引用格式：莫太平, 郑德广, 孙鹏, 等. 融合图通道注意力与多阶门控残差网络的换衣行人重识别模型 [J]. 电子科技大学学报, 2026, 55(3): 374-384.
MO T P, ZHENG D G, SUN P, et al. Cloth-changing person re-identification model integrating graph channel attention and multi-order gated residual network[J]. Journal of University of Electronic Science and Technology of China, 2026, 55(3): 374-384.

融合图通道注意力与多阶门控残差网络的换衣行人重识别模型

莫太平，郑德广，孙 鹏*，覃汉岳，张向文，张 斌

（桂林电子科技大学电子工程与自动化学院，桂林 541004）

摘要：行人重识别（Re-ID）作为计算机视觉的重要任务，旨在跨摄像头实现行人身份的精准匹配。然而，换衣场景下的衣物特征变化导致传统 ReID 方法在特征提取和信息捕捉方面面临挑战。为此，提出一种融合图通道注意力与多阶门控残差网络的换衣行人重识别模型（AMGA-ResNet50）。首先，自适应图通道注意力模块（AGA）将图卷积理论引入换衣通道注意力机制中，将通道依赖关系建模为特征节点连接，通过图结构处理有效捕捉换衣行人的关键身份特征；其次，多阶门控聚合卷积模块（MOGA）通过多阶博弈机制和门控聚合的策略增强特征信息提取的能力。最后，改进三元组损失函数（ITL），通过引入正样本对的绝对距离约束，拉近正样本对之间的距离。实验结果表明，该模型在 PRCC、Celeb-reID 和 DeepChange 这 3 个公开换衣行人重识别数据集上的 Rank-1 和 mAP 分别达到了 67.3% 和 64.7%、66.2% 和 19.0% 以及 59.4% 和 20.7% 的性能，验证了其在换衣场景下的有效性。

关键词：换衣行人重识别；图通道注意力；多阶门控卷积；改进三元组损失

中图分类号：TP39 **文献标志码：**A DOI: 10.12178/1001-0548.2025092

Cloth-changing person re-identification model integrating graph channel attention and multi-order gated residual network

MO Taiping, ZHENG Deguang, SUN Peng*, QIN Hanyue, ZHANG Xiangwen, and ZHANG Bin

(School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: Person re-identification (Re-ID), an important task in computer vision, aims to achieve precise matching of pedestrian identities across cameras. However, significant variations in clothing features under clothing-changing scenarios pose challenges to traditional Re-ID methods in feature extraction and information capture. To address this issue, a clothing-changing person Re-ID model named AMGA-ResNet50, integrating graph channel attention and multi-order gated residual network is proposed. First, in the adaptive graph channel attention (AGA) module, the graph convolution theory is introduced into the channel attention mechanism for clothing-changing scenarios, the channel dependency relationships are modeled as connections between feature nodes, and then the key identity features of clothing-changing pedestrians can be effectively captured through graph-structured processing. Second, the multi-order gated aggregation convolution (MOGA) module is employed to enhance the ability of feature information extraction through a multi-order gaming mechanism and gated aggregation strategy. Finally, an improved triplet loss (ITL) is proposed, which introduces an absolute distance constraint for positive sample pairs to reduce the distance between them. Experimental results show that the proposed model achieves Rank-1 and mAP scores of 67.3%/64.7%, 66.2%/19.0%, and 59.4%/20.7% on three public clothing-changing person Re-ID datasets, PRCC, Celeb-reID, and DeepChange, respectively, verifying the effectiveness of the model in clothing-changing scenarios.

Key words: cloth-changing person re-identification; figure channel attention; multi-order gating convolution; improved triplet loss

收稿日期：2025−04−30
基金项目：国家自然科学基金（62263006）；广西高校中青年教师科研基础能力提升项目（2025KY0258）
作者简介：莫太平，博士，教授，主要从事机器学习和计算机视觉等方面的研究。
*通信作者 E-mail: wxsunpeng@163.com

行人重识别是计算机视觉领域的重要任务，其目标是在不同摄像机中识别同一行人[1]。现有研究多假设行人衣着恒定或要求穿着一致[2]。然而，在实际应用中，行人日常活动中的换装行为较为普遍；此外，在犯罪场景中，犯罪分子可能通过换衣躲避监管部门的追查[3]。传统 ReID 方法在面对这些场景时往往会面临换衣后特征难以提取和信息难以捕捉等问题，导致识别性能显著下降。因此，换衣行人重识别的研究日趋重要，且对公共安全、智能监控等领域具有关键应用价值。

对比传统的 ReID 方法，换衣 ReID 方法面临着行人衣着不同和姿态变化的挑战[2]。因此，如何降低换衣对视觉表征的影响，提升辨识能力，成为亟待解决的难题。现有换衣 ReID 方法多基于残差网络（residual network, ResNet）架构[4]，这主要得益于 ResNet 对注意力机制的良好兼容性。作为一种模拟人类视觉认知的选择性聚焦特性，注意力机制通过模拟认知过程突出重要信息并抑制无关信息，是提升视觉任务性能的有效手段。许多研究探索了这一方向，如文献 [3] 提出 IGCL 方法，通过衣物注意力退化流减少衣物干扰，借助人类语义注意力、身体拼图流突出语义与姿势信息实现端到端协同学习；文献 [5]采用多头自注意力网络构建跨像素高阶关联的全局依赖模型，通过多尺度特征协同优化，实现细粒度细节保持与整体语义感知的平衡；文献 [6] 通过卷积块注意模块（CBAM）增强通道和空间特征，实现对判别性特征区域的自适应聚焦；文献 [7] 提出语义引导自注意力网络，利用语义信息指导双分支网络，帮助网络学习完整的行人特征并避免衣物干扰；文献 [8] 提出基于图卷积通道注意力的解耦模型，通过构建衣物与身份特征的二分图，动态调整通道级注意力权重以抑制衣物变化的影响；文献 [9]提出时空门控注意力网络，针对长期换衣场景引入时间维度注意力，增强跨时段特征的时序一致性建模；文献 [10] 提出基于多重注意力与空间变换的方法，通过融入三重注意力和自适应特征提取模块强化细粒度特征提取。然而，这些方法仍存在一定局限性。

1）由于行人外观的显著变化，现有算法对衣着剧烈变化的适应性较差，传统外观特征（如纹理、颜色）的失效导致特征匹配精度显著下降。同时换衣使得传统行人特征模糊，尤其在多个视角和不同环境下，如何有效捕捉并匹配具有显著外观差异的换衣特征，仍是亟待解决的核心挑战。

2）多数方法依赖卷积层的局部感受野进行细粒度特征提取，虽通过通道加权机制强化特定纹理特征，但受限于卷积操作的局部性本质，难以建立跨区域的全局空间关联。现有注意力模型虽尝试通过自注意力捕捉长程依赖，但对多层次特征的动态融合机制仍不完善，导致复杂场景下的特征表达完备性不足。

