**文章编号**:1008-1534(2026)04-0319-09

# 基于改进 YOLOv12 的多场景老人跌倒检测

王建霞,田楠楠,李　璇,张晓明

(河北科技大学信息科学与工程学院,河北石家庄　050018)

**摘　要**:为了解决多场景老人跌倒检测中模型计算参数量大与复杂场景精度不足的问题,提出一种基于改进 YOLOv12 的多场景老人跌倒检测算法。首先,在 YOLOv12 主干网络嵌入分组混洗卷积(grouped shuffle convolution,GSConv),同时引入 Triplet 注意力机制,并将颈部结构替换为轻量化颈部结构 Slim Neck;其次,通过自建跌倒数据集对改进 YOLOv12 模型进行训练与验证;最后,在测试集上进行检测实验,评估改进模型的检测精度与鲁棒性。结果表明:改进后的模型相比原 YOLOv12 模型,参数量从 2.56 MB 降低至 2.34 MB,减少了约 8.59%;FLOPs 从 6.3 G 降低至 5.4 G,减少了约 14.29%;mAP@50 从 83.8%提升至 87.6%,mAP@50-95 从 51.3%提升至 55.4%,Recall 从 76.5%提升至 81.3%。改进后的 YOLOv12 模型实现了轻量化与检测精度的有效平衡,可为多场景老人跌倒检测提供一种有效的解决方案。

**关键词**:计算机图像处理;YOLOv12;跌倒检测;Slim Neck;分组混洗卷积;Triplet 注意力机制

**中图分类号**:TP391　　**文献标识码**:A　　**DOI**: 10.7535/hbgykj.2026yx04004

# Multi-scenario elderly fall detection based on an improved YOLOv12

WANG Jianxia, TIAN Nannan, LI Xuan, ZHANG Xiaoming

(School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China)

**Abstract**: To address the problems of a large parameter size and insufficient detection accuracy in complex scenarios for multi-scenario elderly fall detection, a multi-scenario elderly fall detection based on an improved YOLOv12 was proposed. First, grouped shuffle convolution (GSConv) was embedded into the backbone network of YOLOv12, while the Triplet attention mechanism was introduced and the neck network was replaced with the lightweight Slim Neck structure. Second, the improved YOLOv12 model was trained and validated on a self-built fall detection dataset. Finally, detection experiments were conducted on the test set to evaluate the detection accuracy and robustness of the improved model. Experimental results show that, compared with the original YOLOv12 model, the parameter size of the improved model is reduced from 2.56 MB to 2.34 MB, a decrease of 8.59%, and the FLOPs are reduced from 6.3 G to 5.4 G, a decrease of 14.29%. Meanwhile, mAP@50 increases from 83.8% to 87.6%, mAP@50-95 increases from 51.3% to 55.4%, and Recall increases from 76.5% to 81.3%. The improved YOLOv12 model achieves a favorable balance between lightweight architecture and detection accuracy, providing

收稿日期:2025-09-17;修回日期:2026-04-10;责任编辑:丁军苗

基金项目:河北省自然科学基金(F2022208002)

第一作者简介:王建霞(1970—),女,河北临城人,教授,硕士,主要从事遥感图像处理方面的研究。

通信作者:田楠楠。E-mail:15081896020@163.com

王建霞,田楠楠,李璇,等.基于改进 YOLOv12 的多场景老人跌倒检测[J].河北工业科技,2026,43(4):319-327.

Wang Jianxia, Tian Nannan, Li Xuan, et al. Multi-scenario elderly fall detection based on an improved YOLOv12[J]. Hebei Journal of Industrial Science and Technology, 2026, 43(4): 319-327.

an effective solution for multi-scenario elderly fall detection.

**Keywords**: computer image processing; YOLOv12; fall detection; Slim Neck; grouped shuffle convolution (GSConv); Triplet attention mechanism

在养老院、社区养老中心等场景中,老人跌倒具有突发性强、危害性高等特点[1]。传统的跌倒监测主要依赖人工巡视或佩戴式传感器,不仅消耗大量人力,还存在监测效率较低和隐私保护不足等问题[2]。随着计算机视觉与深度学习技术的发展,基于目标检测的人体行为识别方法能够同时完成目标定位与类别判别,已被广泛应用于跌倒检测与安全监测领域[3]。

作为单阶段目标检测算法的典型代表,YOLO系列模型凭借较高的检测精度与实时性能,在目标检测任务中展现出显著优势[4]。针对模型轻量化与检测效率问题,相关学者开展了广泛研究。杨贤志等[5]通过分组卷积与特征图级联结构增强特征表达能力,在降低计算开销的同时,保持较好的检测性能;Cai 等[6]在 YOLOv8 中融合分组混洗卷积(grouped shuffle convolution,GSConv)、注意力机制与智慧交并比(wise intersection over union,WIoU)损失函数,实现了检测精度与推理效率的协同优化;田泉泉等[7]提出基于多尺度特征的轻量级目标检测方法,通过多尺度特征融合提升模型对不同尺度目标的适应能力。

随着目标检测技术的不断发展,YOLO 系列模型在跌倒检测领域中的应用逐渐增多。Qin 等[8]针对复杂场景下跌倒目标存在遮挡与姿态变形等问题,在 YOLOv8 基础上通过引入动态卷积增强特征提取能力。Hwang 等[9]针对实时跌倒监测需求,提出跌倒检测 YOLO 网络(fall detection YOLO,FD-YOLO)模型,通过优化网络结构提升检测效率与识别性能,实现了高精度的实时跌倒检测。

