

团体标准

T/CESA XXXX—202X

石化行业智能制造 生产区人员 安全视频分析系统通用要求

Intelligent manufacturing in the petrochemical industry—General requirements for
safety video analysis system for production area personnel

征求意见稿

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

已授权的专利证明材料为专利证书复印件或扉页，已公开但尚未授权的专利申请证明材料为专利公开通知书复印件或扉页，未公开的专利申请的证明材料为专利申请号和申请日期。

202X-XX- XX 发布

202X-XX- XX 实施

中国电子工业标准化技术协会 发布



版权保护文件

版权所有归属于该标准的发布机构，除非有其他规定，否则未经许可，此发行物及其章节不得以其他形式或任何手段进行复制、再版或使用，包括电子版，影印件，或发布在互联网及内部网络等。使用许可可于发布机构获取。

目次

前言..... II

1 范围..... 1

2 规范性引用文件..... 1

3 术语和定义..... 1

4 缩略语..... 1

5 一般要求..... 1

6 系统构成要求..... 1

 6.1 概述..... 2

 6.2 硬件设备..... 2

 6.3 软件系统..... 2

7 视频分析技术实现..... 3

 7.1 视频分析流程..... 3

 7.2 识别阶段..... 3

 7.3 训练阶段..... 4

8 软件系统功能..... 5

 8.1 概述..... 5

 8.2 综合统计..... 5

 8.3 报警信息管理..... 6

 8.4 视频管理..... 6

 8.5 异常识别管理..... 6

 8.6 模型管理..... 6

9 系统性能..... 6

 9.1 响应时间..... 6

 9.2 吞吐量..... 7

 9.3 资源使用率..... 7

 9.4 并发数..... 7

附录 A（资料性）安全视频分析系统架构实例..... 8

附录 B（资料性）场景描述及设计..... 12

参考文献..... 14

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由中国石油化工集团有限公司提出。

本文件由中国电子工业标准化技术协会归口。

本文件起草单位：。

本文件主要起草人：。

石化行业智能制造 生产区人员安全视频分析系统通用要求

1 范围

本文件规定了石化企业生产区人员安全视频分析系统（以下简称“安全视频分析系统”）的一般要求、系统构成要求、技术实现、软件系统功能、系统性能。

本文件适用于石化企业智能工厂建设中生产区内人员安全视频分析系统的规划、设计和应用。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

本文件没有需要界定的术语和定义。

4 缩略语

下列缩略语适用于本文件。

5G：第五代移动通信技术（5th Generation Mobile Communication Technology）

AI：人工智能（Artificial Intelligence）

CPU：中央处理器（Central Processing Unit）

FPS：帧每秒（Frames Per Second）

GPU：图形处理器（Graphics Processing Unit）

HTML：超文本标记语言（HyperText Markup Language）

NAS：网络附属存储（Network Attached Storage）

PTZ：水平方向移动/垂直方向移动/变焦（Pan/Tilt/Zoom）

SLB：服务器负载均衡（Server Load Balancer）

SSL：安全套接层（Secure Socket Layer）

UI：用户界面（User Interface）

5 一般要求

视频采集设备的布置应综合考虑监控目标、要害部位、关键装置和人员密集场所等风险因素，统筹考虑安全生产、反恐防范及安全环保视频建设需求，做到优先重点、兼顾全局。

防爆区内应选择对应等级的防爆设备，布线应满足防爆、防雷、防静电要求。

所有硬件设备应在-40℃～60℃环境中稳定工作。

系统应支持离线视频导入和联网视频接入两种方式。

系统框架可参照GB/T 43782—2024的相关要求，附录A给出了安全视频分析系统的架构设计实例。

6 系统构成要求

6.1 概述

安全视频分析系统由硬件设备和软件系统组成，硬件设备由视频采集设备、存储设备、处理单元、通信设备/器件组成，软件系统由图像处理、视频识别、数据库系统、用户接口等模块组成。