针对上述问题，本文提出一种融合图通道注意力（adaptive graph channel attention, AGA）与多阶门控（multi-order gated aggregation, MOGA）残差网络的换衣行人重识别模型（AMGA-ResNet50），该模型通过 AGA 模块增强空间注意力，并结合 MOGA 模块提供的多层次目标关注，使得模型在复杂场景中既能保留对局部细节的敏感性，又能提升全局特征的辨识能力。

1　本文方法

1.1　总体框架

换衣行人重识别的核心挑战是在换衣场景下提取稳定特征，同时抑制衣着纹理等易变信息。现有方法多依赖局部或单一层级特征，难以建模通道全局依赖，且特征融合存在冗余，导致鲁棒性不足。为此，本文提出 AMGA-ResNet50 整体框架如图 1 所示，通过模块化设计提升特征提取的精准性与鲁棒性。

以预训练的 ResNet50 作为骨干网络，针对换衣场景对精细空间特征（如肢体轮廓）的需求，调整 layer4 下采样步长（last stride=1）以避免特征图分辨率损失，并移除末端全局池化层与全连接层，保留特征空间分布信息，为捕捉体型、步态等全局特征提供基础。在特征提取阶段，结合换衣场景“局部与全局特征需协同判别”的需求，在 layer3 与 layer4 后分别引入平均池化和最大池化，前者强化局部纹理细节的聚合能力，后者增强全局语义特征的显著性，通过层级互补融合提升特征表达能力，生成特征向量 $\boldsymbol{F}_1$。在特征处理阶段，其模块针对核心难点设计：AGA 针对通道特征贡献差异问题，改进传统注意力机制仅依赖全连接层的局限，通过图卷积网络构建通道关联图，将通道作为节点学习拓扑关系实现动态权重分配（生成 $\boldsymbol{F}_2$），从而强化稳定特征通道并抑制易变通道；MOGA 则针对层级特征交互不足问题，通过门控机制对不同阶数特征选择性交互，保留层级独立性并过滤冗余信息（生成 $\boldsymbol{F}_3$），实现局部与全局的

特征协同优化。同时为增强特征多样性并提升计算效率，采用并行融合策略，通过逐元素相乘实现 $\boldsymbol{F}_2$ 与 $\boldsymbol{F}_3$ 的非线性融合，生成 $\boldsymbol{F}_4$：

$$\boldsymbol{F}_4 = \boldsymbol{F}_2 \odot \boldsymbol{F}_3 \tag{1}$$

式中，⊙表示逐元素相乘操作，$\boldsymbol{F}_4$ 为 $\boldsymbol{F}_2$ 与 $\boldsymbol{F}_3$ 融合后的特征向量输出。

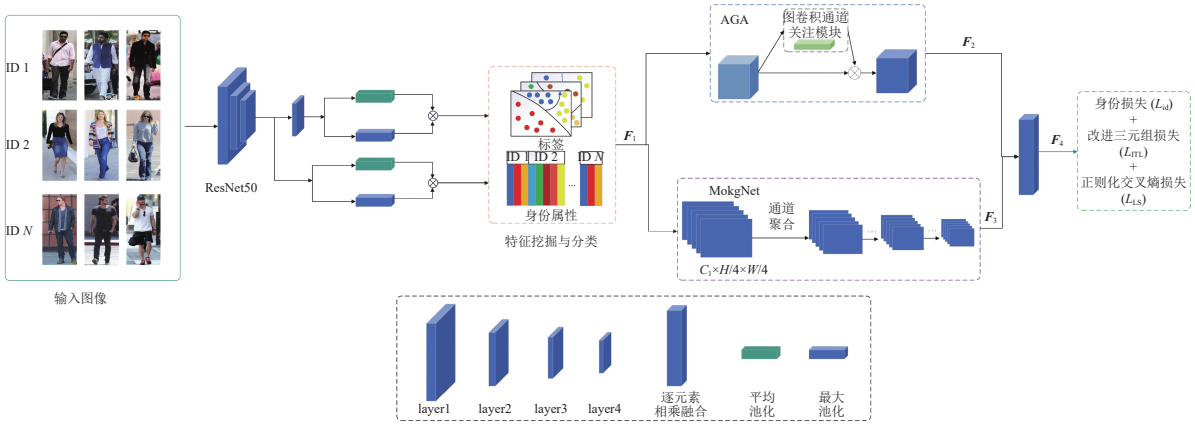


图 1　AMGA-ResNet50 框架图

逐元素相乘操作既强化了特征间的非线性交互，又通过抑制低响应区域有效降低了噪声干扰。在训练阶段，针对换衣场景中类内差异大和类间相似性高的样本分布特点，采用身份损失[11]（L_{id}）、改进三元组损失（L_{ITL}）和正则化交叉熵损失[12]（L_{LS}）进行联合优化，通过多损失协同约束提升特征的判别能力与类内紧凑性。

1.2　AGA

在换衣行人重识别任务中，由于服装变化、跨视角挑战、数据稀缺和计算复杂度等问题，特征获取相比传统行人重识别更困难，这使得特征处理和增强更具挑战性。为此，本文设计了 AGA 模块，该模块可集成于骨干网络的特征输出阶段，通过图卷积建模通道间长距离依赖关系，对卷积层生成的特征图进行通道级权重优化，实现对衣物变化敏感通道的抑制与身份判别通道的强化，其架构如图 2 所示。

由图 2 可知，AGA 模块由两部分组成。第一部分是由两个线性嵌入函数 F_r 与 F_r' 组成的特征映射层。第二部分是自适应图卷积模块（AGCM）。首先，对输入特征图 $\boldsymbol{X} \in R^{C\times H\times W}$ 应用全局平均池化（GAP）操作，其计算公式为：

$$\overline{\boldsymbol{X}} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \boldsymbol{X}(:, i, j) \tag{2}$$

式中，$\overline{\boldsymbol{X}} \in R^{C\times 1\times 1}$；$C$ 表示通道数量，也是 AGA 中特征点的数量。

