

电梯门异常开关检测模型研究

柯家辉 朱康华 何 新
南京理工大学 江苏南京 210094

摘要: 对基于改进YOLOv8-Pose算法进行了简要阐述, 针对电梯环境的封闭性和资源受限特征, 提出了轻量级改进算法MAY (MobileNet-Attention-YOLO) 并引入自适应阈值焦点损失函数(ATFL)。在电梯门异常检测方面, 采用多种边缘检测方法对比并引入PPHT直线检测算法, 设计基于动态前景提取和直线数量判断的开关门判定流程, 优化GhostFaceNet结构, 提出融合E-Ghost Module、SE模块、空间注意力机制及跨层残差连接的ResSpatial_GhostFaceNet模型, 实现高精度、低资源消耗的嵌入式识别。实验结果表明, 本文方法在不同复杂场景下均具有良好的适应性和实用性, 为多场景智能视觉感知系统的构建提供了有效参考。

关键词: 视觉识别; YOLOv8-Pose; GhostFaceNet; 边缘检测; 电梯监测; 电梯门识别

引言

随着智慧城市与公共安全需求的快速增长, 计算机视觉技术在交通、医疗、电梯安全等领域得到广泛应用。电梯作为人员高频进出的封闭空间, 其乘员行为、设备运行状态需要被精准监测。同时, 在公共环境中, 电梯门异常开关的识别难题, 传统模型难以兼顾精度与运行效率。近年来, YOLO系列、GhostNet系列等轻量化深度学习网络在嵌入式视觉识别方面展现出广阔前景。现有研究主要集中在单一任务场景, 如姿态估计、电梯门识别、物体边缘检测等, 往往忽视了算法的部署性能、资源占用与复杂环境下的鲁棒性。本文综合考虑精度、速度与可嵌入性, 分别针对电梯门状态检测、电梯门异常动作识别等问题, 展开系统性研究。

一、方法

(一) 基于改进YOLOv8-Pose算法

基于改进YOLOv8是一款集成了多种先进功能的端到端单阶段目标检测算法, 它由输入层、主干网络、颈部网络和头部网络组成, 属于YOLO系列算法的最新成员。这款算法不仅能够进行目标检测, 还能胜任图像分类、实例分割和姿态估计等复杂的视觉任务。在设计上, YOLOv8的主干网络和颈部网络借鉴了YOLOv7 ELAN的设计理念, 优化了YOLOv5中的C3结构, 引入了梯度流更丰富的C2f结构, 并针对不同尺寸的模型调整了通道数以优化性能。头部网络相较于YOLOv5有两大显著改进: 首先, 它采用了当前流行的解耦头结构, 将分类和

检测任务分离; 其次, 它从基于锚点(Anchor-Based)的方法转变为无锚点(Anchor-Free)的检测方式, 进一步提升了检测的灵活性和准确性。YOLOv8-pose则是在YOLOv8的基础上进一步扩展, 实现了目标检测框和关键点检测的功能^[1]。

(二) Canny算法

Canny算法的实现分为下面几个步骤:

(1) 噪声降低: Canny算法会对输入的图像进行平滑处理, 以减少图像中的噪声。这是通过高斯滤波器实现的, 它有助于去除高频噪声, 同时保留图像的低频信息, 即能够保持边缘的清晰度。高斯滤波函数如公式(1)所示:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (1)$$

(2) 梯度计算: 计算图像中每个像素点的梯度幅度和方向。梯度是图像强度变化的度量, 边缘通常对应于梯度的局部最大值。一般基于Sobel滤波器来实现。根据Sobel算子的卷积核, 对图像上的每个像素点 $I(x, y)$, 计算其在水平、垂直方向上的梯度值(G_x 、 G_y), 如式(2)和式(3)所示:

$$G_x(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 I(x+i, y+j) \cdot G_x(i+1, j+1) \quad (2)$$

$$G_y(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 I(x+i, y+j) \cdot G_y(i+1, j+1) \quad (3)$$