6.2 硬件设备

6.2.1 视频采集设备

视频采集设备包括固定式摄像机和移动式摄像机，摄像机分辨率应不低于1080p，帧率应不低于25FPS。

6.2.2 存储设备

存储设备包括带存储的计算机、硬盘录像机、NAS等实现存储功能硬件设备。存储设备应自动持续存储采集设备输出的视频流，存储设备容量支持的连续存储视频流时间应不少于30天，若个别地区对存储时限有更高要求的应按当地要求执行，并配置相应容量的存储设备。

6.2.3 处理单元

处理单元包括具备图像识别和目标检测算法所需计算资源的服务器、GPU加速器等计算终端。单进程执行时，CPU主频高于2.8GHz，核心数量不低于8个，线程数不低于16，缓存空间不小于16MB。显卡核心频率超过1300MHz，显存空间不低于8GB，显存位宽不低于256bit。

6.2.4 通信设备/器件

通信设备/器件包括构建通信网络所需的交换机、路由器、光纤转换器、网线、光纤等通信设备或元器件，实现图像采集设备、存储设备、处理单元之间的数据通信传输功能。宜采用有线通信模式组建局域网络，实现采集设备、存储设备、计算单元之间的稳定连接。在不具备架设通讯线缆的场景下，可采用无线通信方式。传输带宽应满足采集所得视频流实时传输的需求，网络延迟不应高于100ms，丢包率不应超过1%。

6.3 软件系统

6.3.1 图像处理

图像处理模块宜具备图像增强功能，可对图像或视频流进行降噪、对比度增强、分辨率调整，应保证处理速度满足25FPS帧频的实时性要求。

图像处理模块应将所有视频采集设备的输出转化为统一格式的视频流，帧频不低于25FPS。

图像处理模块应具备图像标注功能。

6.3.2 视频识别

使用深度学习算法，根据生产区内各类人员安全场景进行图像分析和识别，获得所需目标特征的模块。

视频识别模块应具备附录B中各场景的识别能力。

视频识别模块的单个场景识别精确度不应低于80%，召回率不应低于80%，F1分数不应低于85%。视频识别模块应具备对采集设备所得视频流实时识别分析的能力，处理速度应高于25FPS。

注：精确度是指模型预测为正样本的实例中真正为正样本的比例，用于衡量模型预测正样本的准确性；召回率是指实际为正样本的实例中，模型预测为正样本的比例，用于衡量模型能够预测出多少真正的正样本；F1分数是

