

基于强区分性特征挖掘的无监督行人重识别

徐 锐, 王晓红, 赵 徐, 徐世豪

上海理工大学出版印刷与艺术设计学院, 上海

收稿日期: 2024年3月28日; 录用日期: 2024年5月2日; 发布日期: 2024年5月8日

摘 要

目前, 基于深度学习的无监督行人重识别正在通过改进聚类方法来提高生成伪标签的准确性和鲁棒性。然而, 行人的固有物理特征, 如肢体和身体, 尚未充分利用。此外, 遮挡和姿势导致了在行人图像中准确匹配和识别局部区域的困难, 而行人服装颜色引起的噪声干扰也对行人重识别准确性产生了重要影响。为了解决这些问题, 本文提出了一种判别性特征挖掘网络(DFMN), 通过引入Transformer注意力机制来突出显示行人的有效物理特征; 为了深入挖掘不同样本之间的细粒度局部特征, 本文采用了基于最短路径的部分对齐分割机制; 同时, 针对行人服装颜色引起的冗余信息, 本文融合了一个alpha通道, 可以有效消除噪声干扰。实验结果表明, 本文所提出的方法在Market1501数据集和MSMT17数据集上的map指标上分别实现了4.0%和6.4%的提升。

关键词

特征挖掘, 无监督学习, 伪标签, 聚类

Unsupervised Person Re-Identification Based on Discriminative Feature Mining

Kun Xu, Xiaohong Wang, Xu Zhao, Shihao Xu

College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 28th, 2024; accepted: May. 2nd, 2024; published: May. 8th, 2024

Abstract

At present, Unsupervised Person Re-Identification based on deep learning promotes the accuracy and robustness of generating pseudo labels mainly by improving the clustering method. However, for inherent physical features of pedestrians, such as limbs, and bodies, which are not fully utilized, occlusion, posture pose difficulties in accurately matching and recognizing local areas in

pedestrian images, and noise interference caused by pedestrian clothing color have essential impact on person re-identification accuracy. In this paper, we propose a Discriminative Feature Mining Network (DFMN) by introducing a Transformer attention mechanism to highlight the pedestrians' effective physical features. To deeply mining the fine-grained local features among different samples, a part-aligned segmentation mechanism based on the shortest path is used. At the same time, as for the redundant information resulting from the color of pedestrian clothing, we incorporate an alpha channel, which can effectively eliminate noise interference. The experimental results show that the proposed method has achieved improvements of 4.0% and 6.4% on the Market1501 dataset and the MSMT17 dataset for the map metric.

Keywords

Feature Mining, Unsupervised Learning, Pseudo Labels, Clusters

Copyright © 2024 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

大部分行人重识别模型都使用深度学习和标记数据进行训练。然而，随着实际监控需求增加和标记成本上升，无监督行人重识别成为研究的重点。这种方法通过聚类生成伪标签，不需要标记数据。

一种方法是无监督领域自适应(UDA)行人重识别[1]-[6]，该方法首先在源标记数据集上训练模型，然后在目标未标记数据集上对模型进行微调。然而，UDA方法的训练过程复杂且需要源域和目标域相似的限制，这限制了其成为主流研究趋势。另一种方法是纯无监督学习(USL)行人重识别[2] [7] [8] [9] [10] [11]，通常利用完全未标记的数据来生成伪标签。因此，本文侧重于进一步提高USL的性能。

目前最先进的无监督学习(USL)方法[9] [10] [11]使用内存字典和聚类生成的伪标签来训练神经网络。通过聚类算法：如DBSCAN [12]、K-means [13]或Infomap [14]，为每个图像生成伪标签。尽管这些方法取得了良好的性能，但在获取足够具有辨识力的行人特征方面仍存在显著挑战。

本文的主要贡献如下：

- 1) 针对行人的固有物理特征挖掘不充分。本文通过加入Transformer将行人的固有物理特征作为全局信息源，放大和强调其在识别过程中的重要性。
- 2) 针对个体的遮挡和姿势变化可能导致局部特征丢失。本文通过使用部件分割对齐机制来解决这些不稳定和频繁丢失的局部信息。
- 3) 针对行人服装颜色的干扰信息。因此本文加入alpha通道来剔除颜色干扰。

2. 相关工作

目前，无监督行人重识别方法通常采用ResNet50 [9] [10] [11]作为网络的主要框架。先前的工作在某些或整个框架上取得了显著的改进。具体来说，Lin等人[9]将每个样本视为一个簇，然后逐渐将相似的样本分组到一个簇中以生成伪标签。MMCL [10]的预测包括相似性计算和循环一致性的质量伪标签生成。SPCL [11]提出了一种新颖的自适应对比学习框架，逐步创建更可靠的聚类以精炼内存字典特征。然而，它们仅获得了行人的浅层特征，因此，聚类的准确性受到影响。