针对养老院等场景老人跌倒检测的实际需求,研究者进一步开展了相关研究。王晓宇[10]基于人体姿态估计实现跌倒行为识别,通过提取人体关键点信息分析动作特征;吴迪[11]结合轻量化 YOLOv5s 与时序建模方法,提高了跌倒识别精度并降低了类跌倒动作误检率;王润之[12]融合视觉与声学信息,实现多源信息协同的跌倒检测;崔杰[13]针对复杂背景与小目标问题,对实时检测 Transformer(real-time detection transformer, RT-DETR)与 YOLOv8 进行优化,在提升检测性能的同时,保持较高的运行效率。

尽管上述研究在多场景老人跌倒检测方向取得了一定进展,但面对复杂姿态变化、目标尺度差异以及类跌倒动作易混淆等挑战,模型的检测性能仍有提升空间。同时,在满足边缘设备部署条件下,如何在保证模型轻量化的同时增强关键特征的表达能力,并提升复杂场景下老人跌倒检测模型的精度,仍是亟待解决的问题。

本研究在 YOLOv12 模型的基础上引入 GSConv 和 Triplet 注意力机制,并将其颈部结构替换为轻量化颈部结构 Slim Neck,对模型特征提取与特征融合能力进行优化,以实现模型检测精度与轻量化之间的有效平衡。

## 1 改进 YOLOv12 模型设计

### 1.1 整体网络架构

为提高多场景下老人跌倒检测模型的精度并降低资源消耗,本文对原 YOLOv12 模型的主干网络和颈部结构进行针对性改进,包含以下 3 个核心部分:

1)在主干网络的特征提取阶段嵌入 GSConv,通过分组卷积降低卷积计算量,并利用通道洗牌机制增强跨通道信息交互,从而在保证表达能力的同时减少冗余计算;

2)在主干网络中引入 Triplet 注意力机制,分别从空间行、空间列及通道 3 个维度建立特征关联,有效提升模型对不同跌倒姿态(如侧倒、蜷缩等)的敏感度与判别力,降低复杂场景下的误检与漏检;

3)将原有的颈部结构替换为 Slim Neck 轻量化颈部结构,以多尺度分组混洗卷积跨阶段局部(variety of view grouped shuffle convolution cross stage partial,VoV-GSCSP)模块为核心,结合通道混洗机制实现高效特征融合,在显著减少参数量与计算量的同时,保持多尺度特征的传递能力。

改进的 YOLOv12 模型网络架构图如图 1 所示,图中灰色模块为本文模型的创新部分。主干网络嵌入 GSConv 模块与 Triplet 注意力机制,以高效提取并增强跌倒姿态的判别性特征;颈部结构采用轻量化 Slim Neck,基于 VoV-GSCSP 模块完成特征融合与通道重构,在降低计算复杂度的同时,提升了多尺度特征融合效率;检测头沿用 YOLOv12 的高性能解耦检测头设计,通过将类别分类与边界框回归任务解耦独立处理,实现在多场景下老人跌倒行为的精准定位与类别识别。

