

运营商视联网融合运营平台基础架构研究与创新实践

梁秉豪, 张传刚*, 袁明明, 肖红梅

浪潮通信信息系统有限公司, 山东 济南

收稿日期: 2025年9月12日; 录用日期: 2025年10月15日; 发布日期: 2025年10月27日

摘要

当前, 视联网系统在终端设备接入、算法场景适配和数据价值挖掘等方面均面临着较大挑战。本文围绕视频监控产业的数字化、智能化升级需求, 从电信运营商视角系统梳理了视联网系统的发展现状和问题, 并在此基础上设计了面向运营商的视联网融合运营平台。该平台基于接入层、网络层、平台层和应用层的分层解耦架构, 实现了算网资源调度、基础视频处理、视频智能分析和运营数据分析等功能, 并在运营商营业厅管理、餐饮连锁门店管理、市容市貌管理等场景开展落地验证。实践证明, 通过引入该平台可有效降低设备接入成本、提升多场景适配能力。通过大模型与算力网络的深度融合, 实现视频数据的智能分析与价值挖掘, 为社会治理和传统产业升级提供高效支撑。

关键词

视联网, 大模型, 运营商, 算力网络, 视频分析

Research and Innovative Practice on the Basic Architecture of the Video Surveillance Integrated Operation Platform for Operators

Binghao Liang, Chuangang Zhang*, Mingming Yuan, Hongmei Xiao

Inspur Communication Information System Co., Ltd., Jinan Shandong

Received: September 12, 2025; accepted: October 15, 2025; published: October 27, 2025

Abstract

Currently, video surveillance systems face significant bottlenecks in terminal device access, algorithm

*通讯作者。

文章引用: 梁秉豪, 张传刚, 袁明明, 肖红梅. 运营商视联网融合运营平台基础架构研究与创新实践[J]. 计算机科学与应用, 2025, 15(10): 151-162. DOI: 10.12677/csa.2025.1510257

scenario adaptation, and data value mining. To promote the digital and intelligent upgrade of the video surveillance industry, this paper systematically reviewed the current development status and problems of video surveillance systems from the perspective of telecommunications operators. Based on this, a video surveillance integrated operation platform for operators is designed. This platform, based on a layered decoupling architecture of the access layer, network layer, platform layer, and application layer, realized functions such as computing and network resource scheduling, basic video processing, video intelligent analysis, and operation data analysis. We carried out innovative practices in scenarios such as the management of operator business halls, the management of chain restaurant stores, and the management of urban appearance and environment. Practical implementation demonstrates that this platform can effectively reduce device access costs, enhance multi-scenario adaptability, and through the deep integration of large models and computing power networks, achieve intelligent analysis and value mining of video data, providing efficient support for social governance and the upgrade of the traditional industries.

Keywords

Internet of Videos, Large Model, Operator, Computing Power Network, Video Analysis

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着我国城市化进程的加快,城市规模不断扩张,航空、铁路等运输网络不断健全,人口流动日益复杂,社会治安防控压力不断增大,传统依赖人力防范的方式难以全面覆盖和快速应对各类社会事件。近年来,随着智慧城市和雪亮工程政策不断落地,视频监控作为“城市神经末梢”已经实现了较大范围的覆盖。然而,随着设备数量的指数增长,传统视频监控系统和业务流程正面临着存储成本高、事件回溯和检索效率低等问题。当前,随着城市治理逐步精细化,业务需求已经从事后回溯向事前预警和全局化管理方向全面升级。视频监控系统需适配更加复杂和敏捷的业务场景(如视频结构化分析和检索、人流密度统计等),推动其向智能化、融合化演进。新一代信息技术正在重塑视频监控的业务流程,在大模型技术驱动下,实现视频数据智能分析和算法按需调度,深度挖掘视频内容价值,是视联网产业数智化升级的核心,同时为社会治理和各类管理决策提供新路径。

2. 视联网系统发展历程和趋势

我国视联网系统建设主要包括了硬件设备厂商、电信运营商、系统集成商和算法厂商等上下游企业。其中电信运营商利用其网络基础设施(光纤网络和移动通信网络)优势,积极发展视联网业务,提供快速、稳定的视频监控数据传输能力,确保视联网业务的流畅运行。如图 1 所示,本文系统梳理了视联网相关技术的发展趋势,将其发展历程总结为以下三个阶段。

