1. 数据集来源

目前我们已经对来自Maher的数据集进行了评估，如果需要引入其他的gunshot数据集

需要重新进行评估。数据增强方面已经进行了调研，有几个较新的数据增强方法也有现成的代码可以参考。

1. 算法部分

分为先验模型与后续数据后验

Context-dependent sound event detection中提到了使用先验模型的方法，下为原文

除了声学特征和分类方案之外，还研究了各种方法，以将事件的先验知识包括到检测过程中。可以使用帧级n- gram 定义用于环境分类的声学同质段，其中n- gram用于基于先前观察到的帧来建模帧的先验概率[ 28 ]。在具有许多重叠事件的复杂声学环境中，可能的组合的数量太多，以致于无法定义这种声学上均质的片段并无法对它们之间的过渡进行建模。在[ 3提出了一种用于检测关键音效和音频场景类别的分层概率模型。使用HMM对声音效果进行建模，并使用更高级别的模型通过类似于语音识别中语言模型的语法网络连接各个声音效果模型。通过使用概率潜在语义分析来计算先验并学习声音事件之间的关联，在[ 29 ]中已解决了一种对重叠事件先验建模的方法。本文提出的上下文识别阶段将通过根据上下文将事件集分为子集来解决声音事件的关联。此外，从培训材料中估计的基于计数的先验可用于为每个上下文内的声音事件提供概率分布。

<http://www.cs.tut.fi/~heittolt/research-sound-event-detection>

这个网站介绍了几种现有的利用模型先验的声学事件检测方法

1. 算法评估

这部分工作在之后补充