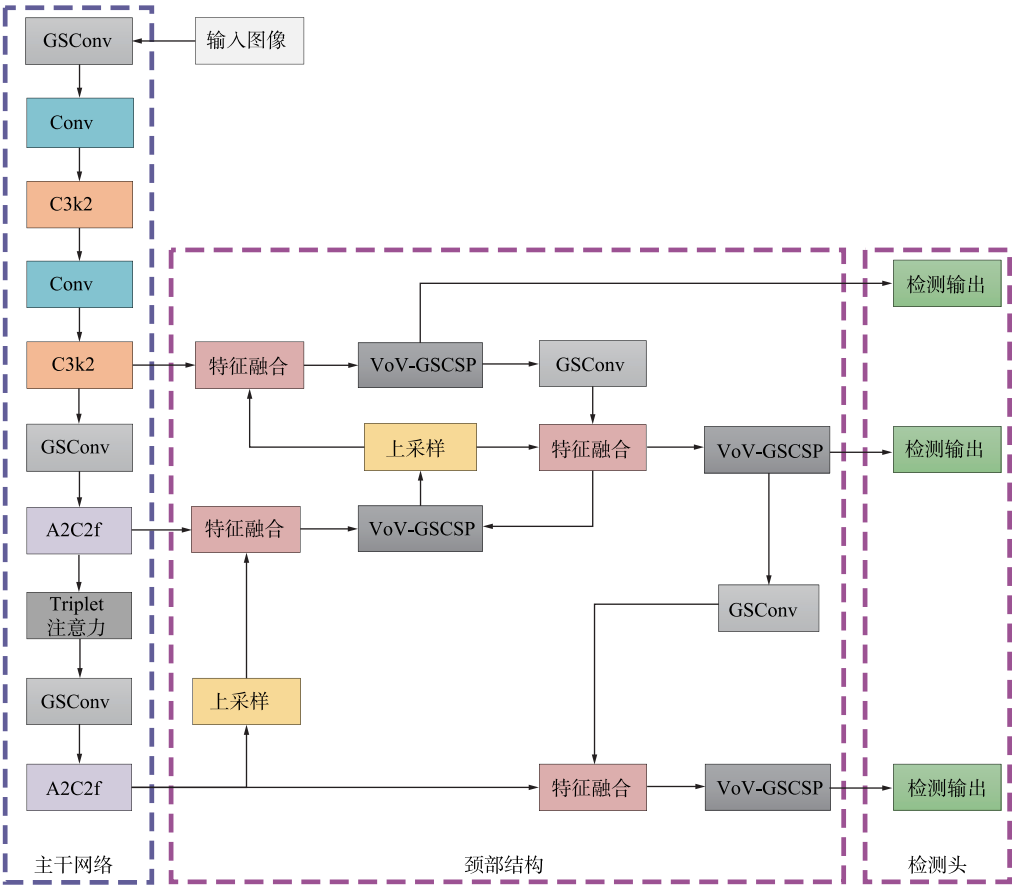


图 1　改进的 YOLOv12 模型网络架构图

Fig. 1　Network architecture of the improved YOLOv12 model

## 1.2　基于 GSConv 的主干网络优化设计

### 1.2.1　GSConv 结构及原理

GSConv 是一种轻量化卷积结构，通过双分支设计与通道洗牌机制，在降低计算开销的同时，提升特征表达能力[14]，适配养老场景中边缘设备对高效计算和高精度特征提取的需求。GSConv 由标准卷积（standard convolution，SConv）分支与深度可分离卷积（depthwise separable convolution，DSConv）分支构成，其结构图如图 2 所示。

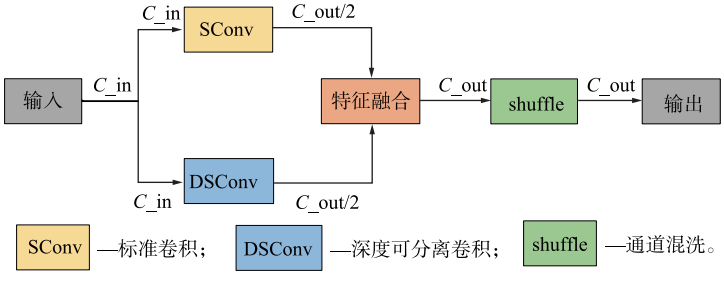


图 2　GSConv 结构图

Fig. 2　Structure of the GSConv module

输入特征分别经过标准卷积分支和深度可分离卷积分支进行特征提取，随后对 2 路特征进行特征融合，并通过通道混洗机制实现通道信息交互，最终得到输出特征。SConv 分支与 DSConv 分支的工作原理如下。

1）SConv 分支

SConv 分支结构图如图 3 所示。该分支采用全通道输入的标准卷积模式，通道数为 $C_in$ 的输入特征图通过 3×3 卷积核进行计算，将通道数精简为目标输出维度的一半（即 $C_out/2$）的同时，增强了通

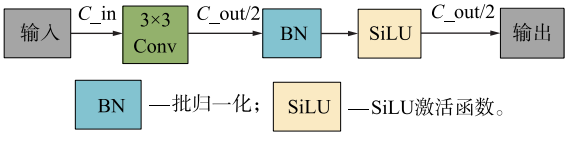


图 3　SConv 分支结构图

Fig. 3　SConv branch structure diagram

道间的特征交互与表达能力，通过批归一化与 SiLU 激活函数的处理，输出该分类的最终特征。

2）DSConv 分支

DSConv 分支结构图如图 4 所示。首先，通道数为 $C_in$ 的输入特征图被输入深度卷积层（又称逐通道卷积层），利用 3×3 卷积核进行逐通道的独立空间特征提取，并输出特征图；然后，输出的特征图直接送入 1×1 的逐点卷积层，在进行跨通道信息融合的同时，将通道数精简为 $C_out/2$；最后，特征图依次通过批归一化与 SiLU 激活函数的处理，输出该分支的最终特征。

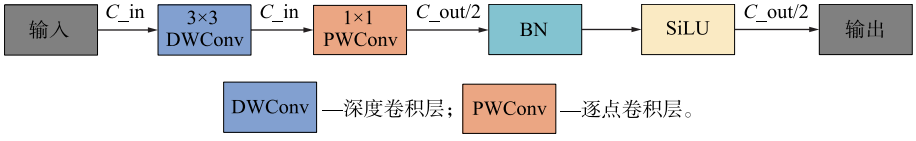


图 4 DSConv 分支结构图

Fig. 4 DSConv branch structure diagram

1.2.2 GSConv 主干网络嵌入

本文将 GSConv 嵌入主干网络的第 0 层（前期层）、第 5 层（中期层）、第 8 层（后期层）。在前期层插入 GSConv 能够有效减少计算量并增强特征提取能力，从而提高模型的效率和精度。在中、后期层引入 GSConv 有助于减少冗余计算，并通过通道混洗操作增强不同组间的特征交互，从而提升模型对复杂跌倒姿态的识别能力。GSConv 通过分组卷积与通道混洗操作降低了模型计算的复杂度，并通过轻量化特征交互机制与 C3k2 模块、A2C2f 模块的深层上下文建模能力相结合，实现了模型效率与特征表达能力的协同提升。

## 1.3 基于 Triplet 注意力机制的主干网络优化设计

1.3.1 Triplet 注意力模块构成与工作原理

Triplet 注意力机制通过对特征图的行、列及通道 3 个维度生成注意力权重，实现跨维度特征交互，Triplet 注意力模块结构图如图 5 所示。该机制主要通过维度置换与零维池化实现轻量化注意力建模[15]。维度置换打破通道维度与空间维度的独立性，建立两者的关联；零维池化则同时提取特征图在特定维度上的最大值与平均值，并将两者进行特征融合，在保留关键特征的前提下，降低后续计算量。零维池化计算公式为

$$Z\text{-Pool}(X)=[\text{MaxPool}_0(X),\text{AvgPool}_0(X)], \tag{1}$$

式(1)中：$X$ 表示输入特征图；$Z$-Pool(•) 表示经过零维池化后的输出特征；$\text{MaxPool}_0(\cdot)$ 和 $\text{AvgPool}_0(\cdot)$ 分别表示在第 0 维（通道维度）方向上的最大池化与平均池化操作；[,] 表示特征在目标维度上进行拼接操作。