精确度和召回率的调和平均数，用于综合评估模型在精确度和召回率上的表现，F1 分数取值范围在 0-1 之间，值越大越好。

6.3.3 数据库系统

数据库系统应包含附录B中各场景特征标签和情景标签等信息的分类存储模块。

6.3.4 用户接口

用户接口为提供操作界面的模块，用于安全管理人员查看识别结果和进行数据管理。

7 视频分析技术实现

7.1 视频分析流程

安全视频分析实现过程分为识别阶段和训练阶段。图1给出了安全视频分析的流程。

识别阶段包括数据获取、数据预处理、智能识别分析、识别结果展示、异常/隐患判定、报警信息推送6个环节。各环节应符合本文件7.2的要求。

训练阶段包括样本库构建、模型训练、模型评估、模型迭代4个环节。各环节应符合本文件7.3的要求。样本库构建、模型训练、模型评估、模型迭代应形成闭环。

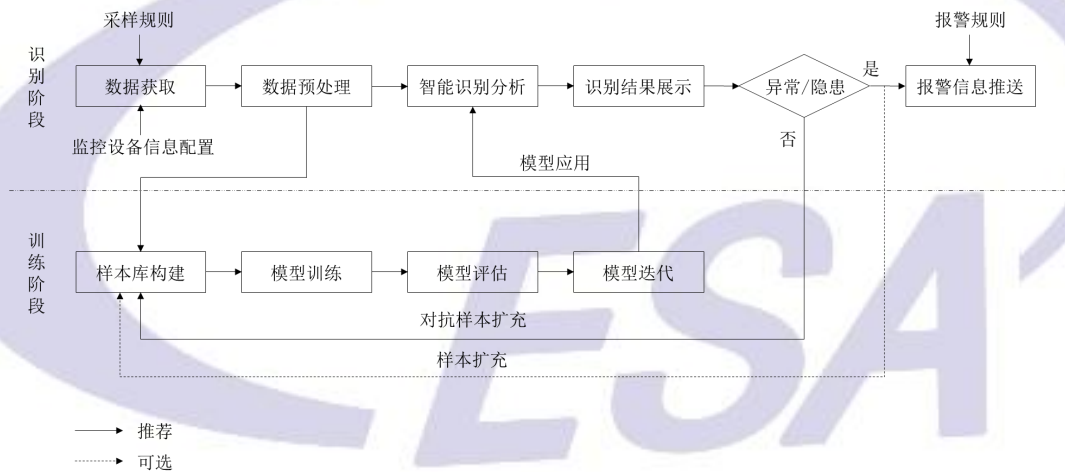


图 1 安全视频分析流程

7.2 识别阶段

7.2.1 数据获取

数据获取包括下列步骤：

- a) 摄像机基本信息获取——包括但不限于摄像机安装位置、摄像机名称、摄像机画质、视频流地址；
- b) 视频截取——按照一定时间间隔，对离线摄像机、在线摄像机视频数据进行截取，并保存为图片格式；
- c) 图片数据收集——将截取的离线摄像机、在线摄像机图片存储至数据库。

7.2.2 数据预处理

支持对获取的数据进行下列预处理：

- a) 图像增广——支持旋转、平移、翻转、缩放、剪切的一种或多种功能；
- b) 视频图像增强——支持对比度增强、低照度增强等功能中的一种或多种功能；
- c) 视频图像复原——支持去雾、去模糊、超分辨率重建、几何畸变校正、偏色校正、去噪等功能中的一种或多种功能。

经预处理后，图像像素分辨率不低于1080p；人员尺寸不小于80×120像素，灭火器尺寸不小于50×100像素；火花尺寸不小于100×200像素。

7.2.3 智能识别分析

将采集到的生产现场视频数据输入识别模型中，计算出附录B中的场景概率，识别出属于附录B中的一个或多个场景。

7.2.4 识别结果展示

可按照附录B中的场景分类展示识别的结果。

7.2.5 异常/隐患判定

当识别结果出现异常后，系统将识别结果推送至安全管理人员进行判定。若判定结果为异常/隐患，人工选择是否补充入样本库；若判定结果为非异常/非隐患，进行人工标记，并补充入对抗样本库。

7.2.6 报警信息推送

参照附录B的报警配置规则进行报警配置，并将判定为异常/隐患的识别结果信息进行报警，并推送至相关人员。

7.3 训练阶段

7.3.1 样本库构建

7.3.1.1 样本集建立

搜集生产区人员安全相关视频图像信息，开展特征工程研究，分类建立样本集。通过对现场各类场景数据的搜集、整理、清洗、特征提取和分拣，分级分类建立附录B中各类安全识别场景样本集。将样本集按8:2随机划分为训练集和测试集。

训练集为面向不安全行为/因素识别算法进行神经卷积训练的数据集。通过收集具有某一类不安全行为/因素特点的视频，对其分别标注姿势、动作、形状等，以备算法能够加载这些数据进行识别训练。可以将同类不安全行为/因素的训练集进行复制与合并，生成新的训练集。

测试集用于保存测试不安全行为/因素识别算法效能的视频库。库中的视频需要包含能够满足全部待识别的不安全行为/因素内容，并建立对应不安全行为/因素在视频中出现时间、出现位置、代表截图的索引，能够用于新算法效能测试评价。算法在版本迭代过程中的测试评价宜使用相同视频，以获得效能提升的对比效果。

7.3.1.2 样本标注

将训练集中图片按照识别场景进行标注，包括图片中目标边界框和类别标签，形成各个场景的图像识别样本库。训练集的标注方式可采取正样本标注或负样本标注的形式完成。标注的过程可采用人工标注和自动标注，并允许人工干预自动标注的样本。完成标注的训练集具有可迁移性和累加性。样本标注过程控制参考GB/T 42755—2023的相关要求。

