

## 面向高效巡检任务推理的边缘辅助无人机机载视频压缩与传输

杨鹏<sup>1</sup>, 梁雨欣<sup>1</sup>, 孔雨新<sup>1</sup>, 刘鸣柳<sup>2</sup>

(1. 华中科技大学电子信息与通信学院, 湖北 武汉 430074; 2. 国网湖北省电力有限公司电力科学研究院, 湖北 武汉 430077)

**摘要:** 面向无人机 (UAV, unmanned aerial vehicle) 巡检任务采集数据的高效推理, 研究了移动边缘计算 (MEC, mobile edge computing) 辅助的 UAV 机载视频中感兴趣区域 (RoI, region of interest) 提取和高效传输问题, 以提升 UAV 在一般巡检任务中采集和分析数据性能。由于 UAV 机载计算资源有限, 提出了一种基于类激活映射 (CAM, class activation mapping) 的轻量级 RoI 提取方法, 以快速定位包含潜在目标的区域, 并将这些 RoI 高效卸载至边缘服务器进行推理。为应对 UAV 动态轨迹与网络环境的变化, 进一步通过自适应 RoI 边界框选择算法对 UAV 采集的 RoI 进行有效筛选, 并利用量化参数 (QP, quantization parameter) 自适应调整机载视频编码质量, 以进一步压缩传输数据量。在此基础上, 构建了一个联合 RoI 边界框选择与自适应编码配置的优化问题, 并采用启发式算法求解该优化问题。实验结果表明, 该方案能够有效提升检测精度, 减少传输数据量, 并显著降低系统时延, 在基于 UAV 的一般巡检任务中表现出优异的性能。

**关键词:** 无人机通信; 移动边缘计算; RoI 提取; 视频编码

**中图分类号:** TP183

**文献标志码:** A

**doi:** 10.11959/j.issn.2096-3750.2024.00418

## Edge-assisted UAV onboard video compression and transmission for efficient inference of patrolling tasks

YANG Peng<sup>1</sup>, LIANG Yuxin<sup>1</sup>, KONG Yuxin<sup>1</sup>, LIU Mingliu<sup>2</sup>

1. School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China

2. Electric Power Research Institute, State Grid Hubei Electric Power Co., Ltd., Wuhan 430077, China

**Abstract:** The problem of region of interest (RoI) extraction and transmission of video frames captured in edge-assisted unmanned aerial vehicle (UAV) systems was investigated to improve the inference performance of patrolling tasks. Due to the limited UAV onboard computational resources, a lightweight RoI extraction method based on class activation mapping (CAM) was proposed, which was able to rapidly locate areas containing patrolling targets. Those RoIs were then transmitted to edge servers for further processing. To address the challenges from dynamic UAV trajectories and fluctuating network conditions, the RoIs collected by UAVs were properly chosen through an adaptive RoI box selection algorithm, followed by adaptive configuration of quantization parameters (QP) of video codec, in order to further compress the transmitted data volume. A joint optimization problem was thus formulated for RoI box selection and adaptive coding configuration, which was solved via a heuristic algorithm. Experimental results demonstrate that, the proposed approach can effectively improve the detection accuracy of patrolling tasks, reduce data transmission volume, and significantly lower system latency, indicating great potential in UAV-based patrolling applications.

**Key words:** UAV communication, mobile edge computing, RoI extraction, video coding

收稿日期: 2024-10-08; 修回日期: 2024-10-19

通信作者: 杨鹏, yangpeng@hust.edu.cn

基金项目: 国家自然科学基金青年科学基金项目 (No. 62001180); 中国科协青年人才托举工程项目 (No. 2022QNRC001)

**Foundation Items:** The Young Scientists Fund of the Nation Natural Science Foundation of China (No. 62001180), The Young Elite Scientists Sponsorship Program by CAST (No. 2022QNRC001)

## 0 引言

随着物联网技术的迅猛发展,越来越多的智能设备接入互联网,产生了海量数据。这些数据具有高频次、多样性和实时性的特点,要求网络基础设施能够支持高速数据传输、低时延处理和及时响应。然而,传统云计算架构由于其集中式处理模式,在面对如此庞大的数据量时,往往会面临带宽消耗大、处理时延高以及数据传输瓶颈等问题。这些局限性使得传统云计算架构在需要超低时延和高性能计算的应用场景中难以满足需求。因此,移动边缘计算(MEC, mobile edge computing)作为一种新兴的分布式计算架构应运而生<sup>[1-2]</sup>。通过将计算、存储和网络资源部署到网络边缘,MEC能够将终端设备的计算任务卸载至离用户最近的边缘服务器或基站,实现本地化的数据处理,从而显著提升计算效率,减少数据传输时延,满足物联网应用对实时性和低时延的需求<sup>[3]</sup>。

在MEC架构下,无人机(UAV, unmanned aerial vehicle)巡检任务成为其应用的一个重要领域。相比于传统的人力巡检和地面固定设备巡检,边缘计算辅助的UAV巡检在成本、效率和安全性等方面具有显著优势。首先,人工巡检往往需要高昂的人力成本和专业的任务培训,特别是在偏远地区或危险地区执行任务时,不仅面临人力资源短缺和流失风险,还可能存在严重的安全隐患。此外,人工巡检和地面固定设备巡检的范围通常有限,人工难以在短时间内完成大面积、高频次的实时监测任务,而固定设备巡检的性能受限于设备视角的局限性和覆盖范围的不足<sup>[4]</sup>。相比之下,集成摄像头、激光雷达、红外传感器等多样化传感器于一体的UAV不仅能够凭借其高度的机动性和灵活性在复杂的地理环境下执行实时的数据采集、监测和巡检任务,避免地面障碍物的影响,还能在MEC的辅助下突破机载计算资源的限制,实现更高效和稳定的数据处理与决策支持,为河道、输电线路和交通运输线路等大型基础设施的智能化巡检赋能<sup>[5-9]</sup>。

尽管边缘辅助的UAV巡检技术应用前景广阔,但如何充分利用其优势实现低时延的实时巡检任务,仍然面临着一系列技术挑战。首先,UAV高空大范围监测时捕获的图像往往密布小目标,这些小目标在图像中不够显著,易被检测算法忽略或误

判,从而降低整体检测精度<sup>[10-12]</sup>。UAV自身计算和存储资源往往有限,难以在本地完成小目标的高精度实时检测任务<sup>[13]</sup>。因此,UAV通常需要将采集到的图像或视频数据传回边缘服务器进行处理。然而,这些数据中包含大量冗余信息,如不必要的背景数据等<sup>[14]</sup>,亟须开发一种能在UAV上高效运行且保证基于巡检数据的推理任务准确性的轻量级感兴趣区域(RoI, region of interest)提取方法。其次,UAV的高机动性带来了通信链路不稳定的问题,尤其是随着UAV逐渐远离基站,其与基站间的通信速率会随着飞行轨迹、飞行距离,以及与地面环境的变化而波动<sup>[15]</sup>。如果将未处理的原始视频数据直接回传,很难在动态网络环境下满足实时检测的需求。因此,需要开发一种自适应的视频帧压缩编码方案,以优化UAV巡检中数据传输流程<sup>[16-17]</sup>。

