

# 一种广角 - 变焦协同的视觉感知装置应用设计

刘德洋<sup>1</sup>, 罗中华<sup>2</sup>, 沈克永<sup>3</sup>, 阮英兰<sup>2</sup>, 徐梦溪<sup>4</sup>, 郑胜男<sup>4</sup>

<sup>1</sup>南京工程学院创新创业学院, 江苏 南京

<sup>2</sup>南昌理工学院电子信息学院, 江西 南昌

<sup>3</sup>南昌理工学院计算机信息工程学院, 江西 南昌

<sup>4</sup>南京工程学院计算机工程学院, 江苏 南京

收稿日期: 2025年12月22日; 录用日期: 2026年1月25日; 发布日期: 2026年3月4日

## 摘要

针对视频监控中全景覆盖与细节捕捉难以兼顾、计算资源及成本受限、带宽限制与数据传输滞后, 以及对广域分散部署的固定或车载移动视频监控复杂场景适应性不足等问题, 本文提出一种广角 - 变焦协同的视觉感知装置(简称: WAZC-VPD)应用设计方案。该方案通过创新设计广角 + 变焦双镜头协同成像组件与三核异构轻量化CPU架构, 集成轻量化Smart侦测技术及5G通信模块, 构建“双镜成像 + 三核计算 + 5G传输”一体化轻量化架构, 实现了“全景覆盖 + 细节捕捉”双模式联动的视觉智能感知及视频数据5G实时传输。结合城市物流视觉感知应用场景的现场测试结果也表明, WAZC-VPD具备合理性与有效性, 为城市物流视频监控场景提供了一种可选的先进视觉感知手段。

## 关键词

城市物流, 视觉感知, 双镜头成像, 智能计算, 轻量化技术

# Application Design of a Wide-Angle and Zoom Collaborative Visual Perception Device

Deyang Liu<sup>1</sup>, Zhonghua Luo<sup>2</sup>, Keyong Shen<sup>3</sup>, Yinglan Ruan<sup>2</sup>, Mengxi Xu<sup>4</sup>, Shengnan Zheng<sup>4</sup>

<sup>1</sup>School of Innovation and Entrepreneurship, Nanjing Institute of Technology, Nanjing Jiangsu

<sup>2</sup>College of Electronics and Information, Nanchang Institute of Technology, Nanchang Jiangxi

<sup>3</sup>School of Computer and Information Engineering, Nanchang Institute of Technology, Nanchang Jiangxi

<sup>4</sup>School of Computer Engineering, Nanjing Institute of Technology, Nanjing Jiangsu

Received: December 22, 2025; accepted: January 25, 2026; published: March 4, 2026

文章引用: 刘德洋, 罗中华, 沈克永, 阮英兰, 徐梦溪, 郑胜男. 一种广角-变焦协同的视觉感知装置应用设计[J]. 仪器与设备, 2026, 14(1): 19-30. DOI: 10.12677/iae.2026.141004

## Abstract

Aiming at the problems in video surveillance, such as the difficulty in balancing panoramic coverage and detail capture, limited computing resources and costs, bandwidth constraints and data transmission delay, as well as insufficient adaptability to complex scenarios of wide-area distributed fixed or vehicle-mounted mobile video surveillance, this paper proposes an application design scheme for a Wide-Angle and Zoom Collaborative Visual Perception Device (abbreviation: WAZC-VPD). By innovatively designing a wide-angle + zoom dual-lens collaborative imaging module and a triple-core heterogeneous lightweight CPU architecture, integrating lightweight Smart detection technology and 5G communication module, the scheme constructs an integrated lightweight architecture of “dual-lens imaging + triple-core computing + 5G transmission”. This architecture enables visual intelligent perception with dual-mode linkage of “panoramic coverage + detail capture” and real-time 5G transmission of video data. Field test results combined with the application scenario of urban logistics visual perception show that the WAZC-VPD is reasonable and effective, providing an optional advanced visual perception method for the urban logistics video surveillance scenario.

## Keywords

Urban Logistics, Visual Perception, Dual-Lens Imaging, Intelligent Computing, Lightweight Technology

Copyright © 2026 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

视频监控作为一种实时感知与场景记录手段,已广泛应用于公共安全、工业生产、交通管理等各个领域,成为保障各类场景有序运行的关键技术支撑。在城市物流领域,视频监控的作用尤为突出,从仓库内货物分拣与存储监管,到运输途中车辆调度与货物安全追踪,再到末端配送环节的包裹交接确认,其贯穿物流作业全流程,是实现现场情况实时掌控、保障作业效率与安全的重要基础。

然而,城市物流场景具有多样性和复杂性。仓库内部货架密集、光线多变,运输车辆行驶中面临动态光影和复杂路况,末端配送点人流密集、环境杂乱,传统单镜头视频采集设备难以兼顾全景覆盖与细节捕捉。若覆盖全景,会因视场角过大导致局部图像分辨率下降,货物条码、车辆牌照等关键信息易出现模糊;若调整参数聚焦局部细节,又会缩小监控范围,不能捕捉周边环境的整体动态和及时发现区域内的异常情况,难以满足全场景清晰监控的需求。另外,城市物流中需要识别的目标多样,包括不同规格的货物、各类运输与作业设备、作业人员及他们的行为等。这些目标在复杂背景下常存在遮挡、运动模糊等情况,传统现场侧视频采集设备的智能识别能力有限[1]-[3]。

