号 | 1903602 |

1. 实验目的

掌握一种书画人识别/语音识别工具包的使用方法，如Kaldi

能够读懂，并按照要求修改说话人确认算法的源代码

能够利用该工具包，完成指定的实验方案，并优化实验结果

能够在实验报告中条理清晰地、详实专业地汇报实验过程，总结实验结果

1. 实验过程
2. 下载实验指导书中给出的kaldi源码以及数据集，并按照给定的连接进行环境的配置与安装
3. 进入实验指导书中给出的egs/voxceleb/v2文件夹，阅读run.sh文件。该文件是脚本文件，用于批处理，描述了实验的训练与测试过程
4. 该run.sh文件的流程如下：

首先准备数据集，然后对数据集提取MFCC特征

对数据进行增强，利用回响和三种噪声增强了训练数据，并测试增强效果

通过归一化的方法除去静音帧

生成的egs用于训练x-vector模型

用embedding层提取特征

用LDA降维

用PLDA打分

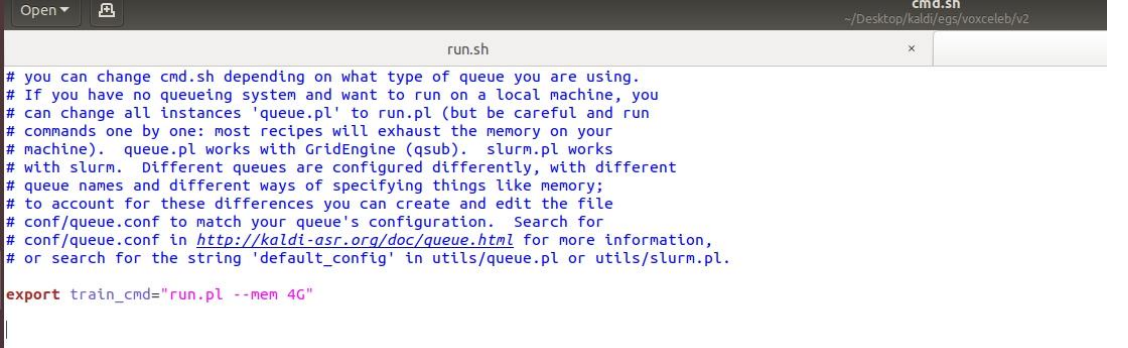
计算EER

1. 由于实验指导书中已经给定模型和数据集，即要求修改此文件使得能在所给的数据集上进行复现测试，因此对其改动如下：

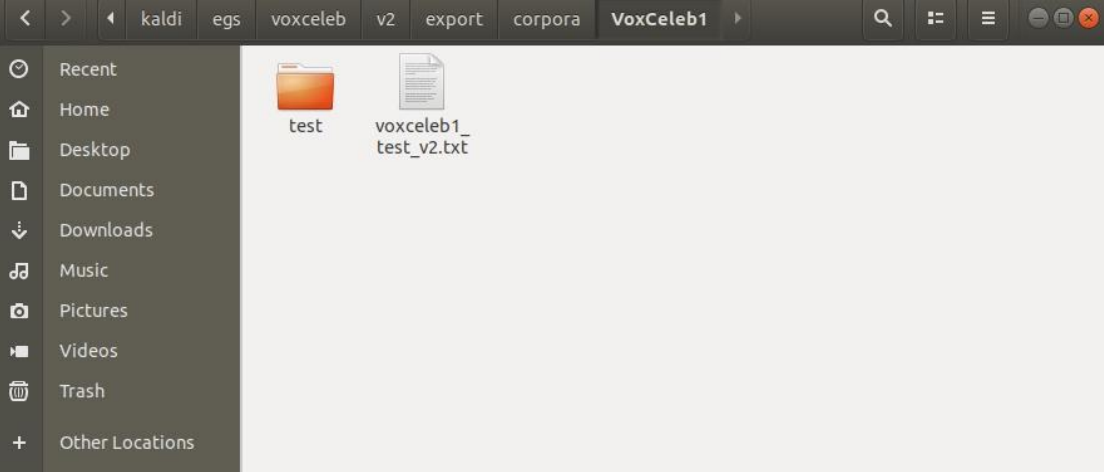
只保留stage1对测试数据集的处理，保留stage9中的“Extract x-vectors used in the evaluation”部分，保留stage11和stage12部分用于进行评分测试。其他部分可以删除。详见sh文件

1. 测试过程如下：

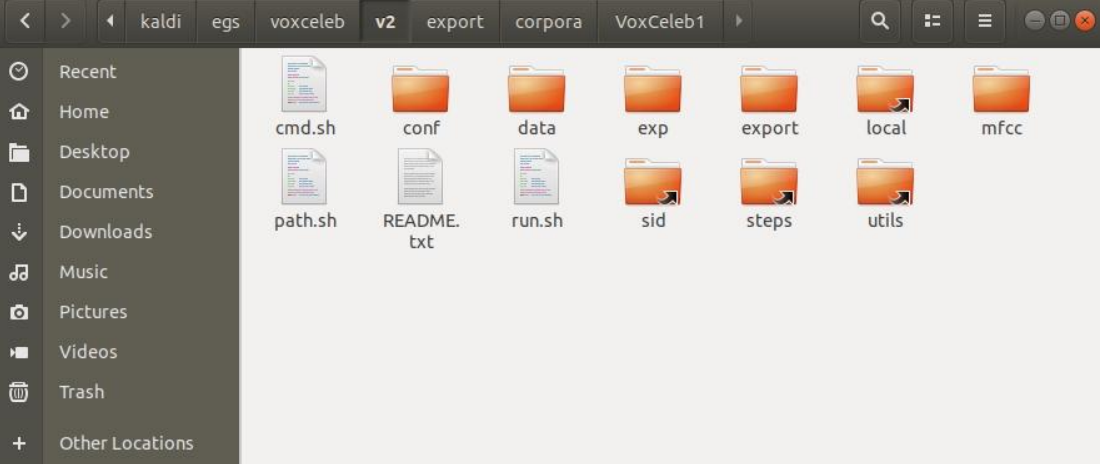
首先修改cmd.sh文件如下：



分别下载模型与数据，测试数据放在如下文件夹中：



模型放在如下文件：



然后手动调用make\_voxceleb1\_v2.pl对测试集数据进行处理，并生成后续所需的scp文件。

在根目录下执行bash run.sh命令，即可以运行。

1. 实验结果

最终得到的实验结果如下所示：

