

基于多任务学习的智慧城市场景人群计数研究

郭禹辰¹, 刘 梦²

¹河北金融学院河北省科技金融重点实验室, 河北 保定

²河北金融学院金融博物馆, 河北 保定

收稿日期: 2025年10月12日; 录用日期: 2025年11月11日; 发布日期: 2025年11月19日

摘 要

为解决城市场景摄像头高度与视角差异导致的人群图像中强透视、多尺度问题, 提出一种以多任务学习为框架的多尺度人群计数算法。在点监督框架上引入一种一体化多尺度金字塔模块, 提升对不同大小人头的特征提取能力, 同时将点标注经多尺度高斯扩散与自适应阈值自动生成人群前景分割标签, 作为辅助任务与计数任务联合训练, 以计数损失和分割损失构成多任务目标, 抑制背景干扰并实现多尺度人群的识别。选取以城市场景为代表的ShanghaiTech数据集A、B和UCF-QNRF数据集进行实验, MAE分别为57.8, 7.6和86.2, 在强透视和密度不均场景下均取得较好的效果, 表现出较高鲁棒性。所提方法以零额外标注成本显著提升了智慧城市场景下人群计数的准确性与可部署性, 适用于跨视角的城市监控设备。

关键词

人群计数, 透视差异, 多尺度特征, 前景分割, 多任务学习框架

Multi-Task Learning-Based Crowd Counting for Smart-City Scenes

Yuchen Guo¹, Meng Liu²

¹Science and Technology Finance Key Laboratory, Hebei Finance University, Baoding Hebei

²Financial Museum, Hebei Finance University, Baoding Hebei

Received: October 12, 2025; accepted: November 11, 2025; published: November 19, 2025

Abstract

To address the strong perspective distortion and multi-scale variations in crowd images caused by differences in camera height and viewpoint in urban settings, we propose a multi-scale crowd counting method under a multi-task learning (MTL) framework. On top of a point-supervised paradigm, we introduce an integrated multi-scale pyramid module that enhances feature extraction for heads of

different sizes. Meanwhile, point annotations are converted—via multi-scale Gaussian diffusion and adaptive thresholding—into crowd-foreground segmentation labels, which serve as an auxiliary task jointly trained with the counting task. The overall objective combines counting loss and segmentation loss, suppressing background interference and enabling reliable recognition across scales. Experiments on urban-scene representative datasets—ShanghaiTech Part A, ShanghaiTech Part B, and UCF-QNRF—yield MAEs of 57.8, 7.6, and 86.2, respectively, demonstrating strong performance under severe perspective and uneven density conditions and indicating high robustness. The proposed approach improves accuracy and deployability for smart-city crowd counting without any additional annotation cost, making it suitable for cross-view urban surveillance systems.

Keywords

Crowd Counting, Perspective Difference, Multi-Scale Features, Foreground Segmentation, Multi-Task Learning Framework

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

视觉感知系统对城市人群态势的持续感知对于智慧城市的安全治理、公共服务与交通组织十分重要。地铁站厅、商圈步行街、校园和综合交通枢纽等场景部署大量的视频摄像头用于人流密度监测、客流引导与异常拥挤预警[1]。然而,在城市级摄像头部署的条件下,人群计数任务存在两个难点:(1)摄像头安装位置与姿态差异显著,导致图像存在透视畸变与跨场景尺度变化;(2)人群拥挤时严重遮挡与邻近目标粘连使局部纹理模糊、边界不清[2][3]。因此,如何在已有标注条件下,构建对尺度与视角敏感的点监督计数模型,对单帧图像内的人群进行计数并估计空间位置,要求模型在跨摄像头、跨区域的条件下保持稳定,对于智慧城市中人群计数的落地十分关键[4]。

城市摄像头高度与俯仰角不同,导致远近区域的头部表观尺寸差异大、同一图中尺度出现长尾分布问题,直接加大了点监督方法中重复检测和稀疏区漏检问题,同时,纯粹的点监督学习方法缺乏人群区域的先验,难以在复杂场景中抑制非人头纹理的干扰[5]。现有人群计数研究主要分为密度图回归、点监督两种方式,以计数任务为主,部分研究引入多任务学习以增强模型的场景泛化能力[6][7]。以MCNN为代表的密度图回归方式,通过在点标注处渲染高斯核,学习连续密度分布并积分得到计数,具有良好的拥挤鲁棒性,但在精确定位与去重方面依赖后处理与阈值,且对跨尺度的自适应仍受限[8]。点监督方法直接预测人的离散位置,通过一对一匹配与偏移回归结合的方式提升定位精度,减少后处理依赖[9][10]。然而,在强透视与多尺度条件下,固定感受野或单尺度表示容易产生邻近冲突与尺度失配[11]。因此,部分工作引入透视图或尺度注意缓解视角差异问题,部分研究利用弱分割先验为计数提供区域约束[12]。