图 5 中，Triplet 注意力模块通过 3 条并行分支（通道-高度维度注意力分支、通道-宽度维度注意力分支、空间注意力分支）实现。其中，$C$、$H$ 和 $W$ 分别表示输入特征图的通道数、空间上的高度和宽度。

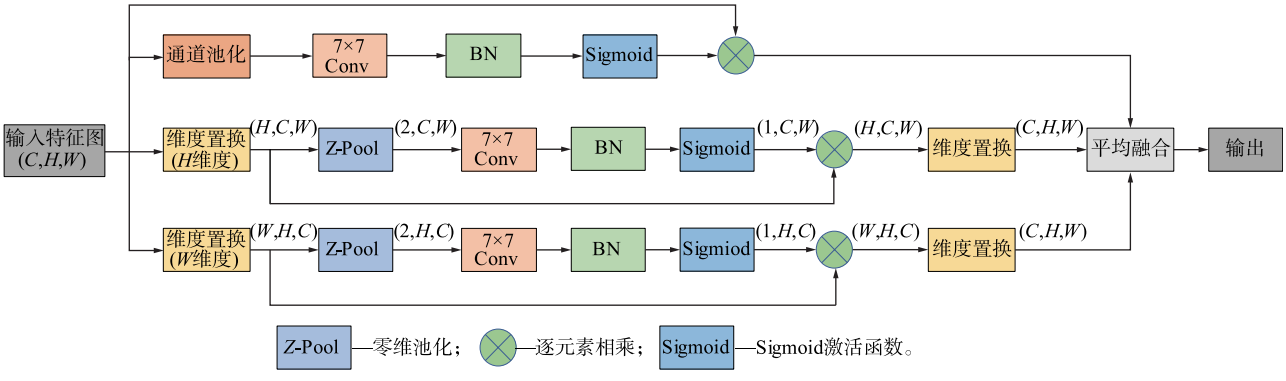


图 5 Triplet 注意力模块结构图

Fig. 5 Architecture of the Triplet attention module

1）跨维度特征交互注意力分支

该分支包含通道-高度维度注意力分支和通道-宽度维度注意力分支，通过跨维度交互建模，强化关键特征响应，提升特征表征能力。首先，对输入特征图进行维度置换，使通道维度分别与空间高度维度或宽度维度建立关联，从而实现跨维度特征交互；其次，采用零维池化，将通道维度进行压缩，以降低计算复杂度；然后，通过 7×7 卷积对跨维度特征进行建模，并经批归一化与 Sigmoid 激活函数生成对应的注意力权重；最后，将生成的注意力权重与维度置换后的特征图进行逐元素相乘，并再次通过维度置换恢复原始特征维度。

2）空间注意力分支

该分支直接对输入特征进行空间维度建模，以突出关键区域的特征响应。首先，对输入特征图进行通道池化操作，以突出重要通道特征；其次，通过 7×7 卷积进行特征建模，并经批归一化与 Sigmoid 激活函数生成通道注意力权重；最后，将注意力权重与原始输入特征逐元素相乘。

3）特征融合与输出

3 条分支的增强特征通过平均融合得到最终输出，不仅能够平衡不同维度注意力的贡献，还能够避免单一分支过度关注局部特征的问题，从而增强模型对于多场景下老人跌倒检测的鲁棒性与泛化能力。

Triplet 注意力机制在不显著增加计算量的前提下，通过多维度特征交互增强特征表达能力，解决单维度注意力的信息片面性问题，有效提升对复杂跌倒姿态的识别精度。

1.3.2　Triplet 注意力的集成方法

首先，本文提出的改进 YOLOv12 模型将 Triplet 注意力模块插入主干网络的第 6 层和第 8 层之间，该位置处于特征初步提取与变换的关键阶段。该模块在特征提取阶段对通道与空间信息进行联合建模，从而增强关键区域特征并抑制背景干扰。经过该模块处理后，特征图的关键信息得到有效增强，有助于后续颈部结构和检测头更准确地进行特征融合和目标检测，提高模型对复杂场景的检测能力。

### 1.4　Slim Neck 颈部结构设计

1.4.1　Slim Neck 颈部结构及工作原理

Slim Neck 颈部结构以 VoV-GSCSP 模块为核心架构单元，在保障特征表达能力的同时，实现模型的轻量化与性能提升[16]。

VoV-GSCSP 模块结构图如图 6 所示，其采用双分支结构，输入通道数为 $C_in$。其中一个分支依次经 SConv 模块、GSbottleneck 模块进行特征提取与精炼；另一个分支经 SConv 模块保留基础特征。最终，通过特征融合增强模型的特征表达能力。GSbottleneck 模块结构图如图 7 所示，其输入通道数为 $C$。该模块的一个分支由串联的 2 个 GSConv 模块构成，用于逐级增强特征表达；另一个分支通过 SConv 模块保留基础特征信息，实现对主干特征的补充。2 个分支的输出特征通过逐元素相加进行特征融合，得到最终输出特征。