首先，从整体视角分析和定义图像中的全局特征。作为常见的全局特征，行人姿势特征可以直观地

从行人图像中感知,有助于区分不同行人的姿势,尽管遮挡会带来一定限制。Transformer 模型[15]提出,用于处理自然语言处理中的序列数据。许多研究也表明了其在计算机视觉任务中的有效性。Han 等人[16]和 Salman 等人[17]调查了 Transformer 在计算机视觉领域的应用。L. Zheng [18]等人首次成功将 ViTs 应用于重识别任务,提出了 TransReID,其在人员和车辆重识别上实现了最先进的性能。这些学者验证了 Transformer 在视觉全局特征获取以及其在重识别任务中的应用中的重要作用。

另一种方法涉及提取个体的动态特征,例如行走姿势、手臂摆动、头部和肩部运动。然而,这些动态特征的提取和处理更加复杂。例如,姿势不变嵌入(PIE) [19]将行人对齐到标准姿势以减少姿势变化的影响。SUN [20]强化了图中信息的部分。通过注意力机制,CHO [21]从整体身体层次挖掘了视觉相似性,并同时从上半身和下半身的局部层次挖掘了视觉相似性。然而,局部特征和全局特征之间没有明确的约束作为判断的共同基础。

在图像内容的部分被环境干扰遮挡的情况下,全局特征的利用和分析会极大地影响实际结果。与全局特征相比,局部特征提供了图像中对象的细粒度描述,涉及对图像特定区域的特征提取。获取局部特征的标准方法包括基于块、基于分割和基于网格的方法。CHO [21]提出了以无监督方式从未标记的域学习局部判别特征,而不是从全局图像中学习。PAST [22]通过使用局部特征改进了行人特征表示。这些学者有效地利用了局部特征作为辅助判别条件。然而,尽管利用局部特征作为判别标准,这些研究人员并未提出解决行人服装颜色对结果的影响的方法。

作为图像中表面特征的另一个直观体现,行人服装的颜色特征从行人图像中可感知。然而,由颜色导致的干扰噪声也存在于行人图像中。当不同身份的行人穿着相同颜色的衣服时,可能会出现误判。因此,抑制这种干扰信息并仅强调行人姿势和外观的主要特征至关重要。

3. 模型框架

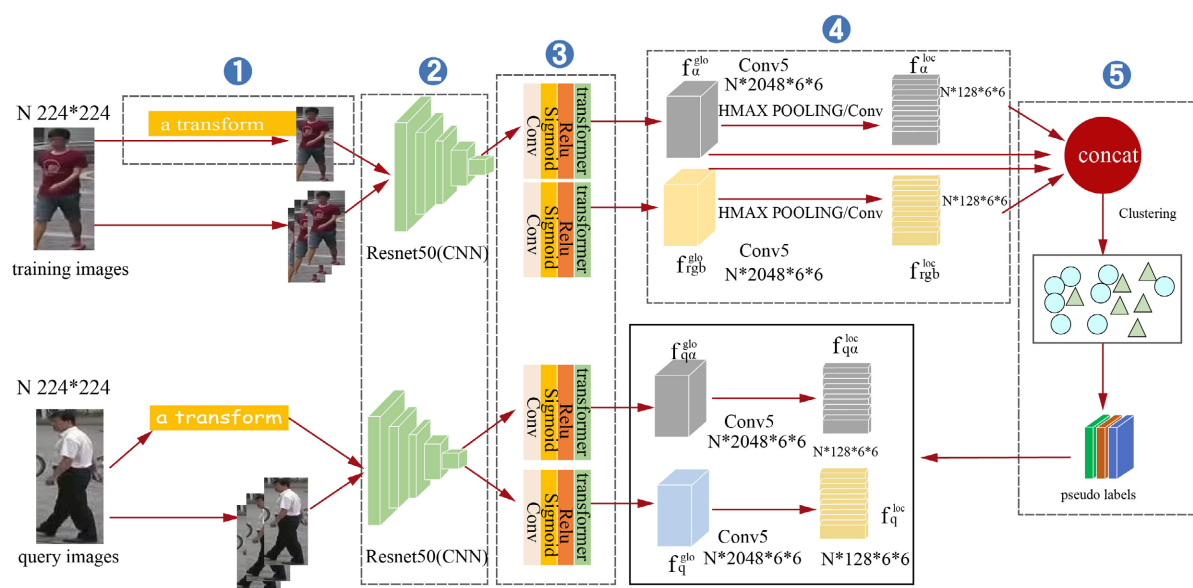


Figure 1. Framework diagram of pedestrian robustness discriminative feature mining algorithm
图 1. 行人强区分性判别特征挖掘算法框架图