(3) 非极大值抑制: 此步骤是为了消除边缘检测过程中的虚假响应。对于每个像素点, 根据之前计算得到

的梯度方向 θ ，确定其梯度方向。然后检查其在梯度方向上的邻近像素点的梯度幅度。对于每个像素点，算法会检查以下四个方向的梯度幅度：梯度方向的正方向、梯度方向的反方向、梯度方向的正方向的对角线方向、梯度方向的反方向的对角线方向。如果当前像素点的梯度幅度不是这四个方向中的最大值，则抑制该点。

(4) 执行双阈值检测：设定两个阈值，一个最小值 L_{min} ，一个最大值 L_{max} 。低于低阈值的梯度被认为不是边缘，高于高阈值的被认为是边缘，而在两者之间的像素被认为是可能的边缘，此时选取其领域像素点的梯度幅值进行判断。

(5) 形成边缘：将孤立的边缘点连接成连续的边缘，形成完整的边缘轮廓。Canny算法在边缘检测领域表现卓越，相较于Sobel和Prewitt等算子，它的优势在于更好地利用了边缘的梯度方向信息，并通过双阈值的滞后阈值处理机制有效减少了误检，从而显著提高了边缘检测的准确性和鲁棒性。

(三) GhostFaceNet算法原理

GhostNet和GhostFaceNet是近年来轻量化神经网络设计领域的重要成果，在资源受限环境中展现了很好的应用潜力。GhostNet通过引入Ghost Module，实现了在减少参数量和计算复杂度的同时，保留较强的特征提取能力。这种设计特别适合移动设备和嵌入式系统等计算资源有限的场景。Ghost Module原理如图1所示。

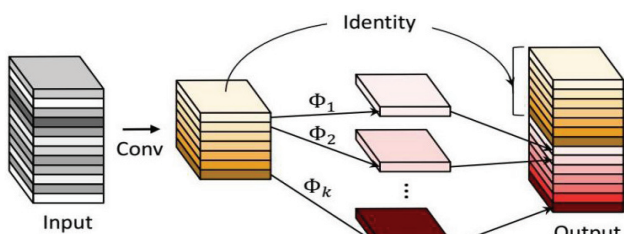


图1 Ghost Module示意图

GhostNet的核心设计思想是通过引入Ghost Module，在较低计算量下生成高效特征。Ghost Module是GhostNet的核心模块，通过减少传统卷积操作中的冗余计算，实现轻量化和高效的特征提取。Ghost Module模块结构如图2所示。

GhostNet注重分类任务，GhostFaceNet在GhostNet的基础上，针对状态识别特定需求进行优化。注重细粒度特征的捕获和判别能力，增加了对特定任务的适配调整，如调整网络深度等。GhostFaceNet网络模型的初始卷积层负责接收标准化的电梯门图像输入，通常尺寸为

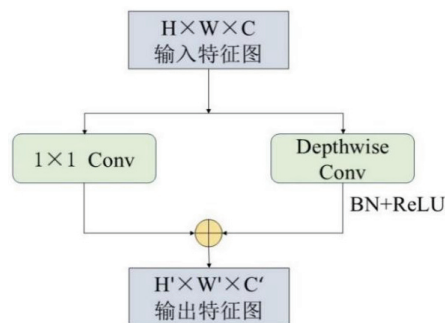


图2 Ghost Module模块结构图

$H \times W \times C$ (H 、 W 、 C 分别为图像的像素宽、高和通道数，例如 $112 \times 112 \times 3$)。通过对原始数据进行归一化处理，并调整图像尺寸以适配网络需求，该部分为后续网络模块提供了高质量的初始特征表示，确保网络输入的规范性和一致性^[2]。