2.1. 第一代视联网系统:视频远程监控系统

第一代视联网系统主要基于网络基础设施,完成了视频监控前端设备的远程接入,同时提供了基础的设备控制、视频编解码[1]和查看等能力。在 1996 年,瑞典企业安讯士发明了全球第一款网络摄像机产品,标志着视频远程监控系统的正式出现。随后在 1998 年,MPEG-4 标准正式发布,通过帧重建技术,完成数据压缩和传输,利用很窄的带宽实现了较高图像质量的实时数据传输。国际标准化组织和国际电

信联盟在 2003 年共同提出了 H.264 标准,是继 MPEG-4 之后的新一代数字视频压缩方法,进一步提升了传输效率,该标准目前仍在广泛使用。在网络加速方面,产业界开始通过内容分发网络[2] (CDN)对音视频内容传输效率进行优化,将静态的音视频资源进行分布式缓存,有助于用户从网络中多个分布式节点选择最优节点获取资源。2004 年,中国电信完成了 CDN 流媒体分发网络的建设,标志着该技术的大范围商业化推广。

2.2. 第二代视联网系统：智能视频分析系统

第二代视联网系统主要利用了计算机视觉算法,对视频内容进行简单的分析,实现了包括目标检测、人脸识别、车牌识别等功能。其标志性事件为在 2005 年的 CVPR 会议上,Navneet Dalal 和 Bill Triggs [3] 提出了利用方向梯度直方图(HOG)进行视频特征的提取,然后再利用线性支持向量机(SVM)作为分类器,从而在平衡速度和识别效果的前提下实现了行人检测。随后, Girshick 等人在 2014 年提出第一个基于深度学习的目标检测算法 R-CNN [4],在识别效果和泛化能力上,相比传统基于特征工程的方法有极大的提升。2016 年, Redmon 等人提出了 YOLOv1 [5]目标检测算法,具有非常快的识别速度和较高的识别精度,由于其大幅降低了视频分析的算力成本,很快就实现了产业化应用,奠定了智能视频分析系统的主流技术路线。

2.3. 第三代视联网系统：视频融合运营系统

第三代视联网系统主要通过大规模预训练模型实现视频内容的结构化分析和存储,随着万物识别[6]、万物分割[7]和视频理解等通用能力的不断提升,使得深度挖掘视频内容价值成为可能。在 2020 年,谷歌公司提出了 Vision Transformer (ViT) [8],成功将 Transformer 架构应用到了视觉任务中,展示出了非常强的迁移学习能力。在视觉理解和推理方面,多模态视觉语言模型 VLMs 如: CLIP [9]、Qwen [10]和 GPT-4V [11]等,同时采用了图像编码器和文本编码器,基于对比学习在共享潜在空间中将图像和文本信息进行对齐,有效提升了对视频内容的理解能力。随着知识检索的增强[12]和智能体技术的不断发展,为传统的视联网系统提供了“自主决策能力”与“外部知识支撑”,有效解决了传统视觉算法在未知样本适应、复杂场景理解、知识依赖性任务中的局限性。

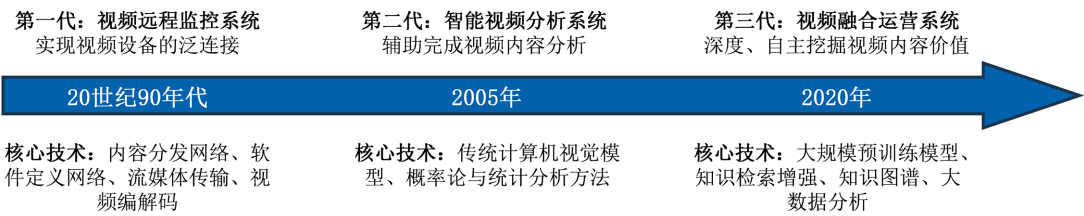


Figure 1. Development history of video surveillance system
图 1. 视联网系统发展历程

2.4. 视联网系统未来发展趋势

2.4.1. 当前业务痛点

痛点 1：多厂商、跨协议设备接入成本高
视联网系统涉及的终端设备种类繁多,不同硬件厂商设备接入方式和传输协议差异明显。政企行业客户通常需要在兼容老旧设备的前提下接入新增设备,对于视联网系统兼容性和服务质量[12]有较高要求。此外,涉及网络环境复杂,包括企业专线、6G/5G 移动通信网络、卫星通信和工业互联网等,不同类型网络的带宽、时延、稳定性差异显著。