7.3.1.3 样本质量审核

针对标注完成的样本，可以采用人工或自动的方式进行标注质量审核，特别是针对目前识别技术在评测过程后遗留的负样本，需要分析产生因素，从而决定该样本重新标注还是舍弃。

7.3.2 模型训练

采用经过标注的训练集中样本进行模型训练。训练集中样本数量不少于1000张特征图像。将训练集中样本输入到目标检测框架进行训练，基于卷积网络构建各场景的图像识别模型，通过调整学习率、激活函数等参数，验证模型的识别精度，直至模型达到预期的效果。

7.3.3 模型评估

应采用精确度、召回率、F1分数评估模型/算法性能。模型/算法性能评估计算方法参见YD/T 4396.3—2023的相关公式。若精确度、召回率、F1分数未达到要求，需要调整各模型的参数重新进行训练，直到精确度、召回率、F1分数符合要求。

7.3.4 模型迭代

- 每个模型迭代版本应分配唯一标识。
- 每个模型迭代版本应有对应的样本训练集、测试集。

8 软件系统功能

8.1 概述

安全视频分析系统应包含综合统计、报警信息管理、视频管理、异常识别管理、模型管理功能。图2给出了安全视频分析系统的功能组成。其中，综合统计功能包括但不限于首页综合展示、监控报警统计、违章处理统计、历史对比子功能。报警信息管理功能包括但不限于报警配置、报警推送、识别结果确认、违章处置记录子功能。视频管理功能包括但不限于在线视频管理、离线视频管理、视频检索、历史监控复盘子功能。异常识别管理功能包括但不限于异常的标记管理、异常的任务管理子功能。模型管理功能包括但不限于模型版本管理、模型更新、模型信息维护、模型测试子功能。

安全视频分析系统				
综合统计	报警信息管理	视频管理	异常识别管理	模型管理
首页综合展示	报警配置	在线视频管理	异常的标记管理	模型版本管理
监控报警统计	报警推送	离线视频管理	异常的任务管理	模型更新
违章处理统计	识别结果确认	视频检索		模型信息维护
历史对比	违章处置记录	历史监控复盘		模型测试

图 2 安全视频分析系统功能组成

8.2 综合统计

- 综合统计的子功能应符合下列要求：
- a) 首页综合展示——汇总并展示各项关键指标，如最新报警信息、近期违章事件统计等，为用户提供实时的全局视图；

- b) 监控报警统计——统计各个监控点或时间段的报警次数,帮助用户发现可能存在安全问题的区域或时间段;
- c) 违章处理统计——统计违章事件的处理情况,包括异常核验统计情况,已处理、未处理的事件数量,以及近期违章事件的处理结果;
- d) 历史对比——对比不同时间段的统计数据,帮助用户了解安全状况的变化趋势。

8.3 报警信息管理

报警信息管理的子功能应符合下列要求:

- a) 报警配置——对监控点的报警模式进行配置,设定报警条件、报警推送方式及报警接收者配置;
- b) 报警推送——通过视频分析模型服务,针对在线实时监控视频的即时分析,找到不安全行为/因素,获取问题关键帧及问题标记,并将问题推送至相关监控点位的安全管理人员;
- c) 识别结果确认——安全管理人员可查看并确认系统的识别结果,并对识别结果进行评判;
- d) 违章处置记录——安全管理人员可配置已确认的不规范/不安全问题识别记录的违章事件处置情况,包括处理时间、处理结果等。

8.4 视频管理

视频管理的子功能应符合下列要求:

- a) 在线视频管理——对实时的监控视频数据进行管理;
- b) 离线视频管理——对历史的监控视频数据进行管理;
- c) 视频检索——根据指定条件,如时间、地点等检索相关视频;
- d) 历史监控复盘——通过选定模型版本,对选定时间范围和摄像点位号的监控进行违规行为分析复盘。

8.5 异常识别管理

异常识别管理的子功能应符合下列要求:

- a) 异常的标记管理——对报警结果的标记进行人工修改、删除,或人工添加标记及信息。若标记结果为典型情况,可由人工添加至样本库,进行样本扩充,作为模型更新的训练集;
- b) 异常的任务管理——对任务进行维护,可添加、删减任务,构建新任务后,可在该任务下添加标记识别结果作为该任务的训练集。

注:任务即分析模型所要识别的目标,如不安全行为/因素的种类等。

8.6 模型管理

模型管理的子功能应符合下列要求:

- a) 模型版本管理——管理模型的各个版本,包括查看版本历史、回滚到旧版本等;
- b) 模型更新——根据新的数据更新模型;
- c) 模型信息维护——维护模型的相关信息,如模型的描述、创建日期、使用情况等;
- d) 模型测试——使用测试集,测试当前模型或历史模型的性能。

9 系统性能

9.1 响应时间

系统平均响应时间小于3秒；简单查询页面小于3秒，复杂查询页面小于10秒。在系统最大并发情况下，响应时间小于5秒。

注：最大并发用户数是指系统设计用户总数的10%。

9.2 吞吐量

吞吐量应满足下列要求：

- a) 静态数据指标——能对 100 个用户数据、状态和应用进行管理；
- b) 动态数据指标——并发访问用户数量应达到 50 个以上。

9.3 资源使用率

应用服务器和数据库服务器CPU占有率峰值均不应大于80%。

9.4 并发数

对同一应用，并发访问用户数量应大于50人，同时在线用户数不小于100人。



附录 A
(资料性)
安全视频分析系统架构实例

A.1 系统架构

借助云计算、5G、AI等新一代信息技术建设石化厂区视频智能识别系统。整个系统采用多层架构设计，包括数据采集层、服务层、应用层、访问层和用户层等组成。图A.1给出了系统架构组成。