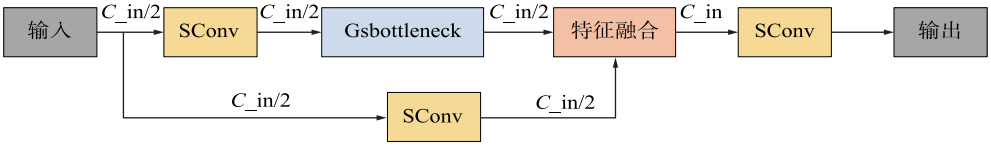


图 6　VoV-GSCSP 模块结构图

Fig. 6　Structure of the VoV-GSCSP module

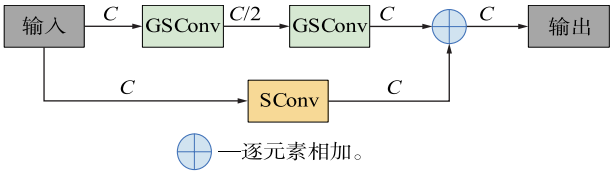


图 7　GSbottleneck 模块结构图

Fig. 7　Structure of the GSbottleneck module

1.4.2　Slim Neck 颈部结构集成方法

Slim Neck 颈部结构能够实现通道精简与高效特征融合，其 VoV-GSCSP 模块能够依托分组卷积与残差连接的轻量化设计大幅削减计算开销，并完整保留主干网络提取的关键特征信息。与 YOLOv12 模型的原始颈部结构相比，该设计用更轻量的分组卷积与残差连接替代了复杂的卷积堆叠，用更少的参数维持特征融合的有效性。将不同尺度的特征在 Slim Neck 颈部结构中进行级联融合，能够显著提升模型对不同尺度跌倒目标的检测鲁棒性。

## 2　实验及结果分析

### 2.1　数据集

为开展面向多场景的老人跌倒检测研究，考虑到养老院等监控环境的隐私合规性，构建了专门针对室内场景的跌倒检测数据集。

该数据集是通过对公开可获取的 Le2i 跌倒检测视频数据集进行帧级抽取与重新标注，并结合特定室内场景进行合规化补充而完成。同时，针对养老院、养老社区中心等特定室内环境进行了补充采集，涵盖了老人站立、坐下以及跌倒的行为动作。

整个数据集共包含 8 394 张图像，按照 7∶1.5∶1.5 的比例分配训练集、验证集和测试集。其中，训练集 5 862 张，验证集 1 310 张，测试集 1 222 张。针对研究任务需求，将图像中的人体行为明确标注为 3 个类别，分别是站立、坐下以及跌倒，且数据集内部的样本比例分布均衡。训练集中站立、坐下与跌倒的样本数量分别约占总体的 40%、30%和 30%，验证集与测试集的人体行为类别分布比例与训练集保持一致。

## 2.2 实验配置

用于本文实验的主要硬件配置：显卡为 NVIDIA RTX 4060(由英伟达公司提供)；GPU 为 NVIDIA GeForce RTX 4060 Laptop(由英伟达公司提供)。深度学习框架采用 PyTorch 2.7.1，底层并行计算依赖 CUDA 11.8 工具包，开发与脚本编写采用 Python 语言。在此基础上，本文对改进 YOLOv12 模型进行训练，实验超参数设置如表 1 所示。

表 1 实验超参数设置

Tab. 1 Configuration of experimental hyperparameters

| 参 数 | 取 值 |
|---|---|
| 优化器 | SGD |
| 批次大小/个 | 32 |
| 训练周期/轮 | 150 |
| 权重衰减系数 | 0.000 5 |
| 初始学习率 | 0.01 |
| 学习率衰减系数 | 0.01 |

## 2.3 评价指标

为了系统地评估本文模型的综合性能，实验选取精确率(Precision)、召回率(Recall)以及平均精度均值(mAP)作为核心准确性评价指标。其计算公式分别如式(2)—式(5)所示。

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}}, \tag{2}$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}}, \tag{3}$$

$$\text{AP}=\int_0^1 P(R)\,\text{d}(R), \tag{4}$$

$$\text{mAP}=\frac{1}{m}\sum_{i=1}^{m}\text{AP}_i, \tag{5}$$

式(2)—式(5)中：TP 表示真正例数量，即模型正确检测到的目标数量；FP 表示假正例数量，即模型误检的目标数量；FN 为假负例数量，即模型漏检的真实目标数量；$m$ 为目标类别总数(本文取 $m=3$，对应跌倒、站立、坐下 3 个人体行为类别)；$\text{AP}_i$ 为第 $i$ 个类别的平均精度，通过计算 $P$-$R$ 曲线下的面积获取。

在本文实验中，采用 mAP@50 和 mAP@50-95 作为核心指标对模型检测性能进行评估。其中，mAP@50 指交并比(IoU)阈值设为 0.50 时的平均精度均值；mAP@50-95 则表示 IoU 阈值在 0.50～0.95 之间(步长为 0.05)的平均精度均值。

此外，为综合衡量模型的空间存储复杂度和计算开销，实验进一步引入参数量(Params)与浮点运算次数(FLOPs)分别作为评估模型结构复杂度与计算复杂度的性能指标。

## 2.4 结果分析

### 2.4.1 对比实验

为验证本文所提改进 YOLOv12n 模型在多场景下老人跌倒检测任务中的有效性，基于自建数据集，将改进模型与经典轻量级模型(YOLOv5n、YOLOv8n、YOLOv11n、NanoDet-m)及原始 YOLOv12n 模型进行对比。上述模型均聚焦于边缘设备部署与高实时性检测，在参数规模与适用场景上与本文模型具备高度一致性，从而保障了对比结果的客观与公平；同时，作为跌倒检测研究中广泛使用的模型，上述模型的检测结果能够有力凸显本文改进方案的先进性。本文模型与其他经典轻量模型的对比实验结果如表 2 所示。