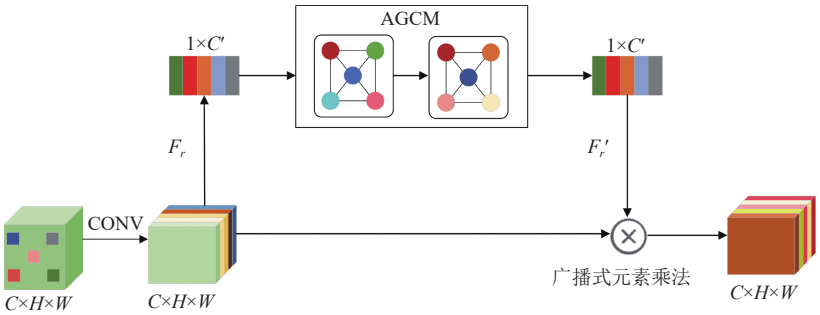


图 2　自适应图通道注意力模块

随后，特征映射层通过一个线性映射函数来提升表示能力，设特征映射层的输入为 $\overline{\boldsymbol{X}}$，输出为：

$$\boldsymbol{Y} = \boldsymbol{W}\overline{\boldsymbol{X}} \tag{3}$$

式中，$\boldsymbol{W} \in R^{C\times C}$ 是一个可学习的权重矩阵。这个过程可以通过 1×1 的卷积核实现。为了简化模型，两个卷积层被替换为瓶颈结构，这是因为瓶颈结构[13]不仅支持全通道交互，还能有效减少特征冗余[14]。由于图卷积网络要求特征顶点互斥，导致 AGA 模块中的特征顶点对特征冗余极为敏感，相似或相同的特征点难以通过邻接矩阵准确表示。因此，瓶颈结构对 AGA 至关重要，可将参数从 $C\times C$ 压缩至 $C\times(C/r)$，同时将邻接矩阵的尺寸缩减为 $(C/r)\times(C/r)$。

在此基础上，AGCM 通过 $C\times C$ 维度的邻接矩阵计算每个特征点的权重，并将得到的权重作为通道权重映射回原始特征图。此映射过程与前述线性嵌入映射相似。每个特征点的权重计算为：

$$\boldsymbol{Z}=\boldsymbol{A}\boldsymbol{Y} \tag{4}$$

式中，$\boldsymbol{A}\in R^{C\times C}$ 为邻接矩阵；$\boldsymbol{Z}\in R^{C\times 1\times 1}$ 表示加权后的特征向量。

最后，AGA 的整体计算过程如下：

$$\boldsymbol{X}=F_r(\mathrm{GAP}(x)) \tag{5}$$

$$y=x\cdot\mathrm{sigmoid}(F'_r(\mathrm{ReLU}(G(\boldsymbol{X},\boldsymbol{A})))) \tag{6}$$

式中，$\cdot$ 表示广播元素乘法；sigmoid 表示激活函数；$F_r(\cdot)$ 和 $F'_r(\cdot)$ 表示嵌入函数，形成瓶颈结构；G 是本文提出的 AGCM；其中 $\boldsymbol{X}$ 表示输入；GAP 表示全局平均池化操作；$\boldsymbol{A}$ 为表示特征关系的邻接矩阵。

1.3　MOGA

本文设计了一个 4 阶段的 MOGA 模块，如图 3 所示。每个阶段 i 中，输入图像或特征首先通过嵌入系统调整分辨率并映射到 C_i 维空间。假设输入分辨率为 $H\times W$，4 个阶段的特征分辨率分别为 $H/4\times W/4$、$H/8\times W/8$、$H/16\times W/16$ 和 $H/32\times W/32$。随后，嵌入的特征进入 Moga 模块，该块由空间和通道聚合模块组成，用于进一步聚合上下文信息。最终输出阶段则通过全局平均池化（GAP）和线性层完成分类任务。此外，4 个级的输出均选择采用瓶颈模块来降低网络复杂度。

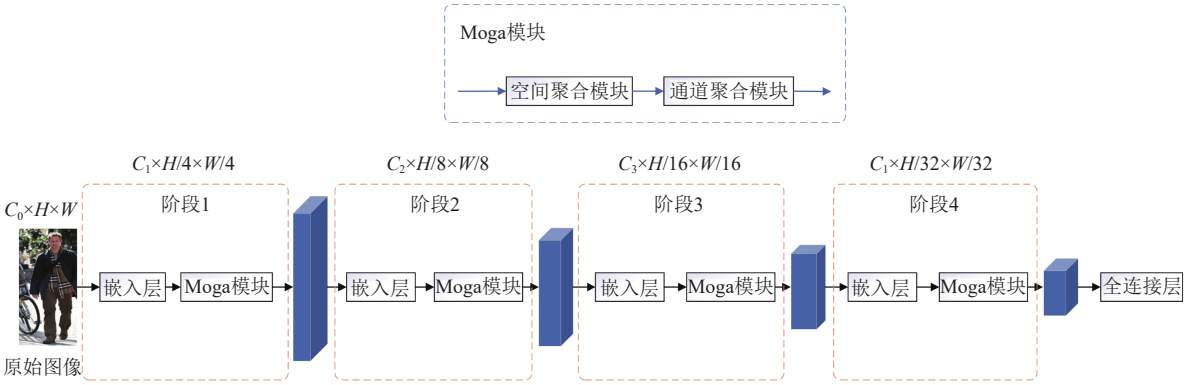


图 3　多阶门控聚合卷积网络模块

针对多阶通道重新分配问题，现有方法通常依赖 2 个线性投影进行通道混合，但存在信息冗余和跨通道交互不足等问题，需大量参数调节，导致计算效率较低。为此，采用了一种参数高效型通道聚合模块（channel aggregation, CA）[15]，如图 4 所示，通过简化映射过程并引入自适应尺度调节机制，实现高效的通道特征重新分配，其计算如下：

$$\boldsymbol{Y}=\mathrm{GELU}(\mathrm{DW}_{3\times 3}(\mathrm{Conv}_{1\times 1}(\mathrm{Norm}(\boldsymbol{X})))) \tag{7}$$

式中，GELU[16] 是一种激活函数；Norm($\boldsymbol{X}$) 表示规范化层处理。CA($\boldsymbol{X}$) 与 GELU($\boldsymbol{X}$) 的具体实现分别为：

$$\mathrm{CA}(\boldsymbol{X})=\boldsymbol{X}+y_c\cdot((\boldsymbol{X}-\mathrm{GELU}(\boldsymbol{X}\boldsymbol{W}_r)) \tag{8}$$

$$\mathrm{GELU}(\boldsymbol{X})=0.5\boldsymbol{X}\left(1+\tanh\left[\sqrt{2/\pi}(\boldsymbol{X}+0.044\,715\boldsymbol{X}^3\right]\right) \tag{9}$$

式中，y_c 是初始化为零的通道尺度因子，旨在通过动态学习实现通道重要性的自适应调节，使模型逐步聚焦关键通道。双曲正切函数 tanh 将特征映射至 [−1,1] 区间，配合常数共同作用使函数呈现类似高斯分布的特性，增强对显著通道特征的选择性，解决现有方法跨通道交互粗糙的问题；同时 tanh 与 y_c 的互补作用可实现更精细的通道特征分配。式 (9) 中，GELU 通过概率分布对特征进行随机正则化，与 CA($\boldsymbol{X}$) 的确定性通道调节形成协同，既保留特征多样性，又通过 y_c 的尺度控制避免跨通道干扰，最终实现高效且鲁棒的通道聚合。