针对上述需求,现有视频分析领域的相关研究在图像分割技术和视频编码技术等方面已经提出了多种解决方案<sup>[18-21]</sup>。然而,这些研究未能充分解决UAV在巡检任务中的特殊挑战,如小目标检测和通信链路的动态变化<sup>[22-25]</sup>。此外,现有的边缘辅助UAV巡检系统研究多集中于通信层面的优化,如带宽分配、能耗管理和飞行路径规划等<sup>[26-30]</sup>,对UAV巡检采集数据(如视频)的处理优化和高效传输缺乏系统性的研究与考量。为此,本文提出了一种面向高效巡检任务推理的边缘辅助无人机机载视频压缩与传输策略,以在服务时延与检测质量之间实现有效的权衡。

在UAV执行巡检任务时,受限于机载资源的有限性以及网络的波动性,传统视频处理方案难以满足复杂的RoI提取任务和大规模采集传输数据的需求。因此,设计出高效且轻量级的RoI提取方案显得尤为关键。为解决这些问题,本文提出了一种基于类激活映射(CAM, class activation mapping)技术的RoI提取策略。该策略利用CAM技术从视频帧中智能识别并提取关键区域,从而避免传输无关的背景数据,显著减少了数据传输的时延,同时缩小了目标检测范围,提升了预测的准确性和处理效率。与传统关键区域提取方法相比,CAM技术更适用于资源受限且对计算效率和检测精度要求严格的UAV巡检任务。传统方法(如帧差法和背景去除法)在特定环境下表现较好,但在动态巡检场景中存在显著的局限性。帧差法通过对连续帧的差

分计算检测运动目标，适用于场景变化较少的简单环境。然而，在背景复杂或动态变化较大的场景中，帧差法容易产生较高的误检率。背景去除法则通过建立静态背景模型识别前景目标，在固定摄像机环境中表现较好，但当无人机移动导致环境背景变化时，背景去除法的效果往往会显著下降。相较之下，CAM技术依赖卷积神经网络生成的特征图，能够更加智能地聚焦于视频帧中的关键目标区域，而无须依赖固定背景模型或场景的静态特征。此外，CAM技术不需要像素级标注，并且模型轻量化，计算开销较低，能够在UAV受限的计算环境下高效运行。输入原始视频帧与基于CAM的RoI提取视频帧的目标检测实验对比结果如图1所示，基于CAM的RoI提取方法在检测精度方面显著优于直接处理原始视频帧的方式。这一方法在减少数据传输时延和提升检测精度方面展现了显著优势，成为资源受限且需快速响应的UAV巡检任务中的理想选择。

RoI提取完毕后，需要将这些信息回传给边缘服务器进行后续处理。在UAV巡检任务中，由于视频帧包含大量像素信息，回传过程往往需要消耗大量带宽资源。为了减少带宽资源的消耗，通常需要在信息回传前对视频帧进行压缩。然而，压缩可能会导致信息丢失，进而影响检测效果。因此，在本研究中，为不同传输视频帧选择最佳的量化参数(QP, quantization parameter)等级，在保证检测精度的前提下进行高效压缩。具体而言，QP值越小，即量化步长越小，量化越精细，保留的图像细节越多，从而提升检测精度。同时，这也意味着更大的传输数据量，增加了带宽消耗。相反，较大的QP值虽然能减少传输带宽的消耗，但会丢失更多的图像细节，导致检测精度下降。因此，在实际UAV巡检任务中，合理选择QP值可以在减少视频

回传带宽消耗的同时，保证较高任务推理性能。

综上所述，本文所提方案充分考虑UAV巡检过程中采集数据和传输链路的高动态特点，提出RoI的自适应提取与视频帧的最佳压缩编码策略。具体而言，UAV首先利用CAM技术在每个时隙内确定视频帧中的RoI，然后动态地调整需要传输的RoI边界框及其编码参数。此外，本文进一步将RoI边界框的选择和编码等级的自适应配置建模为一个非线性广义分配问题。为了求解这一问题，本文提出了高效的启发式算法以优化RoI提取及编码配置选择，进而达到UAV巡检任务中检测质量和时延的最佳平衡。

## 1 系统模型和问题构建

### 1.1 系统概述

为了在动态网络环境下，实现高效、精确的视频帧传输和巡检推理任务，本文提出了边缘辅助UAV实时巡检视频传输和分析系统，如图2所示，该系统以河道巡检为例，由4个主要模块组成：视频帧特征提取模块、基于轻量级CAM的RoI提取模块、RoI边界框选择模块，以及自适应RoI边界框编码配置选择模块。

首先，本系统通过视频帧特征提取器，使用固定模式的特征分割方法对视频帧进行初步处理。接下来，本系统利用基于轻量级CAM的RoI提取技术剔除无关的背景信息，使视频帧聚焦目标物体所在区域。为了提高传输和检测效率，本系统在传输RoI边界框信息前会先依据上一次目标检测的反馈结果，从提取到的所有信息中移除无效的RoI边界框，确保仅传输包含高置信度目标的区域。在此基础上，本系统还通过自适应的RoI边界框编码配置选择策略，根据当前UAV的航迹和带宽资源，动态地调整每帧传输的RoI边界框的视频编码配置，