本文提出了一个详细的网络结构,用于基于强特征挖掘的无监督行人再识别,如图 1 所示。该网络架构的这一部分主要包括五个不同的组件,其中①涉及通过 alpha 变换从训练图像中提取 α 通道。组件②

包括利用 ResNet50 对行人图像进行浅层特征提取,而组件③将这些浅层特征组合起来进行深度特征挖掘,并将其映射到统一的特征空间。在组件④中,获取的全局特征被分类到 α 通道和可见光中,然后进行归一化。通过水平最大化(HMAX) [23]实现对行人的局部特征表示,并利用这些局部特征的对齐来增强整体网络学习。最后,组件⑤涉及对融合特征进行聚类以生成伪标签,从而促进查询特征的匹配。

3.1. 基于 Transformer 的全局特征挖掘

为了突出特征图中的关键信息,减少无关细节,并从行人图像中提取更具区分性的全局特征,该组件采用了基于 Transformer 的全局特征卷积模块,如图 2 所示。

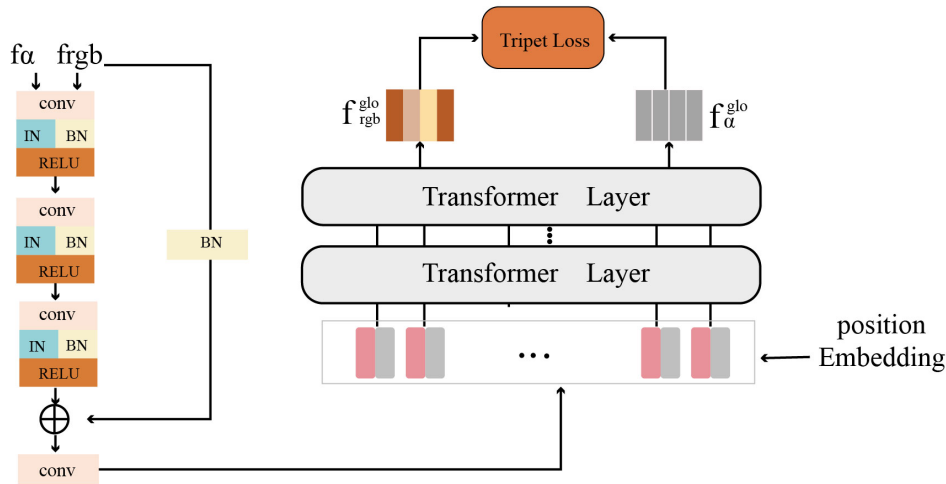


Figure 2. Component ③: Global feature convolution module based on Transformer

图 2. 组件③: 基于 Transformer 的全局特征卷积模块

通过组件①提取行人图像进行 α 通道提取后,本文利用网络中的组件②和③来挖掘更具有区分性的行人图像的全局特征。经过实验证明,带有堆叠的 BN 和 ReLU [23] [24] [25]层的卷积块显著提高了卷积的效率。具体来说,在 Resnet50 的浅层卷积层之后,输入到 Transformer 层的输入序列可以表示为:

$$T_0 = \left[F(f_{rgb}^1); \dots; F(f_{rgb}^N) \right] + \left[F(f_\alpha^1); \dots; F(f_\alpha^N) \right] + \rho \quad (1)$$

其中 T_0 表示输入序列嵌入, $\rho \in \mathbb{R}^{(N+1) \times D}$ 表示位置嵌入。 F 是一个线性投影,将图像分块映射到维度为 D 的空间。此外,Transformer 层用于学习特征表示。基于 CNN 的方法存在的有限感受视野问题得到解决,因为所有的 Transformer 层都具有全局感受视野,此外,没有下采样操作,因此详细信息得以保留。

本文使用滑动窗口来生成具有重叠像素的补丁。将步长表示为 S , 补丁的大小表示为 P , 那么两个相邻补丁重叠的区域形状为 $(P-S) \times P$ 。具有分辨率 $H \times W$ 的输入图像将被分割为 N 个补丁,如公式所示:

$$N = N_H \times N_W = \left\lfloor \frac{H+P-S}{S} \right\rfloor \times \left\lfloor \frac{W+P-S}{S} \right\rfloor \quad (2)$$

