



(12) 发明专利申请

(10) 申请公布号 CN 119229337 A

(43) 申请公布日 2024. 12. 31

(21) 申请号 202411180548.9

(22) 申请日 2024.08.27

(71) 申请人 江苏金海星导航科技有限公司

地址 212000 江苏省镇江市京口区宗泽路
108号二楼

(72) 发明人 金文 马国团 李浩 易康扬

孙中强 熊超 童强

(74) 专利代理机构 深圳市联江知识产权代理事

务所(特殊普通合伙) 44939

专利代理师 陈红运

(51) Int. Cl.

G06V 20/40 (2022.01)

G06V 10/80 (2022.01)

G06Q 50/26 (2024.01)

权利要求书4页 说明书21页 附图3页

(54) 发明名称

一种基于执法督察用的多模态智能分析方法及系统

(57) 摘要

本发明涉及执法智能告警技术领域,尤其涉及一种基于执法督察用的多模态智能分析方法及系统。所述方法包括以下步骤:通过从执法督察记录仪平台中获取执法现场视频文件并进行音视频切分和执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据;对执法现场视频文件进行时间戳及空间位置抽取处理,并进行多模态时空同步关联分析,同时进行执法行为序列连接处理和事件行为意图识别分析,得到执法现场事件行为意图信息数据;获取预设的执法事件行为知识图谱并进行行为规范性评分计算和不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。本发明能够提高了执法不规范行为监管告警的准确性。



1. 一种基于执法督察用的多模态智能分析方法,其特征在于,包括以下步骤:

步骤S1:通过从执法督察记录仪平台中获取执法现场视频文件,并对执法现场视频文件进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

步骤S2:对执法现场视频文件进行时间戳及空间位置抽取处理,以得到执法现场过程时间戳以及执法现场过程空间位置信息;基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,得到多模态时空同步执法行为关联事件数据;

步骤S3:对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据;

步骤S4:获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

2. 根据权利要求1所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S1包括以下步骤:

步骤S11:通过从执法督察记录仪平台中获取执法现场视频文件;

步骤S12:对执法现场视频文件进行音画同步处理,得到执法现场音画同步视频;

步骤S13:对执法现场音画同步视频进行视频内容提取处理,得到执法现场视频内容信息数据;基于执法现场视频内容信息数据对执法现场音画同步视频进行不同执法事件开始及结束时间识别分析,以得到执法现场视频内不同执法事件的开始时间以及结束时间;

步骤S14:根据执法现场视频内不同执法事件的开始时间以及结束时间对执法现场音画同步视频进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;

步骤S15:分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据。

3. 根据权利要求2所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S15包括以下步骤:

步骤S151:利用Speechbrain技术对执法现场音频片段进行执法音频声纹特征分析,得到执法现场音频声纹特征数据;

步骤S152:基于执法现场音频声纹特征数据对执法现场音频片段内的不同说话者进行音频语音信号通道解构处理,得到执法现场音频内不同说话者的音频语音信号子通道;

步骤S153:对执法现场音频内不同说话者的音频语音信号子通道进行执法行为事件触发识别分析,得到执法现场音频执法行为事件触发点;

步骤S154:基于执法现场音频执法行为事件触发点对执法现场音频片段进行执法音频行为信息识别分析,得到执法现场音频行为信息数据;

步骤S155:利用物体检测算法Yolov8以及跟踪技术Bytetrack对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

4.根据权利要求3所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S155包括以下步骤:

利用物体检测算法Yolov8对执法现场图像视频帧片段进行执法现场物体目标识别分析,得到执法现场图像帧物体检测目标;

对执法现场图像帧物体检测目标进行物体边界区域标记处理,得到执法现场物体边界区域标记目标;

利用跟踪技术Bytetrack对执法现场物体边界区域标记目标在执法现场图像视频帧片段内的运动轨迹进行跟踪识别分析,以生成执法现场物体标记目标运动轨迹;

对执法现场物体标记目标运动轨迹进行物体目标行为模式分析,得到不同执法现场物体目标的执法行为模式;基于不同执法现场物体目标的执法行为模式对相对应的执法现场物体边界区域标记目标进行执法行为动作分类处理,以得到不同执法现场物体目标的执法行为动作;

基于不同执法现场物体目标的执法行为模式以及执法行为动作对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

5.根据权利要求1所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S2包括以下步骤:

步骤S21:对执法现场视频文件进行时间戳抽取处理,以得到执法现场过程时间戳;

步骤S22:对执法现场视频文件中的每一帧视频图像进行三维空间坐标系转换,得到执法现场视频中每一帧图像的三维空间坐标系;

步骤S23:对执法现场视频中每一帧图像的三维空间坐标系进行执法现场空间坐标提取处理,以得到执法现场视频中每一帧图像的执法现场空间坐标;

步骤S24:对执法现场视频中每一帧图像的执法现场空间坐标进行空间位置标记处理,以得到执法现场过程空间位置信息;

步骤S25:基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,得到多模态时空同步执法行为关联事件数据。

6.根据权利要求5所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S25包括以下步骤:

步骤S251:对执法现场过程时间戳以及执法现场过程空间位置信息进行时空同步基准构建,以得到执法现场过程时空数据同步基准框架;

步骤S252:基于执法现场过程时空数据同步基准框架对执法现场音频行为信息数据进行音频事件时序同步排列处理,得到执法现场音频事件时序同步序列数据;

步骤S253:基于执法现场过程时空数据同步基准框架对执法现场图像行为信息数据进行图像行为空间映射处理,得到执法现场图像行为空间位置映射数据;

步骤S254:对执法现场音频事件时序同步序列数据以及执法现场图像行为空间位置映射数据进行多模态事件行为交互关联分析,得到执法现场音频事件与图像行为之间的多模态事件行为交互关联关系;

步骤S255:基于执法现场音频事件与图像行为之间的多模态事件行为交互关联关系对执法现场音频行为信息数据以及执法现场图像行为信息数据进行执法行为关联事件提取处理,得到多模态时空同步执法行为关联事件数据。

7.根据权利要求1所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S3包括以下步骤:

步骤S31:对多模态时空同步执法行为关联事件数据进行执法人员及当事人行为序列抽取处理,得到执法行为事件执法人员行为序列以及执法行为事件当事人行为序列;

步骤S32:对执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点之间进行行为描述匹配分析,以得到执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系;

步骤S33:基于执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系对执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点之间进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;

步骤S34:对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据。

8.根据权利要求1所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S4包括以下步骤:

步骤S41:获取预设的执法事件行为知识图谱;

步骤S42:对执法事件行为知识图谱进行行为规范性规则检索抽取处理,以得到执法事件行为规范性规则标准;

步骤S43:对执法现场事件行为意图信息数据进行行为意图语义特征分析,得到执法现场事件行为意图语义特征数据;

步骤S44:基于执法事件行为规范性规则标准对执法现场事件行为意图语义特征数据进行规范性规则适配及行为规范性评分计算,以得到执法现场行为意图规范性评分值;

步骤S45:基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

9.根据权利要求8所述的基于执法督察用的多模态智能分析方法,其特征在于,步骤S45包括以下步骤:

步骤S451:根据预设的规范性评分阈值对执法现场行为意图规范性评分值进行比较判断,当执法现场行为意图规范性评分值大于或等于预设的规范性评分阈值时,则将其执法现场事件行为意图信息数据中相对应的执法现场事件判定为规范行为事件;当执法现场行为意图规范性评分值小于预设的规范性评分阈值时,则将其执法现场事件行为意图信息数据中相对应的执法现场事件判定为不规范行为事件;

步骤S452:根据预设的执法行为不规范告警规则对被判定为不规范行为事件进行不规范行为智能告警处理,以生成本次监督不规范行为事件告警信息;

步骤S453:对本次监督不规范行为事件告警信息进行执法行为问题识别分析,得到本次监督不规范执法行为问题报告;

步骤S454:对本次监督不规范执法行为问题报告进行执法行为优化建议处理,以生成执法现场事件不规范行为优化建议报告。

10.一种基于执法督察用的多模态智能分析系统,其特征在于,用于执行如权利要求1所述的基于执法督察用的多模态智能分析方法,该基于执法督察用的多模态智能分析系统包括:

执法音视频行为识别分析模块,用于通过从执法督察记录仪平台中获取执法现场视频文件,并对执法现场视频文件进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,从而得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

执法行为多模态时空关联分析模块,用于对执法现场视频文件进行时间戳及空间位置抽取处理,以得到执法现场过程时间戳以及执法现场过程空间位置信息;基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,从而得到多模态时空同步执法行为关联事件数据;

执法事件行为意图识别分析模块,用于对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,从而得到执法现场事件行为意图信息数据;

执法事件不规范行为智能告警模块,用于获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

一种基于执法督察用的多模态智能分析方法及系统

技术领域

[0001] 本发明涉及执法智能告警技术领域,尤其涉及一种基于执法督察用的多模态智能分析方法及系统。

背景技术

[0002] 在现代执法领域,多模态智能分析技术的应用正变得越来越重要,通过集成视频、音频和数据日志等多种信息源,并利用先进的机器学习和深度学习算法,对采集集成的数据集进行实时处理和融合分析,以提取出有用的信息和模式,同时通过智能分析模型识别其中的异常行为、违规操作或其他需要关注的情况,并生成详细的分析报告,根据分析结果,提供实时反馈和改进建议,从而帮助执法人员提升执法效率和规范性。然而,传统的执法督察数据分析方法在覆盖面和时效性上均显不足,往往无法及时反应并全面覆盖所有的执法现场行为关键监控点。

发明内容

[0003] 基于此,本发明有必要提供一种基于执法督察用的多模态智能分析方法及系统,以解决至少一个上述技术问题。

[0004] 为实现上述目的,一种基于执法督察用的多模态智能分析方法,包括以下步骤:

[0005] 步骤S1: 通过从执法督察记录仪平台中获取执法现场视频文件,并对执法现场视频文件进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

[0006] 步骤S2: 对执法现场视频文件进行时间戳及空间位置抽取处理,以得到执法现场过程时间戳以及执法现场过程空间位置信息;基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,得到多模态时空同步执法行为关联事件数据;

[0007] 步骤S3: 对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据;

[0008] 步骤S4: 获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0009] 进一步的,步骤S1包括以下步骤:

[0010] 步骤S11: 通过从执法督察记录仪平台中获取执法现场视频文件;

[0011] 步骤S12: 对执法现场视频文件进行音画同步处理,得到执法现场音画同步视频;