表 2 本文模型与其他经典轻量模型的对比实验结果

Tab. 2 Comparative experimental results between the proposed model and other classical lightweight models

| 模 型 | mAP@50/% | mAP@50-95/% | Recall/% | Params/MB | FLOPs/G |
|---|---|---|---|---|---|
| YOLOv5n | 86.4 | 53.4 | 80.7 | 2.50 | 7.1 |
| YOLOv8n | 84.3 | 49.5 | 79.5 | 3.00 | 8.1 |
| YOLOv11n | 84.4 | 51.0 | 79.1 | 2.58 | 6.3 |
| YOLOV12n | 83.8 | 51.3 | 76.5 | 2.56 | 6.3 |
| NanoDet-m | 81.8 | 44.1 | 66.5 | 0.95 | 2.9 |
| 本文模型 | 87.6 | 55.4 | 81.3 | 2.34 | 5.4 |

由表 2 可知，在自建跌倒数据集上，本文模型在 mAP@50、mAP@50-95 和 Recall 指标上表现最优，分别为 87.6%、55.4% 和 81.3%，YOLOv5n 在 mAP@50、mAP@50-95 和 Recall 指标上与本文模型最相近，分别为 86.4%、53.4% 和 80.7%，但其 Params 和 FLOPs 均高于本文模型，分别为 2.50 MB 和 7.1 G，模型复杂度较高。YOLOv8n 在精度上表现一般，其 mAP@50 为 84.3%，mAP@50-95 为 49.5%，Recall 为 79.5%，且 Params 与 FLOPs 最大，分别为 3.00 MB 和 8.1 G，在资源受限环境下部署难度较大。YOLOv11n 和 YOLOv12n 的 Params 与 FLOPs 相对较低，但在精度和稳定性方面表现不足，mAP@50 分别为 84.4% 和 83.8%，mAP@50-95 分别为 51.0% 和 51.3%，Recall 分别为

79.1%和 76.5%，均低于本文模型。NanoDet-m 是轻量化表现最优的模型，其 Params 为 0.95 MB，FLOPs 为 2.9 G，但其 mAP@50 为 81.8%，mAP@50-95 为 44.1%，Recall 为 66.5%，远低于本文模型。相比之下，本文模型的 Params 为 2.34 MB，FLOPs 为 5.4 G，明显低于 YOLOv5n、YOLOv8n、YOLOv11n 和 YOLOv12n 模型。与 NanoDet-m 模型相比，本文模型的 Params 与 FLOPs 较高，但是其 mAP@50 提升了 5.8 个百分点，mAP@50-95 提升了 11.3 个百分点，Recall 提升了 14.8 个百分点，实现了高精度与低复杂度的有效统一。

综上所述，本文模型在检测精度与边缘端部署可行性之间取得了最佳平衡，具有更强的实际应用价值。

### 2.4.2　消融实验

为了系统评估不同改进模块对多场景下老人跌倒检测的影响，本文设计了 4 种实验模型进行消融实验，验证各个模块的有效性。模型 1 为基线模型 YOLOv12n；模型 2 在基线模型的主干网络中嵌入 GSConv；模型 3 在基线模型的主干网络中引入 Triplet 注意力机制；模型 4 将基线模型的原颈部结构替换为轻量级的 Slim Neck 颈部结构；模型 5 综合了模型 2、模型 3 和模型 4 的改进点，为最终改进模型。消融实验结果如表 3 所示。

表 3　消融实验结果

Tab.3　Ablation experiment results

| 模型编号 | 模　型 | mAP@50/% | mAP@50-95/% | Recall/% | Params/MB | FLOPs/G |
|---|---|---|---|---|---|---|
| 1 | YOLOv12n | 83.8 | 51.3 | 76.5 | 2.56 | 6.3 |
| 2 | YOLOv12n+GSConv | 85.2 | 50.0 | 78.5 | 2.34 | 6.0 |
| 3 | YOLOv12n+Triplet | 86.8 | 54.1 | 77.8 | 2.57 | 6.4 |
| 4 | YOLOv12n+Slim Neck | 86.5 | 51.4 | 76.7 | 2.67 | 6.3 |
| 5 | YOLOv12n+GSConv+Triplet+Slim Neck | 87.6 | 55.4 | 81.3 | 2.34 | 5.4 |

由表 3 分析可知：1）将 GSConv 融入主干网络后，模型在 mAP@50 上由 83.8%提升至 85.2%，mAP@50-95 略有下降，从 51.3%降低至 50.0%，Recall 从 76.5%提升至 78.5%。同时，Params 由原来的 2.56 MB 降低到 2.34 MB，FLOPs 由 6.3 G 降低到 6.0 G，基本保持稳定。结果表明，GSConv 在减少冗余计算的同时，有效提升了主干网络的特征表达能力，对轻量化目标检测具有积极作用。2）将 Triplet 注意力机制引入基线模型的主干网络后，模型检测性能得到较显著的提升。mAP@50 由 83.8%提升至 86.8%，mAP@50-95 由 51.3%提升至 54.1%，Recall 由 76.5%提升至 77.8%。同时，模型的 Params 由 2.56 MB 增加至 2.57 MB，FLOPs 由 6.3 G 增加至 6.4 G。结果表明，该机制能够有效增强通道与空间维度之间的特征相关性，提高特征表达的判别能力，从而提升模型对跌倒目标的识别精度。3）将基线模型的颈部结构替换为 Slim Neck 轻量化颈部结构后，模型在检测性能方面表现出一定提升。其中，mAP@50 由 83.8%提升至 86.5%，mAP@50-95 由 51.3%提升至 51.4%，Recall 由 76.5%提升至 76.7%。同时，模型的 Params 由 2.56 MB 增加至 2.67 MB，而 FLOPs 保持在 6.3 G，基本未发生变化。结果表明，Slim Neck 轻量化颈部结构在一定程度上增强了多尺度特征表达能力，从而带来了检测性能的小幅提升。4）在融合了 GSConv、Triplet 注意力机制与 Slim Neck 轻量化颈部结构后，模型整体性能进一步提升，取得了各项指标的最优结果。本文模型相比原 YOLOv12n 模型，参数量从 2.56 MB 降低至 2.34 MB，减少了约 8.59%；FLOPs 从 6.3 G 降低至 5.4 G，减少了 14.29%；mAP@50 从 83.8%提升至 87.6%，mAP@50-95 从 51.3%提升至 55.4%，Recall 从 76.5%提升至 81.3%。结果表明，3 种改进策略在整体架构中具有良好的协同作用。