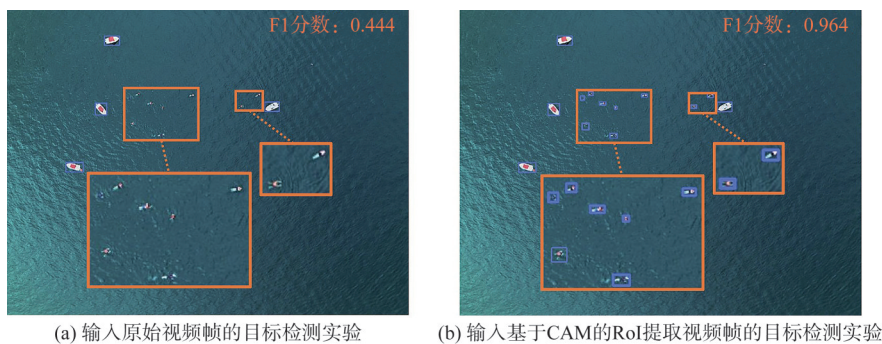


图1 输入原始视频帧与基于CAM的RoI提取视频帧的目标检测实验对比结果

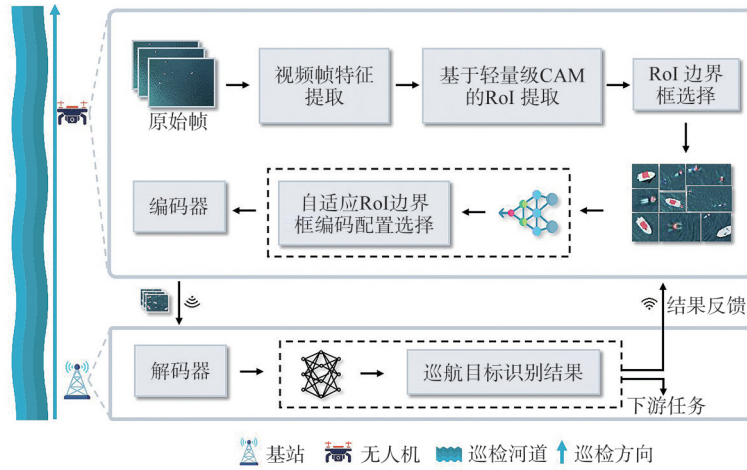


图2 边缘辅助UAV实时巡视视频传输和分析系统

进而平衡传输成本与检测精度。在本节其余部分中，将介绍各组成模块的具体细节。

### 1.2 视频帧特征提取

CAM技术在无须逐像素级标签的条件下拥有出色的目标定位能力。具体而言，通过卷积神经网络最后一层的特征图与分类层权重的结合，CAM可以生成类激活热图，显示模型在分类时所重点关注的图像区域，进而表征视频帧的特征。

对于一张维度为  $Z \in \mathbf{R}^{P \times (A \times B)}$  的视频帧，其通道数为  $P$ ，空间分辨率为  $A \times B$ 。在视频帧特征提取过程中，通过卷积神经网络最后一层提取出的特征图可以表示为  $H_p(x, y)$ ，即通道  $p$  在位置  $(x, y)$  上的激活值。接下来，通过全局平均池化将每个通道特征进行空间维度上的聚合，可以得到该通道的平均激活值  $h_p$ ，即

$$h_p = \frac{1}{A \times B} \sum_{x=1}^A \sum_{y=1}^B H_p(x, y) \quad (1)$$

这些平均激活值代表了每个通道特征的整体响应强度，汇总了该通道在空间维度上的激活情况。

对于类别  $k$  的类激活图  $L_k(x, y)$ ，记为

$$L_k(x, y) = \sum_p \theta_p^k H_p(x, y) \quad (2)$$

其中， $\theta_p^k$  代表通道  $p$  对类别  $k$  的权重，反映了通道  $p$  在类别  $k$  上的重要性。这些权重来自于分类层的权重矩阵，体现了网络在进行类别  $k$  的预测时，某个通道特征的影响力。 $L_k(x, y)$  显示了图像中每个像素位置  $(x, y)$  对类别  $k$  的贡献值。较高的  $L_k(x, y)$  值表明该区域更有可能包含与类别  $k$  相关的目标物体，即视频帧的特征更为明显。

### 1.3 基于CAM的RoI提取

不失一般性，将整个视频处理过程划分为若干个长度为  $\tau$  的时隙，记为  $\mathcal{T} = \{1, 2, \dots, T\}$ 。基于CAM的RoI提取和边界框选择的流程和处理结果如图3所示。图3(a)展示了某个时隙内未经处理的原始视频帧。在每个时隙内，系统首先采用  $2 \times 2$  的等比例划分方法对视频帧进行分割，并对每个子区域的特征图应用类激活映射技术，生成如图3(b)所示的热力图。该热力图利用颜色编码标识出潜在目标物体的位置，其中较高的激活值  $L_k(x, y)$  对应类别  $k$  相关目标出现的高概率区域，即热力图中的暖色区域。为了从热力图中提取感兴趣区域，设定了一个激活值阈值  $\epsilon$ ，将激活值大于阈值的区域设为1，其他区域设为0，以生成二值掩码  $M \{L_k(x, y) \geq \epsilon\}$ ，如图3(c)所示。基于热力图像素值的掩码生成过程可表示为

$$M(x, y) = \begin{cases} 1, & L_k(x, y) \geq \epsilon \\ 0, & L_k(x, y) < \epsilon \end{cases} \quad (3)$$

通过在原始视频帧上应用该掩码，可以去除视频帧的背景信息，并实现逐像素点地保留感兴趣区域的像素值，此过程可以表示为

$$I_{\text{RoI}}(x, y) = I_{\text{original}}(x, y) \times M(x, y) \quad (4)$$

其中， $I_{\text{original}}(x, y)$  为原始视频帧的像素值， $I_{\text{RoI}}(x, y)$  是掩码应用后的结果图像。通过这种方式，系统仅保留与目标物体相关的区域。掩码处理后的结果如图3(d)所示。

### 1.4 RoI边界框选择

从图3(b)的左半部分可以明显看出，CAM在处理背景区域时的定位性仍存在一定局限性，导致

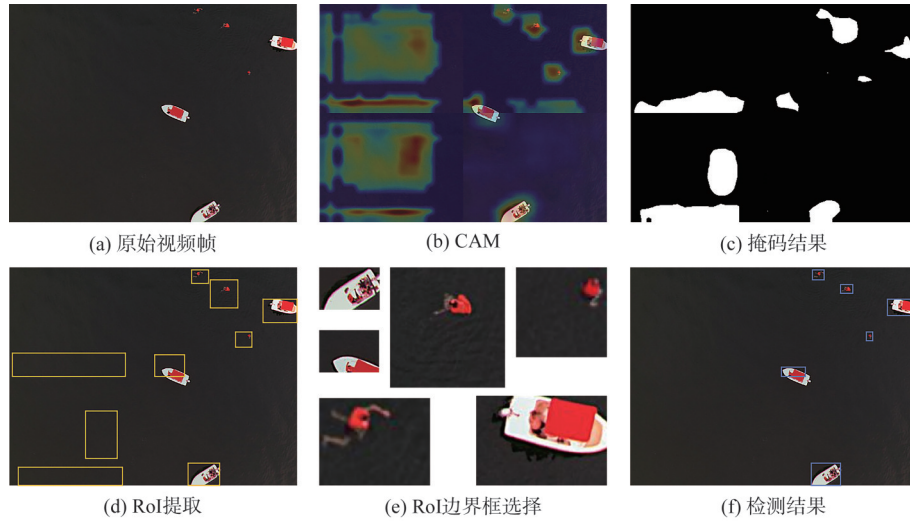


图3 基于CAM的RoI提取和边界框选择的流程和处理结果

部分提取的RoI边界框无效，即这些区域不包含任何目标物体。这类无效的RoI边界框如果传输至服务器，除了浪费带宽资源外，还会增加计算设备的处理时延。因此，在UAV巡检系统中，必须设计和采取动态的RoI边界框选择策略，以最大限度地减少无效边界框的传输和处理。