[0012] 步骤S13:对执法现场音画同步视频进行视频内容提取处理,得到执法现场视频内容信息数据;基于执法现场视频内容信息数据对执法现场音画同步视频进行不同执法事件开始及结束时间识别分析,以得到执法现场视频内不同执法事件的开始时间以及结束时间;

[0013] 步骤S14:根据执法现场视频内不同执法事件的开始时间以及结束时间对执法现场音画同步视频进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;

[0014] 步骤S15:分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据。

[0015] 进一步的,步骤S15包括以下步骤:

[0016] 步骤S151:利用Speechbrain技术对执法现场音频片段进行执法音频声纹特征分析,得到执法现场音频声纹特征数据;

[0017] 步骤S152:基于执法现场音频声纹特征数据对执法现场音频片段内的不同说话者进行音频语音信号通道解构处理,得到执法现场音频内不同说话者的音频语音信号子通道;

[0018] 步骤S153:对执法现场音频内不同说话者的音频语音信号子通道进行执法行为事件触发识别分析,得到执法现场音频执法行为事件触发点;

[0019] 步骤S154:基于执法现场音频执法行为事件触发点对执法现场音频片段进行执法音频行为信息识别分析,得到执法现场音频行为信息数据;

[0020] 步骤S155:利用物体检测算法Yolov8以及跟踪技术Bytetrack对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

[0021] 进一步的,步骤S155包括以下步骤:

[0022] 利用物体检测算法Yolov8对执法现场图像视频帧片段进行执法现场物体目标识别分析,得到执法现场图像帧物体检测目标;

[0023] 对执法现场图像帧物体检测目标进行物体边界区域标记处理,得到执法现场物体边界区域标记目标;

[0024] 利用跟踪技术Bytetrack对执法现场物体边界区域标记目标在执法现场图像视频帧片段内的运动轨迹进行跟踪识别分析,以生成执法现场物体标记目标运动轨迹;

[0025] 对执法现场物体标记目标运动轨迹进行物体目标行为模式分析,得到不同执法现场物体目标的执法行为模式;基于不同执法现场物体目标的执法行为模式对相对应的执法现场物体边界区域标记目标进行执法行为动作分类处理,以得到不同执法现场物体目标的执法行为动作;

[0026] 基于不同执法现场物体目标的执法行为模式以及执法行为动作对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

[0027] 进一步的,步骤S2包括以下步骤:

[0028] 步骤S21:对执法现场视频文件进行时间戳抽取处理,以得到执法现场过程时间戳;

[0029] 步骤S22:对执法现场视频文件中的每一帧视频图像进行三维空间坐标系转换,得到执法现场视频中每一帧图像的三维空间坐标系;

[0030] 步骤S23:对执法现场视频中每一帧图像的三维空间坐标系进行执法现场空间坐标提取处理,以得到执法现场视频中每一帧图像的执法现场空间坐标;

[0031] 步骤S24:对执法现场视频中每一帧图像的执法现场空间坐标进行空间位置标记处理,以得到执法现场过程空间位置信息;

[0032] 步骤S25:基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,得到多模态时空同步执法行为关联事件数据。

[0033] 进一步的,步骤S25包括以下步骤:

[0034] 步骤S251:对执法现场过程时间戳以及执法现场过程空间位置信息进行时空同步基准构建,以得到执法现场过程时空数据同步基准框架;

[0035] 步骤S252:基于执法现场过程时空数据同步基准框架对执法现场音频行为信息数据进行音频事件时序同步排列处理,得到执法现场音频事件时序同步序列数据;

[0036] 步骤S253:基于执法现场过程时空数据同步基准框架对执法现场图像行为信息数据进行图像行为空间映射处理,得到执法现场图像行为空间位置映射数据;

[0037] 步骤S254:对执法现场音频事件时序同步序列数据以及执法现场图像行为空间位置映射数据进行多模态事件行为交互关联分析,得到执法现场音频事件与图像行为之间的多模态事件行为交互关联关系;

[0038] 步骤S255:基于执法现场音频事件与图像行为之间的多模态事件行为交互关联关系对执法现场音频行为信息数据以及执法现场图像行为信息数据进行执法行为关联事件提取处理,得到多模态时空同步执法行为关联事件数据。

[0039] 进一步的,步骤S3包括以下步骤:

[0040] 步骤S31:对多模态时空同步执法行为关联事件数据进行执法人员及当事人行为序列抽取处理,得到执法行为事件执法人员行为序列以及执法行为事件当事人行为序列;

[0041] 步骤S32:对执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点之间进行行为描述匹配分析,以得到执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系;

[0042] 步骤S33:基于执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系对执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点之间进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;

[0043] 步骤S34:对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据。

[0044] 进一步的,步骤S4包括以下步骤:

[0045] 步骤S41:获取预设的执法事件行为知识图谱;

[0046] 步骤S42:对执法事件行为知识图谱进行行为规范性规则检索抽取处理,以得到执法事件行为规范性规则标准;

[0047] 步骤S43:对执法现场事件行为意图信息数据进行行为意图语义特征分析,得到执法现场事件行为意图语义特征数据;

[0048] 步骤S44:基于执法事件行为规范性规则标准对执法现场事件行为意图语义特征

数据进行规范性规则适配及行为规范性评分计算,以得到执法现场行为意图规范性评分值;

[0049] 步骤S45:基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0050] 进一步的,步骤S45包括以下步骤:

[0051] 步骤S451:根据预设的规范性评分阈值对执法现场行为意图规范性评分值进行比较判断,当执法现场行为意图规范性评分值大于或等于预设的规范性评分阈值时,则将其执法现场事件行为意图信息数据中相对应的执法现场事件判定为规范行为事件;当执法现场行为意图规范性评分值小于预设的规范性评分阈值时,则将其执法现场事件行为意图信息数据中相对应的执法现场事件判定为不规范行为事件;

[0052] 步骤S452:根据预设的执法行为不规范告警规则对被判定为不规范行为事件进行不规范行为智能告警处理,以生成本次监督不规范行为事件告警信息;

[0053] 步骤S453:对本次监督不规范行为事件告警信息进行执法行为问题识别分析,得到本次监督不规范执法行为问题报告;

[0054] 步骤S454:对本次监督不规范执法行为问题报告进行执法行为优化建议处理,以生成执法现场事件不规范行为优化建议报告。

[0055] 进一步的,本发明还提供了一种基于执法督察用的多模态智能分析系统,用于执行如上所述的基于执法督察用的多模态智能分析方法,该基于执法督察用的多模态智能分析系统包括:

[0056] 执法音视频行为识别分析模块,用于通过从执法督察记录仪平台中获取执法现场视频文件,并对执法现场视频文件进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,从而得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

[0057] 执法行为多模态时空关联分析模块,用于对执法现场视频文件进行时间戳及空间位置抽取处理,以得到执法现场过程时间戳以及执法现场过程空间位置信息;基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,从而得到多模态时空同步执法行为关联事件数据;

[0058] 执法事件行为意图识别分析模块,用于对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,从而得到执法现场事件行为意图信息数据;

[0059] 执法事件不规范行为智能告警模块,用于获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0060] 本发明的有益效果:

[0061] 1、本发明所提出的基于执法督察用的多模态智能分析方法,与现有技术相比,本申请的有益效果在于通过从执法督察记录仪平台获取执法现场的视频文件,可以确保对执法活动有全面的记录和审查,这些视频文件为执法过程提供了真实、全面的视觉证据,获取这些视频文件使得执法过程的每一个细节都能被记录下来,无论是执法人员的行为,还是执法对象的反应,都能够被完整捕捉,这种详细的记录不仅有助于后续的审查和调查,增强执法活动的可信度和合法性,从而能够帮助识别执法过程中的潜在问题,确保执法操作符合规范和标准。通过对执法现场视频文件进行音视频切分处理,可以将视频数据分解为更小的音频片段和图像视频帧片段,这种切分处理可以将整个视频内容划分为具有独立意义的部分,每一部分对应一个具体的执法事件,这样可以更方便地进行针对性的分析和审查。通过将音频和图像数据分开处理,可以更清晰地了解每个事件中的具体情况,例如,音频片段可以用于分析对话内容和语音情感,而视频帧片段则可以用于观察视觉细节和行为表现,这种方法提高了数据分析的精确度和效率,使得每个执法事件能够被独立审查和评估,从而更全面地了解执法过程。同时,通过分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,可以提取出具体的执法行为信息数据,其中对音频片段的分析能够识别出其中的语言、声音特征和对话内容,帮助理解执法过程中的沟通交流和行为动机,而图像视频帧片段的分析则能够识别出视觉上的行为特征,例如执法动作、现场环境和人员位置等,这种行为信息识别可以为执法活动提供详细的实证数据,有助于支持对执法过程的全面评估。通过综合音频和图像数据的分析,能够更准确地理解执法人员的行为和执法效果,确保执法过程的透明性和合规性,这种多角度的行为信息识别不仅提升了执法过程的审查能力,还加强了对执法质量的控制和优化。其次,通过对执法现场视频文件进行时间戳抽取处理,能够准确地记录和追踪视频中执法事件的时间节点,其中时间戳为执法过程的事件提供了准确的时间记录,使得后续的事件分析可以根据时间序列进行,还原事件的发生顺序,还可以帮助将视频中的执法行为与其他数据源(如音频、传感器数据)进行同步,对比分析不同模态数据的时间一致性,确保多模态数据之间的精确对齐,从而为后续处理过程的透明度和可靠性提供了基础数据保障。还通过对执法现场视频文件中的每一帧视频图像进行空间坐标位置的标记处理,能够为每一帧图像中的空间坐标添加详细的标记和注释,提供关于执法现场的直观信息,通过空间位置标记,执法人员和分析人员可以清晰地了解图像中各个物体和事件的位置及其在现场的具体意义,这种处理可以帮助识别重要位置,如犯罪现场、证据位置或相关人员的位置,使得事件分析和调查变得更加系统化和高效,空间位置标记还支持生成可视化的空间布局图,便于直观展示现场的空间结构和物体分布。并且,还通过基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,能够结合了时间戳和空间位置信息,将音频、图像等多种数据源进行时空同步,可以提供全面的执法事件视图,从而能够将视频图像中的视觉信息与音频中的声音信息进行有效对齐,使得分析人员能够在时间和空间上综合考虑不同数据源,从而更准确地理解事件的发生和发展。通过时空同步,能够实现不同模态数据之间的关联,识别出事件中的关键行为和互动,支持对复杂执法事件的深入分析和理解,通过多模态数据的综合利用可以提升了执法行为信息的全面性和准确性,有助于从不同角度分析执法行为事件,为后续的处理过程提供更为丰富的依据,能够提高了执法行为事件分析的覆盖面和时效性,从而可以及