近年来,国内外在视频监控技术领域持续投入研发,在成像清晰度、基础智能识别等方面已形成较强的技术积累。在成像技术方面,其高清化、宽动态范围镜头显著提升了静态场景的图像质量[1] [4]-[6],但面向城市物流的车载移动、仓库堆叠等动态场景时,单镜头难以同时覆盖货车装卸全景与货物标签细节,多镜头方案协同成像设计也常出现视野重叠冗余或数据断层,全景与细节的联动感知仍有改进空间[2] [3]。

在智能识别方面,主流方案的深度学习算法在标准化场景中表现优异[7]-[10],但城市物流场景中货

物遮挡、光线突变、货运车辆颠簸等干扰,会导致识别准确率下降;同时,现有方案的智能计算多依赖云端或高性能终端,难以适配现场侧视频采集设备有限的计算资源,实时性保障不足[1][5][11]。

在传输与处理方面,5G技术已在主流监控方案中应用,但在广域分散的物流仓库、分拨中心、车辆移动等场景的规模化部署应用中,现场侧设备功耗与成本较高,结合物流应用实际需求做轻量化设计还有待完善[2][12]。

综上,在城市物流“全景与细节联动感知”“复杂场景智能识别的实时性”“轻量化与5G适配”等核心需求上仍有优化空间。本文针对目前城市物流视觉感知场景应用存在的问题,提出一种广角-变焦协同的视觉感知装置(简称:WAZC-VPD)应用设计方案。本文的主要贡献如下:

1) 提出“双镜成像 + 三核计算 + 5G传输”一体化轻量化架构。创新融合广角 + 变焦双镜头协同成像、三核异构轻量化计算及5G通信,解决传统物流视觉感知设备全景覆盖与细节捕捉难以兼顾、车载移动视频流传输滞后等瓶颈问题。

2) 以RK3506J处理器为硬件基础,提出三核异构分工协作的本地计算设计方案。通过明确三核CPU的预处理、算法运行、数据压缩与筛选的三者分工,解决计算资源、成本和带宽受限问题。

3) 基于“双镜成像 + 三核计算 + 5G传输”一体化轻量化架构,设计适配城市物流场景的WAZC-VPD装置,满足物流监控场景实际应用需求。

本文后续的内容组织安排:介绍WAZC-VPD系统架构、广角 + 变焦双镜头协同成像组件设计、三核异构轻量化CPU架构设计、5G模组设计,以及WAZC-VPD在物流现场的模拟测试等内容。

## 2. WAZC-VPD 系统架构

WAZC-VPD采用“感知-计算-传输”协同设计模式,构建“双镜成像 + 三核计算 + 5G传输”一体化轻量化架构,精准匹配城市物流现场侧视觉感知应用需求。针对全景与细节难兼顾问题,通过广角 + 变焦双镜头协同成像实现联动感知;针对计算资源及成本限制,采用三核异构轻量化CPU架构设计平衡性能与成本;针对带宽限制与传输滞后,集成5G模块实现无线实时传输。

技术层面,广角-变焦双镜头采集、三核异构架构(采用RK3506J处理器设计)、5G传输形成闭环,实现“感知-计算-传输”协同连接。应用场景层面,适配固定(仓库、分拨中心)与车载移动监控场景,通过结构优化保证复杂环境下的稳定性,满足实际应用需求。一体化轻量化WAZC-VPD系统架构如图1所示。

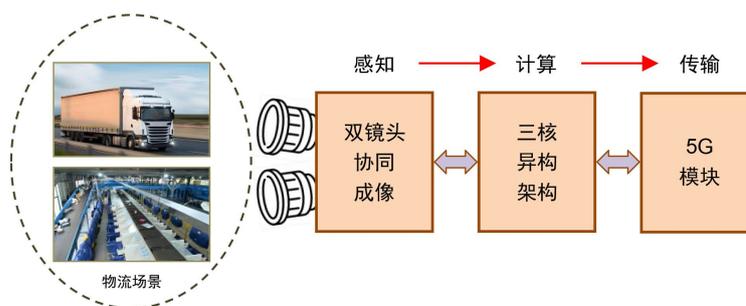


Figure 1. Integrated lightweight WAZC-VPD system architecture  
图 1. 一体化轻量化 WAZC-VPD 系统架构

## 3. WAZC-VPD 设计

### 3.1. 广角 + 变焦双镜头协同成像组件

该组件通过“全局视野覆盖 + 局部细节捕捉”的协同设计,满足城市物流场景中“动态监测与即时

追踪”的应用需求,其联动响应时间经硬件优化与算法协同控制在1秒内,具体设计如下:

#### 1) 组件选型与设计构成。

① 广角镜头:选用湖北云泰时代光学仪器公司 2.8 mm 定焦广角镜头,采用无畸变光学设计,120°大视场角可覆盖仓库全区域、货车周边 360°环境等大范围场景。其光学结构经优化,在画面边缘处的畸变率控制在3%以内,确保全景监控中物体形态无明显失真。