为了实现这一目标，UAV巡检系统需要动态选择有效的RoI边界框进行传输，以过滤不包含目标物体的边界框。将第 $t$ 个时隙内提取的所有RoI边界框表示为 $\mathcal{C}_t = \{c_{1,t}, c_{2,t}, \dots, c_{|\mathcal{C}_t|,t}\}$ ， $|\mathcal{C}_t|$ 为RoI边界框总数。本文设计了一种基于动态调整策略，该策略选择回传的RoI边界框可表示为集合 $\mathcal{M}_t = \{m_{1,t}, m_{2,t}, \dots, m_{|\mathcal{M}_t|,t}\}$ ，其中 $|\mathcal{M}_t|$ 为在第 $t$ 个时隙处理和传输的视频帧中提取到的RoI边界框总数。需要注意的是， $\mathcal{M}_t$ 是 $\mathcal{C}_t$ 的子集并且任意两个RoI边界框之间不应有重叠，可记为

$$\mathcal{M}_t \in \mathcal{C}_t, \forall t \in \mathcal{T} \quad (5)$$

$$m_{j,t} \cap m_{k,t} = \emptyset, \forall j, k \in |\mathcal{M}_t| \forall j \neq k, t \in \mathcal{T} \quad (6)$$

接下来，进一步将 $\mathcal{M}_t$ 中所有RoI边界框进行适当地放大，填满整张帧的画幅以便传输，如图3(e)所示。通过逐步优化RoI的选择，可以确保传输数据尽可能聚焦于包含目标物体的有效区域，从而能够避免带宽资源和计算资源的浪费。

### 1.5 自适应RoI边界框编码配置选择

在UAV巡检任务中，采用三维直角坐标系对UAV的飞行轨迹进行建模。假设基站位于二维坐标系的原点 $(0, 0)$ ，但具有一定的高度 $h$ ，UAV从基站出发，以固定高度 $H$ 和速度 $v$ 沿着巡检路线飞

行。由于巡检任务通常针对特定的基础设施，如河道、输电线路等，其地理位置往往是已知的，并且能够通过坐标表示。这些巡检线路通常表现为较长且相对线性的路径。因此，在这些场景下，UAV飞行路径可以简化为一条直线，便于后续的计算和分析。在第 $t$ 个时隙，UAV三维坐标可表示为 $(vt, 0, H)$ ，基站坐标则为 $(0, 0, h)$ ，且UAV飞行高度 $H > h$ 。基于欧几里得距离公式，基站与UAV之间在时隙 $t$ 时刻的传输距离 $D_t$ 可表达为

$$D_t = \sqrt{(vt)^2 + (H - h)^2} \quad (7)$$

基于这一距离公式，进一步推导出UAV与基站之间的无线传输速率 $r_t$ ，表示为

$$r_t = B \log \left( 1 + \frac{pGD_t^{-\alpha}}{\gamma B} \right) \quad (8)$$

其中， $B$ 表示通信信道带宽， $G$ 是信道增益， $p$ 为UAV发射功率， $\gamma$ 代表链路中的背景噪声密度， $\alpha$ 为路径损耗指数，描述了信号随传播距离衰减的速率，通常在自由空间中取值为2，在复杂环境下可能取值为3或更大的数值。在该场景中，由于UAV与基站之间的传输距离不断变化，视频回传的网络条件是动态波动的。如果采用固定的视频帧编码配置，可能无法适应这种波动的网络环境，导致性能下降。具体而言，视频压缩程度的选择直接影响系统的传输时延和边缘节点上的处理性能。

当使用较高的压缩程度时，尽管可以显著降低视频传输的时延，但会导致视频质量下降，特别是在复杂场景中（如巡检场景存在大量小目标），检测精度会显著下降。如果选择较低的压缩程度，虽

然可以保持较高的视频质量，但随着 UAV 逐渐远离基站，视频数据量变大，传输时间增加，从而失去实时性。因此，为了在不同网络环境下实现实时视频传输和高效目标检测，UAV 巡检系统需要自适应地调整视频编码配置，以应对网络状况的波动。

在视频压缩过程中，假设有  $n$  个量化参数 QP 等级， $q_n \in \{0, 1, 2, \dots, n-1\}$ ，用于控制视频压缩程度。为了选择最适合当前时隙视频帧的压缩等级，定义二进制变量  $x_{i,n}^t$  来表示在时隙  $t$  中第  $i$  个 RoI 边界框对第  $n$  个 QP 的选择决策。向量  $\mathbf{Q}_t = [q_{1,t}, q_{2,t}, \dots, q_{|\mathcal{M}_t|,t}]$  表示当前时隙内所有传输的 RoI 边界框的量化参数选择。为保证每个边界框只选择一个 QP 等级  $q_n$ ，各 RoI 边界框的量化参数选择  $q_i^t$  需要满足以下约束条件

$$q_i^t = \sum_{n=1}^N x_{i,n}^t q_n, \forall i \in \mathcal{C}_t, t \in \mathcal{T} \quad (9)$$

$$\sum_{n=1}^N x_{i,n}^t = 1 \quad (10)$$

在整个实时 UAV 巡检视频传输和分析过程中，影响带宽资源消耗和检测精度的主要因素包括 RoI 边界框传输数量  $|\mathcal{M}_t|$  和  $\mathcal{M}_t$  对应的量化参数选择决策  $\mathbf{Q}_t$ 。不失一般性地，定义  $A_t(\mathcal{M}_t, \mathbf{Q}_t)$  为第  $t$  个时隙中的解析精度函数， $S_t(\mathcal{M}_t, \mathbf{Q}_t)$  表示编码后的视频传输数据大小。当  $q_i^t$  减小时，传输数据大小  $S_t(\mathcal{M}_t, \mathbf{Q}_t)$  增加，同时解析精度  $A_t(\mathcal{M}_t, \mathbf{Q}_t)$  也会随之提高；当  $|\mathcal{M}_t|$  减小时，传输数据大小  $S_t(\mathcal{M}_t, \mathbf{Q}_t)$  减小，同时解析精度  $A_t(\mathcal{M}_t, \mathbf{Q}_t)$  也会随之降低。

## 1.6 问题构建

将视频帧传输和处理的总时延分为 4 个部分。首先， $l_1$  表示基于 CAM 的 RoI 提取  $\mathcal{C}_t$  和选择  $\mathcal{M}_t$  的处理时间，该处理步骤在 UAV 上本地运行。由于 CAM 的轻量化特性， $l_1$  可被视为常数。其次， $l_2(\mathcal{M}_t, \mathbf{Q}_t)$  为编码时间，当 RoI 边界框数量  $|\mathcal{M}_t|$  增加时，编码时间也会相应增长，且当  $\mathbf{Q}_t$  选择的 QP 越小，量化越精细，编码时间越长，可以表示为

$$l_2(\mathcal{M}_t, \mathbf{Q}_t) = \sum_{i=1}^{|\mathcal{M}_t|} \frac{B(m_i)}{r_{q_{m_i}^t}} \quad (11)$$