时反应并全面覆盖所有的执法现场行为状况。然后,通过对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,从而连接生成事件行为序列连接图,这一过程的主要优点在于通过可视化的方式将行为序列中的匹配关系呈现出来,帮助更好地理解事件的动态过程,该行为序列连接图能够直观地展示事件中各方行为的时间顺序、交互关系和关键节点,使得事件的整个过程变得更加清晰。通过连接图,可以识别出行为的关联性和因果关系,发现事件中的关键事件点和重要行为链条,这种图形化的展示不仅有助于分析人员快速掌握事件全貌,从而为后续的处理过程提供直观的数据支持。还通过对执法现场事件行为序列连接图进行事件行为意图识别分析,旨在解读执法现场事件行为序列连接图中各行为的潜在意图和目的,这一分析的主要优点在于能够深入挖掘和理解行为背后的动机和意图,提供对事件发生原因和背景的深刻见解。通过对行为序列的意图识别,可以揭示执法人员和当事人在事件中的真实目的,判断行为是否符合既定的规范和法律要求。例如,识别执法人员的意图是否是为了合法执行任务,还是存在过度行为;同时,也可以分析当事人是否有防御、逃避或其他意图,这种分析不仅提升了对事件的理解深度,还能帮助评估执法过程中的合规性和合法性,为改进执法行为策略和培训提供数据支持。最后,通过获取预设的执法事件行为知识图谱,这一过程涉及整合和建立一个包含各种执法事件及其行为的系统化知识库,其中包括历史执法行为案例数据、行为规范规则、执法程序等信息,通过构建这样一个知识图谱,可以实现对相应的执法行为事件的深度分析,并为执法人员提供直观的参考模型,这不仅有助于识别和理解各种执法行为的规范性和趋势,从而能够为后续制定更加科学的执法标准和优化执法策略提供依据。通过基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,能够实现对执法行为的标准化评估,通过将行为意图的语义特征与知识图谱内的规范性规则标准进行匹配,可以客观地评估行为的规范程度,并计算出相应的规范性评分值,这一评分能够帮助快速识别出不符合规范的行为,并为执法人员提供明确的改进方向。此外,还通过基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,可以自动检测和识别出不符合规范的行为,并生成针对性的优化建议报告,这一过程不仅能够实时地提醒执法人员注意存在的问题,还能够提供改进的具体措施和建议,优化分析报告将帮助执法人员调整和改进执法行为,确保其符合既定的规范标准,这样将显著提高执法过程的规范性和效率,减少错误和不规范执法行为的发生。

[0062] 2、本发明所提出的基于执法督察用的多模态智能分析系统,整体上由执法音视频行为识别分析模块、执法行为多模态时空关联分析模块、执法事件行为意图识别分析模块以及执法事件不规范行为智能告警模块组成,能够实现本发明所述任意基于执法督察用的多模态智能分析方法,用于联合各个模块上运行的计算机程序之间的操作实现基于执法督察用的多模态智能分析方法,系统内部结构互相协作,这样能够大大减少重复工作和人力投入,能够快速有效地提供更为准确、更高效的基于执法督察用的多模态智能分析过程,从而简化了基于执法督察用的多模态智能分析系统的操作流程。

附图说明

[0063] 通过阅读参照以下附图所作的对非限制性实施所作的详细描述,本发明的其它特征、目的和优点将会变得更明显:

- [0064] 图1为本发明基于执法督察用的多模态智能分析方法的步骤流程示意图；
- [0065] 图2为图1中步骤S1的详细步骤流程示意图；
- [0066] 图3为图2中步骤S15的详细步骤流程示意图。

具体实施方式

[0067] 下面结合附图对本发明的技术方法进行清楚、完整的描述,显然,所描述的实施例是本发明的一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域所属的技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0068] 此外,附图仅为本发明的示意性图解,并非一定是按比例绘制。图中相同的附图标记表示相同或类似的部分,因而将省略对它们的重复描述。附图中所示的一些方框图是功能实体,不一定必须与物理或逻辑上独立的实体相对应。可以采用软件形式来实现功能实体,或在一个或多个硬件模块或集成电路中实现这些功能实体,或在不同网络和/或处理器方法和/或微控制器方法中实现这些功能实体。

[0069] 应当理解的是,虽然在这里可能使用了术语“第一”、“第二”等等来描述各个单元,但是这些单元不应当受这些术语限制。使用这些术语仅仅是为了将一个单元与另一个单元进行区分。举例来说,在不背离示例性实施例的范围的情况下,第一单元可以被称为第二单元,并且类似地第二单元可以被称为第一单元。这里所使用的术语“和/或”包括其中一个或更多所列出的相关联项目的任意和所有组合。

[0070] 为实现上述目的,请参阅图1至图3,本发明提供了一种基于执法督察用的多模态智能分析方法,所述方法包括以下步骤:

[0071] 步骤S1: 通过从执法督察记录仪平台中获取执法现场视频文件,并对执法现场视频文件进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

[0072] 步骤S2: 对执法现场视频文件进行时间戳及空间位置抽取处理,以得到执法现场过程时间戳以及执法现场过程空间位置信息;基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,得到多模态时空同步执法行为关联事件数据;

[0073] 步骤S3: 对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据;

[0074] 步骤S4: 获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0075] 本发明实施例中,请参考图1所示,为本发明基于执法督察用的多模态智能分析方法的步骤流程示意图,在本实例中,所述基于执法督察用的多模态智能分析方法包括以下

步骤:

[0076] 步骤S1: 通过从执法督察记录仪平台中获取执法现场视频文件, 并对执法现场视频文件进行音视频切分处理, 以得到执法现场音频片段以及执法现场图像视频帧片段; 分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析, 得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

[0077] 在本发明实施例中, 通过从执法督察记录仪平台获取相应的执法视频原件, 该平台通常具备存储和管理执法记录的功能, 通过用户认证和权限验证后, 登录到平台的文件管理系统中, 在系统界面上, 选择“视频记录”栏目, 在其中浏览和筛选所需的执法视频原件, 利用平台提供的检索功能, 可以根据日期、时间、案件编号或其他关键词找到特定的视频记录, 从而获取得到执法现场视频文件。并且, 通过确定先前获取得到的执法现场视频文件中每个执法事件音频和视频的开始时间和结束时间, 还通过依据先前识别得到的开始时间以及结束时间对执法现场视频文件内相对应的执法事件进行音视频的切分处理, 以通过使用视频编辑工具(如FFmpeg或Adobe Premiere Pro), 将视频文件导入并根据开始时间以及结束时间设定切割参数来分割图像视频轨道, 从而生成多个图像视频片段, 每个片段对应一个执法事件, 同时, 通过对音频轨道进行相同时间段的切割, 确保每个视频片段都有对应的音频片段, 从而得到执法现场音频片段以及执法现场图像视频帧片段。然后, 通过使用语音识别技术(例如Speechbrain技术)对先前经过切分后得到的执法现场音频片段进行行为信息的识别分析, 以将音频内容转化为文字, 分析语言中的关键词和语境, 并从中提取相应的执法行为信息, 从而得到执法现场音频行为信息数据。同时, 通过使用图像识别算法(例如物体检测算法Yolov8以及跟踪技术Bytetrack)分析执法现场图像视频帧片段内的每帧图像, 以从中识别执法人员以及当事人的动作、对象和场景变化, 并将分析结果整理为结构化数据, 包括每个事件的行为描述、参与人员、关键动作等信息, 这些数据提供了详细的执法行为记录, 最终得到执法现场图像行为信息数据。

[0078] 步骤S2: 对执法现场视频文件进行时间戳及空间位置抽取处理, 以得到执法现场过程时间戳以及执法现场过程空间位置信息; 基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析, 得到多模态时空同步执法行为关联事件数据;

[0079] 在本发明实施例中, 通过使用视频解析工具, 例如OpenCV, 逐帧读取执法现场视频文件, 并通过调用视频流的get方法, 获取每帧的时间信息, 以精确记录每一帧的时间戳, 从而得到执法现场过程时间戳。同时, 通过使用相机标定技术对先前提取得到的执法现场视频文件中的每一帧视频图像进行三维空间坐标系转换, 需进行相机的内参和外参标定, 确保能够准确映射图像坐标与实际世界坐标之间的关系, 并为每帧图像生成相应的三维坐标系, 还通过从先前转换得到的三维空间坐标系中提取出与视频中标识物体相对应的空间位置, 从而得到执法现场过程空间位置信息。然后, 通过使用数据融合工具对先前提取得到的执法现场过程时间戳以及执法现场过程空间位置信息进行时空同步处理, 以将时间戳和空间位置信息对齐, 使得每个执法事件的时间和空间位置能够同步对应, 从而整合构建成一个时空数据同步基准框架, 并通过利用先前构建的时空数据同步基准框架对执法现场音频行为信息数据进行时序上的同步排列处理, 以将音频事件的时间戳与同步框架中的时间戳进行匹配, 确保音频事件在时间上的顺序与现场实际情况一致, 并通过结合先前构建的时

空数据同步基准框架将执法现场图像行为信息数据中的图像行为特征与空间位置信息进行匹配,确保图像数据中的空间信息与实际场景中的位置相符,还通过综合使用数据分析和机器学习工具对先前确定得到的音频事件时序同步数据以及图像行为空间位置匹配数据进行交互的关联分析,以将音频事件与图像行为数据进行融合,从中识别并提取音频事件与图像行为之间的交互关联关系,随后,通过结合先前分析得到的音频事件与图像行为之间的交互关联关系对相对应的执法现场音频行为信息数据以及执法现场图像行为信息数据进行执法行为关联事件的关联连接提取处理,以通过建立行为关联模型,将音频事件和图像行为数据中的交互信息进行结合,并通过使用机器学习算法(如分类器或聚类算法)进行处理,以识别出音频事件和图像行为之间的关联事件,例如特定音频事件(如警报声)与相应的图像行为(如人员移动)之间的执法关联事件,最终得到多模态时空同步执法行为关联事件数据,其中包含了执法人员以及当事人在当前执法事件下的执法行为情况。

[0080] 步骤S3:对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据;