其中 $\lfloor \cdot \rfloor$ 表示向下取整函数, S 设置为小于 P 的值, N_H 和 N_W 分别表示高度和宽度方向上的分割补丁的数量。 S 越小,图像被分割成的补丁数量就越多。直观地说,更多的补丁通常会带来更好的性能,但会增加计算量。本文可以得到 RGB 行人全局特征 f_{rgb}^{glo} 和 α 行人全局特征 f_α^{glo} 通过以下表示:

$$f_{rgb}^{glo} = N \times \sum_{i=1}^N F(f_{rgb}^i) \quad (3)$$

$$f_{\alpha}^{glo} = N \times \sum_{i=1}^N F(f_{\alpha}) \quad (4)$$

本文通过构建全局特征模块的三元损失来优化组件②和③，如下所示：

$$L_T = \log \left(\frac{1 + e^{\|f_{rgb}^{glo} - f_{\alpha}^{glo}\|}}{e^{\|f_{rgb}^{glo} - f_{\alpha}^{glo}\|} + e^{\|f_{\alpha}^{glo} - f_{rgb}^{glo}\|}} \right) \quad (5)$$

3.2. 基于最短路径的局部特征对齐机制

在获得全局特征之后，RGB 图像的全局特征为 f_{rgb}^{glo} ， α 图像的全局特征为 f_{α}^{glo} 。局部对齐机制是基于人体结构特征的行人再识别中广泛使用的方法。它涉及将特征图分解成几个部分，提取它们，然后利用最短路径计算来找到更多匹配的局部特征。局部对齐机制可以更好地处理平移、旋转、缩放等几何变化。因此，在 DFMN 的组件④中，本文使用基于最短路径模块的部分分割对齐机制来挖掘局部特征，如图 3 所示。

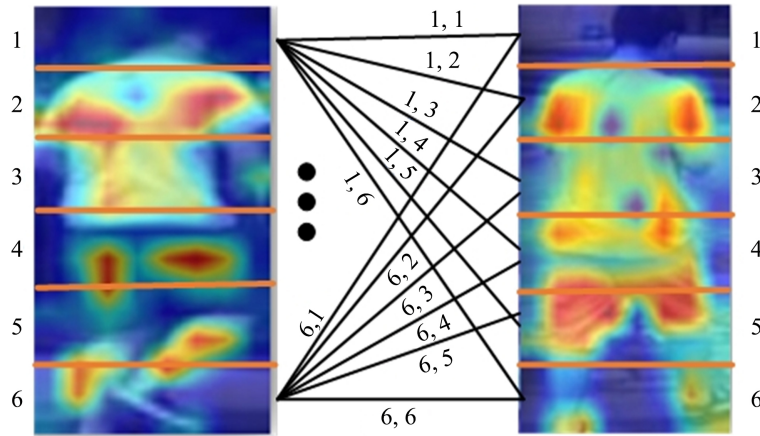


Figure 3. Component ④: Part segmentation alignment mechanism based on shortest path

图 3. 组件④：基于最短路径的部分分割对齐机制

以 RGB 的特征图为例，本文对这些特征进行了 HMAX 池化，包括水平分割。共设定了 M 个水平部分，对于每个部分，提取了一个局部特征图，确保每个局部特征图对应其相应的水平段。然后，本文定义行人局部特征表示的存在为 $f_{rgb}^{loc} = (f_{rgb}^1, f_{rgb}^2, \dots, f_{rgb}^m)$ ， $f_{\alpha}^{loc} = \{f_{\alpha}^1, f_{\alpha}^2, \dots, f_{\alpha}^m\}$ 。 M 是水平分割块的数量。

为了计算两幅图像的局部特征之间的距离，本文动态地从上到下匹配局部部分，找到具有最小总距离的局部特征。本文通过计算相似度来计算两个行人局部特征之间的距离，公式如下：

$$d_{rgb}^{m,n} = \frac{\|f_{rgb}^m - f_{rgb}^n\| - 1}{\|f_{rgb}^m + f_{rgb}^n\| + 1} \quad (6)$$

$$d_{\alpha}^{m,n} = \frac{\|f_{\alpha}^m - f_{\alpha}^n\| - 1}{\|f_{\alpha}^m + f_{\alpha}^n\| + 1} \quad (7)$$

上述公式表示第一幅图像的第 m 个垂直部分与第二幅图像的第 n 个垂直部分之间的距离，和分别表示两幅图像中局部特征的索引。计算绝对距离的平方。对于图像的每个分段部分，本文计算与另一图像的每个分段的局部距离。基于这些距离，形成一个 6×6 的距离矩阵。接下来，本文使用动态规划算法在

矩阵中找到距离和最小路径，并定义距离和计算公式如下：

$$S_{m,n} = \min(d^{m,n} + w(m,n)) \tag{8}$$

在上述的公式中，表示这个矩阵的最短路径的总距离。 $w(m,n)$ 表示从水平部分 m 到 n 的权重。

为了提高具有高背景相似性的行人图像的区别性，本文利用部分级别的距离约束来辅助局部特征学习，优化了组件④中的局部特征的损失，如公式所示：

$$L_M^{PA} = \sum_{i=1}^P \sum_{b=1}^K \sum_{j=2}^E \left(\left[ZG(f_{i,j}^{kb}) - f_{i,j}^{kp} \right]^2 + \left[f_{i,j}^{kn} - ZG(f_{i,j}^{kb}) \right]^2 \right) \tag{9}$$