为解决尺度和场景差异问题,本文提出一种基于多尺度模块的多任务学习人群计数框架,引入一体化多尺度特征金字塔模块进行多尺度特征提取,并通过尺度融合机制,缓解透视导致的尺度差异,在学习目标方面,通过多尺度高斯扩散与自适应阈值,将稀疏点标注转化为符合头部空间分布规律的前景区域,构造分割任务标签图,将计数与人群前景分割任务作为多任务联合优化的学习目标,在不额外增加标注成本的前提下增强监督信号,使网络学习稳定的人群区域先验。

2. 多任务学习人群计数数据集

本文在三个人群计数基准上验证方法对城市场景的适用性。ShanghaiTech 数据集包含两部分共计 1198 张图像, A 和 B 分别采集于高密度的互联网场景与贴近日常街景的中低密度监控画面, 能同时检验强透视下的精定位与城市道路和商圈中背景干扰抑制能力; UCF-QNRF 以大规模人群和超高分辨率图像为主, 覆盖从近景到远景的剧烈尺度变化, 可模拟城市大型活动, 如广场集会、演唱会、体育赛事中跨视角的人群计数能力测试, 训练集和测试集分别为 1201 和 334 [13]。

在点监督框架下引入人群区域前景先验, 依据已有的人头点标注自动构造与图像相同尺寸的二值分割任务标签, 用于训练时作为分割任务分支的辅助监督, 将稀疏点与人群连通区域进行融合。首先, 采用多尺度高斯扩散方式, 将点标注从点扩展为头部的邻域[14]。对每个标注点, 为兼顾近景大头与远景小头的感受范围, 分别以小、中、大多个核渲染二维高斯响应, 对各尺度相应做逐像素最大融合, 得到连续响应图 S 。为适应不同摄像头的曝光、画质与场景密度, 采用自适应阈值而非全局常数, 可基于非零响应的分位数或采用类 Otsu 的类间方差准则自动确定阈值, 在生成过程中, 应保持标签与原图相同尺寸, 分割标签图如图 1 所示。为提升跨摄像头鲁棒性, 保持生成核的一致性, 使区域先验与计数主干共享相同的尺度, 同时应保证对强度响应做归一化处理, 防止某一尺度遮盖其他尺度, 兼顾稀疏区与拥挤区的响应均衡。通过上述流程, 由点标注自动获得人群前景标签。在联合训练中, 该标签为计数主干提供与场景结构一致的区域先验, 抑制复杂背景干扰, 同时可增强多尺度特征提取, 有助于在不同安装高度与视角条件下形成可迁移的尺度不变性。