[0081] 在本发明实施例中,通过从先前分析得到的多模态时空同步执法行为关联事件数据中提取涉及的所有执法人员和当事人行为记录,这些记录来自于监控摄像头、传感器、语音记录和其他数据采集设备,其中包括执法人员的接近、询问、检查、处罚等行为以及当事人的停车、解释、配合、反抗等行为,并利用数据标签或事件标识符将行为记录分为执法人员行为序列和当事人行为序列,并通过将先前分析得到的执法人员行为序列内的行为描述节点以及当事人行为序列内的行为描述节点进行比对,每对行为描述节点通过匹配算法计算其相似度分数,确定其匹配关系,例如,通过设置一个阈值,当相似度分数超过该阈值时,认为这两个节点匹配成功,将匹配分析的结果记录到一个匹配关系表中,从而确定每对匹配的执法人员和当事人行为描述节点之间的匹配关系,同时,通过结合先前匹配分析得到的执法人员和当事人行为描述节点之间的匹配关系对执法人员行为序列内相对应的行为描述节点以及当事人行为序列内相对应的行为描述节点之间进行匹配连接处理,以通过使用图论中的图构建技术(如邻接矩阵或邻接表)将匹配的行为节点按照时间顺序连接起来,每个行为节点作为图中的一个节点,相连的行为描述作为边,从而构建生成执法现场事件行为序列连接图。然后,通过构建行为意图模型,该模型基于已知的行为模式和意图模板,可以利用机器学习技术(如分类算法或序列标注模型)训练,以识别不同的行为意图,例如,通过训练一个支持向量机(SVM)分类器来预测每个行为序列的意图类型,并根据先前构建的行为意图模型对执法现场事件行为序列连接图中的每个行为节点和行为序列进行意图识别,以将行为意图模型应用于图中的节点,通过计算每个行为序列的意图概率,确定其对应的行为意图,例如,分析行为序列是否符合特定的执法意图(如警告、劝阻、处罚、逮捕、逃避等),最终得到执法现场事件行为意图信息数据。

[0082] 步骤S4:获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0083] 在本发明实施例中,通过从预设的知识库中提取相应执法事件行为的知识图谱,该图谱包含了多种执法事件的行为模式及其相关的规范性规定,以利用自然语言处理技术将不同来源的执法行为文档、规范性条款和案例数据转换为结构化知识,数据来源可以是法律数据库、执法记录和专家领域知识,通过图谱构建平台的可视化界面,整合并构建行为知识图谱,从而获得得到执法事件行为知识图谱。其次,通过应用规则抽取方法对先前获取得到的执法事件行为知识图谱进行处理,首先,定义行为规范性规则检索的标准,利用规则引擎进行检索,该规则引擎通过解析图谱中节点和边的属性,提取出与行为规范相关的规则,结果生成一份包含详细行为规范标准的规则列表,该列表可以包括执法人员的行为准则、处理程序及相应的规范性依据,并通过利用自然语言处理和机器学习模型对先前分析得到的执法现场事件行为意图信息数据进行行为意图语义特征的统计分析,以实现现场事件的描述文本进行分词、词性标注和句法分析,并通过语义分析工具提取行为意图的特征数据,例如执法行为意图的目的、动机及情境等语义特征,并通过结合先前检索抽取得到的行为规范性规则标准对相对应的执法现场事件行为意图语义特征数据进行规范性规则的适配,以采用匹配算法如余弦相似度或欧氏距离计算行为意图特征向量与规范性规则的相似度,并通过使用规范性规则适配器根据计算得到的相似度评分生成规范性评分,此评分反映了现场行为意图与规范性规则的符合程度,从而得到执法现场行为意图规范性评分值。然后,通过使用先前量化计算得到的执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行分析,以通过设定阈值,将评分值低于阈值的行为标记为不规范行为,并应用智能告警方法对这些不规范行为进行进一步分析,使用异常检测算法识别潜在的问题区域,并对其进行优化分析包括生成详细的优化建议报告,报告内容包括识别出的不规范行为类型、存在的规范性风险以及改进建议,该报告利用数据可视化工具展示分析结果和建议,以帮助执法人员调整其执法行为,确保执法过程的规范化,最终生成执法现场事件不规范行为优化建议报告。

[0084] 进一步的,步骤S1包括以下步骤:

[0085] 步骤S11:通过从执法督察记录仪平台中获取执法现场视频文件;

[0086] 步骤S12:对执法现场视频文件进行音画同步处理,得到执法现场音画同步视频;

[0087] 步骤S13:对执法现场音画同步视频进行视频内容提取处理,得到执法现场视频内容信息数据;基于执法现场视频内容信息数据对执法现场音画同步视频进行不同执法事件开始及结束时间识别分析,以得到执法现场视频内不同执法事件的开始时间以及结束时间;

[0088] 步骤S14:根据执法现场视频内不同执法事件的开始时间以及结束时间对执法现场音画同步视频进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;

[0089] 步骤S15:分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据。

[0090] 作为本发明的一个实施例,参考图2所示,为图1中步骤S1的详细步骤流程示意图,在本实施例中步骤S1包括以下步骤:

[0091] 步骤S11:通过从执法督察记录仪平台中获取执法现场视频文件;

[0092] 在本发明实施例中,通过从执法督察记录仪平台获取相应的执法视频原件,该平

台通常具备存储和管理执法记录的功能,通过用户认证和权限验证后,登录到平台的文件管理系统中,在系统界面上,选择“视频记录”栏目,在其中浏览和筛选所需的执法视频原件,利用平台提供的检索功能,可以根据日期、时间、案件编号或其他关键词找到特定的视频记录,选定目标视频后,点击“下载”按钮,将视频文件保存到本地存储设备中,视频文件通常采用标准的视频格式如MP4或AVI,并包含执法现场的完整录制数据,需确保下载的文件完整且未损坏,最终获取得到执法现场视频文件。

[0093] 步骤S12:对执法现场视频文件进行音画同步处理,得到执法现场音画同步视频;

[0094] 在本发明实施例中,通过将先前获取得到的执法现场视频文件中的音频和视频进行同步处理,以导入从执法督察记录仪平台下载的视频文件到视频处理软件中,选择支持音视频同步的工具,如Adobe Premiere Pro或FFmpeg,在软件中,分离视频和音频轨道,以检查是否存在时间偏差,若有偏差,通过调整音频轨道的起始时间来实现同步,具体操作包括,使用软件的时间轴工具,将音频轨道与视频轨道对齐,确保在播放时音画一致,同步完成后,将调整后的音频和视频合并为一个文件,保存为新的视频文件格式,最终得到执法现场音画同步视频。

[0095] 步骤S13:对执法现场音画同步视频进行视频内容提取处理,得到执法现场视频内容信息数据;基于执法现场视频内容信息数据对执法现场音画同步视频进行不同执法事件开始及结束时间识别分析,以得到执法现场视频内不同执法事件的开始时间以及结束时间;

[0096] 在本发明实施例中,通过对先前经过音画同步后得到的执法现场音画同步视频进行内容提取处理,以通过使用视频分析软件,如OpenCV或自定义的视频分析工具,将视频逐帧解析,提取执法视频内的具体内容信息,从而得到执法现场视频内容信息数据。同时,通过结合先前分析得到的执法现场视频内容信息数据对执法现场音画同步视频进行不同执法事件的识别分类处理,以获取执法现场视频内所含有的执法事件,并根据不同执法事件的视觉特征和音频中的语音内容,标记并记录每个执法事件的时间戳,从中确定每个执法事件的开始时间和结束时间,最终得到执法现场视频内不同执法事件的开始时间以及结束时间。

[0097] 步骤S14:根据执法现场视频内不同执法事件的开始时间以及结束时间对执法现场音画同步视频进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;

[0098] 在本发明实施例中,通过依据先前识别得到的执法现场视频内不同执法事件的开始时间以及结束时间对执法现场音画同步视频内相对应的执法事件进行音视频的切分处理,以通过使用视频编辑工具(如FFmpeg或Adobe Premiere Pro),将视频文件导入并根据开始时间以及结束时间设定切割参数来分割视频轨道,从而生成多个视频片段,每个片段对应一个执法事件,同时,通过对音频轨道进行相同时间段的切割,确保每个视频片段都有对应的音频片段,从而生成音频片段和视频帧片段,每个片段都独立保存为不同的文件,最终得到执法现场音频片段以及执法现场图像视频帧片段。

[0099] 步骤S15:分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,得到执法现场音频行为信息数据以及执法现场图像行为信息数据。

[0100] 在本发明实施例中,通过使用语音识别技术(例如Speechbrain技术)对先前经过

切分后得到的执法现场音频片段进行行为信息的识别分析,以将音频内容转化为文字,分析语言中的关键词和语境,并从中提取相应的执法行为信息,从而得到执法现场音频行为信息数据。同时,通过使用图像识别算法(例如物体检测算法Yolov8以及跟踪技术Bytetrack)分析执法现场图像视频帧片段内的每帧图像,以从中识别执法人员以及当事人的动作、对象和场景变化,并将分析结果整理为结构化数据,包括每个事件的行为描述、参与人员、关键动作等信息,这些数据提供了详细的执法行为记录,最终得到执法现场图像行为信息数据。

[0101] 进一步的,步骤S15包括以下步骤:

[0102] 步骤S151:利用Speechbrain技术对执法现场音频片段进行执法音频声纹特征分析,得到执法现场音频声纹特征数据;

[0103] 步骤S152:基于执法现场音频声纹特征数据对执法现场音频片段内的不同说话者进行音频语音信号通道解构处理,得到执法现场音频内不同说话者的音频语音信号子通道;

[0104] 步骤S153:对执法现场音频内不同说话者的音频语音信号子通道进行执法行为事件触发识别分析,得到执法现场音频执法行为事件触发点;

[0105] 步骤S154:基于执法现场音频执法行为事件触发点对执法现场音频片段进行执法音频行为信息识别分析,得到执法现场音频行为信息数据;

[0106] 步骤S155:利用物体检测算法Yolov8以及跟踪技术Bytetrack对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

[0107] 作为本发明的一个实施例,参考图3所示,为图2中步骤S15的详细步骤流程示意图,在本实施例中步骤S15包括以下步骤:

[0108] 步骤S151:利用Speechbrain技术对执法现场音频片段进行执法音频声纹特征分析,得到执法现场音频声纹特征数据;

