

基本知识

2023年4月4日 上午 07:38

- 1.Data: 如很多十秒的音频剪辑或者各种图像 (Distribution?如男人女人小孩的声音)
- 2.Input: 如一段10s的音频，包含各种词汇（包括激活词）（Resolution?如语音处理的最佳采样率sample rate）
- 3.Output: $y=000010000$ $y=0000000000$ $y=10000000000$ （Last Activation?如Sigmoid函数，对于每个时间步长，使用该函数输出0或1）
- 4.Architecture(架构): 选择的设计，例如逻辑回归、深层浅层神经网络、卷积神经网络、递归神经网络，可以选择它们更深或更浅
- 5.Loss: 对于每一个时间步，都应该用一次对应的损失函数，并对所有时间戳求和

思考：What is critical to the success of this project?（如a.收集数据时需要收集三个数据库：positive words, negative words, background noise. 找到创建数据的正确策略十分重要！！b.架构搜索和参数调整: talk to the experts如找到用于该傅里叶变换的正确超参数，将信号转化为矩阵