图 A.1 系统架构

A.2 应用架构

应用架构由数据层、平台层、应用层组成。应用层中包括综合统计、报警信息管理、视频管理、行为识别管理、模型管理5个功能模块。图A.2给出了应用架构组成。



图 A.2 应用架构

A.3 数据架构

数据架构由视频监控系统、结构数据源、数据类型、数据处理、展示层组成。图A.3给出了数据架构组成。表A.1给出了支撑系统运行数据内容及其说明。

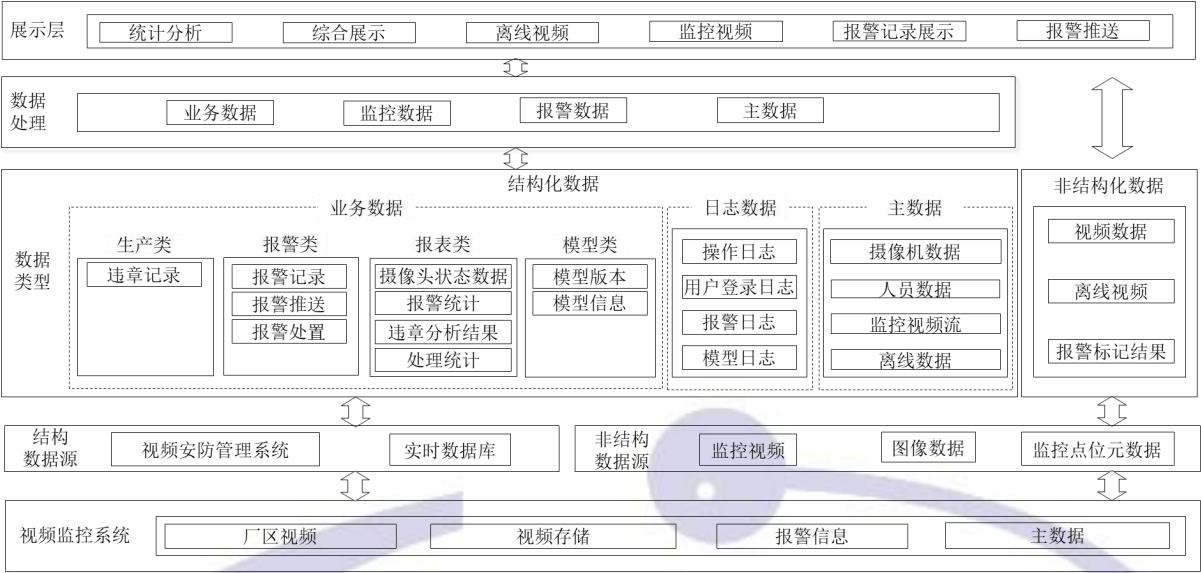


图 A.3 数据架构

表 A.1 数据说明

序号	数据内容	数据源系统	数据类型	数据量	说明
1	实时监控视频流数据	视频安防管理系统	非结构化数据	1080p 视频流单路 1~2MB/s	实时按采样规则采集，需要有监控点位信息，对于 PTZ 摄像机需要有摄像机状态信息
2	历史监控视频、图像数据	视频安防管理系统	非结构化数据	单路历史保存近 30 天	按任务采集，需要有视频的监控信息和对应摄像点位信息
3	监控信息、监控点位信息，摄像机状态信息	视频安防管理系统	结构化数据	2GB	按任务采集或定时采集，历史视频数据需要有监控信息，进行识别分析的实时监控视频需要有监控点位信息，对可转角的 PTZ 摄像机需要有摄像机状态信息
4	设备装置状态	实时数据库	结构化数据	2GB	定时采集和更新，获取生产装置区域开停工状态，需要对装置设备和摄像机进行关联

A.4 技术架构

石化厂区视频智能识别系统基于持续交付平台对开发过程进行管理，充分复用相关平台的身份认证服务、用户权限、消息队列提醒等技术组件。图A.4给出了技术架构组成。



图 A.4 技术架构

A.5 集成架构

石化厂区视频智能识别系统集成架构遵循集成规范，集成人工智能中台、实时数据、视频监控平台，通知系统4项，系统可应用于管控中心大屏展示。图A.5给出了集成架构组成。表A.2给出了系统集成方案。

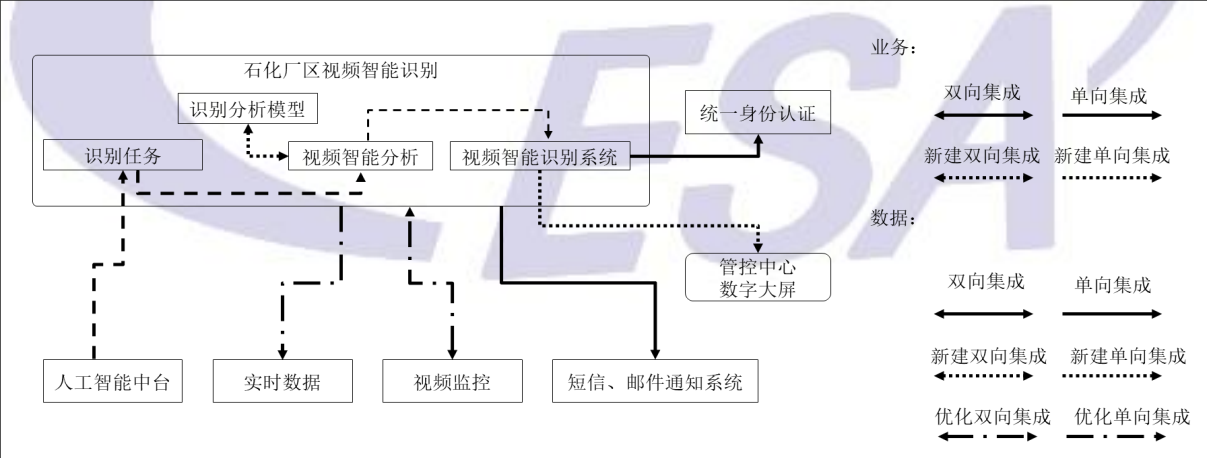


图 A.5 集成架构

表 A.2 集成方案

序号	源系统名称	流向	目的系统名称	交互数据	数据格式	频次
1	视频监控平台	→	人员安全行为识别	监控视频数据流	监控视频流	实时
2	视频监控平台	→	人员安全行为识别	历史监控视频数据	视频数据	按日
3	安全管理系统	←	视频智能分析	不安全分析结果推送	结构化数据	实时
4	安全管理系统	→	视频智能分析	问题确认及处理	结构化数据	实时
5	大屏	←	视频智能分析	界面功能展示	结构化数据、非结构化数据	实时