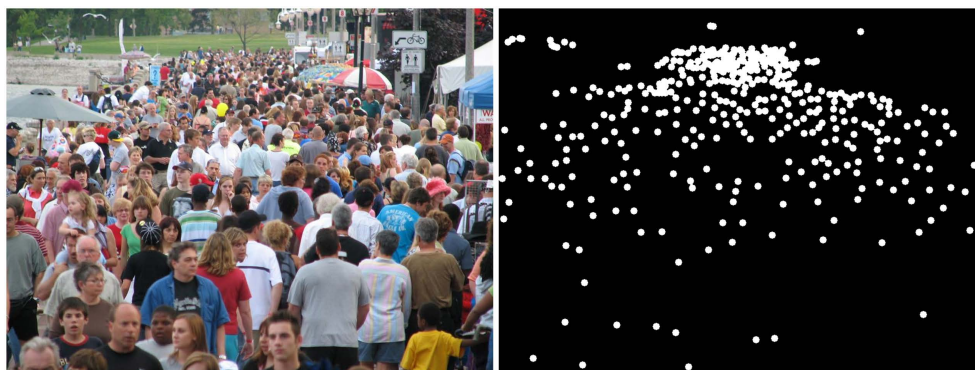


Figure 1. Foreground segmentation label map
图 1. 前景分割标签图

如图 2 所示的散点图以每幅图的人数为横轴, 以归一化局部尺度为纵轴, 整体皮尔逊系数为-0.417, 统计结果呈现负相关关系, 随着人数增加, 个体之间的相对间距变小, 曲线由左至右呈现先陡降后趋于平缓的趋势, 说明当人群规模达到一定程度后, 局部可允许的最小间距接近场景极限。左侧数据在纵向上离散性更大, 反映了在稀疏或小规模人群下, 尺度对局部尺度占主导作用。而中高人数范围内, 尺度方差显著收敛, 分布聚集在较低的归一化距离内。在 ShanghaiTech Part A 部分高人数区间出现长尾离群点, 体现了强透视与跨场景多尺度带来的不均匀性, Part B 则更集中于低至中等人数区间。

3. 轻量化多任务联合学习人群计数框架

3.1. 人群计数网络框架

轻量化的多任务联合学习人群计数框架以图像 $I \in R^{H \times W \times 3}$ 为输入, 采用 VGG 主干提取多层语义特征

$P\{C_3, C_4, C_5\}$, 并通过一体化的多尺度特征金字塔模块获取统一尺度下的增强表达特征 F , 对分层特征 P 进行信息强化后为计数和分割辅助任务提供输入。计数任务沿用点监督方式 P2PNet 的整体架构, 在特征图上划分规则锚点几何 $A = \{a_i\}$ 上 N_i 为 1, 网络同时预测每个锚点的二分类概率 p_i 与坐标偏移 Δ_i , 并将偏移解码为最终的点预测坐标。为了引入人群区域先验并提升多尺度鲁棒性, 额外加入基于 PSPNet 的金字塔池化分割头, 产生前景概率图 M 。整体框架实现端到端优化, 共享主干特征实现监督计数和前景分割的多任务协同, 计数分支受益于分割先验抑制背景干扰, 分割分支利用计数头的点级监督提升定位精度, 整体结构图如图 3 所示。