② 变焦镜头:采用长春博信光电子公司 10~150 mm 连续变焦镜头,支持 0.1 倍~15 倍光学变焦,可灵活聚焦于货架标签(最小识别尺寸 0.5 cm × 0.5 cm)、货车车牌等局部细节。镜头内置温度补偿机制,在 -20℃~60℃环境下仍能保持对焦精度。

③ 镜头固定与校准:双镜头通过定制金属一体化支架固定,支架采用铝合金材质。安装时通过激光校准工具控制光轴平行度误差在 0.5°以内,确保成像区域存在 20%~30%重叠区(如仓库货架与地面交界带),为后续画面坐标校准提供自然基准。

#### 2) 同步触发与驱动模块。

① 同步触发:内置高精度 RTC 时钟芯片(时钟误差 ≤ 1 ppm),通过 GPIO 同步信号控制双镜头曝光时刻,曝光同步误差 ≤ 1 ms。

② 驱动模块:变焦镜头搭配步进电机驱动模组(步距角 0.9°),支持 0.01 mm 级焦距调节精度。接收计算单元指令后,可在 500 ms 内完成从广角到长焦的切换并精准对焦,确保快速捕捉目标细节(如 10 米外货车车牌的字符清晰度达 200 dpi)。

#### 3) 图像传感器适配。

① 广角镜头传感器:搭配安森美 Hyperlux LH 系列 CMOS 图像传感器(200 万像素,1920 × 1080 分辨率),120 dB 动态范围可适应仓库明暗交替环境(如逆光下的货架与阴影区域),30 fps 帧率满足动态场景连续拍摄需求。

② 变焦镜头传感器:适配韩国安森美 Hyperlux SG 系列 CMOS 图像传感器(500 万像素,2592 × 1944 分辨率),高全局快门效率(GSE)可减少运动伪影(如移动货车的车牌无拖影),确保细节画面清晰。

③ 编码支持:两者均支持 H.265 编码,可将原始视频数据压缩至原大小的 1/8~1/10(如 1080P 30 fps 视频压缩后码率约 2 Mbps),大幅降低后续传输与存储压力。

#### 4) 图像采集与控制模块。

引入 FPGA 主要因 RK3506J 仅支持单路 MIPI CSI 接口,无法直接接入双传感器;同时需在 RAW 域完成低延迟预处理(如去噪、运动初筛),避免 CPU 带宽过载。若改用 CSI 桥接芯片虽可省 FPGA,但会增加成本并引入额外延迟,当前方案在成本与实时性间更优。

该模块基于 FPGA 芯片设计,选用 Lattice Semiconductor 公司的 iCE40 HX 系列 FPGA (型号: iCE40HX8K)。模块内部构建了双路图像采集通道(每路最大采集速率 30 fps)。每路通道首先通过专用的数据接口电路与镜头传感器相连,实时接收双镜头传感器输出的 RAW 格式图像数据。数据进入芯片后,先经过预处理逻辑单元,进行去噪、增益调整等初步处理。随后,数据被暂存至 sysMEM™ 嵌入式 RAM 块中,等待进一步处理。场景分析逻辑单元独立于图像采集通道,持续对缓存中的图像数据进行分析。该单元采用基于背景差分与特征匹配的运动目标检测算法以及基于物体轮廓与运动轨迹分析的异常事件检测算法。一旦检测到运动目标(如叉车移动)或异常事件(如货物坠落),场景分析逻辑单元会立即通过 UART 接口向计算单元发送触发信号。同时,根据计算单元反馈的指令,控制变焦镜头聚焦目标的逻辑电路开始工作,该电路通过调整变焦镜头的驱动信号,实现镜头的快速聚焦,响应延迟 ≤ 100 ms,以此实现“全景监测-目标识别-细节追踪”自动联动功能。

### 3.2. 三核异构轻量化 CPU 架构

作为 WAZC-VPD 的计算核心，三核异构轻量化 CPU 架构以 RK3506J 处理器(瑞芯微电子产品)为硬件基础，通过“专核专用”的分工设计与算法优化，实现“高效计算 + 低功耗 + 低成本”的平衡 [13]。RK3506J 通过共享 LPDDR4 内存与片上 Mailbox 机制实现 A7 与 M0 核间通信，关键指令采用中断驱动，实测消息延迟  $\leq 8 \mu\text{s}$ 。在峰值负载下 IPC 队列积压不超过 2 条，M0 可在  $100 \mu\text{s}$  内响应 A7 事件，协同效率满足系统实时性需求，未形成性能瓶颈。

#### 1) RK3506J 处理器功能设计。

RK3506J 采用 BGA 封装( $15 \text{ mm} \times 15 \text{ mm}$ )，集成 3 个 ARM Cortex-A7 架构计算核心(1.5 GHz，典型功耗 1.2 W/核)和 1 个 ARM Cortex-M0 架构控制核(400 MHz，功耗 0.3 W)。各个核的职责分工与数据交互归纳于图 2。RK3506J 硬件成本仅为高端工业处理器的 1/3，可直接集成到 WAZC-VPD 的主板中，适配车载终端、仓库角落的监控盒等设备小型化需求。

