**代码说明：**

该文件夹中的内容是用于生成对抗音频验证码。

其中，

**data/**里面是数据集文件；

**DeepSpeech/**里面是deepspeech源码等相关文件；

**deepSpeech\_model/**里面是调用deepspeech语音识别模型所用到的模型；

**lingvo/**里面是lingvo相关的源码；

**lingvo\_model/**里面是调用lingvo语音识别模型所用到的预训练模型；

**OpenSeq2Seq/**里面是jasper和wav2letter+相关的源码；

**jasper\_model/**里面是调用jasper语音识别模型所用到的预训练模型；

**w2lplus\_model/**里面是调用wav2letter+语音识别模型所用到的预训练模型；

**w2lplus\_model\_new/**里面是调用wav2letter+语音识别模型实际用到的预训练模型(由于集成中w2lplus\_model和jasper\_model参数名重复的问题，需要对其中一个模型参数重命名，w2lplus\_model\_new是经过modelRename.py重命名后的模型)；

**result/**里面存放代码运行结果；

**temp/**里面是临时文件；

**util/**里面是deepspeech相关的文件；

常用语音数据集下载：<https://github.com/jim-schwoebel/voice_datasets>

1. **主实验代码**

主体实验部分训练集与测试集为同一个数据集；

**训练：**

**main\_1asr.py：**deepspeech单独训练的代码；

**main\_1lingvo.py：**lingvo单独训练的代码；

**main\_2.py：**deepspeech和lingvo参与集成训练的代码；

**main\_3.py：**deepspeech、lingvo以及jasper参与集成训练的代码；

**main\_4.py：**deepspeech、lingvo、jasper以及wav2letter+参与集成训练的代码。

**测试：**

**advGenerate2.py：**deepspeech和lingvo参与集成训练，利用最后一步的模型对所有音频生成对抗语音的代码；

**advGenerate3.py：**deepspeech、lingvo以及jasper集成训练相关测试代码；

**advGenerate4.py：**deepspeech、lingvo、jasper以及wav2letter+参与集成训练相关测试代码。

1. **Universal GAN**

universal GAN部分训练集与测试集不同且不重合，论文中只进行了deepspeech参与训练的实验；

**训练：**

**main\_deepspeech.py：**deepspeech单独训练的代码。

**测试：**

**main\_deepspeech\_test.py：**deepspeech训练，利用最后一步的模型对所有测试集音频生成对抗语音的代码。

1. **代码执行具体流程：**

注意所有代码路径需要进行更改。

1. 所有数据集需要先转为为：单声道、int16、16000Hz、wav文件格式。（./data/data\_process\_code/audio\_format.py）；
2. 对于recaptcha v2数据集，需要截掉两边的纯噪声部分。（./data/data\_process\_code/cut\_noise.py）；
3. 对于recaptcha v2截掉两边噪声的数据集，需要在两边填充一小段，填充值为0。（./data/data\_process\_code/padding.py）；
4. 划分数据集，按照一定比例，分出训练集、验证集和测试集；
5. 将音频文件对应的转录存入tran.txt文件，每一行的格式（音频文件名:原始转录），注意，因为我们接下来只顾有目标攻击的情况，所以不会用到原转录，所以原始转录用文件名替代（音频文件名:音频文件名），文件名不包括后缀。（./data/data\_process\_code/audio\_getTran.py）
6. 将所有音频文件分别输入deepspeech、lingvo、jasper以及wavletter+语音识别系统，获得其相应转录，存入txt文件（将获得tran\_d.txt，tran\_d\_l.txt，tran\_d\_l\_j.txt， tran\_d\_l\_j\_w.txt），最终格式（音频文件名: 音频文件名:deepspeech转录:lingvo转录:jasper转录:wavletter+转录），注意，原始转录不一定用到，可以用文件名替代。（分别按顺序执行deepSpeech\_call.py， lingvo\_call.py，jasper\_call\_new.py，w2l\_call\_new.py）
7. 将所有转录统一为相同的长度（最大长度值）。（./data/data\_process\_code/unify\_audio\_tran.py）
8. 执行训练和测试代码，例如想进行universal GAN的实验部分，则

python main\_deepspeech.py

python main\_deepspeech\_test.py