痛点 2：跨行业、多元化场景适配难度大

不同行业对于视联网系统的核心诉求差异极大，例如城市安防需实现 7×24 小时全域覆盖，工业控制和质检要求实现毫米级缺陷识别和毫秒级系统响应，农业养殖需实现视频监控与物联网传感器的融合分析及联动。对于相同行业不同细分场景同样存在较大差异，以交通场景为例，对于高速公路需要实现团雾预警和事故快速定位，对于城市路口需要关注违章行为抓拍，对于景区停车场则需要提供车位引导等功能。

电信运营商作为网络基础设施提供者[13]，具有较为完善的标准化产品和服务能力，然而视联网业务落地高度依赖对细分场景的渗透率，场景多样性、设备复杂性和算法个性化等因素，对视联网领域标准化产品建设和市场推广提出了更大的挑战。

2.4.2. 未来发展趋势

趋势 1：全面集成大模型能力，提升系统智能化等级

视联网系统将通过大小模型融合架构，向自适应进化方向演进。云端大模型将实现跨模态知识的自主迭代，通过联邦学习融合边缘节点场景数据，持续优化通用能力基座，提升预训练模型对新场景的适配效率。边缘小模型支持动态加载任务，实现“端-边-云”三级算力协同，基于视频特征自动分配算力资源。通过隐私计算与模型压缩算法的结合，解决敏感场景数据共享难题。

趋势 2：深度挖掘视频数据价值，提升视联网系统应用场景

依托大模型对视频流进行细粒度特征提取，构建包含群体属性、时空轨迹和交互关系等信息的多维结构化数据集。结合业务场景挖掘视频数据价值[14]，例如通过群体特征和行为数据[15]，提升产品营销成功率、业务办理效率和客户服务体验等。指导各行业业务处理流程优化，构建视频数据要素流通机制，实现价值的规模化释放。

3. 视联网融合运营平台基础架构

3.1. 业务架构

为解决上述业务痛点，提升视联网业务的统一运营能力和数智化水平，本文设计了一套面向运营商视联网业务的视联网融合运营平台，平台业务架构如图 2 所示，主要划分为接入层、网络层、平台层和应用层。



Figure 2. Overall business architecture diagram
图 2. 整体业务架构图

接入层为上层应用提供了基础的资源和运行环境支撑,统一纳管计算设备资源(如通用服务器、AI 服务器、边缘计算设备等)与前端视频采集和存储设备(如摄像头、NVR 等)。网络层作为数据传输数枢纽,提供了视频专网、互联网、企业内网、移动通讯网络等多元异构网络通道,适配多业务场景网络需求,一方面传输接入层采集的视频数据,另一方面对平台运行产生的业务数据进行分发,保障数据在各层间高效、稳定传递。平台层基于云原生环境,提供视频处理、AI 分析基础能力,同时构建了媒体中心、汇聚中心、AI 服务中心、运维中心、运营中心、策略中心等核心模块。应用层主要聚焦城市治理、园区管理、工程管理、安防监控等业务场景,借助平台层的技术能力与资源整合,提供视频分析标准化应用,为不同领域的管理需求提供可视化、智能化的解决方案,实现从数据到业务价值的转化。

3.2. 技术架构

本文基于上述业务架构和业务难点进行技术方案设计(如图 3 所示),构建了“AI 算法 + 视频处理 + 算网调度”三层协同体系,有效支撑视频监控智能分析和业务运营场景。其中 AI 技术层是智能分析大脑,视频技术层保障视频数据全链路可靠处理,算网技术层构建算力与网络底座。通过三者协同联动,既满足垂直场景的精准识别需求,又能应对复杂场景的深度理解与生成,同时通过视频技术优化数据流转效率,算网技术保障算力高效供给,最终为智能视频平台提供视频监控、内容理解 and 多模态交互等关键能力。



Figure 3. Technical implementation architecture diagram
图 3. 技术实现架构图

AI 技术层通过小模型与大模型协同,在成本可控的前提下,实现了对复杂业务场景的精准分析。小模型类主要聚焦垂直场景的精准任务,覆盖了计算机视觉基础任务以及多模态交互与理解任务,为视频内容提供基础智能解析能力。大模型类主要侧重复杂场景的深度理解与生成,涵盖图片理解/生成、视频理解/摘要/检索/生成、音频生成、文本生成,通过多模态融合与泛化能力,解决小模型在复杂语义、动态场景中的分析瓶颈,实现从“识别”到“理解 + 创造”的跨越。

视频技术层围绕视频“采集 - 传输 - 应用”全生命周期,提供核心技术支持,通过视频剪辑、抽帧、转码、智能编解码,实现视频内容的按需处理与高效传输,适配不同终端与网络环境。依托视频加密保