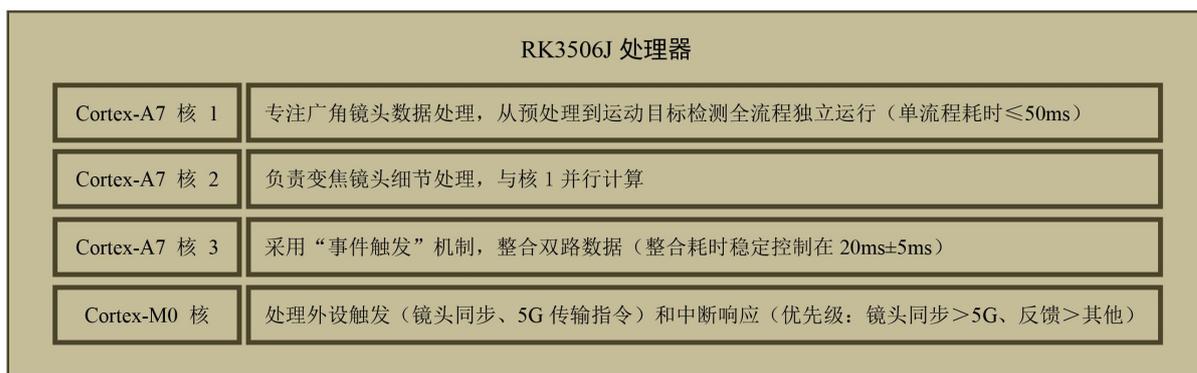


Figure 2. Task division and data interaction of the nuclear system

图 2. 核的任务分工与数据交互

RK3506J 具体功能设计如下：

① 视觉数据处理：3 个 Cortex-A7 核通过任务并行机制分工：2 个核分别处理广角镜头全景画面(运动目标检测，采用改进背景差分法，检测延迟  $\leq 30 \text{ ms}$ )和变焦镜头细节画面(特征提取，基于精简 CNN 模型，提取耗时  $\leq 40 \text{ ms}$ )；剩余 1 个核负责数据整合，通过坐标转换算法(基于重叠区域特征点匹配)生成“全景 + 细节”联动视图，再经小波变换压缩(压缩比 5:1)后输出关键信息。

改进的背景差分法引入以符号  $\Psi_i(u, v)$  表示的动态背景模型，其更新规则定义为：

$$\Psi_i(u, v) = \lambda_i(u, v) \cdot \mathcal{I}_i(u, v) + (1 - \lambda_i(u, v)) \cdot \Psi_{i-1}(u, v)$$

其中， $\mathcal{I}_i(u, v)$  为当前帧图像强度， $\lambda_i(u, v) \in (0, 1)$  为局部自适应学习率，依据像素邻域梯度变化动态调节；同时融合 L-K 光流一致性检验剔除伪前景。该策略能够有效抑制仓库中叉车移动、灯光闪烁等动态干扰。

② 设备协同控制：Cortex-M0 核依托 RTOS 实时操作系统，通过 GPIO 接口控制双镜头同步曝光(同步误差  $\leq 1 \text{ ms}$ )、驱动变焦镜头步进电机(位置控制精度  $\pm 1$  步)，并通过 UART 接口与 5G 模组建立通信，协调数据传输时机。

③ 运行效率优化：采用模型压缩与量化技术，通过敏感度分析裁剪冗余卷积层(模型体积压缩 50%)，结合 INT8 量化将浮点参数转为整数型(计算量减少 60%~70%)。优化后，目标识别算法在保持 95%左右的准确率的同时，端到端时延控制在 80 ms 以内，且单帧处理功耗降低至 0.8 W~0.9 W。