其中， $B(m_i)$  是视频帧  $m_i$  的原始数据量， $r_{q_{m_i}^t}$  是在量化等级  $q_{m_i}^t$  下的编码速度。

第三部分时延  $l_3(\mathcal{M}_t, \mathbf{Q}_t)$  为视频从摄像头传输

至边缘服务器的时间，可以表示为

$$l_3(\mathcal{M}_t, \mathbf{Q}_t) = \sum_{i=1}^{|\mathcal{M}_t|} \frac{S(m_i)}{r_t} \quad (12)$$

最后， $l_4$  表示视频分析任务在边缘节点上的处理时间，与待处理数据的规模相关。与 UAV 采集的原始数据相比，基于 CAM 技术获得的 RoI 数据量小。因此，边缘所需处理的数据量相对较小，处理时间的波动范围也较为有限，故可以将其视为一个常量。综上所述，第  $t$  个时隙的视频处理时延  $L_t(\mathcal{M}_t, \mathbf{Q}_t)$  可表示为

$$L_t(\mathcal{M}_t, \mathbf{Q}_t) = l_1 + l_2(\mathcal{M}_t, \mathbf{Q}_t) + l_3(\mathcal{M}_t, \mathbf{Q}_t) + l_4 \quad (13)$$

由此可见，时延由 RoI 提取与选择时间、编码时间、传输时间和固定分析时间组成。通过仅传输每个视频帧的有效 RoI 边界框，系统能够显著减少传输数据量，并通过调整  $\mathbf{Q}_t$  来降低所需比特率，从而减少传输时间。这一自适应编码配置策略在巡检任务中极大地提升了带宽资源的利用效率，同时保证了实时性和检测精度的平衡。

本文目标是在动态网络环境下实现实时 UAV 巡检视频传输和分析系统的精度和时延权衡。因此，该问题可表示为

$$\begin{aligned} \mathbf{P} : \max_{\mathcal{M}_t, \mathbf{Q}_t} \sum_{t \in \mathcal{T}} A_t(\mathcal{M}_t, \mathbf{Q}_t) - \mu L_t(\mathcal{M}_t, \mathbf{Q}_t) \\ \text{s.t. 式(3), 式(4), 式(7), 式(8)} \\ L_t(\mathcal{M}_t, \mathbf{Q}_t) \leq \tau, \forall t \in \mathcal{T} \end{aligned} \quad (14)$$

其中， $\mu$  为平衡分析精度  $A_t(\mathcal{M}_t, \mathbf{Q}_t)$  和系统时延  $L_t(\mathcal{M}_t, \mathbf{Q}_t)$  的非负权参数，可根据用户的偏好设置或多准则决策理论进行定制<sup>[31]</sup>。优化问题  $\mathbf{P}$  中共有  $|\mathcal{M}_t|$  个需要传输的 RoI 边界框和  $N$  个可选的量化参数。该优化问题属于非线性广义分配问题，其性质为 NP 难的。因此，随着视频中 RoI 数量或量化等级的增加，使用暴力求解法将变得不切实际。为了解决该问题，设计了如下所述的轻量级算法。

## 2 算法设计

为了解决上述优化问题  $\mathbf{P}$ ，提出了基于启发式算法的自适应 RoI 边界框及编码配置选择算法，如算法 1 所示。算法 1 中，步骤 1 提出了一种基于反馈的动态调整策略来选择需要传输的 RoI 边界框。该策略利用视频帧的连续性，根据上一帧的检测结果对当前帧的 RoI 选择进行调整。如果上一帧中某个 RoI 被检测为无效区域，则在后续帧中系统会降

低与该区域交并比较大的 RoI 边界框的传输优先级，减少其被传输的概率。这样可以避免不必要的数据传输，从而降低系统整体时延。该步骤的时间复杂度为  $O(|\mathcal{M}_{t-1}||\mathcal{C}_t|)$ 。

**算法 1** 基于启发式算法的自适 RoI 边界框及编码配置选择算法

输入

第  $t-1$  时隙传输的 RoI  $\mathcal{M}_{t-1} = \{m_{1,t-1}, \dots, m_{|\mathcal{M}_{t-1}|,t-1}\}$

第  $t-1$  时隙检测结果  $\text{Result}_{t-1} = [R_{m_{1,t-1}}, \dots, R_{m_{|\mathcal{M}_{t-1}|,t-1}}]$

量化参数范围  $q_n \in \{0, 1, 2, \dots, n-1\}$

第  $t$  时隙提取的原始 RoI  $\mathcal{C}_t = \{c_{1,t}, c_{2,t}, \dots, c_{|\mathcal{C}_t|,t}\}$

输出

传输 RoI 边界框集合  $\mathcal{M}_t = \{m_{1,t}, m_{2,t}, \dots, m_{|\mathcal{M}_t|,t}\}$

量化参数选择决策向量  $\mathbf{Q}_t = [q_{1,t}, q_{2,t}, \dots, q_{|\mathcal{M}_t|,t}]$

**步骤 1:** #基于反馈的动态调整策略

初始化边界框交并比阈值  $\sigma_1$

初始化无效边界框检测阈值  $\sigma_2$

初始化丢弃的无效边界框集合  $\text{Evict}_t = \{\}$

for all  $i \in \mathcal{C}_t$  do

  for all  $j \in \mathcal{M}_{t-1}$  do

    if  $i \cap j \geq \sigma_1$  do

      if  $R_j \leq \sigma_2$  and  $i \notin \text{Evict}_t$  do

$\text{Evict}_t.$  Add( $i$ )

      end if

    end if

  end for

end for

$\mathcal{M}_t = \mathcal{C}_t - \text{Evict}_t$

**步骤 2:** #基于启发式算法的编码配置选择

设置概率  $p_c$  和概率  $p_m$

初始化计数器  $c = 1$

#生成  $K$  个初始个体

for  $k = 1$  to  $K$  do

  生成  $\mathbf{Q}_k^0 = [q_{1,k}^0, q_{2,k}^0, \dots, q_{|\mathcal{M}_t|,k}^0], \forall q_{x,k} \in q_n$

  计算  $F_k^0 = \sum_{i \in \mathcal{T}} A_i(\mathbf{Q}_k^0) - \mu L_t(\mathbf{Q}_k^0)$

  计算  $P_k^0 = \frac{F_k^0}{\sum_{i=1}^K F_i^0}$

end for

while  $c \leq \text{MAX}$  do

  计算  $P_k^c = \frac{F_k^c}{\sum_{i=1}^K F_i^c}, P^c = [P_1^c, P_2^c, \dots, P_K^c]$

  筛选  $\mathbf{Q}^{c+1} \leftarrow \text{Choose}(\mathbf{Q}^c, P^c)$

  for  $k = 1$  to  $K$  do

$\mathbf{Q}_k^c \leftarrow \text{Crossover}(\mathbf{Q}_k^c, \mathbf{Q}_k^c, p_c)$