障数据安全，通过传输协议适配优化网络传输效率，确保视频流稳定、安全流转。

算网技术层构建了“算力 + 网络”基础设施，保障上层技术高效运行。通过算力池化整合通用服务器、AI 服务器等算力资源，形成弹性可调度的算力池，支撑模型训练与推理任务。基于高通量传输方案优化网络带宽与传输协议，保障视频数据、模型参数的高速流转。通过智能资源编排与硬件加速技术(如 GPU 调度)，提升模型推理效率，降低端到端处理延迟。

3.3. 功能架构

3.3.1. 整体功能架构

视联网融合运营平台基于不同用户和角色的实际使用场景和需求，对平台功能进行了解耦和权限控制，基于四层协同架构支撑视联网业务闭环管理。其中算网资源调度模块面向运维团队，保障底层资源高可用，基础视频处理模块面向视频监控场景实现视频资源可视化，视频智能分析模块面向智能布控等场景实现人工替代，运营数据分析模块面向产品运营者，深度挖掘数据价值和提升产品能力。如图 4 所示，平台从底层资源到上层应用逐层递进、协同联动，既满足视频采集、处理、分析的技术需求，又通过数据运营实现业务闭环，支撑城市治理、园区管理等多场景的智能视频应用。

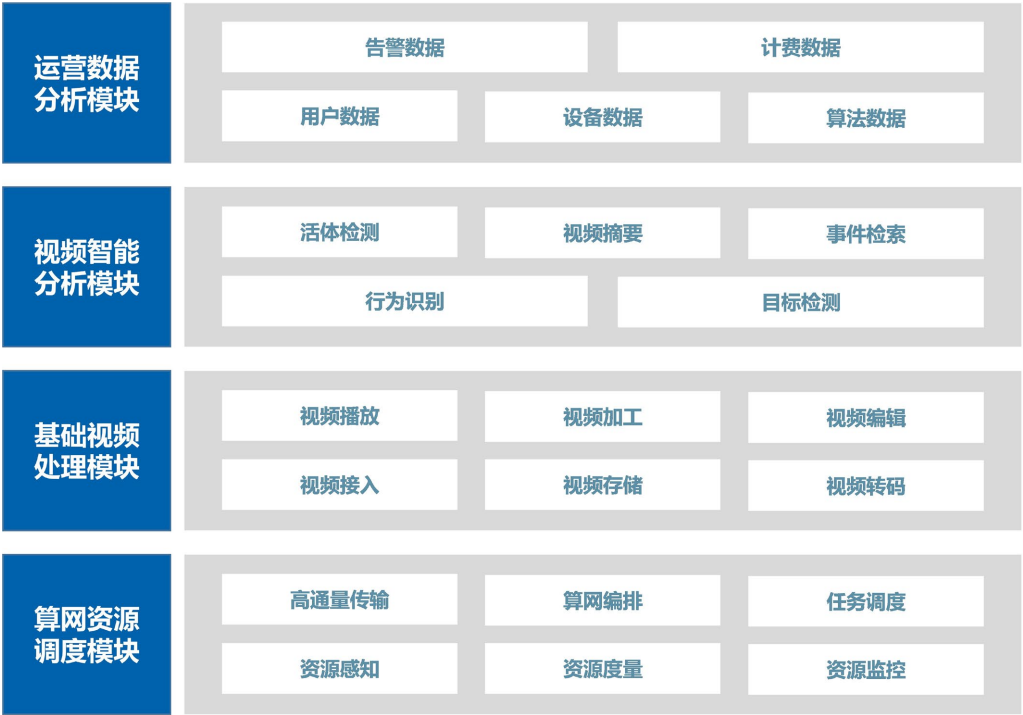


Figure 4. Functional architecture diagram
图 4. 功能架构示意图

算网资源调度模块作为平台算力与网络的调度中枢，负责资源基础管理，通过资源感知、资源度量、资源监控，实时掌握算网资源(服务器、带宽等)的状态、性能与负载。实现智能调度编排，依托高通量传输、算网编排、任务调度，根据业务需求动态分配算力与网络资源，保障视频处理、AI 分析等任务高效运行，解决资源闲置与业务峰值拥堵问题。

基础视频处理模块聚焦视频生产与处理全环节，提供基础接入与存储，通过视频接入、视频存储，支持多类型前端设备(摄像头、NVR)的数据采集与长期留存。通过视频转码、视频加工、视频编辑适配各