2) 程序流程设计(参见图 3)。

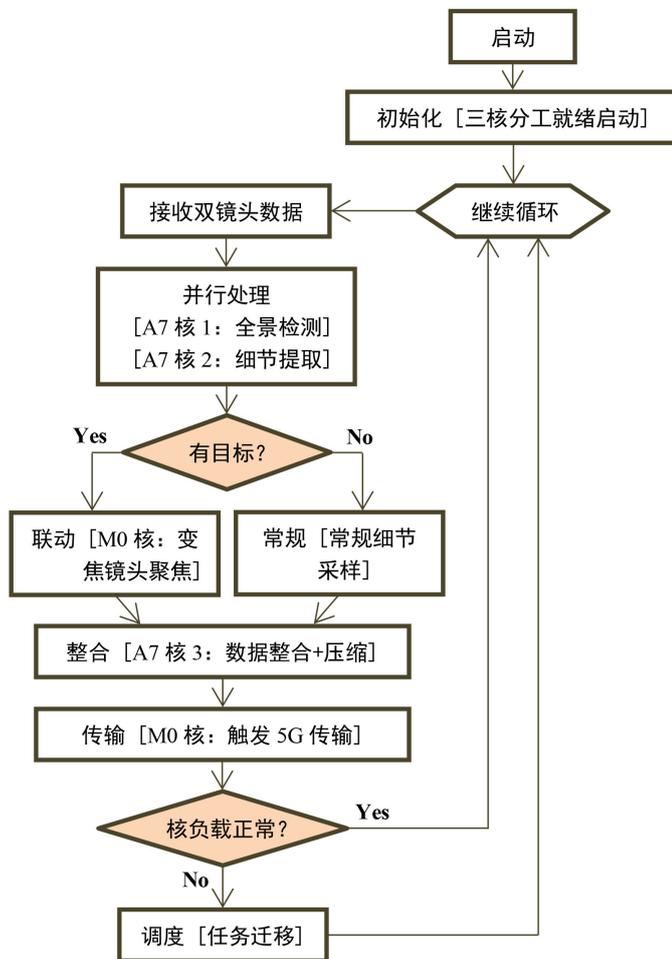


Figure 3. Program flow block diagram  
图 3. 程序流程框图

① 数据接收与预处理：接收图像采集模块输出的双路图像数据，Cortex-A7 核并行执行降噪、增强等预处理(耗时  $\leq 20$  ms)。

② 目标检测与联动：Cortex-A7 核检测目标后，Cortex-M0 核触发变焦镜头聚焦，同步记录目标坐标(定位误差  $\leq 5$  个像素)。

③ 数据输出与传输：Cortex-A7 核生成压缩后的联动数据，Cortex-M0 核检测到数据就绪后，向 5G 模组发送传输指令。

④ 动态资源调度：通过内置调度算法实时监控核的负载，当某一核占用率超过 80%时，自动将部分任务(如非关键特征提取)迁移至空闲核，确保流程稳定性。

3) 任务调度算法设计。

① 视觉数据处理调度：采用基于优先级的动态调度策略，其中运动目标检测(广角镜头)优先级设为最高(P1)，目标特征提取(变焦镜头)设为次高(P2)，确保动态场景下优先捕捉运动目标。当检测到异常事件(如货物被盗)时，自动将变焦镜头任务优先级提升至 P1，优先处理细节画面。

② 数据整合调度：采用“事件触发”机制，仅当双镜头数据均处理完成后启动整合流程，避免无效

等待(整合耗时稳定设计在  $20\text{ ms} \pm 5\text{ ms}$ )。

③ 外设控制调度: Cortex-M0 核采用“中断优先”策略,镜头同步触发信号(中断等级最高)、5G 传输状态反馈(中断等级次高)优先响应,确保外设协同无延迟。

4) 深度学习模型优化设计。

本文所用目标识别模型基于轻量级 MobileNetV2 主干网络,共 14 层,以深度可分离卷积为核心结构,兼顾精度与计算效率。模型训练数据由合作物流企业实地采集,涵盖仓库分拣区、物流园区道路、配送站点等 6 类典型城市物流场景,共计 12,850 张人工标注图像,包含货物、车牌、作业人员、叉车等 17 类目标;其中 30%为夜间或逆光图像,20%存在遮挡、运动模糊或部分残缺,按 8:1:1 比例划分为训练集、验证集与测试集,确保模型在复杂真实环境下的泛化能力。

在此基础上,针对 RK3506J 处理器的资源约束,实施以下两项轻量化优化:

① 模型裁剪:采用通道级敏感度分析策略,逐层评估移除各卷积输出通道对验证集 mAP 的影响,设定阈值为 0.5%,即仅保留移除后导致 mAP 下降超过 0.5%的通道。最终裁剪冗余通道 48.7%,主要集中在中间特征提取层,而输入层与输出层保持完整,以保障原始表征能力和检测头性能。

② 模型量化:将训练后的 32 位浮点权重与激活值统一转换为 8 位整数(INT8),通过校准及控制量化误差  $\leq 2\%$ 。量化后模型推理速度提升约 3 倍,单帧处理时间从 120 ms 降至 40~45 ms,存储占用由 3.2 MB 压缩至 0.8 MB,完全适配 RK3506J 内置的 1GB LPDDR4 内存资源,满足端侧实时运行需求。

### 3.3. 5G 模组

5G 模组作为 WAZC-VPD 装置的数据传输中枢,需满足 4K 视频 200 Mbps 车载传输、复杂环境下网络稳定性及外设协同需求[12],具体设计如下:

1) 5G 模组选型。

选用美格智能技术公司 SRM810 系列 5G 模组,该模组基于高通骁龙 X55 芯片组设计,支持国内三大运营商 5G 频段(n1/n3/n41/n78/n79)及 SA/NSA 双模。其硬件参数适配装置需求:下行速率最高 2 Gbps (满足 4K 视频 200 Mbps 传输余量)、上行速率 900 Mbps,通过 USB 3.1 接口(传输速率 5 Gbps)与 RK3506J 连接,通信延迟  $\leq 10\text{ ms}$ 。

2) 5G 功能设计与实现。

① 数据传输功能:模组接收 RK3506J 输出的 H.265 压缩视频(4K@30fps,码率 200 Mbps)后,通过内置调制解调模块(支持 NR-LTE 混合载波聚合)转换为 5G 信号,经基站传输至云端服务器。传输过程中采用 UDP 协议(配合前向纠错机制),确保视频帧丢失率  $< 0.1\%$ 。

② 网络管理功能:模组实时监测 5G 信号强度(RSRP)与信噪比(SINR),并通过 AT 指令向 RK3506J 反馈(更新周期 100 ms)。当信号弱于  $-105\text{ dBm}$  时,处理器自动触发降级策略(如将 4K 视频降至 1080P,码率降至 50 Mbps),保障传输连续性。