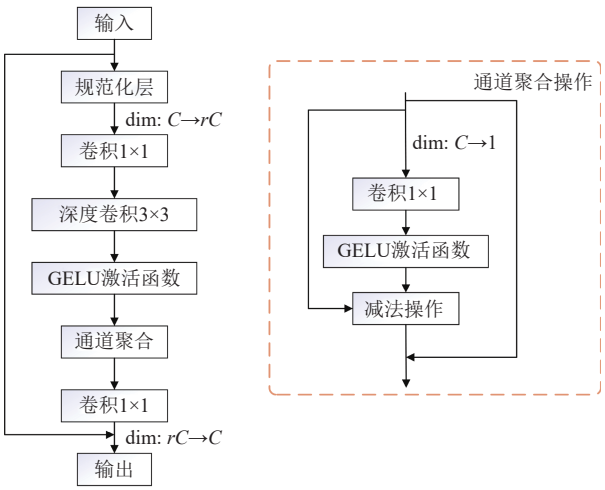


图 4　通道聚合模块

1.4 损失函数

本文模型共有 3 种损失函数应用于训练阶段，分别为身份分类损失[11]、标签正则化交叉熵损失[12]和改进三元组损失函数。

身份分类损失是通过训练模型使其能够正确地识别或分类输入数据对应的身份或类别。一般采用交叉熵损失（cross-entropy loss）来计算，交叉熵损失计算公式为：

$$L_{\mathrm{id}} = -\sum_{i=1}^{N} q_i \log(p_i) \tag{10}$$

式中，p_i 表示 i 类的身份预测逻辑；y 表示身份标签；q_i 为真实标签的编码，对于类别 i，当样本属于该类别时，$q_i = 1$，否则，$q_i = 0$；N 表示样本数量。

三元组损失[17]则是一种用于度量学习的损失函数，通过同时优化锚点样本、正样本和负样本之间的嵌入距离，使同类样本的嵌入空间更接近，不同类样本的嵌入空间更远离。其核心目标是确保锚点与正样本的距离小于锚点与负样本的距离，并引入一个预设的间隔。其计算公式为：

$$L_{\mathrm{tri}} = (d_{a,p} - d_{a,n} + \alpha)_+ \tag{11}$$

式中，$d_{a,p}$和$d_{a,n}$分别为正样本对和负样本对的特征距离；α 为间隔阈值；$(z)_+ = \max(z, 0)$。

然而，传统三元组损失过度依赖随机样本，容易导致简单样本对的过度聚集，而忽视困难样本的优化，从而降低训练效率和模型判别能力。为此，引入困难样本三元组损失[18]，使模型专注于挖掘和优化训练过程中最难区分的样本对，提升模型的收敛速度和特征判别能力。

同时在换衣行人重识别场景中，服装变化作为主要干扰信息，仅依赖困难样本三元组损失仍难以确保模型在处理正样本对（即不同服装的同一行人）时的一致性，从而导致鲁棒性不足。为此，本文进一步改进了三元组损失（ITL），通过结合正样本对的绝对距离与困难样本三元组损失，进一步提升正样本对的相似度，使同一行人的不同图像在服装变化下仍能保持较小距离，从而有效解决了换衣行人重识别中的服装变化问题，提升模型对相同身份行人的判别能力和一致性，增强重识别的准确性和鲁棒性。其计算公式分别为：

$$L_{\mathrm{ITL}} = \omega_m d_{|a,p|} + \frac{1}{PK} \sum a \in \mathrm{batch} L_{\mathrm{th}} \tag{12}$$

$$L_{\mathrm{th}} = (\max_{P \in A} d_{a,p} - \min_{n \in B} d_{a,n} + \alpha)_+ \tag{13}$$

$$d_{|a,p|} = \sqrt{\sum_{i=1}^{N} (a_i - p_i)^2} \tag{14}$$

式中，ω_m为正样本对绝对距离的权重，设置为 0.8；L_{th} 为困难样本三元组损失函数[18]；$d_{|a,p|}$为正样本对之间的绝对距离；P 和 K 分别表示每个训练 batch 随机挑选的行人数量和每个行人的图片数量；超参数 α 通常设置为 0.4。

此外，针对训练过程中模型出现的过拟合现象，引入标签平滑正则化交叉熵损失，通过对传统的硬标签进行平滑处理，从而减少标签依赖，避免了模型过度自信，有助于减少过拟合和缓解类不平衡问题。计算公式为：

$$L_{\mathrm{LS}} = -\sum_{j=1}^{c} \left((1-\sigma) y_j + \frac{\sigma}{c} \right) \log(\hat{y}_j) \tag{15}$$

式中，σ 为标签平滑的超参数，本文设置为 0.1；c 是类别数；y_j是实际标签值，对于正确类别，$y_j = 1$，其他类别$y_j = 0$；$\hat{y}_j$为模型输出的概率。

本文采用上述 3 种损失函数进行训练，在早期训练阶段，模型倾向于通过区分简单样本（相同身份和服装）来学习易识别的身份信息。然而，本文提出的模块和函数旨在帮助模型学习复杂身份特征，过早引入可能导致模型陷入局部最优。因此，采用分阶段训练策略：首先仅使用身份分类损失（L_{id}）进行监督，待模型初步学习掌握复杂身份特征后，再逐步引入改进三元组损失（L_{ITL}）和标签平滑正则化交叉熵损失（L_{LS}），帮助模型学习更鲁棒的身份特征。整体损失函数为：

$$L = \lambda_1 L_{\mathrm{id}} + \lambda_2 L_{\mathrm{LS}} + \lambda_3 L_{\mathrm{ITL}} \tag{16}$$

式中，λ 是每一项损失函数的权重系数，本文将其设置为$\lambda_1 = \lambda_2 = \lambda_3 = 1$。

2 实验与结果分析

2.1 评估数据集

为了验证所提模型的有效性，本文在广泛使用的换衣 ReID 数据集，即 Celeb-reID[19]、PRCC[2]，以及一个大规模的换衣数据集 Deepchange[20] 上进行了训练与评估。3 个数据集的介绍如表 1 所示。

Celeb-reID 是一个从明星街拍中收集的数据集，分为训练集、图库集和查询集。该数据同一行人的衣着有显著变化，超过 70% 的图片展示了不

同的服装风格，但也存在少量重复穿着的情况。此外，Celeb-reID 数据集还包含部分后视图，进一步增加了行人识别的挑战性和多样性。

PRCC 数据集由 3 台室内摄像机捕获而得，其独特性在于多视角与衣物变化的交叉设计，对于图像中的行人，相机 A 和 B 的图像都没有衣服改变，但相机 C 的图像衣着发生了明显变化。

DeepChange 数据集则专注于长期场景下的衣物变化问题，时间跨度超过 12 个月，所有样本均包含显著衣物变化，且附带行人属性（性别、年龄、上衣类型等）和时间戳标注。该数据集模拟了真实世界中衣物、外观和环境的长期动态变化，为研究衣物不变特征提供了理想场景。

表 1　3 个数据集介绍

数据集	身份数	图片数	相机数
Celeb-reID[19]	1 052	34 186	—
PRCC[2]	221	33 698	3
DeepChange[20]	1 121	178 407	17