类前端播放需求，同时支持视频格式转换、画质优化、内容剪辑，是视频数据从“原始采集”到“业务可用”的关键环节。

视频智能分析模块基于 AI 技术层能力，实现视频内容的深度理解与智能识别。通过目标检测、行为识别，完成对视频中物体、人员动作的精准识别。依托活体检测、视频摘要、事件检索，解决复杂场景需求(如安防活体验证、长视频关键信息检索、异常事件快速回溯等)，将视频数据转化为业务决策依据。

运营数据分析模块聚焦数据驱动的产品运营与优化，深度分析用户数据(行为、需求)、设备数据(状态、故障)、算法数据(精度、效率)在时间维度和空间分布等方面的特征，支撑视联网业务和产品的升级迭代。通过告警数据、计费数据(流量、调用量等)精准采集，实现平台运营的风险监控与商业价值管理，让数据反哺业务决策。

3.3.2. 算网资源调度模块

算网资源调度模块(如图 5 所示)主要划分为资源接入层、资源管理层和编排调度层。资源接入层作为基础底座统一汇聚异构算力节点，接入计算、存储和网络资源及终端设备，构建算网资源池，为上层应用提供底层资源支撑。资源管理层聚焦算网路由与供需感知，通过算网路由标识、转发、寻址及路由信息管理，保障资源按需流转；借助算网资源、需求感知，精准掌握资源状态与业务需求。编排调度层实现智能管控，算力度量和建模打造业务、异构硬件、节点服务模型，支撑业务适配；算网编排提供固定规则与智能化编排，算网调度通过算力注册 OAM、智能融合调度，按需分配资源，驱动算网高效协同。

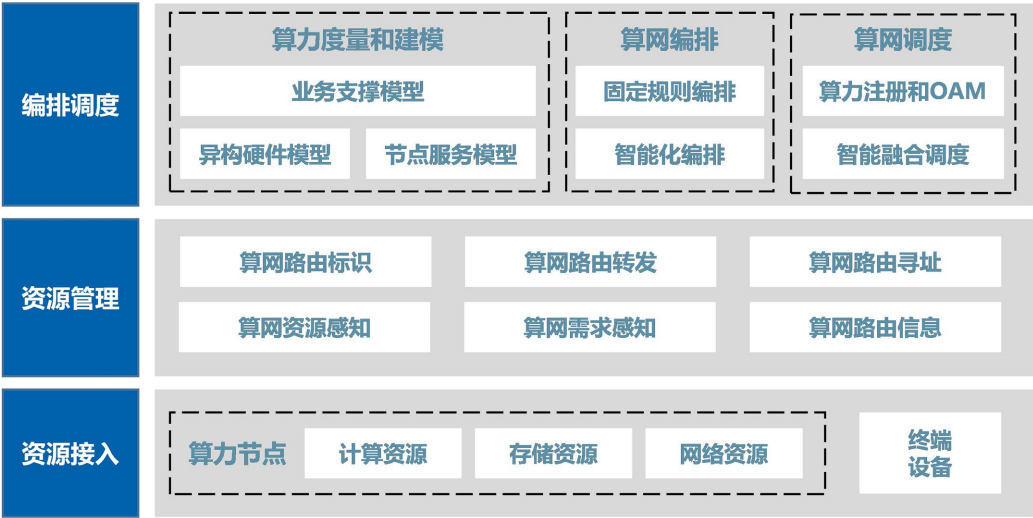


Figure 5. Function architecture of resource scheduling module of the computing network
图 5. 算网资源调度模块功能架构

通过算网资源调度模块可以实现算网资源高效调度，汇聚多元资源(算力、存储、网络等)，资源管理感知供需，编排调度实现智能编排与调度(如算网编排匹配业务需求、算网调度优化资源分配)，充分利用算网资源，保障业务稳定、高效运行，提升资源使用效率与业务响应速度。此外，提供业务适配与创新支撑能力，构建业务支撑模型，适配异构硬件与节点服务，结合智能编排调度，可精准匹配业务对算网资源的需求(如 AI 推理、视频编解码)，为创新应用提供灵活、可靠的算网底座，助力用户拓展业务边界，提升竞争力。