③ 外设协同功能:与定位模块(GPS + 北斗双模)联动,将经纬度信息(定位精度 1 米)嵌入视频流 metadata (元数据,为定位数据提供关键背景信息)字段,实现“视频画面 + 位置信息”同步传输。当检测到异常事件时,模组支持“优先传输”机制,即暂停普通视频流,优先发送警报信息及关键帧(传输延迟  $\leq 500\text{ ms}$ ),确保后端快速响应。

## 4. WAZC-VPD 现场模拟测试

### 4.1. 现场部署模拟测试环境

模拟测试的现场选择江西赣州某物流企业。该企业已建有网络视频监控系統,但已部署的海康威视

摄像机均不具备智能分析及 5G 传输功能。搭建模拟测试环境前,首先在视频监控中心配置安装了一台 5G 工业网关(选用厦门计讯物联公司的 TG463),通过网线连接原视频系统,网关加载海康威视软件开发工具包(Hikvision Software Development Kit)及 WAZC-VPD 的标准 ONVIF (开放型网络视频接口论坛)协议,按规则转发数据实现与原视频系统对接。并配置一台笔记本电脑作为测试用监控中心服务器(配属了上位视频分析软件)。现场的模拟测试环境如图 4 示意。

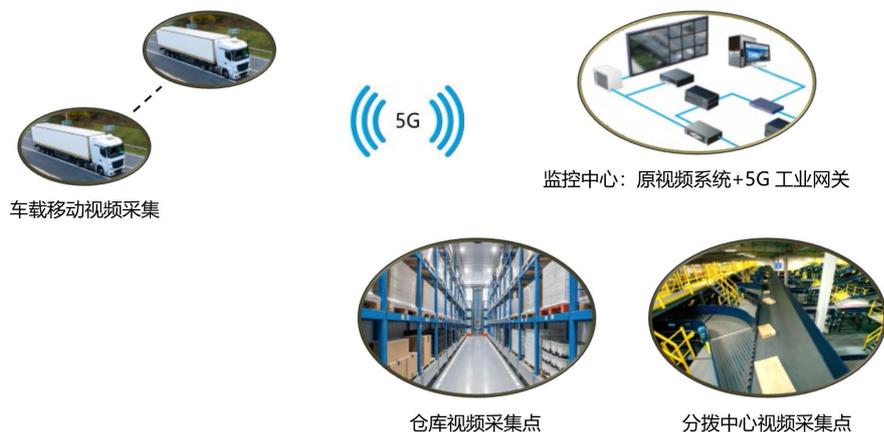


Figure 4. Test environment of the WAZC-VPD on site  
图 4. WAZC-VPD 现场测试环境

## 4.2. 测试内容及结果

### 4.2.1. 基础测试

以 WAZC-VPD 采集的 5 组固定场景测试图像为直接依据,完成“提取并计算基础测试数据”“建立图像与数据的追溯关系”“验证基础功能达标情况”三项测试目标。图 5 给出了 3 组“广角 + 细节”图像数据示例。

① 提取并计算基础测试数据。从 WAZC-VPD 采集的图像及对应图像采集日志中,获取“全景覆盖视场角、细节捕捉分辨率、双模式联动延迟、目标识别准确率”四类关键数据,具体通过“图像中已知尺寸目标反推视场角”“测量细节目标像素换算分辨率”“日志记录模式切换时间”“图像识别结果计算准确率”等方式完成。

② 建立图像与数据的追溯关系。明确每一项测试数据对应的具体图像来源,例如“分拣中心 118°视场角”对应图 5(a)分拣中心广角图,“0.2 cm 级分辨率”对应图 5(b)细节图(机位号“20”),确保数据可溯源。

③ 验证基础功能达标情况。通过上述数据判断装置的全景覆盖、细节捕捉、模式联动、目标识别四项基础功能是否符合预设标准。

本基础测试内容包含“全景覆盖能力、细节捕捉能力、双模式联动效果、目标识别能力”四项,具体操作流程与工具如下:

① 全景覆盖能力测试(验证 120°视场角)。选取分拣中心、物流园区道路、乡间三个典型场景,以图像中已知距离的目标为参考(分拣中心内 12 m 处的工作机台、物流园区道路 70 m 远处的汽车车牌、乡间 110 m 远处的小桥),使用图像 2 分析软件测量目标在广角图像中的成像尺寸,结合目标实际尺寸与拍摄距离,通过三角几何关系反推装置的实际视场角;测试过程中,需确保广角图像完整覆盖各场景的核心区域(分拣中心的机台与作业通道、园区道路的行车道与步道、乡间的小桥与树木),避免视野盲区。

② 细节捕捉能力测试(验证 0.2 cm 级分辨率)。针对上述场景中需重点识别的局部目标, 切换装置至变焦模式拍摄细节图像: 分拣中心内选取 12 m 处的机位编号“20”、园区道路 70 m 处停驶汽车的车牌、乡间 110 m 处小桥栏杆和雨伞; 使用图像分析软件测量这些目标在细节图像中的像素尺寸, 结合目标实际尺寸与拍摄距离, 换算装置的实际分辨率, 判断是否达到 0.2 cm 级标准。