在上述公式中， $ZG(\cdot)$ 表示零梯度函数，即在计算梯度时将变量视为常数， $f_{i,j}^{kb}$ 表示 α /RGB 局部特征的原始样本，表示 RGB/ α 通道中距离原始样本最远的正样本， $f_{i,j}^{kn}$ 表示 RGB/ α 通道中距离原始样本最近的负样本， P 表示行人的数量， K 表示同一行人的图像数量， E 表示行人特征图的长度。

3.3. 特征融合聚类用于生成伪标签

本文使用DFMN模型在 α 通道和RGB下提取全局和局部特征，然后通过串联将这些特征融合起来，得到融合特征，然后将这些融合特征发送到DBSCAN聚类算法，从组件⑤生成伪标签。图4显示了这个过程的流程图。

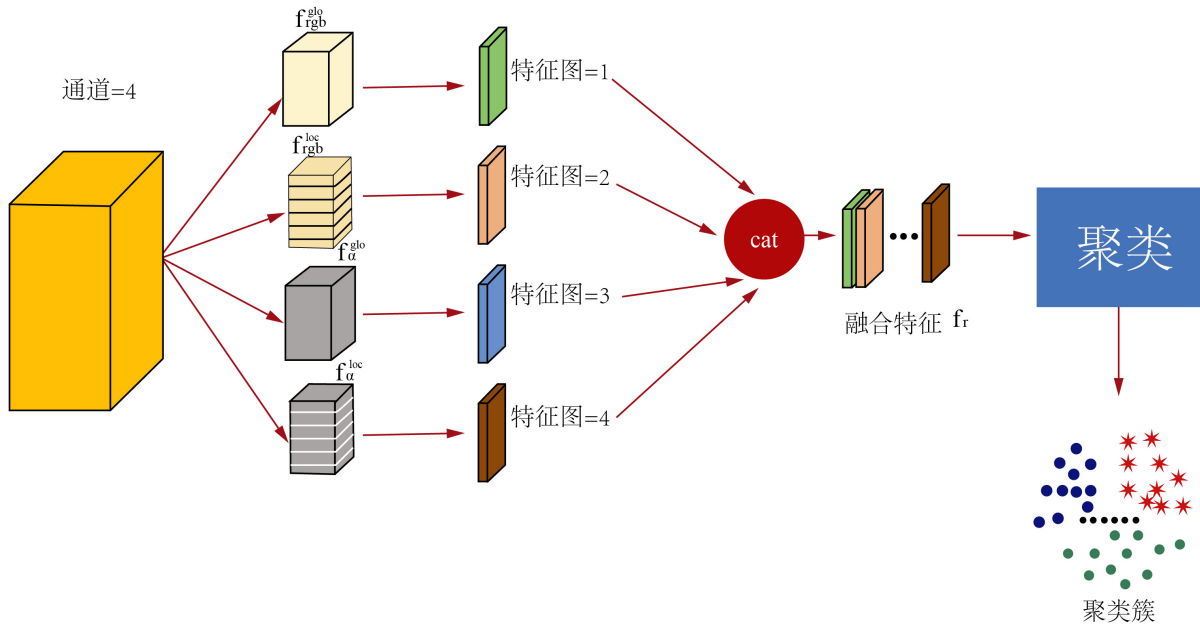


Figure 4. Component ⑤: Feature fusion and clustering

图 4. 组件⑤：特征融合和聚类

本文将融合的特征向量定义为 $f_r = \{f_r^1, f_r^2, \dots, f_r^i\}$ ，那么 F 的信息内容如公式所示：

$$I(f_r^i) = -\log Q(F = f_r^i) \tag{10}$$

其中 Q 代表取某个值的概率 f_r^i ，公式中的信息度量受限于在随机变量取特定值时处理信息度量的能力。在文中，熵被用来描述整个概率分布下的平均信息度量

为了减少类内变异性，本研究缩短了同一类别样本之间的距离，从而增加它们的相似性并将它们更接近样本中心，以更好地匹配查询图像的特征。损失函数定义如公式所示：

$$L_C = \frac{1}{2N} \sum_{i=1}^N |f_T^i - c|_2^2 \quad (11)$$

其中，类中心 c 指的是给定类别内所有样本的平均特征。这是通过将每个样本的特征通过高性能网络传递到特征层来实现的。因此，本文的 DFMN 网络的总损失如公式所示：

$$L_{DFMN} = L_T + L_M^{PA} + L_C \quad (12)$$

4. 实验

4.1. 数据集和参数设置

本文研究在两个大规模行人重识别数据集 Market1501 和 MSMT17 上评估所提出的方法。这些数据集在实际的行人重识别中被广泛使用，输入图像的尺寸为 224×224 。采用 SGD 优化器对 re-ID 模型进行训练，迭代次数为 $5e-6$ 。DBSCAN 聚类方法用于伪标签分配，所有实验使用相同的参数设置。两个样本之间的最大距离 d 设置为 0.6，核心点中的最小邻居数设置为 4。