[0109] 在本发明实施例中,通过使用Speechbrain技术对先前经过切片后得到的执法现场音频片段进行声纹特征分析,以将执法现场音频片段输入到Speechbrain框架中,该框架提供了丰富的音频处理功能,并通过使用其预训练的声纹识别模型,可以提取出音频片段中的声纹特征,这些声纹特征是基于音频信号中的频谱特征和语音样本的特定模式,进行特征向量化处理得到的,具体实施中,首先对音频信号进行预处理,包括去噪、归一化等步骤,以确保输入数据的质量,并利用声纹识别模型提取出每个说话者的特征向量,这些向量表示了说话者的声纹特征,最终得到执法现场音频声纹特征数据。

[0110] 步骤S152:基于执法现场音频声纹特征数据对执法现场音频片段内的不同说话者进行音频语音信号通道解构处理,得到执法现场音频内不同说话者的音频语音信号子通道;

[0111] 在本发明实施例中,通过结合先前分析得到的执法现场音频声纹特征数据对执法现场音频片段中的不同说话者进行信号通道的解构处理,以将声纹特征数据应用于音频信号的分离算法,如基于声纹特征的盲源分离(BSS)技术,并通过对音频信号进行盲源分离,能够将混合的音频信号解构为多个子通道,每个子通道对应一个独立的说话者的音频信号,具体地,算法会依据声纹特征将音频信号中的各个说话者的语音成分提取出来,并将这些成分分配到不同的音频语音信号子通道中,最终得到执法现场音频内不同说话者的音频

语音信号子通道。

[0112] 步骤S153:对执法现场音频内不同说话者的音频语音信号子通道进行执法行为事件触发识别分析,得到执法现场音频执法行为事件触发点;

[0113] 在本发明实施例中,通过对先前解构后得到的执法现场音频内不同说话者的音频语音信号子通道进行执法行为事件触发情况的识别分析,以利用音频事件检测算法对每个子通道中的音频信号进行处理,这些算法可以基于特定的声学模型,识别出音频信号中的关键事件,如口令、呼叫、冲突声等,具体实施时,首先训练一个事件识别模型,使其能够识别特定的行为事件声音特征,并将每个音频语音信号子通道输入到训练好的模型中,模型会输出每个事件的触发时间点及其相关信息,这些触发点被记录下来,用于分析执法现场中的具体行为事件,最终得到执法现场音频执法行为事件触发点。

[0114] 步骤S154:基于执法现场音频执法行为事件触发点对执法现场音频片段进行执法音频行为信息识别分析,得到执法现场音频行为信息数据;

[0115] 在本发明实施例中,通过结合先前识别得到的执法现场音频执法行为事件触发点对相对应的执法现场音频片段进行行为信息的识别分析,以将音频片段按照触发点进行分段,每段音频围绕一个事件进行分析,并通过使用行为信息识别模型对每段音频进行处理,该模型能够识别和分类与特定执法行为相关的音频特征,如声音强度、音调变化等,具体实施中,首先对每个事件片段进行特征提取,包括声音的频率特征、时长特征等,并将这些特征输入到分类器中,分类器将这些特征与预定义的执法行为类别进行匹配,输出每个事件的具体行为信息数据,最终得到执法现场音频行为信息数据。

[0116] 步骤S155:利用物体检测算法Yolov8以及跟踪技术Bytetrack对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

[0117] 在本发明实施例中,通过使用物体检测算法Yolov8以及跟踪技术Bytetrack对先前经过切片后得到的执法现场图像视频帧片段进行图像内执法行为的识别分析,以通过使用Yolov8算法对图像视频帧中的物体进行检测,Yolov8作为一种先进的目标检测算法,能够高效地识别图像中的各类物体,并生成相应的边界框,具体实施中,将每一帧图像输入到Yolov8模型中,模型会输出检测到的物体及其位置信息,并通过应用Bytetrack跟踪技术对这些物体进行跟踪,确保对每个目标在多个帧中的移动进行持续跟踪,Bytetrack结合了目标检测和跟踪技术,通过将物体在连续帧中的位置和运动轨迹进行关联,生成每个物体目标的运动行为信息,其中包括物体的行为类别、时间戳、位置及其他相关的上下文信息,最终得到执法现场图像行为信息数据。

[0118] 进一步的,步骤S155包括以下步骤:

[0119] 利用物体检测算法Yolov8对执法现场图像视频帧片段进行执法现场物体目标识别分析,得到执法现场图像帧物体检测目标;

[0120] 在本发明实施例中,通过使用物体检测算法YOL0v8(You Only Look Once version 8)对执法现场图像视频帧片段中每一帧图像的物体目标进行识别分析,其中YOL0v8是一种深度学习模型,经过训练可以识别图像中的各种物体并进行分类,在此过程中,算法会利用卷积神经网络对图像进行特征提取,生成边界框(bounding box)来标记图像中每个检测到的物体,每个边界框将包含物体的位置坐标以及物体的类别标签,最终得到执法现场图像帧物体检测目标。

[0121] 优选地,对执法现场图像帧物体检测目标进行物体边界区域标记处理,得到执法现场物体边界区域标记目标;

[0122] 在本发明实施例中,通过对先前检测得到的执法现场图像帧物体检测目标进行进一步处理,主要是进行物体边界区域的标记,这一过程涉及将YOLOv8输出的边界框坐标转换为实际图像中的区域标记,具体实施过程中,首先将边界框坐标映射到图像的像素坐标系中,然后,通过绘制边界框和附加标签,在图像中明确标记每个物体的边界区域,边界区域标记的结果将以图像叠加形式呈现,使得每个物体在图像中具有清晰的可视化边界,最终得到执法现场物体边界区域标记目标。

[0123] 优选地,利用跟踪技术Bytetrack对执法现场物体边界区域标记目标在执法现场图像视频帧片段内的运动轨迹进行跟踪识别分析,以生成执法现场物体标记目标运动轨迹;

[0124] 在本发明实施例中,通过使用跟踪技术Bytetrack对先前确定得到的执法现场物体边界区域标记目标在执法现场图像视频帧片段内的运动轨迹进行跟踪的识别分析,其中Bytetrack是一个高效的多目标跟踪算法,能够处理复杂场景中的目标跟踪问题,以将每个物体的初始边界框作为跟踪目标,通过对视频帧进行连续分析,Bytetrack算法会计算物体在每帧之间的运动轨迹,并通过结合物体的外观特征和运动模式来更新每个目标的位置,确保准确追踪物体的移动路径,并将以连续的坐标点和时间戳形式记录下来,最终生成执法现场物体标记目标运动轨迹。

[0125] 优选地,对执法现场物体标记目标运动轨迹进行物体目标行为模式分析,得到不同执法现场物体目标的执法行为模式;基于不同执法现场物体目标的执法行为模式对相对应的执法现场物体边界区域标记目标进行执法行为动作分类处理,以得到不同执法现场物体目标的执法行为动作;

[0126] 在本发明实施例中,通过使用机器学习模型或统计方法对先前跟踪生成的执法现场物体标记目标运动轨迹进行物体行为模式的分析,以通过对物体的运动轨迹数据进行统计和分析,从中提取出物体在执法现场中的运动模式,这包括分析物体的移动速度、方向变化以及轨迹的连续性等,例如移动路径的规律、停留时间及其与其他物体的交互模式,从而得到不同执法现场物体目标的执法行为模式。同时,通过结合先前分析得到的不同执法现场物体目标的执法行为模式对相对应执法现场物体边界区域标记目标的执法动作行为进行分类处理,以基于行为模式分析结果,建立分类模型对物体在视频帧中的行为动作进行分类,分类过程涉及对每个目标的运动轨迹和行为模式进行比对,将其分配到预定义的行为类别中,例如“巡逻”、“静止”或“接触”等,具体实施时,采用特征提取方法从轨迹数据中提取关键行为特征,并利用分类算法(如支持向量机或神经网络)进行处理,从而输出每个物体目标的行为动作类别,最终得到不同执法现场物体目标的执法行为动作。

[0127] 优选地,基于不同执法现场物体目标的执法行为模式以及执法行为动作对执法现场图像视频帧片段进行执法图像行为信息识别分析,得到执法现场图像行为信息数据。

[0128] 在本发明实施例中,通过结合先前分析得到的不同执法现场物体目标的执法行为模式以及执法行为动作对整个执法现场的图像视频帧片段进行全面的执法图像行为信息识别分析,以通过将物体的行为分类信息整合到视频帧数据中,使用信息融合技术将物体的动作信息与图像数据相结合,生成详细的行为信息数据,具体实施中,将每个物体的行为

分类结果与视频帧进行对齐,分析整个场景中物体的行为模式及其相互作用,生成的行为信息数据将包括物体的行为类别、时间戳、位置及其他相关的上下文信息,最终得到执法现场图像行为信息数据。

[0129] 进一步的,步骤S2包括以下步骤:

[0130] 步骤S21:对执法现场视频文件进行时间戳抽取处理,以得到执法现场过程时间戳;

[0131] 在本发明实施例中,通过使用视频解析工具,例如OpenCV,逐帧读取执法现场视频文件,并通过调用视频流的get方法,获取每帧的时间信息,以精确记录每一帧的时间戳,在实际实现过程中,需将执法现场视频文件转换为帧序列,并使用系统时钟记录每一帧的播放时间,最终得到执法现场过程时间戳。

[0132] 步骤S22:对执法现场视频文件中的每一帧视频图像进行三维空间坐标系转换,得到执法现场视频中每一帧图像的三维空间坐标系;

[0133] 在本发明实施例中,通过使用相机标定技术对先前提取得的执法现场视频文件中的每一帧视频图像进行三维空间坐标系转换,需进行相机的内参和外参标定,确保能够准确映射图像坐标与实际世界坐标之间的关系,并通过已知的相机参数,应用透视变换,将图像的像素坐标转换为三维空间坐标系中的坐标,这可以通过使用计算机视觉库中的函数,例如cv2.projectPoints,完成该转换,并为每帧图像生成相应的三维坐标系,最终得到执法现场视频中每一帧图像的三维空间坐标系。

[0134] 步骤S23:对执法现场视频中每一帧图像的三维空间坐标系进行执法现场空间坐标提取处理,以得到执法现场视频中每一帧图像的执法现场空间坐标;

[0135] 在本发明实施例中,通过从先前转换得到的执法现场视频中每一帧图像的三维空间坐标系中提取出与视频中标识物体相对应的空间位置,从而提取与执法行为相关的关键点坐标,确保提取的坐标准确反映执法现场的真实情况,最终得到执法现场视频中每一帧图像的执法现场空间坐标。