A.6 部署架构图

图A.6给出了部署架构组成。表A.3给出了系统配置列表。

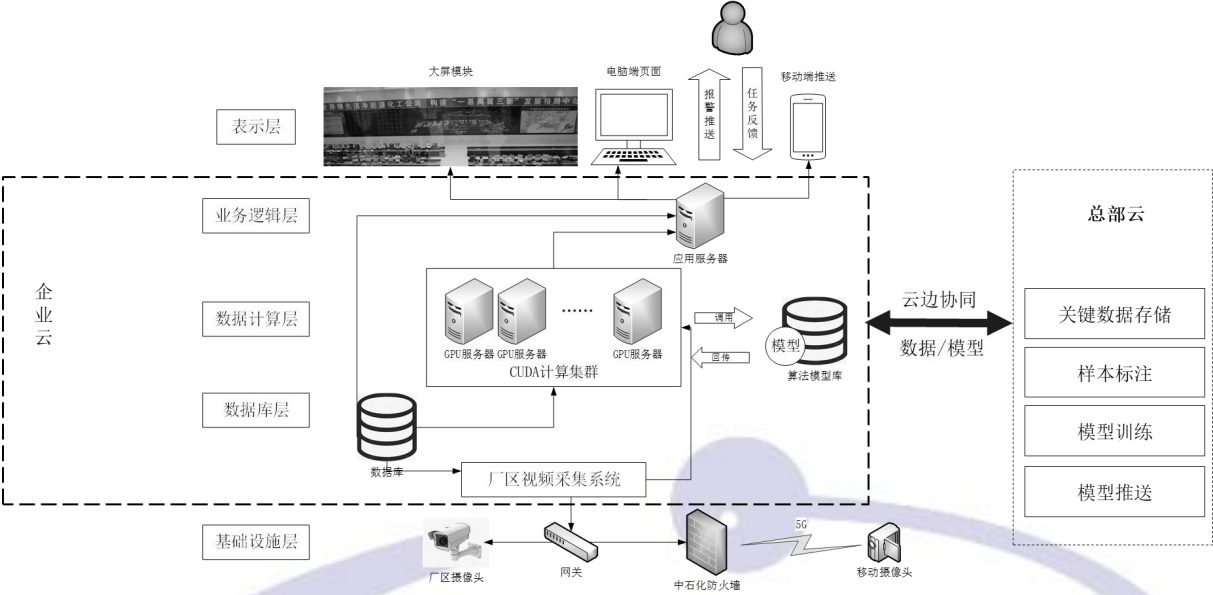


图 A.6 部署架构

表 A.3 系统配置

序号	设备名称	设备数量/台	设备用途
1	应用服务器	2	部署系统应用前后端
2	存储服务器	1	存储系统中非结构化数据，如视频数据、图片数据
3	GPU 服务器	10	提供基于 GPU 的应用于视频编解码、深度学习、科学计算等多种场景的快速、稳定、弹性的计算服务
4	移动布控球	50	施工现场临时作业视频监控
5	万兆网络交换机	2	用于网络传输和服务器间传输
6	其他	按需	用于支持硬件环境搭建，包括线缆、机架、配电盒、三角支架等

附 录 B
(资料性)
场景描述及设计

安全视频分析场景包括但不限于安全帽佩戴识别、工服穿戴识别、吊装区域人员监测、司索/指挥人员在岗分析、监管人员在场检测。安全视频分析场景的描述和设计参见表B. 1。

每个场景应包括两个维度的标签：

- a) 特征标签——用于描述目标场景环境的标签，通过指定的标签项目和标签内容，构建一种用于描述环境的“维度”，丰富目标信息的同时，也可用于针对不同环境对情景进行分类，可以泛化模型对于目标情景的识别。特征标签包括区域场所、监控场景、设备状态、危害分类、重要程度、环境要素、人群描述、时间等要素；
- b) 情景标签——通过视频识别模型对图像进行目标识别检测的标签，对每种情景进行训练和调参，并构建情景的识别、评分、报警规则，用于指导模型对目标的识别和报警。情景标签应与场景名称对应。