3.3.3. 基础视频处理模块

基础视频处理模块(如图 6 所示)主要包括视频接入层、视频处理层和视频管理层。视频接入层兼容多

元采集源，支持海、大、宇等设备厂商平台接入，实现了文件接入及 GB28181、ONVIF、RTSP、T1400 等协议适配，打破设备与协议壁垒，汇聚各类视频流。视频处理层聚焦视频技术加工，具备转码、剪辑、压缩功能，适配不同播放与存储需求。提供本地存储(NVR 等)和云端对象存储(S3 等)，灵活应对实时数据和归档数据存储等场景。视频管理层支撑业务应用落地，覆盖视频直播、回看、图片抓拍、电子地图功能，满足监控与追溯需求。通过设备、平台、策略、运维管理，实现设备管控、系统配置及故障预警，保障视频业务稳定运行。



Figure 6. Function architecture of basic video processing module
图 6. 基础视频处理模块功能架构

通过基础视频处理模块可以实现全流程视频管控，覆盖视频接入(多协议/厂商兼容)、处理(转码/存储灵活)、管理(直播/回看等功能)，一站式解决视频业务全链路需求，减少系统对接成本，提升视频部署、运维效率，让用户高效管理视频资源。此外，可以快速适配多元场景，保障用户体验，支持直播、回看、抓拍等业务场景，结合转码、存储策略，既能满足实时监控需求，又能通过历史视频回溯、图片证据留存辅助决策，同时适配不同网络与设备，保障视频服务稳定、流畅，优化使用体验。

3.3.4. 视频智能分析模块

视频智能分析模块(如图 7 所示)主要包括数据处理层、视觉算法层和智能应用层。数据处理层负责视频数据流转，通过抽帧、抓拍、截取等预处理，实现图片/短视频推送、视频数据回流，为算法训练与应用输出提供数据支撑，保障数据闭环。视觉算法层以目标检测、实例分割等算法为核心，融合 OCR、语音识别，结合视觉大模型，为上层应用提供智能分析能力，让视频内容理解更精准高效。智能应用层聚焦业务场景，提供人车物识别、事件识别等功能，覆盖考勤、统计、检索、告警全流程，如人脸考勤用于人员管理，车流统计辅助交通分析，满足安防、运营等需求。

通过视频智能分析模块可以实现智能场景化应用提效降本，依托视觉算法(目标检测、OCR 等)支撑智能应用(人车物识别、事件识别等)，覆盖安防、考勤、统计等场景，替代人工操作，快速识别、分析视频数据，提升业务流程效率，降低人力成本。此外，可实现标准化视频数据处理和算法接入，视频处理(抽帧、抓拍等)生成视觉数据并通过算法分析后回流到业务应用，通过人工抽检和智能巡检方式相结合，持续提升算法效果。结合自然语言检索、视频结构化，助力用户深度挖掘视频价值，如快速定位关键信息、追溯事件，为决策提供数据支撑，强化视频数据的业务价值。



Figure 7. Function architecture of the video intelligent analysis module
图 7. 视频智能分析模块功能架构

3.3.5. 运营数据分析模块

运营数据分析模块(如图 8 所示)主要包括数据资产、数据分析和产品运营。数据资产子模块汇聚了多维度数据,包括视频结构化、算法告警等业务数据,设备实时、业务订购、平台使用等运营数据,以及用户、设备、套餐、算法等基础信息,构建数据底座。数据分析子模块主要基于数据资产开展场景化分析,通过用户画像洞察深层次需求,基于设备运行历史数据辅助运维工作保障系统稳定,基于人脸考勤与人车流统计辅助企业内部管理。通过自然语言检索、视频结构化存储提升视频数据价值密度,实现数据价值转化。产品运营子模块提供了 AI 应用商城,帮助用户快速选择合适的视频套餐和 AI 应用,完成产品订购和开通,是平台商业变现的重要门户。



Figure 8. Function architecture of operational data analysis module
图 8. 运营数据分析模块功能架构

通过视频智能分析模块可以实现产品数据挖掘,驱动业务增长,通过汇聚视频结构化信息、设备实时数据等多维度数据资产,通过用户画像、车流统计等个性化运营数据,帮助用户精准洞察业务发展情况,优化产品运营策略,以数据赋能业务增长。此外,通过全链路数据采集,提升产品运营效率,覆盖AI应用商城订购、数据资产管理、智能分析场景,形成“运营-数据-分析”闭环。从应用开通便捷化,到数据统一管理,再到智能分析辅助决策(如设备运维、告警处理),全流程提升用户运营效率,降低管理成本。