$\mathbf{Q}_k^c \leftarrow \text{Mutation}(\mathbf{Q}_k^c, p_m)$

    计算  $F_k^c = \sum_{i \in \mathcal{T}} A_i(\mathbf{Q}_k^c) - \mu L_t(\mathbf{Q}_k^c)$

  end for

  更新计数器  $c = c + 1$

end while

  令  $\mathbf{Q}_t = \text{argmax}(F(\mathbf{Q}_t))$

在第  $t$  个时隙，基于上述反馈策略，可以获得选择传输的 RoI 边界框集合  $\mathcal{M}_t$ 。在此基础上，继续求解优化问题中关于量化参数选择的部分，需要为  $\mathcal{M}_t$  中的每个元素选择最佳的量化参数等级  $\mathbf{Q}_t$ 。然而，即使已经确定了  $\mathcal{M}_t$ ，该问题的搜索空间依然十分庞大，其复杂度为  $O(|\mathcal{M}_t|^n)$ 。因此，设计了一种基于启发式算法的轻量级求解策略来解决这一问题，如算法 1 的步骤 2 所示。每个个体代表一种特定的编码配置选择方案。通过定义优化目标，计算每个个体的适应度，并依据适应度对配置决策进行迭代优化。整个迭代过程持续进行，直到满足预设的停止条件，即计数器达到预先设定的阈值 MAX 或结果收敛。该基于启发式算法的自适应 RoI 编码配置选择算法的目标是找到能够在时延和精度之间取得最佳平衡的量化参数组合  $\mathbf{Q}_t$ ，其优化目标可表示为

$$F = \sum_{i \in \mathcal{T}} A_i(\mathbf{Q}_t) - \mu L_t(\mathbf{Q}_t) \quad (15)$$

该步骤的时间复杂度为  $O(cK|\mathcal{M}_t|)$ 。 $c$  为算法迭代次数， $K$  为每次迭代生成的个体数量， $|\mathcal{M}_t|$  为编码参数维度。因此，整个算法的总时间复杂度为  $O((|\mathcal{M}_{t-1}||\mathcal{C}_t| + cK|\mathcal{M}_t|))$ 。在 UAV 巡检任务中，每帧中实际需要重点处理和传输的 RoI 数量通常远小于启发式算法的迭代次数或个体数量，因而算法的总时间复杂度可以近似为  $O(cK|\mathcal{M}_t|)$ 。

### 3 性能评估

为了全面评估本文提出的感兴趣区域提取与传输编码优化策略，在 UAV 机载摄像头捕获的水上场景数据集上进行了目标检测任务的仿真实验。实

验通过在不同的分辨率和网络带宽条件下，对比不同基线算法的性能，验证了该系统在检测准确率、数据传输量和带宽适应性方面的优势。

### 3.1 实验设置

1) 实验设备：为了模拟实际的UAV巡检场景，利用NVIDIA Jetson TX2模拟巡检UAV。边缘服务器运行Ubuntu，配备了NVIDIA GeForce RTX 3080 Ti GPU和Intel i9-10900K CPU，用于模拟边缘计算设备的强大计算能力，以支持RoI提取后的目标检测 and 数据处理任务。

2) 实验模型及数据集：为验证UAV RoI提取的有效性，选取了一个典型的水上巡航目标检测作为实验任务。在边缘服务器上运行YOLOv10x以实现高精度的目标检测。测试数据集来自SeaDronesSec<sup>[32]</sup>，该数据集包括多种场景下的水上目标检测任务，涵盖不同分辨率的图像。仿真参数见表1。

表1 仿真参数

名称	取值
热力图阈值 $\epsilon$	150
边界框交并比阈值 $\sigma_1$	0.6
无效边界框检测阈值 $\sigma_2$	3
带宽	10~80 Mbit/s

3) 实验对比基线：将提议的系统与以下基线算法进行比较。

**本地算法：**在UAV端运行轻量级YOLOv10n模型进行目标检测，仅传输检测结果信息。

**边缘算法：**将UAV采集的原始图像直接回传至边缘服务器进行检测。

**最佳配置算法：**所有RoI边界框经过画幅调整，并以最佳QP进行压缩和传输。

**EdgeDuet<sup>[33]</sup>算法：**将UAV采集的原始图像平均分割成块，仅将包含小目标的块直接回传至边缘服务器进行检测，剩余目标块在UAV端进行检测。

### 3.2 性能分析

通过实验评估和比较了在1080P和4K两种分辨率的图像以及不同的网络带宽条件下，本文提出的优化策略与其他基线算法的性能表现。F1分数结合了查准率和查全率，能够更全面地反映模型检测结果的准确性和可靠性。因此，以F1分数作为目标检测性能的核心评价指标。此外，还考查了在

不同带宽条件下，各算法准确率的变化趋势。不同算法在准确率和每帧传输数据量的比较如图4所示。

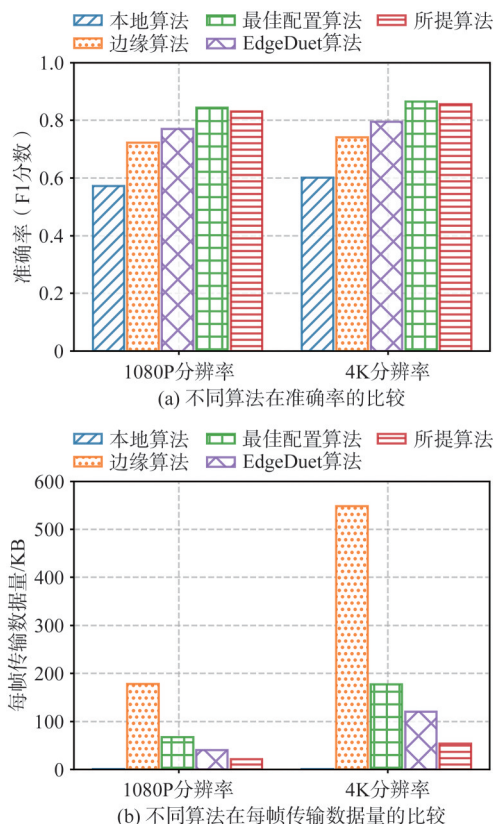


图4 不同算法在准确率和每帧传输数据量的比较

如图4(a)所示，本文所提算法在1080P和4K分辨率的图片下，准确率显著优于本地算法、边缘算法和EdgeDuet。在1080P分辨率下，本文算法准确率相比本地算法提高了42.1%，相比边缘算法提高了15.3%，相比EdgeDuet算法提高了7.79%；在4K分辨率下，本文算法准确率相比本地算法提高了45.6%，相比边缘算法提高了15.5%，相比EdgeDuet提高了7.54%。本地算法由于在UAV端运行轻量级模型，受限于模型规模，检测性能较低；边缘算法依赖边缘服务器端强大的计算能力，但未对图像进行任何优化，尤其在小目标检测中，精度较低；EdgeDuet算法的平均分割方案无法完全避免背景区域的干扰，导致检测精度受限。相比之下，本文提出的RoI提取与优化策略能够完整提取各种目标区域，并在精度与传输效率之间取得较好的平衡，达到与最佳配置算法相近的检测性能。