表B. 1 安全视频分析场景描述及设计

序号	场景名称	场景描述	识别规则	采样规则	报警规则	报警解除
1	安全帽佩戴	在生产区内规定必须佩戴安全帽的场所，通过形状、颜色、位置，对人员安全帽佩戴进行识别。	首先进行人员目标识别，当识别为“人”时，对目标进行安全帽佩戴情况识别，若佩戴情况评分低于监测下限，进行目标跟踪识别，并进行持续评分，若评分低于报警阈值，则进行报警推送。 构建人员穿戴不合规报警加权规则。根据时间和识别置信度、Map函数等，对目标进行评分。	1 帧/秒	单次报警，同类报警只推送一次，推送至关联人员。	人员确认
2	工装穿戴	在生产区内，通过形状、颜色、位置，对人员穿戴工作服进行识别。	首先进行人员目标识别，当识别为“人”时，对目标进行工装穿戴情况识别，若穿戴情况评分低于监测下限，进行目标跟踪识别，并进行持续评分，若评分低于报警阈值，则进行报警推送。 构建人员穿戴不合规报警加权规则。根据时间和识别置信度、Map函数等，对目标进行评分。	1 帧/秒	单次报警，同类报警只推送一次，推送至关联人员。	人员确认
3	火焰监控	对生产区设备故障、泄漏、操作失误或者人为放火等可能导致火灾和其他安全事故的不规则或不期望的火源进行识别。 注： 火焰的识别时，排除正常运行过程中产生的火焰，如燃烧废气的火炬塔发出的火焰。	可配置检测阈值和报警阈值，提高抗干扰能力。检测到超过检测阈值的目标后，进行秒级监测；设置点位报警阈值，当目标分数超过报警阈值，进行火焰报警。 构建火焰目标加权规则。通过前后监控时序图像比对，对时间和轮廓进行加权，持续时间越长、轮廓持续增大，加权越高。 通过危害检测页面，实时监控全厂推送的火焰报警。	1 帧/秒	单次持续报警，推送至关联人员。	人员确认

表B.1 安全视频分析场景描述及设计（续）

序号	场景名称	场景描述	识别规则	采样规则	报警规则	报警解除
4	动火作业 灭火器配备	生产区内规定设置灭火器的场所，如易燃易爆或动火作业场所，通过颜色对灭火设备的种类、数量以及是否在规定安全区域内进行识别。	人工可配置灭火设备缺失报警规则。配置最小报警时间，当超过该时间后，当区域中未识别到灭火设备，或灭火设备目标数量小于报警配置中的数量时，进行报警推送，由人工确认后结束本次报警。可配置图片中灭火设备检测区域。	1 帧/分	单次报警，同类报警只推送一次，推送至关联人员。	人员确认
5	吊装区域内 人员检测	生产区内吊装作业过程中，对吊装区内是否存在人员进行识别。	进行人员目标检测。可配置图片中人群规模检测区域。可配置吊装区域人员报警规则。配置最小报警时间，当超过该时间后，当区域中检测到人员数量超过额定值，进行报警推送。可配置禁止人员区域，当检测到人员时，即可报警。	1 帧/秒	单次报警，同类报警只推送一次，推送至关联人员。	人员确认
6	司索/指挥人员 在岗分析	吊装作业过程中对司索人员/指挥人员是否在吊装现场进行识别。	首先进行人员目标识别，识别区域中缺少“人”时，识别人员标识物，如头盔颜色、袖章位置和形态等，判断是否为司索/指挥人员，根据报警所配置的最少时间规则，超时进行报警推送。可配置图片中检测区域。	1 帧/秒	单次报警，同类报警只推送一次，推送至关联人员。	人员确认
7	施工现场监管 人员在场检测	对施工作业过程中监管人员是否在场进行识别。	首先进行人员目标识别，识别区域中缺少“人”时，识别人员标识物，如头盔颜色、袖章位置和形态等，判断是否为监管人员，根据报警所配置的最少时间规则，超时进行报警推送。可配置图片中检测区域。	1 帧/秒	单次报警，同类报警只推送一次，推送至关联人员。	人员确认

参 考 文 献

- [1] GB/T 42755—2023 人工智能 面向机器学习的数据标注规程
- [2] GB/T 43782—2024 人工智能 机器学习系统技术要求
- [3] YD/T 4396.3—2023 信息内容识别技术 第3部分：基于视频识别的内容检测服务系统指标要求和评估方法