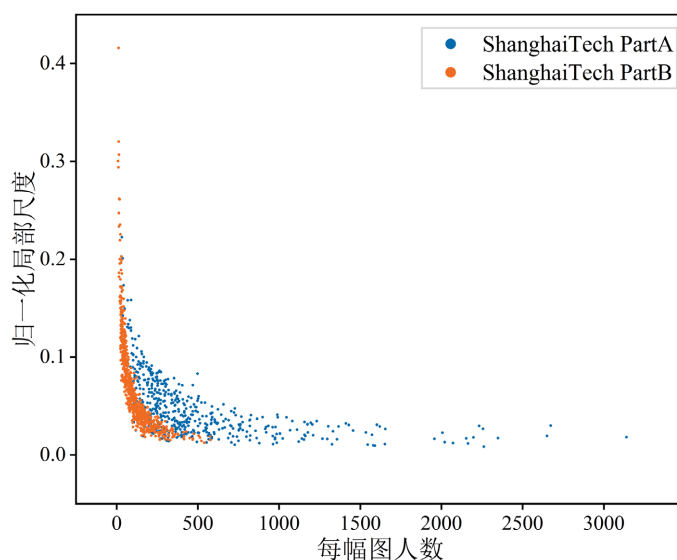


Figure 2. Crowd-scale statistics in ShanghaiTech dataset

图 2. ShanghaiTech 数据集人群尺度统计

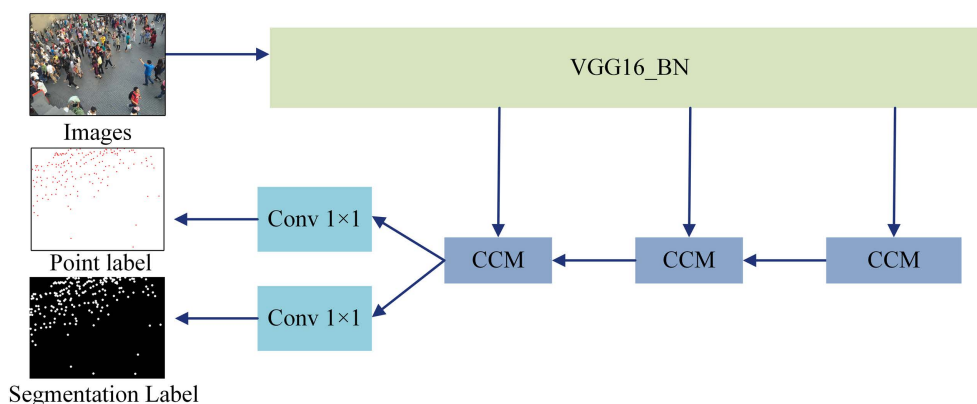


Figure 3. Diagram of the multi-task crowd counting

图 3. 多任务人群计数框架图

3.2. 一体化多尺度特征金字塔模块

为了在单一分辨率特征图上获得兼顾远景、近景人头区域, 将自上而下的金字塔结构与可重参数化的通道-上下文模块融合, 构成一体化多尺度特征金字塔模块。设主干输出特征为 $\{C_3, C_4, C_5\}$, 空间步长为 $\{8, 16, 32\}$, 通道数为 $\{256, 512, 512\}$ 。首先, 使用 1×1 卷积实现各输入特征图的通道对齐到 256,

便于进行跨层融合,在此基础上,构造自上而下的金字塔路径,从最高层 C_5 出发,经上采样并与低层级进行逐像素相加融合,随后经过 3×3 卷积做局部区域细化,得到中低层特征。考虑到不同层级特征对当前尺度的贡献不同,引入一种可学习加权融合策略提升跨层语义对齐效果,如式(1):

$$P_l = \frac{\alpha_l \cdot C_l + \beta_l \cdot \text{Up}(C_{l+1})}{\alpha_l + \beta_l + \varepsilon} \quad (1)$$

其中, ε 为数值稳定项, α 和 β 为可自适应权重,通过对其施加非负约束和归一化处理,在不同数据分布下自适应分配各层权重,避免由传统金字塔结构简单相加造成的语义冲突。选取低层特征 C_3 作为最终的预测分辨率输入,在高分辨率特征图上获得密集且精细的空间响应。

为进一步扩大感受野,在每个层级融合节点后叠加一个通道-上下文模块(Channel-Context Module, CCM),如图 4 所示,该模块由可重参数化的并行分支构成,通过 3×3 深度可分离膨胀卷积提供大感受野,以 1×1 点卷积增强对特征图的通道混合能力,同时连接 2 个带残差连接的 3×3 卷积块提取语义。受 MCNN 多尺度结构设计启发,将三分支并联设计,模型推断时将其重参数转化为单一卷积核加快计算效率。同时,在每个 CCM 模块后,引入 SE 注意力机制进行全局特征的细化。通过注意力机制使网络在密集背景与多尺度目标同时存在的场景下,关注人头区域上下文。

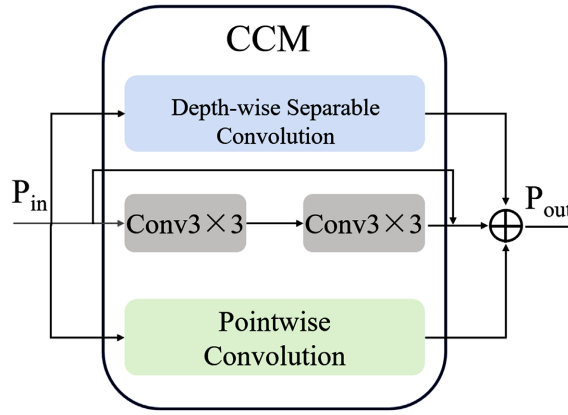


Figure 4. Diagram of the CCM

图 4. CCM 模块结构图

3.3. 多任务联合学习框架

在点监督的人群计数方法下,将任务转化为集合预测,模型在固定锚点集合上同时输出人和背景的二分类概率与预测点集。设计一种以图像分割和点监督融合的多任务人群计数框架,在共享的预测特征图 P_3 上,引入分割子网络预测人头区域的前景概率图 M ,通过由高斯扩散方式生成的掩膜构建分割目标 T ,随后,利用掩膜与点一致性约束与计数损失将分割信息反馈到点监督分支中,形成互补优化。

分割子网络以特征图 C_3 为输入,通过多分支空洞卷积获取多尺度上下文,并以 GroupNorm 提升小批量训练稳定性,空洞率分别设置为 1, 2, 3,在不改变分辨率的前提下,获得最大感受野与细粒度特征响应。