2.2　实验设置

2.2.1　实现细节

本文模型基于 PyTorch 1.12.1 框架实现，并采用 Python 3.8 作为编程语言，GPU 选择 NVIDIA GeForce RTX 3090。

实验选择 ImageNet 上预训练的 ResNet50[4] 作为主干网络。在数据增强方面，使用了随机水平翻转、填充、随机裁剪和随机擦除[21]，概率均设置为 0.5。训练过程中采用 Adam 优化器[22]，其权重衰减设定为 5×10^{-4}，训练周期为 80 个 epoch。采用学习率预热策略，即在前 10 个 epoch 内将学习率从 3.5×10^{-6} 线性提升至 3.5×10^{-4}，之后每 20 个 epoch 降低 10 倍。超参数 λ、τ 和 P 的值分别设置为 0.1、1/16 和 2，批量大小为 64。采用 DBSCAN[23] 进行聚类，最小样本量设为 1，扫描半径设定为 0.4。

2.2.2　评价指标

为了评估模型的有效性，本文采用换衣 ReID 研究中常用的标准指标，即累积匹配特征（cumulative matching characteristic, CMC）曲线和平均精度（mean average percision, mAP）。累积匹配特征用于衡量模型图像检索能力，能够反映查询样本检索结果的匹配排位表现，可直观体现不同排位下正确样本的匹配成功率，本文重点参考首位匹配成功率 Rank-1 指标。平均精度则是一种综合评价指标，通过计算所有查询的平均精度来评估模型的整体检索性能。

2.3　实验结果分析

为了验证各模块在换衣 ReID 任务中的有效性。本文在 3 个数据集上进行了全面的消融实验。实验结果如表 2 所示，其中 Baseline 为基线模型 ResNet50。

表 2　在 3 个数据集上的消融实验结果　　%

方法	PRCC		Celeb-reID		DeepChange	
	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
ResNet50	41.0	37.5	52.9	9.9	41.7	11.0
AGA	66.0	63.1	64.7	18.3	58.5	20.4
MOGA	61.6	59.5	64.4	18.1	57.7	19.5
ITL	66.4	63.3	64.9	18.5	58.7	20.4
AGA+MOGA	66.9	64.1	65.2	18.5	58.6	20.6
AGA+ITL	66.5	64.4	65.6	18.7	59.1	20.6
MOGA+ITL	66.3	64.1	65.5	18.8	58.8	20.5
AGA+MOGA+ITL	**67.3**	**64.7**	**66.2**	**19.0**	**59.4**	**20.7**

由表 2 可知，AGA 模块通过图卷积建模通道依赖关系，在所有数据集上实现稳定提升。以 PRCC 数据集为例，其 Rank-1/mAP 较基线相比分别提高了 25.0% 和 25.6%，验证了其剥离衣物变化噪声、强化身份特征通道的能力，有效解决了换衣场景中通道级的干扰问题。

MOGA 模块则通过多阶门控机制动态聚合跨层特征，在 PRCC、Celeb-reID 和 Deepchange 上 Rank-1 分别提升了 20.6%、11.5% 和 16.0%，该模块通过门控权重自适应选择浅层轮廓与深层语义特征的融合路径，有效缓解了多尺度特征融合中的尺度不匹配问题，尤其在 PRCC 相机 C 的显著衣着变化场景中表现突出。但其单独作用时性能存在瓶颈，这是因为 MOGA 的跨阶特征融合效果高度依

赖输入特征质量：在缺乏通道筛选的情况下，衣着纹理等易变信息具有更强的特征响应，导致门控机制难以区分稳定特征（体型、步态）与易变特征（衣着），门控权重易被冗余信息主导，反而强化了类内差异。

ITL 损失函数通过引入正样本对绝对距离约束，在 PRCC、Celeb-reID 和 DeepChange 上分别提升 Rank-1 至 66.4%、64.9% 和 58.7%，其核心价值在于强制拉近同一身份不同衣着样本的特征距离，显著增强了类内紧凑性，从而减少相似衣物导致的身份混淆。尤其在 Celeb-reID 的少量重复穿着样本中，mAP 从 9.9% 提升至 18.5%，充分证明其在换衣 ReID 任务中的有效性。

值得注意的是，AGA、MOGA 和 ITL 单一模块性能超过多数对比方法，核心是设计针对性与场景适配性的差异。对比方法中存在通用模型（如 PCB、MGN）特征提取依赖衣着等易变信息，在换衣场景中因同身份衣着变化导致特征判别性骤降。而本文模块直击换衣痛点：AGA 以图卷积抑制衣着通道，缓解通用注意力对易变特征的过度响应；MOGA 聚焦稳定特征跨层融合，规避衣着信息干扰交互的问题；ITL 通过绝对距离约束强化类内紧凑性，弥补通用三元组损失对类内差异的约束不足。这种适配设计使其表现更优。

当模块组合时，协同效应显著超越单一组件之和。AGA 与 MOGA 模块的互补性体现在：AGA 通过通道筛选机制有效抑制衣着纹理等易变通道的干扰，为 MOGA 提供了纯净的输入特征，解决了单独使用 MOGA 时因噪声输入导致的门控机制失效问题。而 MOGA 通过门控机制在空间维度强化换衣特征响应，形成“通道筛选–空间聚合”的层级优化协同，即 AGA 负责去除通道级噪声，使 MOGA 能更精准地聚焦体型、步态等稳定特征的跨层交互，在 PRCC 上的 Rank-1、mAP 分别提升提升 25.9% 和 26.6%。AGA 与 ITL 的组合在 DeepChange 上 mAP 提升 9.6%，这得益于通道级身份特征经 ITL 优化后，在长期跨度场景中形成更紧密的簇类分布。当 AGA+ITL 与 MOGA 结合时，协同效应进一步升级，PRCC 的 Rank-1/mAP 提升至 67.3% 和 64.7%，ITL 通过约束正样本对距离优化特征空间几何结构，使 AGA 筛选的稳定特征形成更紧凑的簇类，此时 MOGA 的多阶融合不再仅依赖特征本身的交互，更能基于 ITL 优化后的特征空间几何结构，针对性强化“同簇内”跨层特征关联，避免不同身份的相似特征混淆。这种“通道筛选–特征空间优化–空间聚合”的三级协同，既保留了通道级去噪的精准性，又通过特征空间约束提升了跨层融合的靶向性，最终形成完整技术闭环，切实提升换衣 ReID 方法的识别精度。

2.4 损失函数对比实验

本文对不同的损失函数策略进行了测试，包括 L_{id}、L_{LS}、L_{ITL}。PRCC 数据集上的实验结果如表 3 所示。

表 3 不同损失函数策略在 PRCC 数据集的性能比较

%

损失函数策略	Rank-1	Rank-5	mAP
ResNet50	41.0	45.3	37.5
L_{id}	49.2	55.0	46.2
L_{LS}	62.1	69.7	61.5
L_{ITL}	63.8	70.5	61.8
L_{id}+L_{LS}	65.1	71.2	62.3
L_{id}+L_{ITL}	65.8	71.9	63.6
L_{LS}+L_{ITL}	66.1	73.0	64.0
L_{id}+L_{LS}+L_{ITL}	67.3	74.4	64.7