[0136] 步骤S24:对执法现场视频中每一帧图像的执法现场空间坐标进行空间位置标记处理,以得到执法现场过程空间位置信息;

[0137] 在本发明实施例中,通过对先前提取得的执法现场视频中每一帧图像的执法现场空间坐标进行空间位置的标记处理,以通过定义不同的标记规则,对每一帧图像中的空间坐标进行标记,例如,采用图形处理工具,将空间坐标转化为可视化的点云模型,并对不同的执法行为进行色彩编码或形状标记,最终标记得到执法现场过程空间位置信息。

[0138] 步骤S25:基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,得到多模态时空同步执法行为关联事件数据。

[0139] 在本发明实施例中,通过使用数据融合工具对先前提取得的执法现场过程时间戳以及执法现场过程空间位置信息进行时空同步处理,以将时间戳和空间位置信息对齐,使得每个执法事件的时间和空间位置能够同步对应,并将时间戳与空间位置信息整合构建成一个时空数据同步基准框架,同时,通过利用先前构建的时空数据同步基准框架对执法现场音频行为信息数据进行时序上的同步排列处理,以将音频事件的时间戳与同步框架中的时间戳进行匹配,确保音频事件在时间上的顺序与现场实际情况一致,并通过结合先前

构建的时空数据同步基准框架将执法现场图像行为信息数据中的图像行为特征与空间位置信息进行匹配,确保图像数据中的空间信息与实际场景中的位置相符,还通过综合使用数据分析和机器学习工具对先前确定得到的音频事件时序同步数据以及图像行为空间位置匹配数据进行交互的关联分析,以将音频事件与图像行为数据进行融合,建立一个包含时间和空间信息的多模态数据集,并应用多模态融合算法,如融合神经网络(Multi-modal Fusion Neural Networks),对音频事件和图像行为之间的关系进行分析,从中识别并提取音频事件与图像行为之间的交互关联关系,然后,通过结合先前分析得到的音频事件与图像行为之间的交互关联关系对相对应的执法现场音频行为信息数据以及执法现场图像行为信息数据进行执法行为关联事件的关联连接提取处理,以通过建立行为关联模型,将音频事件和图像行为数据中的交互信息进行结合,并通过使用机器学习算法(如分类器或聚类算法)进行处理,以识别出音频事件和图像行为之间的关联事件,例如特定音频事件(如警报声)与相应的图像行为(如人员移动)之间的执法关联事件,最终得到多模态时空同步执法行为关联事件数据,其中包含了执法人员以及当事人在当前执法事件下的执法行为情况。

[0140] 进一步的,步骤S25包括以下步骤:

[0141] 步骤S251:对执法现场过程时间戳以及执法现场过程空间位置信息进行时空同步基准构建,以得到执法现场过程时空数据同步基准框架;

[0142] 在本发明实施例中,通过使用数据融合工具对先前提取得到的执法现场过程时间戳以及执法现场过程空间位置信息进行时空同步处理,以将时间戳和空间位置信息对齐,使得每个执法事件的时间和空间位置能够同步对应,具体操作包括使用数据对齐算法,如动态时间规整(DTW)或插值算法,将时间戳数据与空间位置数据进行匹配,并构建一个时空数据同步基准框架,将时间戳与空间位置信息整合成一个时空数据同步模型,以确保后续分析中的数据一致性,最终得到执法现场过程时空数据同步基准框架。

[0143] 步骤S252:基于执法现场过程时空数据同步基准框架对执法现场音频行为信息数据进行音频事件时序同步排列处理,得到执法现场音频事件时序同步序列数据;

[0144] 在本发明实施例中,通过利用先前构建的执法现场过程时空数据同步基准框架对执法现场音频行为信息数据进行时序上的同步排列处理,以将音频事件行为数据从记录设备中提取,并通过音频处理软件(如Audacity)将其转化为时间序列数据,并基于执法现场过程时空数据同步基准框架,将音频事件的时间戳与同步框架中的时间戳进行匹配,确保音频事件在时间上的顺序与现场实际情况一致,应用音频事件检测算法,识别并标记音频中的关键事件,例如喊话、警告声或干预声,同时将这些音频事件按照时间顺序排列,最终得到执法现场音频事件时序同步序列数据。

[0145] 步骤S253:基于执法现场过程时空数据同步基准框架对执法现场图像行为信息数据进行图像行为空间映射处理,得到执法现场图像行为空间位置映射数据;

[0146] 在本发明实施例中,通过结合先前构建的执法现场过程时空数据同步基准框架,将执法现场图像行为信息数据中的图像行为特征与空间位置信息进行匹配,确保图像数据中的空间信息与实际场景中的位置相符,具体操作包括图像坐标系统与空间位置坐标系统的转换,以及图像特征的空间映射,使用空间映射算法(如特征点匹配算法),将图像中的行为位置与执法现场过程时空数据同步基准框架中的空间位置数据进行对齐,最终得到执法

现场图像行为空间位置映射数据。

[0147] 步骤S254:对执法现场音频事件时序同步序列数据以及执法现场图像行为空间位置映射数据进行多模态事件行为交互关联分析,得到执法现场音频事件与图像行为之间的多模态事件行为交互关联关系;

[0148] 在本发明实施例中,通过综合使用数据分析和机器学习工具对先前确定得到的执法现场音频事件时序同步序列数据以及执法现场图像行为空间位置映射数据进行交互的关联分析,以将音频事件与图像行为数据进行融合,建立一个包含时间和空间信息的多模态数据集,并应用多模态融合算法,如融合神经网络(Multi-modal Fusion Neural Networks),对音频事件和图像行为之间的关系进行分析,从中识别并提取音频事件与图像行为之间的相关性,例如声音和图像行为是否在相同的时间段内发生,生成多模态事件行为交互关联关系的数据报告,来展示音频事件和图像行为数据的交互关联关系,最终得到执法现场音频事件与图像行为之间的多模态事件行为交互关联关系。

[0149] 步骤S255:基于执法现场音频事件与图像行为之间的多模态事件行为交互关联关系对执法现场音频行为信息数据以及执法现场图像行为信息数据进行执法行为关联事件提取处理,得到多模态时空同步执法行为关联事件数据。

[0150] 在本发明实施例中,通过结合先前分析得到的执法现场音频事件与图像行为之间的多模态事件行为交互关联关系对相对应的执法现场音频行为信息数据以及执法现场图像行为信息数据进行执法行为关联事件的关联连接提取处理,以通过建立行为关联模型,将音频事件和图像行为数据中的交互信息进行结合,并通过使用机器学习算法(如分类器或聚类算法)进行处理,以识别出音频事件和图像行为之间的关联事件状况,例如特定音频事件(如警报声)与相应的图像行为(如人员移动)之间的执法关联事件,最终得到多模态时空同步执法行为关联事件数据,其中包含了执法人员以及当事人在当前执法事件下的执法行为情况。

[0151] 进一步的,步骤S3包括以下步骤:

[0152] 步骤S31:对多模态时空同步执法行为关联事件数据进行执法人员及当事人行为序列抽取处理,得到执法行为事件执法人员行为序列以及执法行为事件当事人行为序列;

[0153] 在本发明实施例中,通过从先前分析得到的多模态时空同步执法行为关联事件数据中提取涉及的所有执法人员和当事人行为记录,这些记录来自于监控摄像头、传感器、语音记录和其他数据采集设备,其中包括执法人员的接近、询问、检查、处罚等行为以及当事人的停车、解释、配合、反抗等行为,并利用数据标签或事件标识符将行为记录分为执法人员行为和当事人行为,可以通过正则表达式匹配或规则引擎实现,将具有特定标签的记录归类到相应的序列中,例如,执法人员行为可以标记为“OfficerAction”,而当事人行为标记为“SuspectAction”,并将每条行为记录分解为具体的行为描述节点,每个节点应包含行为的时间戳、动作类型和相关的上下文信息,同时,根据时间戳对行为描述节点进行排序,生成有序的行为序列,确保每个行为序列准确地反映了执法人员和当事人的行为动态,最终得到执法行为事件执法人员行为序列以及执法行为事件当事人行为序列。

[0154] 步骤S32:对执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点之间进行行为描述匹配分析,以得到执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系;

[0155] 在本发明实施例中,通过将先前分析得到的执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点进行标准化处理,可以使用自然语言处理技术(如词嵌入或文本相似度计算)将不同的行为描述转换为统一的语义表示形式,并通过构建行为描述的向量表示,同时,通过使用匹配算法(如余弦相似度计算、动态时间规整算法或基于模板的匹配方法)对执法人员行为描述节点与当事人行为描述节点进行比对,每对行为描述节点通过算法计算其相似度分数,确定其匹配关系,例如,通过设置一个阈值,当相似度分数超过该阈值时,认为这两个节点匹配成功,将匹配分析的结果记录到一个匹配关系表中,从而确定每对匹配的执法人员和当事人行为描述节点之间的匹配关系,最终得到执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系。

[0156] 步骤S33:基于执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系对执法行为事件执法人员行为序列内的行为描述节点以及执法行为事件当事人行为序列内的行为描述节点之间进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;

[0157] 在本发明实施例中,通过结合先前匹配分析得到的执法人员行为描述节点与当事人行为描述节点之间的执法事件匹配关系对执法行为事件执法人员行为序列内相对应的行为描述节点以及执法行为事件当事人行为序列内相对应的行为描述节点之间进行匹配连接处理,以通过使用图论中的图构建技术(如邻接矩阵或邻接表)将匹配的行为节点按照时间顺序连接起来,每个行为节点作为图中的一个节点,相连的行为描述作为边,从而构建行为序列连接图,展示执法现场的事件发展过程,并确保每对匹配节点能够正确地连接在一起,形成完整的行为序列,最终连接生成执法现场事件行为序列连接图。

[0158] 步骤S34:对执法现场事件行为序列连接图进行事件行为意图识别分析,得到执法现场事件行为意图信息数据。

[0159] 在本发明实施例中,通过构建行为意图模型,该模型基于已知的行为模式和意图模板,可以利用机器学习技术(如分类算法或序列标注模型)训练,以识别不同的行为意图,例如,通过训练一个支持向量机(SVM)分类器来预测每个行为序列的意图类型,并根据先前构建的行为意图模型对执法现场事件行为序列连接图中的每个行为节点和行为序列进行意图识别,以将行为意图模型应用于图中的节点,通过计算每个行为序列的意图概率,确定其对应的行为意图,例如,分析行为序列是否符合特定的执法意图(如警告、劝阻、处罚、逮捕、逃避等),最终得到执法现场事件行为意图信息数据。