③ 广角 - 变焦双模式联动效果测试。在拍摄每组“广角图像 + 细节图像”(如分拣中心广角图 5(a) 与细节图 5(b) (机位编号“20”)、园区道路广角图 5(c) 与车牌细节图 5(d)) 时, 调取装置的图像拍摄日志, 记录从“广角模式开启并完成拍摄”到“变焦模式切换完成并拍摄细节图像”的时间间隔, 该间隔即为双模式联动延迟; 测试需重复 3 次, 取平均值作为最终结果。

④ 目标识别能力测试。针对细节图像中的关键目标(分拣中心的机位编号“20”、园区道路上汽车车牌), 使用 OCR (Optical character recognition) 光学字符识别工具提取图像中的文字信息, 统计识别正确的次数与总测试次数的占比, 即为目标识别准确率; 对于乡间场景中的远处小桥栏杆/雨伞, 通过人工观察细节图像中目标纹理的清晰程度(是否存在模糊、断连), 辅助验证细节识别效果; 测试过程中使用的关键工具包括图像分析软件、OCR 识别工具及装置操作日志系统, 所有工具均提前完成校准, 确保数据准确性。



Figure 5. Example of “wide-angle + detail” image data  
图 5. “广角 + 细节” 图像数据示例

测试结果列于表 1。从表中可以看出，图像与数据形成完整互证，WAZC-VPD 装置全景覆盖、细节捕捉、双模式联动及目标识别功能均达标，所有场景视场角偏差控制 10%以内，分辨率优于预设标准 2 倍以上，联动延迟最长仅 1.0 s，目标识别准确率达 95%以上。

**Table 1.** Basic test content and results

**表 1.** 基础测试内容及结果

测试指标	测试结果	说明
全景覆盖视场角	视场角 120° 偏差均≤10%	采集的“广角 + 细节”图像数据显示，其中，园区(图 5(c))、乡间场景视场角实测值与 120°宣称值偏差均≤5%~8%
细节捕捉分辨率	均优于 0.5 cm 级	其中，分拣中心 0.20 cm 级(图 5(b))、园区 0.20 cm 级(图 5(d))、乡间 0.25 cm 级(图 5(f))
广角 - 变焦 双模式联动延迟	广角→细节延迟 ≤ 1.0 s	其中，分拣中心 0.8 s (图 5(a)→5(b))、园区 0.7 s (图 5(c)→图 5(d))
目标识别准确率	≥95%	其中，分拣中心机位编号“20”(图 5(b))、园区车牌识别率 100%(图 5(d))

#### 4.2.2. 专项测试(货物异常状态智能识别)

验证 WAZC-VPD 对“货物倾斜”“条码模糊”两类核心异常状态的智能识别能力，通过量化识别准确率和响应时间，判断装置能否快速、准确发现货物异常，证明其在物流货物监管中的实用价值。具体测试过程如下：

1) 货物倾斜识别测试：测试场景选取分拣中心工作机台，使用 3 组规格统一的标准货箱(长 × 宽 × 高 = 60 cm × 40 cm × 50 cm)，通过电子倾角仪调整货箱倾斜角度，分别设置为 5°、10°、15°(模拟物流场景中货箱堆叠倾斜的常见角度)；将装置固定在机台正前方 20 m 处，切换至变焦模式拍摄货箱细节图像，每次拍摄后启动装置的智能识别算法，记录“图像采集完成”到“输出‘货物倾斜’判定结果”的过程，统计不同倾斜角度下识别正确的次数，计算识别准确率；测试过程中确保货箱摆放位置与装置拍摄角度的匹配性，沿用细节成像精度(0.2 cm 级)，保障倾斜状态的识别基础。

2) 条码模糊识别测试。测试场景仍为分拣中心货架，制作 3 个模拟物流场景中常见模糊问题的条码样本，样本 1 为“污渍覆盖条码”(用灰色颜料覆盖条码边缘 10%区域)、样本 2 为“边缘褶皱条码”(将条码纸沿边缘折叠后展平，模拟运输中挤压褶皱)、样本 3 为“局部残缺条码”(裁剪条码右侧 15%区域，模拟条码磨损)；将条码样本贴在货架上，装置在 10 m 处通过变焦模式拍摄条码细节图像，使用专业条码识别软件读取图像中的条码信息，记录每次识别的成功与否；重复测试 5 次，统计每个样本的识别成功次数，计算识别成功率；测试中使用图像标注工具标记条码模糊区域，辅助分析算法对不同模糊类型的适配能力。

测试结果列于表 2。从表中可以看出，WAZC-VPD 对物流场景核心货物异常状态识别效果显著：货物倾斜识别准确率超 93%，即使 5°的轻微倾斜也能实现 92%的识别率；模糊条码识别成功率超 92%，仅局部残缺条码的识别率略低(88%)，但仍高于 85%的预设标准；且两类异常识别的平均响应时间 ≤ 1.2 s，可快速触发预警。

**Table 2.** Special test content and results

**表 2.** 专项测试内容及结果

异常类型	测试指标	测试结果
货物倾斜	识别准确率	5°倾斜识别率 92%、10°识别率 98%、15°识别率 100%
条码模糊	识别成功率	污渍条码 95%、褶皱条码 93%、残缺条码 88%
整体异常识别	平均响应时间	从成像到输出异常结果平均耗时 ≤ 1.2 s

