

标注员指南:使用LosslessCut进行视频片段选择

1. 初始设置

步骤 1: 文件夹创建

为每个工作视频创建一个特定的文件夹。

文件夹名称应采用以下格式: `xx_s1_ep1` (请根据实际的会话和集数替换 `s1` 和 `ep1`)。

例如:

对于第2季, 第3集, 应命名为 `xx_s2_ep3`。

所有视频片段和时间戳CSV文件必须保存在这个特定的文件夹中。

步骤 2: 下载LosslessCut

访问LosslessCut的官方GitHub页面[\[link\]](#), 下载适合您的操作系统版本(Windows、macOS或Linux)。

根据您的平台, 按照安装或设置说明进行操作。LosslessCut可能作为便携应用程序(例如Linux的ApplImage, 或者Windows/macOS的独立可执行文件)提供, 无需安装即可直接运行。

步骤 3: 熟悉LosslessCut界面

通过阅读GitHub上的“典型工作流程”来熟悉界面。

主要功能:

- 将视频文件拖放到播放器中。
- 按下空格键(SPACE)播放/暂停, 或者使用◀▶, , 进行快进/倒退。
- 通过移动时间标记并按下I键设置开始时间, 再按下O键设置结束时间来选择片段的开始和结束时间。
建议首先尝试切割一些小的片段并查看效果, 然后再开始批量切割。

2. 使用LosslessCut进行视频片段选择

主要任务:

你需要在LosslessCut界面中观看视频, 选择你认为带有讽刺意味的语句。并为每个讽刺语句选择另一个同一说话者的真诚语句。

在切割过程中, 最好的片段是包含一个完整说话者的完整话语(而不是“我到剧场报名演出, 老板只看了我一眼, 当场签约, 我刚签完字, 他就让...”, 完整话语包括: “我到剧场报名演出, 老板只看了我一眼, 当场签约, 我刚签完字, 他就让助理给我买500万的人身意外保险, 这个老板太有经济

头脑了, 只要是出点啥事儿, 我就是天使轮”), 并且请避免背景音和观众笑声(切割视频片段时请在讲话结束后进行切割, 因为笑声通常出现在讲话后)。话语应该包含完整的对话内容(例如: “要是出点啥事儿, 我就是天使轮”), 而不仅仅是片段。

步骤 1: 加载视频并设置目录

打开LosslessCut。

将视频文件拖放到LosslessCut界面中。

按下“工作目录未设置”按钮, 将保存目录指向预定义文件夹。

步骤 2: 观看并选择片段

使用界面控件播放视频。

找到并标记片段如下:

- 讽刺语句:
当你找到一个讽刺语句时, 按下I键或“开始切割”按钮标记开始时间。
继续播放直到语句结束, 再按下O键或“结束切割”按钮标记结束时间。
- 真诚语句(配对):
找到同一说话者的真诚语句(不一定直接位于讽刺语句之前或之后)。
重复上述步骤, 使用I和O键进行标记。

步骤 3: 导出片段

完成标记后:

点击“export”按钮, 显示一个导出选项概览。

确保使用预定义的文件夹保存片段。

然后再次按“export”确认导出。

双重检查视频片段的数量, 确保其与所选片段相匹配。

步骤 4: 导出时间戳文件

在LosslessCut中, 选择文件菜单中的时间戳数据:

选择“export project”并选择“timestamps(CSV)”, 以生成包含所有片段时间戳的文件。

将此文件保存到相同的文件夹。

3. 质量检查与上传

步骤 1: 审查片段和时间戳

播放每个导出的片段, 确保它包含了意图的语句。

确认片段顺序正确(奇数编号的片段应为讽刺, 偶数编号的片段应为真诚)。

确保时间戳文件完整且准确(检查时间戳是否与所选片段一致)。

步骤 2: 上传文件夹

一旦所有片段正确导出并经过审查, 将整个文件夹上传到指定的文件夹进行共享。

4. 给标注员的建议

- 首先选择少量的视频片段, 以确保你了解流程并获得反馈, 然后再进行大批量的标注。
- 如果遇到任何问题或对标注和导出片段有任何疑问, 请随时联系。

5. 故障排除