[0160] 进一步的,步骤S4包括以下步骤:

[0161] 步骤S41:获取预设的执法事件行为知识图谱;

[0162] 在本发明实施例中,通过从预设的知识库中提取相应执法事件行为的知识图谱,该图谱包含了多种执法事件的行为模式及其相关的规范性规定,以利用自然语言处理技术将不同来源的执法行为文档、规范性条款和案例数据转换为结构化知识,数据来源可以是法律数据库、执法记录和专家领域知识,通过图谱构建平台的可视化界面,整合并构建行为知识图谱,形成一个包含执法行为节点及其关系的知识图谱,最终获取得到执法事件行为知识图谱。

[0163] 步骤S42:对执法事件行为知识图谱进行行为规范性规则检索抽取处理,以得到执

法事件行为规范性规则标准；

[0164] 在本发明实施例中,通过应用规则抽取方法对先前获取到的执法事件行为知识图谱进行处理,首先,定义行为规范性规则检索的标准,利用规则引擎进行检索,该规则引擎通过解析图谱中节点和边的属性,提取出与行为规范相关的规则,而在检索过程中,将图谱中的行为模式与规范性条款进行比对,从中抽取符合规范的行为规则,结果是生成一份包含详细行为规范标准的规则列表,该列表可以包括执法人员的行为准则、处理程序及相应的规范性依据,最终得到执法事件行为规范性规则标准。

[0165] 步骤S43:对执法现场事件行为意图信息数据进行行为意图语义特征分析,得到执法现场事件行为意图语义特征数据；

[0166] 在本发明实施例中,通过利用自然语言处理和机器学习模型对先前分析得到的执法现场事件行为意图信息数据进行行为意图语义特征的统计分析,以实现到现场事件的描述文本进行分词、词性标注和句法分析,并通过语义分析工具提取行为意图的特征数据,例如执法行为意图的目的、动机及情境等,而机器学习模型可以使用预训练的语言模型(如BERT或GPT)进行特征提取,并通过特征选择算法筛选出最具代表性的语义特征,最终得到执法现场事件行为意图语义特征数据。

[0167] 步骤S44:基于执法事件行为规范性规则标准对执法现场事件行为意图语义特征数据进行规范性规则适配及行为规范性评分计算,以得到执法现场行为意图规范性评分值；

[0168] 在本发明实施例中,通过结合先前检索抽取得到的执法事件行为规范性规则标准对相对应的执法现场事件行为意图语义特征数据进行规范性规则的适配,以采用匹配算法如余弦相似度或欧氏距离计算行为意图特征向量与规范性规则的相似度,并通过使用规范性规则适配器根据计算得到的相似度评分生成规范性评分,此评分反映了现场行为意图与规范性规则的符合程度,计算过程涉及对每个规则的权重和匹配度的加权,从而量化计算得到的行为规范性评分值用以评价现场行为意图的规范程度,最终得到执法现场行为意图规范性评分值。

[0169] 步骤S45:基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0170] 在本发明实施例中,通过使用先前量化计算得到的执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行分析,以通过设定阈值,将评分值低于阈值的行为标记为不规范行为,并应用智能告警方法对这些不规范行为进行进一步分析,使用异常检测算法识别潜在的问题区域,并对其进行优化分析包括生成详细的优化建议报告,报告内容包括识别出的不规范行为类型、存在的规范性风险以及改进建议,该报告利用数据可视化工具展示分析结果和建议,以帮助执法人员调整其执法行为,确保执法过程的规范化,最终生成执法现场事件不规范行为优化建议报告。

[0171] 进一步的,步骤S45包括以下步骤:

[0172] 步骤S451:根据预设的规范性评分阈值对执法现场行为意图规范性评分值进行比较判断,当执法现场行为意图规范性评分值大于或等于预设的规范性评分阈值时,则将其执法现场事件行为意图信息数据中相对应的执法现场事件判定为规范行为事件;当执法现

场行为意图规范性评分值小于预设的规范性评分阈值时,则将其执法现场事件行为意图信息数据中相对应的执法现场事件判定为不规范行为事件;

[0173] 在本发明实施例中,通过根据预先设置的规范性评分阈值(其是根据规范行为的最低要求设定的,例如设定为80分)对先前量化计算得到的执法现场行为意图规范性评分值进行比较判断,如果执法现场行为意图规范性评分值大于或者等于这个规范性评分阈值80分,说明其相对应的执法现场行为事件较为规范化,则将该执法现场事件行为意图信息数据中相对应的执法现场事件判定为规范行为事件;而如果执法现场行为意图规范性评分值小于这个规范性评分阈值80分,说明其相对应的执法现场行为事件不是规范的,则将该执法现场事件行为意图信息数据中相对应的执法现场事件判定为不规范行为事件。

[0174] 步骤S452:根据预设的执法行为不规范告警规则对被判定为不规范行为事件进行不规范行为智能告警处理,以生成本次监督不规范行为事件告警信息;

[0175] 在本发明实施例中,通过从先前判断结果中提取所有被判定为不规范行为事件的相关记录,这些记录包括执法行为事件的详细信息,如执法行为事件时间、地点、规范性评分值等,并根据预设的执法行为不规范告警规则对提取的不规范行为事件进行处理,其执法行为不规范告警规则包括如行为严重性、重复违规次数等指标,会分析这些数据,并响应生成相应的告警信号,同时将其告警信号转化为标准化的告警信息,其中包括执法行为事件的详细情况、告警等级、建议的处理措施等,例如,如果某个事件规范性评分值为75分且涉及重复违规行为,则会生成一个高等级的告警信息,最终生成本次监督不规范行为事件告警信息,这些信息包含了本次经过监督生成不规范行为事件的详细描述和告警等级。

[0176] 步骤S453:对本次监督不规范行为事件告警信息进行执法行为问题识别分析,得到本次监督不规范执法行为问题报告;

[0177] 在本发明实施例中,通过对先前告警生成的本次监督不规范行为事件告警信息进行执法行为问题的详细分析,会对告警事件的频率、类型、涉及的执法人员等因素进行统计,以识别潜在的问题,从而从中识别出频繁出现的问题或系统性的违规行为问题,并将其识别到的问题整理成相应的问题报告,其中包括不规范行为问题的类型、涉及的执法事件、频率及其潜在原因,最终得到本次监督不规范执法行为问题报告。

[0178] 步骤S454:对本次监督不规范执法行为问题报告进行执法行为优化建议处理,以生成执法现场事件不规范行为优化建议报告。

[0179] 在本发明实施例中,通过从先前分析得到的本次监督不规范执法行为问题报告中提取出具体的不规范执法行为问题,例如某类当事人违规行为的处理流程不当或执法人员培训不足,通过使用流程优化工具(如业务流程管理软件或流程建模工具)对现有执法流程进行建模,并根据识别出的问题提出改进建议,其中包括重新设计流程步骤、增加培训环节或调整执法资源分配,从而编写详细的优化建议报告,报告内容包括问题描述、优化建议、预期效果及实施步骤,将报告提交至管理层,作为改进执法行为的参考依据,最终生成执法现场事件不规范行为优化建议报告。

[0180] 进一步的,本发明还提供了一种基于执法督察用的多模态智能分析系统,用于执行如上所述的基于执法督察用的多模态智能分析方法,该基于执法督察用的多模态智能分析系统包括:

[0181] 执法音视频行为识别分析模块,用于通过从执法督察记录仪平台中获取执法现场

视频文件,并对执法现场视频文件进行音视频切分处理,以得到执法现场音频片段以及执法现场图像视频帧片段;分别对执法现场音频片段以及执法现场图像视频帧片段进行执法行为信息识别分析,从而得到执法现场音频行为信息数据以及执法现场图像行为信息数据;

[0182] 执法行为多模态时空关联分析模块,用于对执法现场视频文件进行时间戳及空间位置抽取处理,以得到执法现场过程时间戳以及执法现场过程空间位置信息;基于执法现场过程时间戳以及执法现场过程空间位置信息对执法现场音频行为信息数据以及执法现场图像行为信息数据进行多模态时空同步关联分析,从而得到多模态时空同步执法行为关联事件数据;

[0183] 执法事件行为意图识别分析模块,用于对多模态时空同步执法行为关联事件数据进行执法行为序列连接处理,以生成执法现场事件行为序列连接图;对执法现场事件行为序列连接图进行事件行为意图识别分析,从而得到执法现场事件行为意图信息数据;

[0184] 执法事件不规范行为智能告警模块,用于获取预设的执法事件行为知识图谱,并基于执法事件行为知识图谱对执法现场事件行为意图信息数据进行行为规范性评分计算,以得到执法现场行为意图规范性评分值;基于执法现场行为意图规范性评分值对执法现场事件行为意图信息数据中相对应的执法现场事件进行不规范行为智能告警优化分析,以生成执法现场事件不规范行为优化建议报告。

[0185] 因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在申请文件的等同要件的含义和范围内的所有变化涵括在本发明内。

[0186] 以上所述仅是本发明的具体实施方式,使本领域技术人员能够理解或实现本发明。对这些实施例的多种修改对本领域的技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是要符合与本文所发明的原理和新颖特点相一致的最宽的范围。