### 4.2.3. 移动目标响应能力测试

为评估 1 秒联动延迟对动态场景的实际影响,在分拣中心部署电动滑轨平台,搭载标准货箱以 0.5 m/s、1.0 m/s 和 2.0 m/s 三种速度横向匀速移动。WAZC-VPD 部署于 15 米外,启动广角监测模式,一旦检测到运动目标即触发变焦追踪。系统记录每次抓拍是否成功(以能否清晰识别货箱标签 OCR 为准)。测试结果表明:在 0.5 m/s 时抓拍成功率达 98%,1.0 m/s 时为 92%,2.0 m/s 时降至 85%;但所有成功抓拍图像均满足 OCR 识别所需的清晰度要求(字符分辨率  $\geq 200$  dpi),说明 1 秒内的联动延迟在典型城市物流移动场景中具备实用性。

## 5. 结论

本文针对城市物流视觉感知场景中全景与细节失衡、计算成本约束、传输效率不足等问题,设计了一种广角-变焦协同的视觉感知装置(WAZC-VPD),通过固定场景与专项测试验证了其设计合理性与有效性。

研究表明,WAZC-VPD 基于“双镜成像+三核计算+5G 传输”架构实现了多维度改进设计:① 广角-变焦双镜头协同解决全景与细节兼顾难题,120°视场角与 1.0 cm 级细节捕捉的联动响应  $\leq 1$  秒;② 基于 RK3506J 处理器的三核异构硬件设计,将硬件成本控制为高端工业处理器的 1/2,同时通过轻量化算法优化,目标识别时延  $\leq 80$  ms,满足现场侧实时需求;③ 集成的 5G 模块可稳定实现 200 Mbps 速率的视频数据传输,网络波动下仍能保障数据传输连续性;④ 专项测试证实其对“货物错放”“掉落”等异常识别准确率均  $\geq 92\%$ ,满足场景监管精度要求优异。

后续的研究,可进一步优化双镜头在极端光照(如强光直射、夜间低照度)下的成像效果,完善现场测试场景与样本覆盖(如增加暴雨、大雾等恶劣天气测试),并探索多装置协同感知模式,以提升对广域分散物流场景的整体监控能力。

## 基金项目

本文得到国家自然科学基金(62563030)、江西省自然科学基金(20242BAB25049)的资助。

## 参考文献

- [1] Zhou, W., Zhao, L., Zhang, R.Y., Cui, Y.F., Huang, H.P., Qie, K. and Wang, C. (2024) Vision Technologies with Applications in Traffic Surveillance Systems: A Holistic Survey. arXiv: 2412.00348.
- [2] 华经产业研究院. 2025-2031 年中国物联网安防行业发展前景预测及投资方向研究报告[EB/OL]. <https://www.huaon.com/channel/security/1076702.html>, 2025-07-18.
- [3] 湖南贝哲斯信息咨询有限公司. 2023 年全球与中国人工智能视频监控行业前景预测报告[EB/OL]. <https://www.shangyexinzhi.com/article/9990378.html>, 2023-07-18.
- [4] Lu, Y., Wang, Z., Liu, M., Wang, H. and Wang, L. (2023) Learning Spatial-Temporal Implicit Neural Representations for Event-Guided Video Super-Resolution. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, 17-24 June 2023, 1557-1567. <https://doi.org/10.1109/cvpr52729.2023.00156>
- [5] 徐梦溪,沈克永,涂宏斌,等.一种视频与物流业务数据融合的效率-安全双目标协同优化系统设计模式[J].软件工程与应用,2025,14(4):897-905.
- [6] 徐梦溪,刘妹怡,程晓玲,等.基于多光谱成像与边缘计算的物流安全风险预警模式及系统实现[J].计算机科学与应用,2025,15(10):85-96.
- [7] 曹行健,张志涛,孙彦赞,等.面向智慧交通的图像处理与边缘计算[J].中国图象图形学报,2022,27(6):1743-1767.
- [8] Cheng, X., Xu, M., Yan, X., Yang, Y., Xu, Y. and Ruan, Y. (2024) A Design Pattern of IAPVS Platform Based on Distributed Edge Computing. *Journal of Physics: Conference Series*, **2732**, Article ID: 012001. <https://doi.org/10.1088/1742-6596/2732/1/012001>

- [9] Zhang, X., Yoon, J., Bansal, M. and Yao, H. (2024) Multimodal Representation Learning by Alternating Unimodal Adaptation. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 27446-27456. <https://doi.org/10.1109/cvpr52733.2024.02592>
- [10] Zou, Z., Chen, K., Shi, Z., Guo, Y. and Ye, J. (2023) Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*, **111**, 257-276. <https://doi.org/10.1109/jproc.2023.3238524>
- [11] 王晨, 周威, 章世祥. 一种特征融合的视频事故快速检测方法[J]. *交通运输工程与信息学报*, 2022, 20(1): 31-38.
- [12] 徐梦溪, 罗中华, 程晓玲, 等. 基于双镜头视野协同成像的无线视频传感器网络构建[J]. *传感器技术与应用*, 2024, 12(1): 54-62.
- [13] 徐梦溪, 刘姝怡, 刘梓莹, 等. 一种基于 DVFS 特性曲线的异构计算单元低功耗协同调度方法[J]. *软件工程与应用*, 2025, 14(6): 1219-1230.