在表 3 中，模型使用单一损失时，L_{id} 的性能提升最小，说明单纯依赖身份分类难以克服换衣干扰；L_{ITL} 表现最优，Rank-1 和 mAP 分别达到 63.8% 和 61.8%，原因在于改进三元组损失对换衣判别性特征学习起到核心作用，通过优化样本间的距离度量，能够有效应对衣着变化引发的类内差异。

采用联合损失函数时，L_{id}+L_{LS} 较单一损失提升显著，相较于性能最优的 L_{LS}，Rank-1 和 mAP 分别提升了 3.0% 和 0.8%，这是因为身份分类与标签正则化的互补性，即 L_{id} 提供 ID 判别先验，L_{LS} 通过正则化抑制过拟合，提升模型泛化能力。L_{id}+L_{ITL} 的 mAP 高于 L_{id}+L_{LS}1.3%，验证了三元组损失对排序质量的强化作用，但其缺少正则化，可能限制模型鲁棒性，Rank-1 低于 L_{LS}+L_{ITL} 组合。L_{LS}+L_{ITL} 获得最优双损失性能，Rank-1 和 mAP 分别达到 66.1% 和 64.0%，体现出独特的协同优势，前者保障特征空间泛化性，后者优化细粒度判别，提高了模型的鲁棒性。L_{id}+L_{LS}+L_{ITL} 的三损失融合性能达到最优，较最优双损失组合的 Rank-1 和 mAP 分别提升 1.2% 和 0.7%，这是因为三者存在层次化互补：L_{id} 提供高层语义监督，L_{LS} 约束特征分布平滑性，L_{ITL} 驱动局部判别优化，共同构建对换衣干扰的鲁棒表示。实现换衣特征表示的最优

平衡。

2.5　超参数消融实验

为验证关键超参数对模型性能的影响，本节在 PRCC 数据集上针对 DBSCAN 聚类扫描半径 eps、温度系数 τ 及损失函数 L_{ITL} 中系数 λ_3 进行消融实验，分析不同取值对模型性能的影响，结果如图 5 所示。

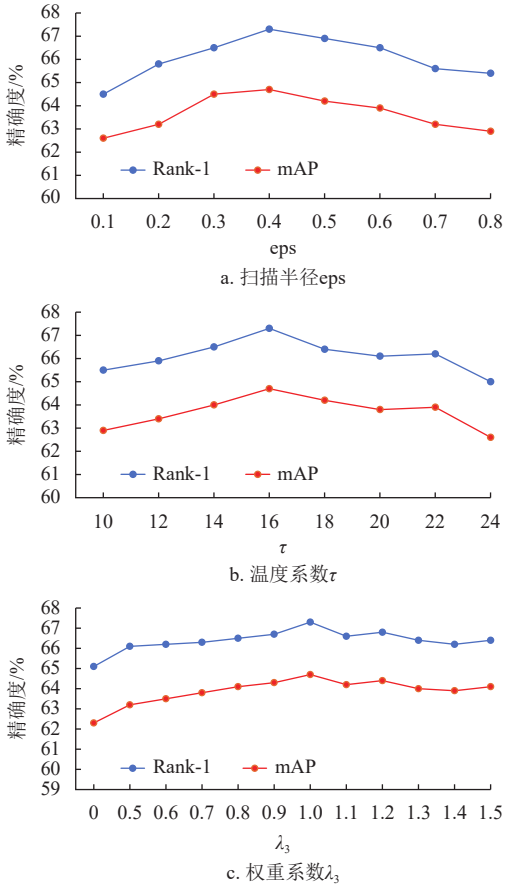


图 5　超参数消融研究

2.5.1　DBSCAN 聚类扫描半径 eps 的影响

eps 作为 DBSCAN 算法的核心参数，用于控制聚类的空间邻域范围，直接影响换衣场景中同身份样本簇的聚合效果。由图 5a 可知，当 eps 过小时（如 0.1～0.3），模型对样本空间的划分过于严格，导致同一身份换衣后的样本因特征差异被拆分，类内紧凑性不足；eps=0.4 时性能最优，此时聚类既能将换衣前后的同一身份样本聚合为紧凑簇，又能有效区分不同身份的相似体型样本；当 eps 过大（如 0.5～0.8）时，过宽的邻域范围导致不同身份的相似特征样本被错误合并（如相似体型的不同行人），类间混淆加剧，性能显著下降。故实验中 eps 设为 0.4。

2.5.2　温度系数 τ 的影响

温度系数 τ 调节特征分布软化度，影响模型对细微特征的敏感度。图 5b 显示，τ 为 10～14 时，模型过度关注局部差异，忽略换衣场景的稳定特征（体型、步态）；τ 为 18～24 时，特征分布过度平滑，身份区分度下降；τ 为 16 时性能最优，此时 τ 通过适度软化特征分布，保留身份判别性特征的响应强度，同时避免对衣着过度敏感，增强了模型的换衣干扰能力。因此实验中 τ 设为 16。

2.5.3　损失函数 L_{ITL} 系数 λ_3 的影响

λ_3 为 L_{ITL} 的权重系数，控制其在多损失优化中的贡献度，影响特征空间的类内紧凑性与类间分离度。由图 5c 可知，当 λ_3 较小（0～0.9）时，L_{ITL} 的约束作用不足，难以有效缩小同身份换衣样本的特征距离，类内离散度较高；当 λ_3=1.0 时性能最优，此时 L_{ITL} 与 L_{id}、L_{LS} 形成协同，既通过 L_{ITL} 强化同身份换衣样本的类内紧凑性，又通过其他损失保证不同身份样本的类间分离度，避免单一损失片面性；当 λ_3 过大（1.1～1.5）时，L_{ITL} 的过度约束导致特征空间压缩，反而模糊了不同身份的边界，性能下降。故实验中 λ_3 设为 1.0。

2.6　与现有方法比较

为验证本文解决方案的有效性，将本文模型与其他具有相似实验设置的现有先进方法进行比较。评价指标为 Rank-1 和 mAP。在 3 个换衣 ReID 数据集 PRCC、Celeb-reID 和 DeepChange 上的性能比较结果如表 4 所示。