图4(b)展示了不同算法在每帧传输数据量的比较。与其他基线算法相比，本文所提算法在传输数

据量上展现了显著优势。本地算法由于在 UAV 端完成目标检测，无须传输图像数据，数据量最小；边缘算法需要传输整张图像，数据量最大；最佳配置算法传输以最高 QP 参数压缩的所有 RoI 边界框，虽然在保证检测精度的同时减少了部分传输数据量，但仍未能显著降低传输开销；EdgeDuet 算法虽然只传输小目标的分块，但未进行有效的数据压缩优化，因此传输数据量仍较大。相比之下，本文所提算法通过自适应选择 RoI 边界框并动态调整 QP 参数，使得每帧传输数据量相比边缘算法降低了 88.2%~90.1%，相比最佳配置算法降低了 68.7%~69.5%，相比 EdgeDuet 算法降低了 47.5%~55.1%。

不同算法的准确率随网络带宽波动的变化趋势如图 5 所示。对于本地算法，由于其不依赖于图像传输，准确率保持稳定。而边缘算法在带宽较低时无法保证图像能够及时传回，准确率明显下降。最佳配置算法虽然能保持较高的检测精度，但其 RoI 边界框数据量较大，带宽受限时会出现传输时延，导致检测精度下降。EdgeDuet 算法同样面临类似问题，由于其未对传输块进行 QP 压缩，导致在带宽有限时传输时延，进而影响准确率。本文提出的 RoI 提取与优化策略通过自适应调整传输数据量和任务负载，在带宽较差的情况下仍能保持较高的检测精度。在带宽为 10 Mbit/s 时，所提算法能够保证平均 0.65 的准确率，相比其他算法提高了 5.1%~14.3%。值得注意的是，在网络带宽达到 80 Mbit/s 时，最佳配置算法因带宽不再受限，超越了所提算法。此时带宽资源足够充足，不再限制视频帧的传输。

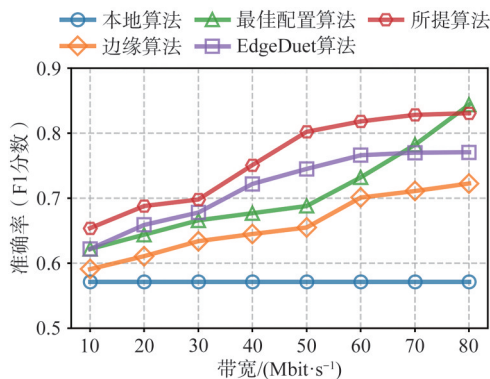


图5 不同算法的准确率随网络带宽波动的变化趋势

最后，深入分析边界框交并比值  $\sigma_1$  和无效边界框检测阈值  $\sigma_2$  对系统准确率的影响。边界框交并

比值和无效边界框检测阈值对准确率的影响如图 6 所示， $\sigma_1$  过高或  $\sigma_2$  过低时，UAV 回传的 RoI 边界框数量增多，导致传输压力增加；而  $\sigma_1$  过低或  $\sigma_2$  过高时，回传的 RoI 边界框减少，可能导致漏检。因此， $\sigma_1$  和  $\sigma_2$  值的合理设置，对于系统的带宽消耗和检测准确率至关重要。实验结果表明，合理的参数选择能够在保证检测准确率的同时有效控制传输数据量，最终实现精度与传输效率的最优平衡。

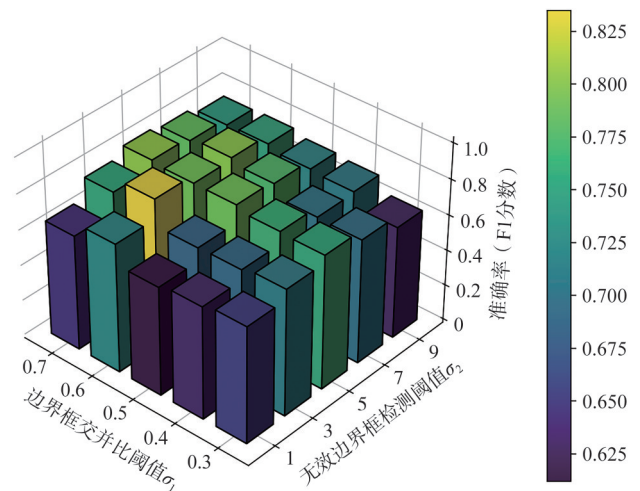


图6 边界框交并比值和无效边界框检测阈值对准确率的影响

#### 4 结束语

鉴于巡检任务中边缘计算辅助的 UAV 在低时延和高效数据处理方面面临的挑战，本文提出了一种轻量级的感兴趣区域提取与传输优化策略。首先，详细分析了 UAV 巡检任务特点和 UAV 在小目标检测、计算资源受限和通信链路不稳定等方面的困难；接着，提出了一种自适应 RoI 提取和编码等级选择的优化方法，以达到检测质量与服务时延的平衡；最后，通过采用启发式算法，有效解决了 RoI 提取与编码配置选择中的复杂问题，并验证了该方法在 UAV 河道巡检中的实际应用效果。研究表明，本文提出的感兴趣区域提取与传输编码优化策略能够显著降低河道巡检任务中 UAV 采集视频的传输数据量，并提高小目标检测的准确性。未来的研究工作可以进一步关注 UAV 在更为复杂的水上、电力线路等场景中的动态感知能力，以及在多 UAV 协同环境中的应用扩展，以持续提升边缘辅助的 UAV 巡检系统的检测性能和系统效率。