4. 视联网融合运营平台创新实践

4.1. 创新实践思路

基于上述视联网融合运营平台,本文主要围绕产业数智化和数智产业化两个方面探索创新场景和落地实践。在产业数智化方面,传统制造业和服务业仍有大量业务环节需要人工参与,如何通过智能视频技术提升相关产业数智化水平是目前有待解决的问题之一。传统视频监控方案受限于“技术复杂度高、投入成本大”,尤其对中小企业不够友好。本平台通过云服务模式提供了开箱即用的一站式智能视频服务,大幅降低了传统产业数智化升级门槛。通过智能视频编解码技术降低了视频传输与存储所需的网络和硬件成本,使中小微企业也能使用“高清监控+云端分析”的数智化方案。基于异构算力的智能调度实现“边缘轻量分析+云端深度处理”的分层部署策略,可以满足中型企业管理需求,同时支持按需扩容云端服务,降低初期投资规模。基于多模态大模型的动态适配能力,满足垂直行业、特殊场景的视频分析需求,无需专业技术团队或者算法定制开发投入,即可快速实现设备运营异常和未按要求施工等个性化分析服务。在数字产业化方面,通过打造标准化视联网数据产品形成规模化供给能力,实现了“场景标签化、要素可复用”。平台基于标准化数据产品开发个性化应用,满足不同行业的数据需求。提供基础统计类数据产品,针对城市管理、零售、交通等领域,提炼“客流密度时序分布”、“区域人群特征(年龄/性别占比)”、“交通运行状态统计”等结构化数据,形成数据接口服务,基于数据调用量或订阅周期进行收费。

4.2. 运营商营业厅管理

运营商营业厅地理位置分布较广,通常由省公司统一管理,需保障视频数据传输安全,实时把控运营状况,优化资源调度与管理效率,同时要监控门店内外部环境、员工形象及营销展示的规范性与专业性。保障营业厅外墙广告完整清晰、墙壁清洁,店内地板干净、商品陈列规范,员工着装行为合规,以此塑造良好的品牌形象,提升客户体验。此外,需要精准了解店内广告推广效果,基于客户关注时长优化内容和陈列方式,提升营销转化,助力业务推广。

本文基于视联网融合运营平台统一接入、配置各营业厅视频监控摄像机,利用视觉算法实现摄像机状态和成像质量的监测。通过国密协议对视频数据进行加密传输保障业务安全,借助视频分析营业厅实时运营数据(如客户流量、业务办理时长),辅助运营决策。基于大模型识别营销广告完整性、实时性、清晰度及合规性,监测门店环境清洁度,生成保洁工作任务提醒。同时,对员工着装行为规范进行智能识别,及时通知责任人,保障服务标准化。分析客户对店内广告的关注时长、行为反应,评估广告效果并优化投放。结合人脸识别与行为分析,了解不同客群关注情况,为橱窗展示策略调整(如内容、布局)提供数据,提升品牌宣传与营销转化。

4.3. 餐饮企业连锁门店管理

餐饮连锁企业门店分布全国各地,同时包括自营门店和加盟门店,管理难度较大。需要保障后厨操

作、食材储存等流程规范,实时把控各门店运营状况(如出餐效率、食材加工规范),优化人力、食材等资源调度与管理效率。同时需要保障门店内外环境卫生和整洁、员工形象(着装、操作规范)及餐品展示(明档陈列、菜单推广)的规范性与专业性,塑造食品安全可靠、服务优质的品牌形象,提升顾客用餐体验。

本文依托视联网融合运营平台统一接入各门店后厨及前厅监控摄像机,通过云端 AI 算法对门店进行智能监管,采集各门店实时运营数据,如后厨出餐高峰期和食材加工耗时等,为人员排班、食材采购及加工流程优化提供数据支撑,同时基于潮汐调度能力节省视频分析所需要的算力资源。基于计算机视觉大模型,自动识别后厨明厨亮灶区域的清洁度,监测食材储存是否合规,一旦发现油污残留、食材变质风险等问题,生成并推送提醒。对员工着装(工作服、帽子、口罩佩戴)、操作行为(是否遵循标准化烹饪、备餐流程)进行识别,若有违规,迅速通知责任人整改,保障服务与操作标准化,守护食品安全与品牌形象。

4.4. 市容市貌管理

城市辖区范围广,涵盖街道、商圈、公园、建筑楼宇等众多区域,各区域分布分散,城市管理部门需对市容市貌进行全面且高效的监管。责任部门需要掌握城市各区域的环境状态(如街道清洁情况、公共设施完好程度),优化环卫、市政维修等城市管理资源的调度与管理效率。同时,需维护城市外观的规范性与美观性,包括户外广告的合规清晰、建筑外立面的整洁、公共区域的环境卫生、市政设施的完好,以及规范占道经营、乱停乱放等行为,以此塑造整洁有序、宜居宜游的城市形象,提升市民生活幸福感和城市整体形象。此外,需要精准了解城市重点区域的人流情况和市民活动偏好,基于反馈优化城市管理和服