4.2. 与无监督方法的比较

该实验首先将本文提出的方法与当前最先进的无监督学习(USL)方法进行比较，这是本文工作的主要焦点。如表 1 和表 2 所示，本文的方法在所有现有的无监督方法中表现出优越的性能，从而证明了其有效性。具体来说，使用相同的流程和 DBSCAN 聚类方法，本文提出的方法在 Market-1501 和 MSMT17 上的平均准确率(mAP)都超过了当前最先进的 USL 方法。

Table 1. Comparisons with state-of-the-art methods on Market1501 for USL Methods

表 1. 在 Market1501 数据集上与最先进方法的比较(无监督学习方法)

方法	Market1501			
	mAP	Rank-1	Rank-5	Rank-10
Buc [26]	38.3	66.2	79.6	84.5
SSL [27]	37.8	71.7	83.8	87.4
MMCL [10]	45.5	80.3	89.4	92.3
HCT [28]	56.4	80.0	91.6	95.2
SPCL [11]	73.1	88.1	95.1	97.0
ICE [19]	78.9	91.7	97.1	97.7
GSL [29]	79.2	92.3	96.6	97.8
CCL [30]	83.0	92.9	96.7	97.9
PPRL [21]	84.4	94.3	97.8	98.6
DFMN	88.4	94.6	97.9	98.7

Table 2. Comparisons with state-of-the-art methods on MSMT17 for USL Methods
表 2. 在 MSMT17 数据集上与最先进方法的比较(无监督学习方法)

方法	MSMT17			
	mAP	Rank-1	Rank-5	Rank-10
MMCL [10]	11.2	35.4	44.8	49.8
TAUDL [31]	12.5	28.4	-	-
UTAL [32]	13.1	31.4	-	-
UGA [5]	21.7	49.5	-	-
SPCL [11]	19.1	42.3	55.6	61.2
ICE [19]	22.7	48.4	61.1	67.0
GSL [29]	24.6	56.2	67.3	71.6
CCL [30]	31.2	61.5	71.8	76.7
PPRL [21]	42.2	73.3	83.5	86.5
DFMN	48.6	74.1	84.4	87.3

本文提出的方法展现了良好的结果。通过可视化和比较聚类结果与先前方法的结果，本文观察到本文的方法通过提取更具辨别力的特征增强了后续聚类步骤的稳健性，如图 5 所示。

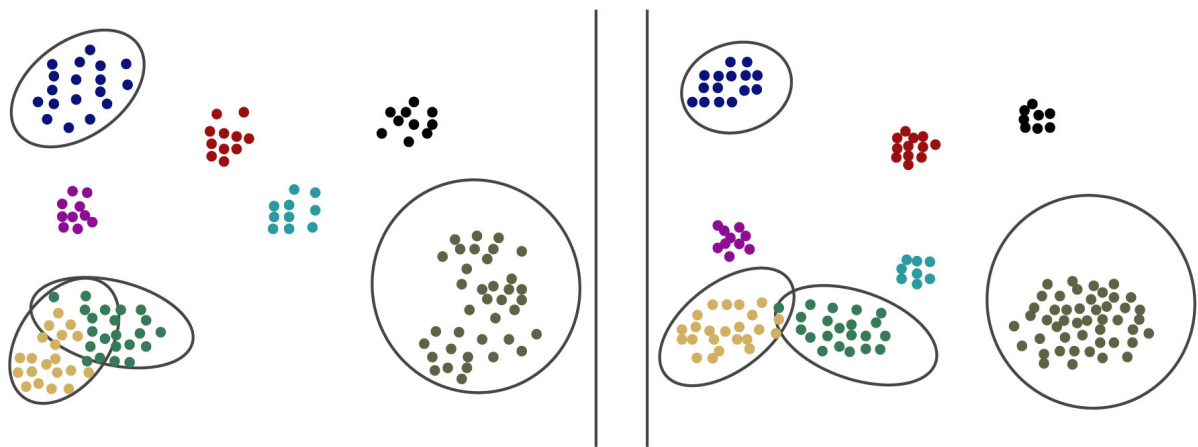


Figure 5. Our approach (Right) has been found to exhibit greater robustness than the unimproved method (Left), as demonstrated through visual analysis and comparison of the clustering results