## 参考文献:

- [1] MAO Y Y, YOU C S, ZHANG J, et al. A survey on mobile edge computing: the communication perspective[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(4): 2322-2358.
- [2] LYU F, YANG P, SHI W S, et al. Online UAV scheduling towards throughput QoS guarantee for dynamic IoVs[C]//*Proceedings of the ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. Piscataway: IEEE Press, 2019: 1-6.
- [3] NOMIKOS N, GKONIS P K, BITHAS P S, et al. A survey on UAV-aided maritime communications: deployment considerations, applications, and future challenges[J]. *IEEE Open Journal of the Communications Society*, 2022, 4: 56-78.
- [4] LU W J, LAN C Z, NIU C Y, et al. A CNN-transformer hybrid model based on CSWin transformer for UAV image object detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16: 1211-1231.
- [5] ULLAH Z, AL-TURJMAN F, MOSTARDA L. Cognition in UAV-aided 5G and beyond communications: a survey[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 6(3): 872-891.
- [6] YANG T T, JIANG Z, SUN R J, et al. Maritime search and rescue based on group mobile computing for unmanned aerial vehicles and unmanned surface vehicles[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(12): 7700-7708.
- [7] LYU H G, SHAO Z Y, CHENG T, et al. Sea-surface object detection based on electro-optical sensors: a review[J]. *IEEE Intelligent Transportation Systems Magazine*, 2023, 15(2): 190-216.
- [8] 邵瑰玮, 刘壮, 付晶, 等. 架空输电线路无人机巡检技术研究进展[J]. *高电压技术*, 2020, 46(1): 14-22.
- SHAO G W, LIU Z, FU J, et al. Research progress in unmanned aerial vehicle inspection technology on overhead transmission lines[J]. *High Voltage Engineering*, 2020, 46(1): 14-22.
- [9] BENJDIRA B, KOUBAA A, AZAR A T, et al. TAU: a framework for video-based traffic analytics leveraging artificial intelligence and unmanned aerial systems[J]. *Engineering Applications of Artificial Intelligence*, 2022, 114: 105095.
- [10] FENG Q H, XU X Z, WANG Z X. Deep learning-based small object detection: a survey[J]. *Mathematical Biosciences and Engineering*, 2023, 20(4): 6551-6590.
- [11] ZHANG X D, IZQUIERDO E, CHANDRAMOULI K. Dense and small object detection in UAV vision based on cascade network[C]//*Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. Piscataway: IEEE Press, 2019: 118-126.
- [12] RAZAAK M, KERDEGARI H, ARGYRIOU V, et al. Multi-scale feature fused single shot detector for small object detection in UAV images[C]//*Proceedings of the Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2019: 778-786.
- [13] ZHOU R T, WU X Y, TAN H S, et al. Two time-scale joint service caching and task offloading for UAV-assisted mobile edge computing[C]//*Proceedings of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. Piscataway: IEEE Press, 2022: 1189-1198.
- [14] 叶刚. 基于视觉注意的立体视频感兴趣区域提取[D]. 杭州: 浙江大学, 2013.
- YE G. ROI Extraction for stereoscopic video based on visual attention [D]. Hangzhou: Zhejiang University, 2013.
- [15] 申帅. 无人机边缘计算中的动态资源分配算法研究[D]. 成都: 电子科技大学, 2024.
- SHEN S. Research on dynamic resource allocation algorithm in unmanned aerial vehicle aided edge computing[D]. Chengdu: University of Electronic Science and Technology of China, 2024.
- [16] WANG C Z, YANG P, HOU J W, et al. Dependence-aware multi-task scheduling for edge video analytics with accuracy guarantee[J]. *IEEE Internet of Things Journal*, 2024, 11(16): 26970-26983.
- [17] KONG Y X, YANG P, CHENG Y. Adaptive on-device model update for responsive video analytics in adverse environments[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, PP(99): 1.
- [18] JAIN S, ZHANG X, ZHOU Y H, et al. Spatula: efficient cross-camera video analytics on large camera networks[C]//*Proceedings of the 2020 IEEE/ACM Symposium on Edge Computing (SEC)*. Piscataway: IEEE Press, 2020: 110-124.
- [19] 期治博, 杜磊, 霍如, 等. 基于边缘计算的多摄像头视频协同分析方法[J]. *通信学报*, 2023, 44(8): 14-26.
- QI Z B, DU L, HUO R, et al. Multi-camera video collaborative analysis method based on edge computing[J]. *Journal on Communications*, 2023, 44(8): 14-26.
- [20] JIANG J C, ANANTHANARAYANAN G, BODIK P, et al. Chameleon: scalable adaptation of video analytics[C]//*Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*. New York: ACM Press, 2018: 253-266.
- [21] YANG P, CHENG Y, ZHANG N, et al. Adaptive network configuration for efficient and accurate neural video inference[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10(1): 263-276.
- [22] ZHAO C J, LIU R W, QU J X, et al. Deep learning-based object detection in maritime unmanned aerial vehicle imagery: review and experimental comparisons[J]. *Engineering Applications of Artificial Intelligence*, 2024, 128: 107513.
- [23] SUN W, DAI L, ZHANG X R, et al. RSOD: real-time small object detection algorithm in UAV-based traffic monitoring[J]. *Applied Intelligence*, 2022, 52(8): 8448-8463.
- [24] WANG Y T, CHEN W W, LUAN T H, et al. Task offloading for post-disaster rescue in unmanned aerial vehicles networks[J]. *IEEE/ACM Transactions on Networking*, 2022, 30(4): 1525-1539.
- [25] DAI X X, XIAO Z, JIANG H B, et al. UAV-assisted task offloading in vehicular edge computing networks[J]. *IEEE Transactions on Mobile Computing*, 2024, 23(4): 2520-2534.

- [26] 江雪, 赵亮. 无人机辅助移动边缘计算网络中轨迹设计和带宽分配策略[J]. 物联网学报, 2023, 7(4): 123-131.  
JIANG X, ZHAO L. Trajectory design and bandwidth allocation strategy in UAV-assisted MEC network[J]. Chinese Journal on Internet of Things, 2023, 7(4): 123-131.
- [27] XU S, ZHANG X Y, LI C G, et al. Deep reinforcement learning approach for joint trajectory design in multi-UAV IoT networks[J]. IEEE Transactions on Vehicular Technology, 2022, 71(3): 3389-3394.
- [28] 冯建新, 栾帅帅, 刘俊梅, 等. 一种高吞吐量的无人机轨迹规划方法[J]. 计算机工程, 2021, 47(1): 172-181.  
FENG J X, LUAN S S, LIU J M, et al. An unmanned aerial vehicle trajectory planning method with high throughput[J]. Computer Engineering, 2021, 47(1): 172-181.
- [29] LIN X H, SU G C, CHEN B, et al. Striking a balance between system throughput and energy efficiency for UAV-IoT systems[J]. IEEE Internet of Things Journal, 2019, 6(6): 10519-10533.
- [30] LI J, YI C Y, CHEN J Y, et al. Joint trajectory planning, application placement, and energy renewal for UAV-assisted MEC: a triple-learner-based approach[J]. IEEE Internet of Things Journal, 2023, 10(15): 13622-13636.
- [31] WALLENIUS J, DYER J S, FISHBURN P C, et al. Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead[J]. Management Science, 2008, 54(7): 1336-1349.
- [32] VARGA L A, KIEFER B, MESSMER M, et al. SeaDronesSee: a maritime benchmark for detecting humans in open water[C]//Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE Press, 2022: 3686-3696.
- [33] YANG Z, WANG X, WU J H, et al. EdgeDuet: tiling small object detection for edge assisted autonomous mobile vision[J]. IEEE/ACM Transactions on Networking, 2023, 31(4): 1765-1778.

## [作者简介]



杨鹏(1991-), 男, 博士, 华中科技大学副教授, 主要研究方向为人工智能及其在下一代网络、移动边缘计算、VR和视频分析中的应用等。



梁雨欣(2001-), 女, 华中科技大学硕士生, 主要研究方向为移动边缘计算、视频分析和移动 AIGC 网络等。



孔雨新(2000-), 男, 华中科技大学硕士生, 主要研究方向为移动边缘计算、视频分析和移动 AIGC 网络等。



刘鸣柳(1992-), 女, 国网湖北省电力有限公司电力科学研究院工程师, 主要研究方向为物联网和边缘计算等。