务,助力城市精细化治理。

本文依托视联网融合运营平台,实时监测摄像机的在线状态和画面清晰度。分析城市不同路段的垃圾堆积高峰时段、重点区域的人流密度等,为环卫人员排班、市政资源调配等提供数据支撑,辅助管理决策。基于计算机视觉大模型,自动识别公共区域是否存在垃圾堆积、市政设施是否完好。对占道经营、车辆乱停乱放等违规行为进行智能识别,及时通知执法人员进行处置,保障城市秩序,维护城市良好形象。

5. 结束语

本文围绕运营商视联网融合运营平台的基础架构与创新实践展开研究,系统梳理了视联网从远程监控到智能分析再到融合运营的发展脉络,揭示了技术迭代与场景需求驱动下视联网系统的演进趋势。针对当前多厂商设备接入复杂和跨行业场景适配困难等核心痛点,提出了涵盖业务、技术、功能的多层次融合运营平台架构,通过算法能力、视频能力和算网能力协同机制,实现了设备统一接入、视频高效处理、数据智能分析与资源优化调度的闭环管理。未来将持续深化大模型、算力网络与视联网业务的融合,探索隐私计算、联邦学习在敏感场景的应用,拓展在工业制造、智慧农业等更多落地场景,助力视联网产业实现更高效的数智化转型,为社会治理现代化与产业高质量发展提供更强支撑。

基金项目

泰山产业领军人才工程项目(TSCX202312006); 山东省博士后创新项目(SDCX-ZG-202400307)。

参考文献

- [1] 周建同, 杨海涛, 刘东, 等. 视频编码的技术基础及发展方向[J]. 电信科学, 2017, 33(8): 16-25,
- [2] Farahani, R., Timmerer, C. and Hellwagner, H. (2024) Towards Low-Latency and Energy-Efficient Hybrid P2P-CDN Live Video Streaming. arXiv: 2403.16985. <https://arxiv.org/abs/2403.16985>

-
- [3] Dalal, N. and Triggs, B. (2005) Histograms of Oriented Gradients for Human Detection. 2005 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, 20-25 June 2005, 886-893. <https://doi.org/10.1109/CVPR.2005.177>
- [4] Girshick, R., Donahue, J., Darrell, T., *et al.* (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/CVPR.2014.81>
- [5] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [6] Cheng, T., Song, L., Ge, Y., *et al.* (2024) YOLO-World: Real-Time Open-Vocabulary Object Detection. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 16901-16911. <https://doi.org/10.1109/CVPR52733.2024.01599>
- [7] Kirillov, A., Mintun, E., *et al.* (2023) Segment Anything. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 1-6 October 2023, 3992-4003. <https://doi.org/10.1109/ICCV51070.2023.00371>
- [8] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale. arXiv: 2010.11929. <https://doi.org/10.48550/arXiv.2010.11929>
- [9] Radford, A., Kim, J.W., Hallacy, C., *et al.* (2021) Learning Transferable Visual Models from Natural Language Supervision. arXiv: 2103.00020. <https://api.semanticscholar.org/CorpusID:231591445>
- [10] Yang, A., Yang, B., Zhang, B., Hui, B., Zheng, B., Yu, B., Li, C., Liu, D., *et al.* (2024) Qwen2.5 Technical Report. arXiv: 2412.15115. <https://arxiv.org/abs/2412.15115>
- [11] Zhu, D., Chen, J., Shen, X., Li, X. and Elhoseiny, M. (2023) MiniGPT-4: Enhancing Vision-Language Understanding with Advanced Large Language Models. arXiv: 2304.10592. <https://api.semanticscholar.org/CorpusID:258291930>
- [12] Zhao, P., Zhang, H., Yu, Q., *et al.* (2024) Retrieval-Augmented Generation for AI-Generated Content: A Survey. arXiv: 2402.19473.
- [13] 韩建亭, 张凤. 基于智能终端的视频通信业务服务质量评测模型研究[J]. 电信科学, 2013, 29(4): 27-32.
- [14] 胡敏达, 徐泽华, 杨东鹏. 运营商视联网产业发展分析: 现状、挑战与未来路径[J]. 通信企业管理, 2025(1): 46-48.
- [15] 郝鹏. 视联网数据管理的挑战与机遇[J]. 中国战略新兴产业, 2025(5): 120-122.