Ghost Bottleneck特征提取层作为GhostFaceNet网络模型的核心部分，由多个Ghost Module模块堆叠构成。Ghost Module模块通过减少计算量生成高效特征，其核心机制在于分离主特征与冗余特征，利用廉价的线性变换生成后者，从而在计算开销降低的同时保持了强大的特征表达能力。在完成特征提取后，特征聚合层通过全局平均池化GAP (Global Average Pooling) 进一步压缩特征图。全局平均池化通过计算每个特征图上所有像素值的平均值，将每个特征图压缩成一个标量，从而将高维特征图压缩为低维的特征向量。该部分通过对空间信息的全局描述，将高维特征映射为低维特征表示，同时减少了参数量和计算开销，有助于降低过拟合的风险^[3]。这一设计不仅提升了模型的运行效率，还为后续的嵌入特征生成提供了固定长度的向量表示，便于后续任务的统一处理。

全连接层将聚合后的高维特征映射为低维嵌入表示，如960D或512D的特征向量，其中每个元素代表网络在该层的某个特定特征的强度。在目标是电梯门开关识别的情况下，这个向量通常会包含面部的高维特征表示，包括电梯信息、电梯门状态、等方面的综合特征。低维嵌入特征具有较强的区分性和紧致性，能够有效支持相似度计算和身份验证等任务需求。此外，全连接层通过进一步的特征优化确保了特征在不同类别之间的差异性，为模型的最终性能提供了保障。

最后Softmax层的作用主要是将网络最后一层的归一化的输出转化为概率值，使得每个类别的预测值介于0和1之间，并且所有类别的概率之和等于1。Softmax会对

网络的每个输出进行指数运算，然后归一化，对于一个多类分类问题，假设网络输出为向量 $Z=[z_1, z_2, \dots, z_K]$ ，其中 K 为类别数， z_{ij} 是模型对于样本 i 类别属于类别 j 的原始输出（未经过 Softmax 函数的 logit 值）。Softmax 函数计算的公式如下：

$$\hat{y}_{ij} = \frac{e^{z_{ij}}}{\sum_{k=1}^C e^{z_{ik}}} \quad (4)$$

GhostFaceNet 不仅在结构设计上充分结合了轻量化与高效特性，还支持通过引入任务专用损失函数进一步优化特征分布。损失函数通过在特征空间内最大化类间差异和最小化类内差异，优化特征向量的几何分布，使得同一身份的特征更加紧凑，不同身份的特征更加分离。常用的识别损失函数如 ArcFace 和 SphereFace，通过结合任务专用损失函数，GhostFaceNet 在资源受限的嵌入式环境中能表现出更强的任务适配能力，在遮挡和复杂场景下的电梯门异常开关中展现出优异的鲁棒性和精度。

二、实验设计与结果

在 COCO-Pose 与定制电梯数据集上验证关节检测性能，引入消融实验评估 ATFL 与 MAY 模块改进效果。电梯门状态识别部分通过真实视频序列验证边缘与直线检测准确率，在不同检测算子下统计开关门识别准确率。电梯门开关识别部分基于公开数据集测试 ResSpatial_GhostFaceNet 识别率，相较 GhostFaceNet 提升显著，验证所提结构在遮挡复杂度与资源开销间的权衡优势。

三、讨论

本文模型均在轻量化设计基础上通过模块级优化提升感知能力。YOLOv8-Pose 与 ATFL 结合关注小目标；MAY 结构适应封闭空间场景部署；Canny 与 PPHT 兼顾边缘提取与形态判断；ResSpatial_GhostFaceNet 网络则在高遮挡识别任务中兼具紧凑性与鲁棒性。此外，多数模型设计均可灵活裁剪以适配不同终端算力，具备良好的迁移与推广能力。

四、结论与展望

本文针对三类典型嵌入式视觉识别任务提出定制化轻量模型优化，涵盖电梯门状态检测、电梯门异常开关识别。未来研究可拓展至多模态传感融合、多场景模型自适应、软硬件协同优化方向，以推动视觉识别系统在智能安防、电梯监控等领域的深入应用。

参考文献

- [1]Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [2]Tan M, Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. PMLR, 2019: 6105–6114.
- [3]Han K, Wang Y, Tian Q, et al. GhostNet: More features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1580–1589.