考虑人群图像中存在的人头前景稀疏问题,对 BCE 损失进行类别加权,引入 Dice 与边界一致项对分割任务进行共同优化:

$$L_{bce} = -\frac{1}{N} \sum_{i=1}^N T(x) \log M(x) + (1 - T(x)) \log (1 - M(x)) \quad (2)$$

$$L_{\text{dice}} = 1 - \frac{2\langle M, T \rangle + \varepsilon}{\|M\|_2^2 + \|T\|_2^2 + \varepsilon} \quad (3)$$

$$L_{\text{edge}} = \|\nabla M - \nabla T\|_1 \quad (4)$$

其中, $\langle, \rangle$ 表示内积, ∇ 表示空间梯度, 可用 Sobel 算子进行计算, $\|\cdot\|_1$ 表示 L_1 范数, ε 为数值稳定项。

综合以上损失, 融合分割和点监督任务, 构造联合损失, 如式(5):

$$L_{\text{total}} = \lambda_{\text{seg}} (L_{\text{bce}} + \lambda_{\text{dice}} L_{\text{dice}} + \lambda_{\text{edge}} L_{\text{edge}}) + L_{\text{P2P}} \quad (5)$$

其中, L_{P2P} 表示 P2PNet 点监督框架损失, λ_{seg} 、 λ_{dice} 和 λ_{edge} 为各项权重参数。

多任务联合框架用无标注成本且可解释的软掩膜填补像素级监督的缺位, 并通过对齐把分割先验回注到点匹配与定位中, 从而在不改变 P2P 框架的前提下[5], 提高遮挡重叠场景的召回与假阳性抑制。与直接把分割当作辅助特征不同, 分割提供形状与连通性, 点监督提供实例的稀疏锚定, 两者相互校正, 最终在单一分辨率上实现稳定、细粒度的人群点级预测。

4. 实验与结果分析

4.1. 实验设置

实验在 NVIDIA RTX 3070 显卡上进行, 深度学习环境为 PyTorch, 对于 ShanghaiTech 数据集采用原图进行输入, UCF-QNRF 数据集中部分超大分辨率图像则缩减为原图尺寸的一半进行输入。训练中采用 Adam 优化器, 学习率为 1×10^{-4} , 批训练大小为 8, 最大训练轮次为 800。对于多任务框架的权重参数设置, 训练时先将 λ_{seg} 与点监督损失固定为相同数值, 再微调其内部各项权重, 保证训练的稳定。

4.2. 评价指标

以人群计数任务中经典指标平均绝对误差(Mean Absolute Error)和均方根误差(Mean Square Error)对所提方法进行评估, 分别用于评价一张图上预测总人数与真实总人数的偏差。

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |C_i - C_i^{GT}| \quad (10)$$

$$\text{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N |C_i - C_i^{GT}|^2} \quad (11)$$

式中, N 表示测试集中测试图像的数量, C_i 和 C_i^{GT} 分别表示网络预测的人群总数和真实人群密度图中的人群总数。

4.3. 实验结果与分析

4.3.1. 模型结果对比

实验在 ShanghaiTech A、B 和 UCF-QNRF 数据集上进行, 从表 1 结果可以看出, MCNN 在三个数据集上误差最高, CSRNet 得益于空洞卷积提升感受野的优势, MAE 和 MSE 明显下降, 后续引入分布建模与更强监督的 BL、CAN、DM-Count, 性能逐步提升, MAE 显著降低, 在 ShanghaiTech 数据集上分别达到 62.8、62.3 和 59.7。在此基础上, 本文所提算法在 ShanghaiTech Part A 上 MAE 为 57.8, 相较 DM-Count 的 59.7 降低了 3.18%, 相较 CSRNet 降低 15.25%。这说明在强透视和高密度的城市场景下, 多尺度和分割辅助能有效抑制背景干扰。

对所提算法在 ShanghaiTech 数据集 A 和 B 上进行测试, 可视化结果如图 5 所示。其中, 第一行 A 数据集上的绝对误差分别为 21、53、20, 第二行的 B 数据集绝对误差分别为 1、0、9。数据集 A 中, 多为