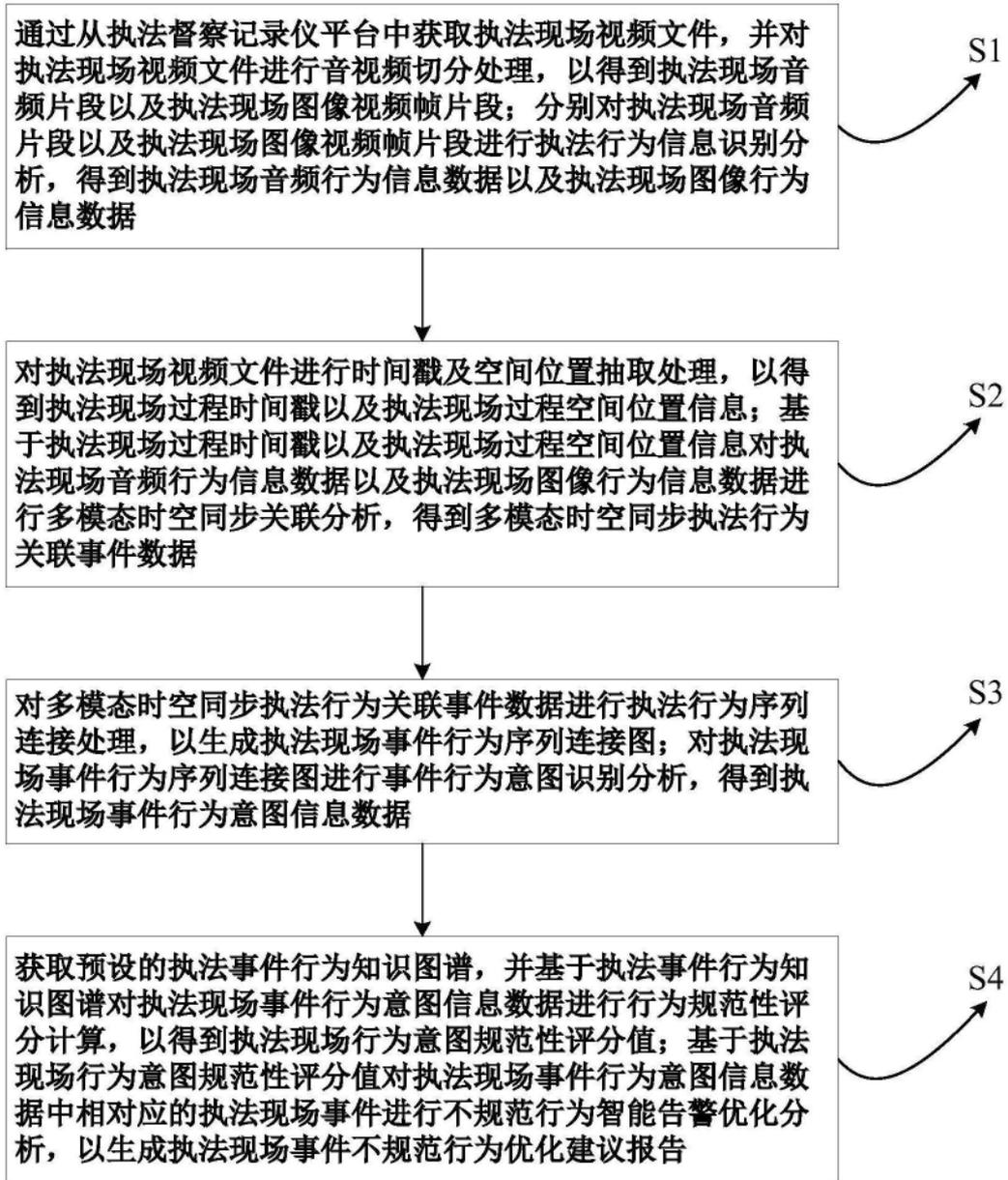


图1

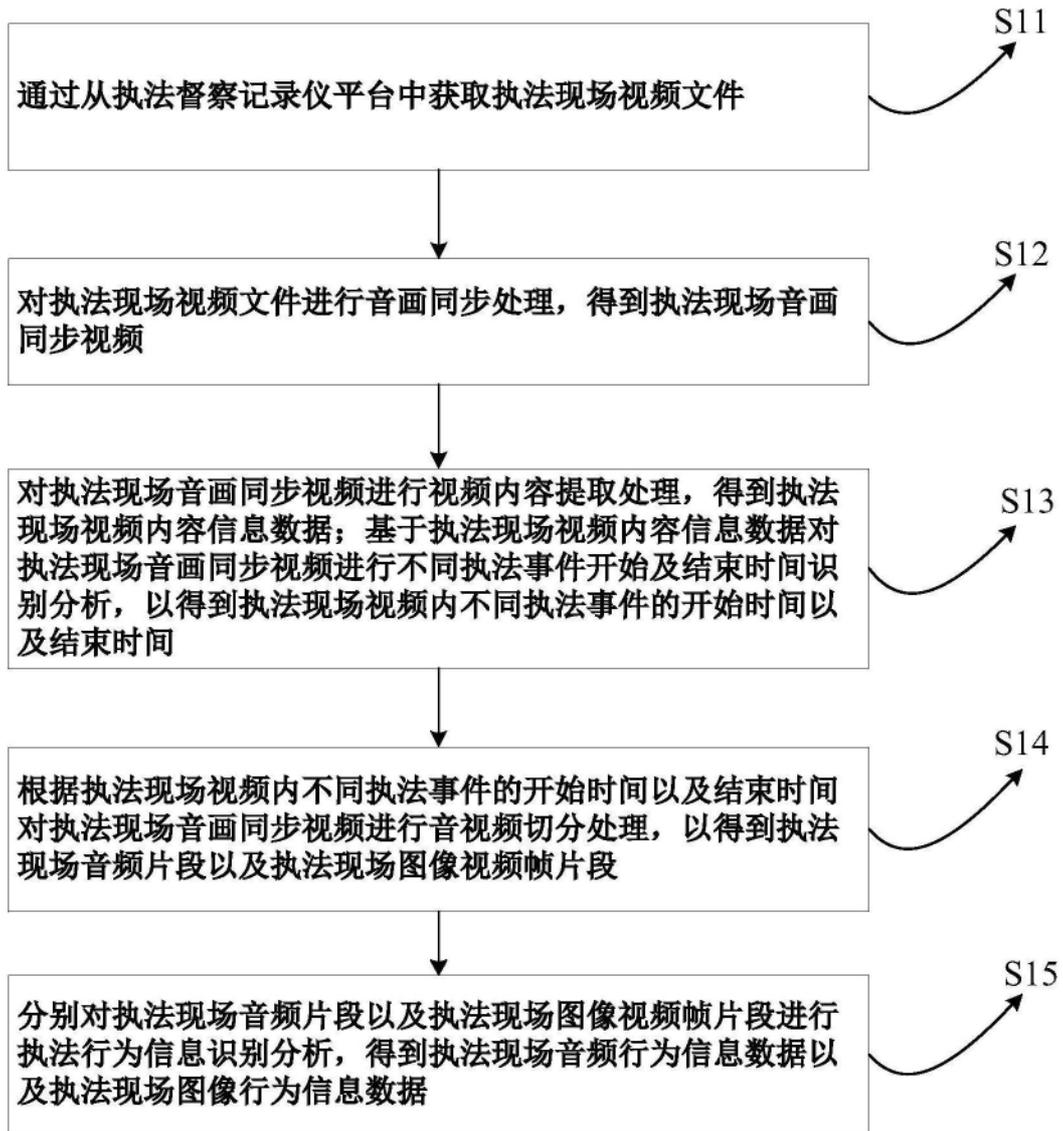


图2

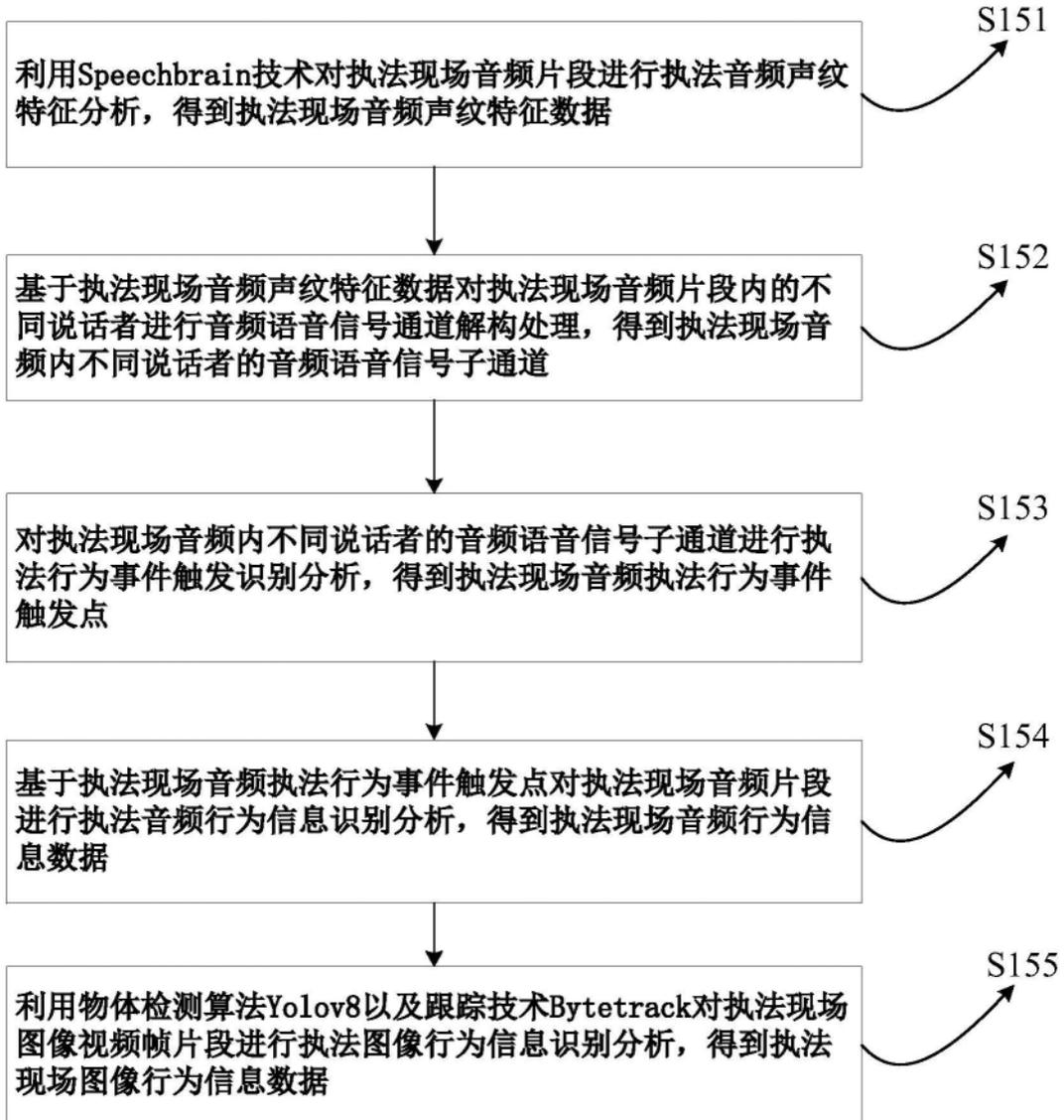


图3