综上所述，GSConv、Triplet 注意力机制和 Slim Neck 轻量化颈部结构均对模型性能提升与轻量化目标起到积极作用。其中 GSConv 有效降低了模型的冗余参数与计算开销，Triplet 注意力机制在不增加模型复杂度的前提下显著提升了检测精度，而 Slim Neck 轻量化颈部结构则通过结构优化改善了特征融合效率。最终改进模型在检测精度、召回率与计算复杂度之间取得了较优平衡，更加契合养老院等实际应用场景对实时性、稳定性与可靠性的综合需求。

### 2.4.3　可视化分析

本文模型对验证集部分样本的跌倒检测可视化结果如图 8 所示。由图可知，改进模型能够较为准确地识别出当前人物的动作状态，并显示出相应的置信度，验证了其检测结果的有效性。

选取样本图 A—样本图 D 作为测试样本，采用 YOLOv5n、YOLOv8n、YOLOv11n、YOLOv12n、NanoDet-m 及本文模型对人物姿态进行检测。本文模型与其他模型检测可视化的对比结果如图9所示。

图 8 本文模型对验证集部分样本的跌倒检测可视化结果

Fig. 8 Visualization results of the proposed model for fall detection on validation set samples

图 9 本文模型与其他模型检测可视化的对比结果

Fig. 9 Visual comparison of detection results between the proposed model and other models

其中，样本图 A—C 中人物对应的姿态分别是跌倒(fall)、跌倒(fall)和站立(stand)。样本图 D 中存在 2 个人物，左边人物称为人物 1，其姿态为坐下(sit)；右边人物称为人物 2，其姿态为站立(stand)。

由图 9 可知，对样本图 C，6 种模型均正确识别人物状态为站立，其中，本文模型的置信度最高；对样本图 A 和样本图 B，人物状态均被正确识别为跌倒，本文模型同样表现最优。在样本图 D 中，左侧人物检测存在不同问题：YOLOv5n 置信度较低，YOLOv8n 和 NanoDet-m 出现漏检，YOLOv11n 和 YOLOv12n 出现状态混淆的误检，而本文模型左侧人物的检测框出现重叠(同时输出站立和坐下 2 类目标检测框)；右侧人物状态均被正确识别为站立。

为更清晰地展示各模型在样本图上的检测效果，将各模型在样本图上的检测结果及置信度列于表 4。由表可见，本文模型在大多数样本中的置信度仍普遍高于其他模型，进一步验证了其在多场景老人跌倒检测任务中的有效性与稳定性。

表 4 各模型在样本图上的检测结果及置信度

Tab. 4 Detection status and confidence of different models on sample images

| 模 型 | 识别状态/置信度 | | | | |
|---|---|---|---|---|---|
| | 样本图 C | 样本图 A | 样本图 B | 样本图 D-人物 1 | 样本图 D-人物 2 |
| YOLOv5n | 站立/0.75 | 跌倒/0.77 | 跌倒/0.74 | 坐下/0.31 | 站立/0.82 |
| YOLOv8n | 站立/0.81 | 跌倒/0.82 | 跌倒/0.76 | 漏检 | 站立/0.82 |
| YOLOv11n | 站立/0.75 | 跌倒/0.81 | 跌倒/0.54 | 站立/0.42 | 站立/0.79 |
| YOLOv12n | 站立/0.76 | 跌倒/0.78 | 跌倒/0.68 | 站立/0.43 | 站立/0.69 |
| NanoDet-m | 站立/0.74 | 跌倒/0.69 | 跌倒/0.57 | 漏检 | 站立/0.72 |
| 本文模型 | 站立/0.84 | 跌倒/0.86 | 跌倒/0.79 | 坐下/0.64，站立/0.28 | 站立/0.84 |

## 4 结 语

本文针对多场景下老人跌倒检测任务，提出了一种基于改进 YOLOv12 的跌倒检测模型，在模型精度与轻量化之间实现了良好平衡。主要研究结论如下。

1) 消融实验表明，GSConv、Triplet 注意力机制

及轻量化颈部结构 Slim Neck 均能够有效提升 YOLOv12n 模型性能。Triplet 注意力机制对检测精度提升显著，GSConv 在降低模型复杂度的同时保持了较好的检测性能，Slim Neck 进一步优化了模型的特征融合能力。各改进模块协同作用，提升了本文模型在多场景老人跌倒检测任务中的综合性能。