本文模型在 PRCC 数据集上与多种方法进行了对比。传统行人重识别方法 HACNN[24] 和 PCB[25] 主要依靠衣物颜色、纹理等外观特征建模，在换衣场景中，由于衣物特征失效，导致它们在 PRCC 数据集上的表现较差；TransReID[26] 通过引入 Transformer 架构，增强了人物重识别任务中的特征表示能力，在一定程度上缓解了衣着变化带来的局部特征漂移问题，但在衣物差异显著的相机 C 视角下，仍存在跨模态对齐能力不足的瓶颈；SPT（spatial polar transformation）[2] 是提出 PRCC 数据集的方法，针对多视角衣着变化设计了视角对齐损失，Rank-1 达到 55.2%，展现出领域适配优势，然而其依赖相机内外参数的先验信息，泛化至无校准多视角场景时性能下降明显；最后几种方法则为近几年性能较优的代表性换衣行人重识别方法，CAL[27] 通过生成衣物对抗网络来训练模型，

RCSANet*[28]、CRE+BSGA[29] 和 FIRe²[1] 则采用其他方式（如轮廓草图、人体姿势），减少衣物的影响，但它们都有着过度依赖额外标注数据和手工设计的特征融合权重的问题。相比之下，本文模型通过动态特征选择模块，实现了无额外数据依赖的特征增强与自适应特征融合，提升了换衣场景下外观变化与视角差异的联合干扰的模型鲁棒性。由表 4 可知，本文模型在 PRCC 数据集中 Rank-1 和 mAP 分别达到 67.3% 和 64.7%，相较于 FIRe²，精度分别提升了 2.3% 和 1.6%，表现优异。

表 4　不同方法的性能比较

%

方法	PRCC		Celeb-reID		DeepChange	
	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
HACNN[24]	21.8	—	47.6	9.5	—	—
PCB[25]	41.8	38.7	37.1	8.2	—	—
TransReID[26]	46.6	44.8	—	—	—	—
MGN[30]	—	—	49.0	10.8	—	—
JLCN [31]	—	—	51.6	10.8	—	—
SPT [2]	34.4	—	—	—	—	—
RCSANet*[28]	50.2	48.6	55.6	11.9	—	—
OSNet[32]	—	—	—	—	39.7	10.3
ReIDCaps[14]	—	—	51.2	9.8	39.5	11.3
CAL[26]	55.2	55.8	—	—	54.0	19.0
CRE+BSGA[29]	61.8	58.7	—	—	—	—
CAMC[33]	—	—	57.5	12.3	—	—
FIRe²[1]	65.0	63.1	64.0	18.2	57.9	20.0
本文	**67.3**	**64.7**	**66.2**	**19.0**	**59.4**	**20.7**

进一步地，本文将所提模型与传统的行人重识别方法及针对 Celeb-reID 设计的先进换衣方法进行了对比。结果如表 4 所示，传统方法 HACNN[24]、MGN[30] 和 PCB[25] 因依赖外观特征而表现较差。相比之下，现有换衣重识别方法 JLCN[31]、CAMC[33] 和 FIRe2[1] 则通过轮廓草图、人体姿态或服装标签减少服装干扰，但其性能仍不及本文模型。本文模型仅使用 RGB 模态输入，通过挖掘更具区分性的通道信息，在 Celeb-reID 上获得了 Rank-1 和 mAP 分别为 66.2 和 19.0% 的优秀结果，相较于 FIRe² 模型，Rank-1 提升了 2.2%，说明了本文模型的优越性。

此外，在 DeepChange 数据集上对本文模型进行了评估，该数据集规模较大，尤其是测试样本数量庞大，因而是最能反映现实长期换衣场景下复杂性和难度的数据集之一。尽管大多数先进方法未在此数据集上报告结果，仍选择其中几种强大的基线方法 OSNet[32]、ReIDCaps[14]、CAL[26] 和 FIRe2[1] 的官方代码与本文模型进行比较，由表 4 可知，本文模型的 Rank-1 和 mAP 分别为 59.4% 和 20.7%，其性能仍然明显优于其他方法。进一步验证了本文模型的有效性。

2.7　可视化分析

为了直观展示本文模型的有效性，在 PRCC 数据集上进行了可视化分析，重点是将 baseline 模型与本文模型进行排序对比，随机选取 4 张查询图像（Query），分别查询 Rank-1～Rank-5 图像，其中绿色和红色边框分别表示正确和错误的匹配结果。

如图 6 所示，图 6a 和图 6b 中的 Rank-1～Rank-5 图像框均显示绿色，即 baseline 模型与本文模型均检索正确，表明本文模型能够有效提取丰富的特征信息。然而，图 6c 中 Baseline 模型在 Rank-1、Rank-4 和 Rank-5 的检索结果中出现了错误检索（红色边框），本文模型则仅有 Rank-3 出现检索错误，体现了本文模型对衣着变化的鲁棒性。图 6d 中，针对脸部遮挡的查询图像，Baseline 模型在 Rank-2 中出现错误检索（红色红框），反映出其对局部可见特征（如面部）的过度依赖；而本文模型的检索结果全部正确，有力验证了本文模型通过全局关系建模，引导网络聚焦衣着之外的判别性身体特征（如体型、步态），从而有效克服面部遮挡的干扰。总体而言，本文模型在换衣场景下的行人识别性能更优，能更准确地识别换衣后的行人身份，从而提高检索精度。