图 5. 本文的方法(右)展现出比未改进的方法(左)更大的稳健性，通过对聚类结果进行可视化分析和比较

通过利用聚类技术，基于具有区分性的行人特征生成更为稳健的伪标签，本文提出的方法显著提高了重新识别的准确性。通过将行人样本匹配结果与未改进方法进行比较，图 6 展示了本文方法的有效性。

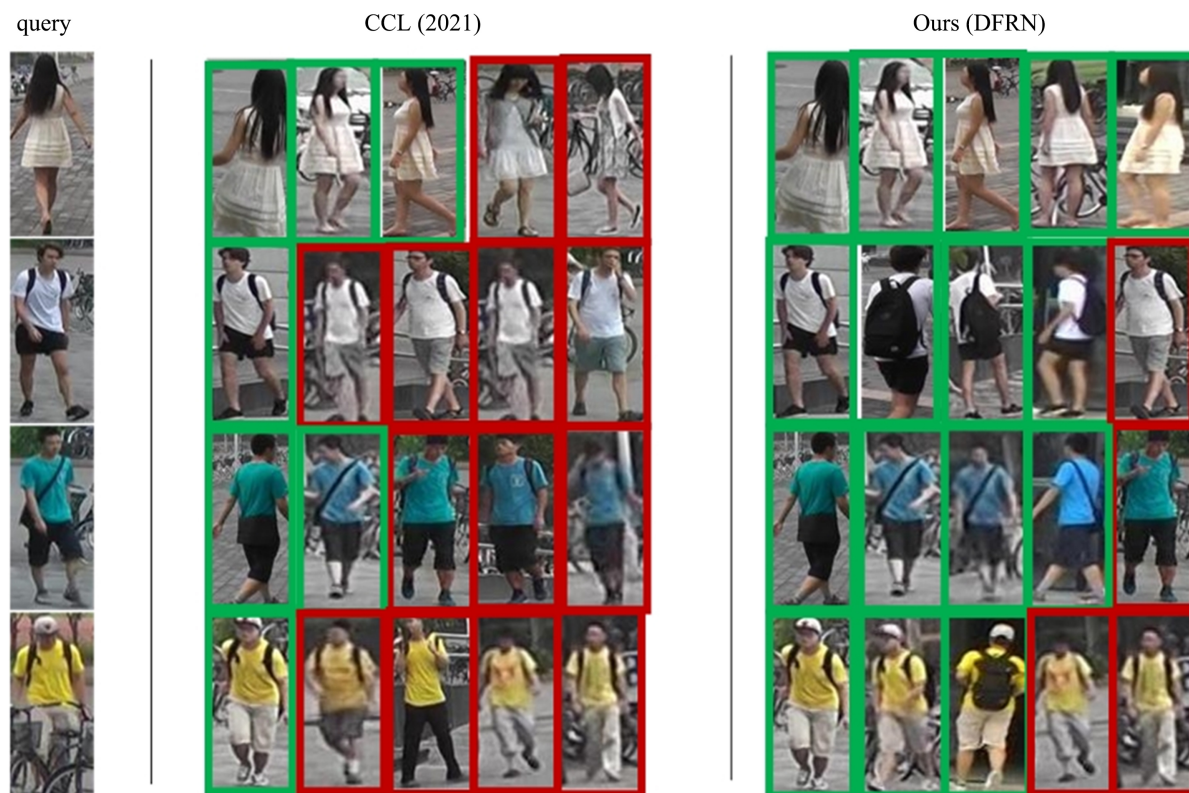


Figure 6. Comparison of top 5 retrieved images on Market1501 between CCL and our method. Green boxes denote correct results, while red boxes denote false results

图 6. 在 Market1501 数据集上, 本文的方法与 CCL 方法之间的前 5 个检索图像比较。绿色框表示正确结果, 红色框表示错误结果

4.3. 模块组合消融实验

本研究在 Market1501 数据集上进行了组件组合消融实验, 分别对四个方面进行了校准: ①、②、③和④。实验结果记录在表 3 中。结果显示, Transformer 注意机制对网络特征学习产生积极影响, 可以有效提高准确性。此外, alpha 通道被证明可以消除行人服装颜色的干扰信息, 突出身体姿势的有效信息。