核心商圈或集会场景, 呈现出高人群密度和强透视特点, 近景头部尺寸大, 远景人头像素占比极小。强遮挡与纹理混杂导致前景、背景边界模糊。相对地, 数据集 B 中多为广场、步道等公共场所, 人群稀疏且少有遮挡, 预测值与标注高度一致。从多尺度角度看, 对数据集 A 测试的图像中, 尺度跨度显著, 固定感受野的点回归方法往往会在远景低分响应而被阈值抑制。数据集 B 测试的图像中, 由于整体尺度分布窄、头部间距相对稳定, 网络的特征尺度与目标尺度更易匹配。

Table 1. Comparison of crowd counting algorithm results on different datasets

表 1. 不同数据集上人群计数算法结果对比

算法	SHTech Part A		SHTech Part B		UCF-QNRF	
	MAE	MSE	MAE	MSE	MAE	MSE
MCNN [8]	110.2	173.2	26.4	41.3	-	-
CSRNet [15]	68.2	115.0	10.6	16.0	-	-
BL [16]	62.8	101.8	7.7	12.7	88.7	154.8
CAN [17]	62.3	100.0	7.8	12.2	107	183
DM-Count [18]	59.7	96.7	7.4	11.8	85.6	148.3
Ours	57.8	92.4	7.6	12.3	86.2	151.7



Figure 5. Crowd counting result of ShanghaiTech Part A and Part B dataset

图 5. ShanghaiTech Part A 和 Part B 的人群计数结果

在 UCF-QNRF 数据集上的可视化结果如图 6 所示。四幅图的绝对误差分别为 13、14、10、14。从场景属性看, UCF-QNRF 更贴近智慧城市中的复杂街景, 既包含游行、集会等高密度强透视场景, 也包含夜间街区与车行干道等光照与背景干扰强的场景, 近景头部尺寸大, 尺度变化较为显著。所提多任务多尺度算法在此类跨视角条件下依然保持稳定, 一体化金字塔结构增强了对多尺度人头的感知, 在强遮挡与密度不均区域仍能获得较好效果。

4.3.2. 参数选择对比

在三个数据集上测试了多任务权重的计数性能影响, 结果如表 2 所示增强分割分支并采用 Dice 和 Edge 约束能在三个数据集上同时降低 MAE, 且在强透视、密度不均匀的 Part A 和 UCF 上较为明显。具体来说, 分割主损权重 λ_{seg} 由 0.5 增长到 1.0 时, MAE 分别由 58.9、8.1 和 94.1 降至 57.8、7.6 和 86.2。

引入 Dice 监督后，MAE 由 62.1、8.6 和 96.9 进一步降为 57.8、7.6 和 86.2，说明在伪掩膜存在噪声的点监督场景下，Dice 对前景连贯性较为重要；同时，边界一致性也有助于抑制背景结构干扰。



Figure 6. Crowd counting result of UCF-QNRF dataset
图 6. UCF-QNRF 数据集结果

Table 2. Comparison of weight parameters
表 2. 权重参数对比

参数选择		SHTech_A		SHTech_B		UCF-QNRF	
		MAE	MSE	MAE	MSE	MAE	MSE
λ_{seg}	0.5	58.9	94.5	8.1	13.2	94.1	162.1
	1.0	57.8	92.4	7.6	12.3	86.2	151.7
λ_{dice}	0	62.1	99.1	8.5	13.7	96.9	164.4
	1	57.8	92.4	7.6	12.3	86.2	151.7
λ_{edge}	0.5	58.3	93.2	8.0	12.6	93.2	158.2
	1.0	57.8	92.4	7.6	12.3	86.2	151.7

5. 结论

本文面向智慧城市监控视角提出多任务多尺度的人群计数框架，引入一体化金字塔模块提取图像中的多尺度人群图像，以多尺度高斯扩散与自适应阈值生成前景分割，实现分割任务辅助的人群计数。所提方法在 ShanghaiTech A、B 和 UCF-QNRF 上的 MAE 分别为 57.8、7.6 和 86.2。相较代表性方法 DM-Count，在 Part A 上 MAE 降低 3.18%，在强透视、密度不均与复杂背景条件下保持稳定，体现出良好的鲁棒性与可部署性。未来工作将面向大规模跨域场景，引入透视先验增强模型的特征提取能力，探索轻量化实时部署方案。