2）对比实验表明，本文模型相比原 YOLOv12n 模型，参数量从 2.56 MB 降低至 2.34 MB，减少了约 8.37%；FLOPs 从 6.3 G 降低至 5.4 G，减少了 14.29%；mAP@50 从 83.8%提升至 87.6%，mAP@50-95 从 51.3%提升至 55.4%，Recall 从 76.5%提升至 81.3%。尽管本文模型的 Params 和 GFLOPs 高于超轻量 NanoDet-m 模型，但其 mAP@50、mAP@50-95 和 Recall 较 NanoDet-m 模型分别提升了 5.8 个百分点、11.3 个百分点和 14.8 个百分点。

本文模型在检测精度和轻量化之间达到了良好平衡，精准契合多场景老人跌倒对检测准确性和计算资源受限的双重需求。尽管如此，本研究仍存在不足：其一，实验验证主要基于对公开可获取的跌倒检测视频数据进行帧级抽取与重新标注而自建的跌倒数据集，极端场景样本（例如多目标、场景遮挡、轮椅辅助跌倒等）相对不足，导致模型在此类场景中的泛化性能仍有待提升；其二，本文模型仅聚焦于跌倒事件的检测，尚未结合时序信息实现跌倒风险预判，难以满足实际养老场景中提前干预的更高需求。因此，未来的研究将从以下两方面展开：首先，扩充和优化数据集，重点补充多目标遮挡与特殊设备辅助下的跌倒样本，并尝试引入迁移学习与领域自适应方法，以增强模型的跨场景适配能力；其次，在模型功能方面，引入 3D 卷积、时序注意力或光流分析等方法，使模型具备对跌倒前动态特征的感知与预测能力，构建集检测与预警于一体的双分支模型，为养老院、社区养老中心等智能安全监控提供更具前瞻性的技术支撑。

## 参考文献/References：


[1] 刘晓红，吴淼，牛茜. 老年人跌倒危害因素分析[J]. 北京医学，2021，43(6)：533-534.
Liu Xiaohong，Wu Miao，Niu Qian. Analysis on risk factors of falls in the elderly[J]. Beijing Medical Journal，2021，43(6)：533-534.

[2] 李平，黄同成，曾文飞，等. 跌倒监测系统研究及应用[J]. 电声技术，2022，46(2)：122-124.
Li Ping，Huang Tongcheng，Zeng Wenfei，et al. Research and application of fall monitoring system[J]. Audio Engineering，2022，46(2)：122-124.

[3] Alam E，Sufian A，Dutta P，et al. Vision-based human fall detection systems using deep learning：A review[J]. Computers in Biology and Medicine，2022，146：105626.

[4] Wang C Y，Bochkovskiy A，Liao H Y M. YOLOv7：Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver：IEEE，2023：7464-7475.

[5] 杨贤志，黄国方，周宁宁. 基于分组卷积和特征图级联的轻量级目标检测[J]. 计算机应用研究，2021，38(5)：1590-1594.
Yang Xianzhi，Huang Guofang，Zhou Ningning. Light-weight object detection network based on group convolution and feature maps cascade[J]. Application Research of Computers，2021，38(5)：1590-1594.

[6] Cai Shaobin，Zhang Xiangkui，Mo Yuchang. A lightweight underwater detector enhanced by attention mechanism，GSConv and WIoU on YOLOv8[J]. Scientific Reports，2024，14(1)：25797.

[7] 田泉泉，张杨. 基于多尺度特征的轻量级交通标志检测模型[J]. 河北工业科技，2026，43(1)：10-20.
Tian Quanquan，Zhang Yang. Lightweight traffic sign detection model based on multi-scale features[J]. Hebei Journal of Industrial Science and Technology，2026，43(1)：10-20.

[8] Qin Yong，Miao Wuqing，Qian Chen. A high-precision fall detection model based on dynamic convolution in complex scenes[J]. Electronics，2024，13(6)：1141.

[9] Hwang H，Kim D，Kim H. FD-YOLO：A YOLO network optimized for fall detection[J]. Applied Sciences，2025，15(1)：453.

[10] 王晓宇. 老人跌倒检测算法的研究及应用[D]. 呼和浩特：内蒙古大学，2024.
Wang Xiaoyu. Research and application of fall detection algorithms for elderly people[D]. Hohhot：Inner Mongolia University，2024.

[11] 吴迪. 基于改进 YOLOv5s 和 GSAlpha-BiGRU 的跌倒检测方法研究[D]. 淮南：安徽理工大学，2025.
Wu Di. Research on fall detection method based on improved YOLOv5s and GSAlpha BiGRU[D]. Huainan：Anhui University of Science and Technology，2025.

[12] 王润之. 融合多源信息的独居老人跌倒检测系统设计与实现[D]. 重庆：西南大学，2025.
Wang Runzhi. The design and implementation of a fall detection system for elderly living alone by integrating multi-source information[D]. Chongqing：Southwest University，2025.

[13] 崔杰. 基于深度学习的老年人跌倒检测算法[D]. 天津：天津理工大学，2025.
Cui Jie. Deep Learning-based fall detection algorithm for the elderly[D]. Tianjin：Tianjin University of Technology，2025.

[14] Wang Jianlong，Qin Congcong，Hou Beibei，et al. LCGSC-YOLO：A lightweight apple leaf diseases detection method based on LCNet and GSConv module under YOLO framework [J]. Frontiers in Plant Science，2024，15：1398277.

[15] Misra D，Nalamada T，Arasanipalai A U，et al. Rotate to attend：Convolutional triplet attention module [C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway：IEEE，2021：3138-3147.

[16] Li Hulin，Li Jun，Wei Hanbing，et al. Slim-neck by GSConv：A lightweight-design for real-time detector architectures [J]. Journal of Real-Time Image Processing，2024，21(3)：62.