

开放场景人群计数方法研究——摘要

随着全世界人口规模的扩大和城市化进程的加速，高效准确地估计一定范围内的人群数量已经成为城市管理的重要问题。近几年，视觉监控设备、网络通信技术、计算机视觉技术的发展催生基于影像的人群计数研究，该研究旨在通过计算机视觉算法自动估计一段影像中的总人数，其在公共安全、交通调度、商业营销等领域具有广阔的应用前景。随着深度学习技术的发展，现有的基于卷积神经网络的人群计数模型已经能在一些简单场景取得较好的性能，然而当把这类模型应用到其它风格发生变化的场景时，计数性能可能很差，即出现“领域偏移”问题。

为了缓解“领域偏移”对计数算法落地应用的不利影响，本文对开放场景人群计数任务展开研究，旨在设计出能以较低成本实现跨领域应用的计数算法。除了“领域偏移”问题，该研究还面临多个难点，分别是对模型的复杂度有限制、训练集中手工标注的数据规模有限、目标场景无人工标注以及模型更新时的算力受限。为了实现预期目标，本文从两方面着手：1) 提出计数性能较好、复杂度较低且具有一定程度领域通用能力的计数模型（第三章和第四章）；2) 提出能将计数模型以较低的成本适配到目标领域的适配算法（第五章）。本文的主要创新点和贡献如下：

第一，提出面向跨领域应用的基于双分支注意力机制的人群计数方法。通过建模帧内前景和背景之间的相关性，能够缓解计数模型难以判别帧内特征的问题。为了在监督信息中引入人头信息和人头周边环境信息，以促进对通用背景特征的处理，提出结合几何自适应和泰森多边形剖分的密度图生成方法。为了增强模型对特征的精炼能力，引入自注意力机制来建模前景特征和背景特征的相关性，并创新性地设计双分支架构来提升对多尺度特征的处理能力。为了缓解现有手工人群数据集规模不足、多样性差等问题，提出大规模仿真人群数据集 CrowdX，并将其作为计数模型的预训练数据集，以进一步提升模型的泛化性。在多个公开数据集上的实验表明，提出的方法能够取得优越的计数性能，能够为跨领域应用提供基础。

第二，提出面向跨领域应用的基于帧间密度图链式关联的人群计数方法。通过建模相邻帧密度图之间的光流运动关联，能插件式地提升图像计数模型的跨领域计数性能。提出的方法采用迭代的前向链式框架，运动先验在帧间显式地传递，可以获得时序上更加稳定的计数结果。该框架由单帧计数模块、运动信息提取模块和密度图融合模块组成，模块之间相互解耦。单帧计数模块用于生成基于视觉特征的密度图、运动信息提取模块用于生成基于运动关联的密度图、密度图融合模块建模这两个密度图之间的相关性并输出增强的密度图。为了更有效地训练该框架，提出结合局部损失和全局损失的混合损失函数。在多个公开数据集上的实验表明，提出的计数方法具有高准确性、高鲁棒性以及良好的泛化性。

第三，提出基于监控场景先验构造中间域的计数模型领域适配方法。本文用目标场景先验显式地构造中间域，通过将模型迁移到中间域，来间接提升模型在目标域的计数性能。该方法由中间域生成和模型更新两阶段构成。在中间域生成阶段，利用目标场景的动态性、连续性等先验信息，逐步提纯出高置信度的目标域人员位置标签。在获得大量人员位置标签后，用模拟的方式自动地构造出中间

域数据样本。在模型更新阶段，为了无偏地更新计数模型，设计出一种协同源域和中间域的模型迭代策略。此外，本文构建了针对特定目标场景的基于仿真策略的计数模型初始化方法，能缓解模型在迁移过程中的冷启动难题。在多个数据集上的实验表明，提出的方法能够在场景先验可知时极大地缓解“领域偏移”问题。

综上所述，本文面向人群计数算法的跨领域应用，分别从提升模型的泛化性能和领域适配能力着手，提出了符合约束条件的计数模型领域迁移方法。本文提出的方法能够为人群计数算法的大规模应用奠定基础。

关键词： 人群计数，跨领域学习，泛化性，领域适配，仿真数据集，信息融合，领域迁移