基金项目

河北省社会科学基金项目“转型金融对河北省高碳企业低碳技术创新的驱动机制研究”（HB24ERJ027）。

参考文献

- [1] 林园园, 杨会成, 胡耀聪. 基于轻量化卷积神经网络的人数估计算法研究[J/OL]. 重庆工商大学学报(自然科学版), 1-13. <https://link.cnki.net/urlid/50.1155.N.20240226.1244.014>, 2025-11-13.
- [2] 介艳良, 郝磊, 闫树军, 等. 基于图像处理的城市轨道交通监控系统设计[J]. 自动化与仪器仪表, 2023(2): 126-130+136.
- [3] Bai, H., Mao, J. and Gary Chan, S.-H. (2022) A Survey on Deep Learning-Based Single Image Crowd Counting: Network Design, Loss Function and Supervisory Signal. *Neurocomputing*, **508**, 1-18. <https://doi.org/10.1016/j.neucom.2022.08.037>
- [4] 蔡一庆, 马振伟, 王庭枢, 等. 面向跨域人群计数的头部感知密度适应网络[J]. 计算机辅助设计与图形学学报, 2021, 33(10): 1514-1523.
- [5] Song, Q., Wang, C., Jiang, Z., Wang, Y., Tai, Y., Wang, C., *et al.* (2021) Rethinking Counting and Localization in Crowds: A Purely Point-Based Framework. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 3345-3354. <https://doi.org/10.1109/iccv48922.2021.00335>
- [6] 王大正, 张涛. 双任务交互下的四段监督人群计数网络[J]. 小型微型计算机系统, 2023, 44(10): 2120-2126.
- [7] 陈训敏, 叶书函, 詹瑞. 基于多任务学习及由粗到精的卷积神经网络人群计数模型[J]. 计算机科学, 2020, 47(S2): 183-187+208.
- [8] Zhang, Y., Zhou, D., Chen, S., Gao, S. and Ma, Y. (2016) Single-Image Crowd Counting via Multi-Column Convolutional Neural Network. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 589-597. <https://doi.org/10.1109/cvpr.2016.70>
- [9] Chen, I., Chen, W., Liu, Y., Yang, M. and Kuo, S. (2024) Improving Point-Based Crowd Counting and Localization Based on Auxiliary Point Guidance. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T. and Varol, G., Eds., *Lecture Notes in Computer Science*, Springer, 428-444. https://doi.org/10.1007/978-3-031-72691-0_24
- [10] Lin, W. and Chan, A.B. (2023) Optimal Transport Minimization: Crowd Localization on Density Maps for Semi-Supervised Counting. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 21663-21673. <https://doi.org/10.1109/cvpr52729.2023.02075>
- [11] 文帅, 蒋勇, 杨丹, 等. 基于多尺度注意力网络的密集人群计数[J]. 计算机应用与软件, 2025, 42(1): 130-136+157.
- [12] 桑军, 刘新悦, 吴志伟, 等. 基于背景辅助的高效人群计数多任务学习网络[J]. 西南师范大学学报(自然科学版), 2022, 47(8): 1-8.
- [13] Idrees, H., Tayyab, M., Athrey, K., Zhang, D., Al-Maadeed, S., Rajpoot, N., *et al.* (2018) Composition Loss for Counting, Density Map Estimation and Localization in Dense Crowds. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 544-559. https://doi.org/10.1007/978-3-030-01216-8_33
- [14] 马圣南, 严华. 基于自适应策略的人群密度图纠正算法[J]. 现代计算机, 2024, 30(10): 23-28.
- [15] Li, Y., Zhang, X. and Chen, D. (2018) CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 1091-1100. <https://doi.org/10.1109/cvpr.2018.00120>
- [16] Ma, Z., Wei, X., Hong, X. and Gong, Y. (2019) Bayesian Loss for Crowd Count Estimation with Point Supervision. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October 2019-2 November 2019, 6142-6151. <https://doi.org/10.1109/iccv.2019.00624>
- [17] Liu, W., Salzmann, M. and Fua, P. (2019) Context-Aware Crowd Counting. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 5094-5103. <https://doi.org/10.1109/cvpr.2019.00524>
- [18] Wang, B., Liu, H., Samaras, D. and Nguyen, M.H. (2020) Distribution Matching for Crowd Counting. *Advances in Neural Information Processing Systems*, **33**, 1595-1607.