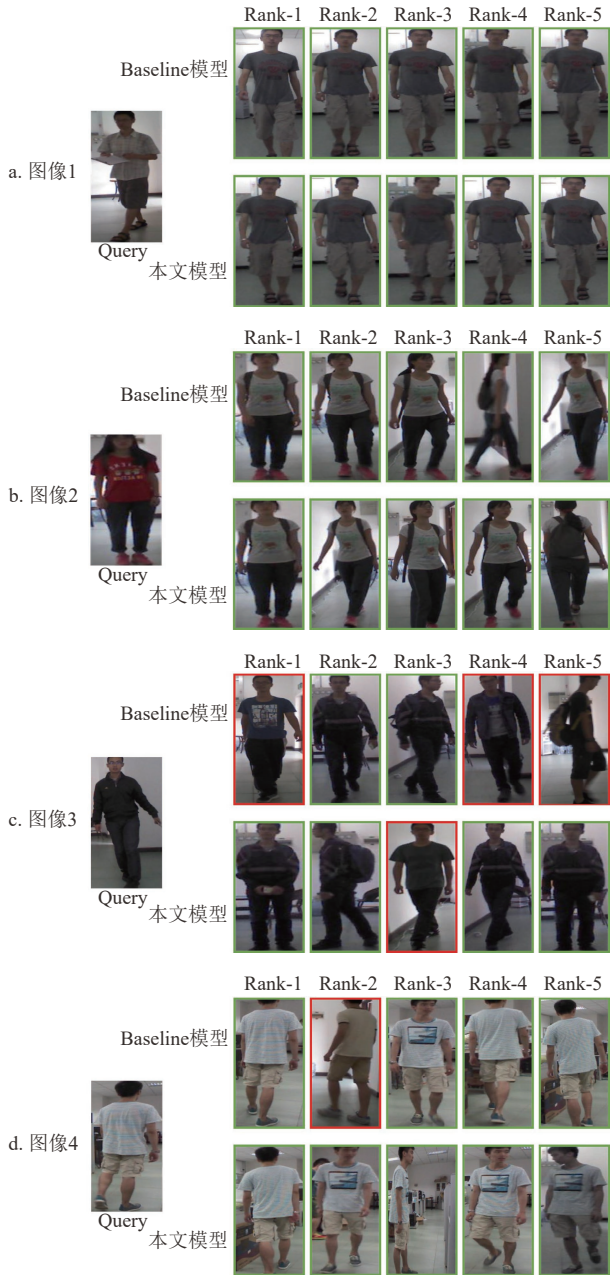


图 6　PRCC 数据集上 Baseline 模型与本文模型的前 5 张检索图像对比

3　结束语

本文提出了一种融合图通道注意力与多阶门控残差网络的换衣行人重识别模型（AMGA-ResNet50）。首先设计了 AGA 模块，将图卷积理论引入通道注意力中，直接操作特征以避免复杂的像素空间编码。然后，提出 MOGA 模块，通过高效的信道聚合模块自适应地重新分配高维隐藏空间中的信道特征，实现全局信息之间的紧密获取。最后，改进了 ITL 函数，以拉近正样本距离。本文模型在多个换衣 ReID 数据集上均达到了较为先进的性能，展现了其在换衣行人重识别任务中的优势。未来将专注于换衣 ReID 任务的多粒度学习研究，以进一步提升模型的性能。

参 考 文 献

[1] WANG Q Z, QIAN X L, LI B, et al. Exploring fine-grained representation and recomposition for cloth-changing person re-identification[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 6280-6292.

[2] GAO Z, WEI S, GUAN W, et al. Identity-guided collaborative learning for cloth-changing person reidentification[J]. IEEE Trans Pattern Anal Mach Intell, 2024, 46(5): 2819-2837.

[3] YANG Q, WU A, ZHENG W S. Person re-identification by contour sketch under moderate clothing change[J]. IEEE Trans Pattern Anal Mach Intell, 2021, 43(6): 2029-2046.

[4] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 770-778.

[5] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[EB/OL]. [2024-10-21]. https://arxiv.org/abs/2010.11929.

[6] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 3-19.

[7] 钟建华, 邱创一, 巢建树, 等. 基于语义引导自注意力网络的换衣行人重识别模型[J]. 计算机应用, 2023, 43(12): 3719-3726.
ZHONG J H, QIU C Y, CHAO J S, et al. Cloth-changing person re-identification model based on semantic-guided self-attention network[J]. Journal of Computer Applications, 2023, 43(12): 3719-3726.

[8] LI X, WANG Y, LI Z. Graph convolutional channel attention for clothing-change person re-identification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 12345-12354.

[9] WANG H, CHEN X, LIU S. Spatio-temporal gated attention network for long-term person re-identification[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2024: 67890-67899.

[10] 李鹏辉, 王洪元, 张继, 等. 基于多重注意力机制和空间变换网络的换衣行人重识别[J]. 南京大学学报（自然科学）, 2025, 61(2): 202-213.
LI P H, WANG H Y, ZAHNG J, et al. Cloth-changing person re-identification based on multiple attention mechanisms and spatial transformer networks[J]. Journal of Nanjing University (Natural Sciences), 2025, 61(2): 202-213.

[11] LUO H, GU Y, LIAO X, et al. Bag of tricks and a strong baseline for deep person re-identification[EB/OL]. [2025-02-23]. https://arxiv.org/pdf/1903.07071.

[12] 吴含笑, 赵倩倩, 朱建清, 等. 基于度量正则化的红外与可见光跨模态行人重识别算法[J]. 计算机科学, 2023, 50(6A): 221100046.
WU H X, ZHAO Q Q, ZHU J Q, et al. Metric regularized infrared and visible cross-modal person re-identification[J]. Computer Science, 2023, 50(6A): 221100046.
[13] ZHOU D, HOU Q, CHEN Y, et al. Rethinking bottleneck structure for efficient mobile network design[C]// Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2020: 680-697.
[14] 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 (CVPR). Piscataway: IEEE, 2020: 1580-1589.
[15] LI S, WANG Z, LIU Z, et al. MOGA: Multi-order gated aggregation network[EB/OL]. [2024-11-21]. https://arxiv.org/abs/2211.03295.
[16] ALTUNTAS B E, AKSU O, CELIK M Y, et al. Deep learning based automatic modulation recognition using GELU activation function[C]//Proceedings of the 2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA). Piscataway: IEEE, 2024: 1-4.
[17] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification[EB/OL]. [2025-03-21]. https://arxiv.org/abs/1703.07737.
[18] ZHAO L, LI X, ZHUANG Y, et al. Deeply-learned part-aligned representations for person re-identification[C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 3219-3228.
[19] HUANG Y, XU J, WU Q, et al. Beyond scalar neuron: Adopting vector-neuron capsules for long-term person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(10): 3459-3471.
[20] XU P, ZHU X. Deepchange: A long-term person re-identification benchmark with clothes change[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2023: 11196-11205.
[21] ZHONG Z, ZHENG L, KANG G, et al. Random erasing data augmentation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020: 13001-13008.
[22] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. [2024-12-19]. https://arxiv.org/abs/1412.6980.
[23] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). New York: ACM, 1996: 226-231.
[24] LI W, ZHU X, GONG S. Harmonious attention network for person re-identification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 2285-2294.
[25] SUN Y, ZHENG L, YANG Y, et al. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline)[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer, 2018: 501-518.
[26] HE S, LUO H, WANG P, et al. TransReID: Transformer-based object re-identification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 15013-15022.
[27] GU X, CHANG H, MA B, et al. Clothes-changing person re-identification with RGB modality only[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 1050-1059.
[28] HUANG Y, WU Q, XU J S, et al. Clothing status awareness for long-term person re-identification[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 11895-11904.
[29] MU J, LI Y, LI J, et al. Learning clothes-irrelevant cues for clothes-changing person re-identification[C]// Proceedings of the British Machine Vision Conference. Durham: BMVA Press, 2022: 337.
[30] WANG G, YUAN Y, CHEN X, et al. Learning discriminative features with multiple granularities for person re-identification[C]//Proceedings of the ACM International Conference on Multimedia. New York: ACM, 2018: 274-282.
[31] 刘乾, 王洪元, 曹亮, 等. 基于联合损失胶囊网络的换衣行人重识别[J]. 计算机应用, 2021, 41(12): 3596-3601.
LIU Q, WANG H Y, CAO L, et al. Cloth-changing person re-identification based on joint loss capsule network[J]. Journal of Computer Applications, 2021, 41(12): 3596-3601.
[32] ZHOU K, YANG Y, CAVALLARO A, et al. Omni-scale feature learning for person re-identification[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 3702-3712.
[33] WANG Q, QIAN X, FU Y, et al. Co-attention aligned mutual cross-attention for cloth-changing perso[C]// Proceedings of the Asian Conference on Computer Vision. Cham: Springer, 2022: 2270-2288.

责任编辑 税 红