Table 3. The ablation experimental results on the Market1501

表 3. Market1501 数据集上的消融实验结果

组件	Market1501			
	mAP	Rank-1	Rank-5	Rank-10
① + ② + ⑤	82.7	92.7	96.3	97.6
② + ③ + ⑤	86.0	94.7	97.6	98.4
② + ④ + ⑤	83.5	93.2	96.9	97.3
① + ② + ③ + ⑤	83.2	92.8	96.8	97.1
② + ③ + ④ + ⑤	87.2	93.5	97.2	98.0
① + ② + ③ + ④ + ⑤	88.5	94.4	98.1	98.8

从表 3 可以看出本文提出的网络框架专注于全局特征挖掘和局部特征挖掘模块，这显著提高了重新识别的准确性。组件③和④对 DMFN 的影响更加显著。

4.4. 局部特征对齐模块消融实验

为验证本文提出的基于块分割的局部特征对齐方法的有效性，本文使用 Market1501 数据集进行了测试实验。采用对齐模块机制对全局特征进行分块并提取局部特征，重点关注局部特征维度和水平分割块数的影响。首先，本文在将块数设置为 6 的情况下，验证了局部特征维度变化对准确性的影响，得到了如表 4 所示的实验结果。

Table 4. The impact of changes in local feature dimensions on accuracy

表 4. 局部特征维度变化对准确性的影响

维度	Market1501			
	mAP	Rank-1	Rank-5	Rank-10
128	84.6	92.5	96.3	97.6
256	84.2	91.3	96.1	97.8
512	83.5	91.4	96.9	97.3
1024	84.5	92.6	96.8	97.1

从表 4 可以看出，在一定数量的切片情况下，随着维度的增加，准确度 mAP 值逐渐下降，然后在达到最大值之前达到最大维度。而排名值显示出一定的波动。总体而言，局部特征维度的准确性取决于切块的数量。

Table 5. Experiment on the number of local feature blocks

表 5. 局部特征切块数实验

切块数量	Market1501			
	mAP	Rank-1	Rank-5	Rank-10
0	82.7	92.7	96.3	97.6
2	83.2	92.8	97.7	98.5
3	83.5	93.2	96.9	97.3
4	83.2	93.3	96.8	97.1
5	83.6	93.3	96.7	97.9
6	87.2	93.5	97.2	98.0
7	85.5	93.2	97.0	97.6
8	84.8	93.0	96.9	97.4
9	83.6	92.8	97.7	98.1

在前述步骤中，在保持相同数量的分割块的情况下，评估了不同维度特征对准确性的影响。在随后的实验中，本节特别选择了特定的特征维度，以研究分割块数对测试结果的影响。在 Market1501 数据集上进行测试时获得的实验数据记录在表 5 中。根据表 5 显示，随着分块数的增加，rank1 和 mAP 值均呈正比增加。然而，超过一定阈值后，随着分块数的继续增加，准确性会降低。当分块数设定为 6 时，达到最佳准确性。

5. 总结与展望

本文针对行人图像测试数据集中由于行人固有物理特征挖掘不充分，遮挡问题以及姿态不同而导致的局部特征信息丢失，以及行人衣服颜色带来的干扰信息这些因素所造成的低识别精度问题，从而提出了一种解决方案。本文的方法基于抑制颜色干扰信息，突出行人全局特征，并利用局部特征辅助判别这三个模块来解决这些问题。

首先，本文使用 Alpha 通道提取解决服装颜色干扰问题。然后，通过在主网络中嵌入 Transformer 来突出行人的固有物理全局特征，提高获取的全局特征鲁棒性。随后，本文使用最短路径部件分割对齐机制，利用全局特征进行水平分割以获取局部特征并约束行人组件。此外，本文采用全局和局部特征的联合学习，最后融合获得的特征并应用到机器学习聚类算法生成伪标签。实验结果的客观定量分析表明，本文提出的算法在各种性能评价指标上优于最新算法，从而验证了本文算法的出色性能。

参考文献

- [1] Deng, W.J., Zheng, L., Ye, Q.X., *et al.* (2018) Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Re-identification. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 994-1003. <https://doi.org/10.1109/CVPR.2018.00110>
- [2] Zhong, Z., Zheng, L., Li, S.Z., *et al.* (2018) Generalizing a Person Retrieval Model Hetero-and Homogeneously. *Computer Vision-ECCV 2018*, Munich, 8-14 September 2018, 176-192. https://doi.org/10.1007/978-3-030-01261-8_11
- [3] Zhong, Z., Zheng, L., Luo, Z.M., *et al.* (2019) Invariance Matters: Exemplar Memory for Domain Adaptive Person Re-identification. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 598-607. <https://doi.org/10.1109/CVPR.2019.00069>
- [4] Wei, L., Zhang, S., Gao, W. and Tian, Q. (2018) Person Transfer GAN to Bridge Domain Gap for Person Re-Identification. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 79-88. <https://doi.org/10.1109/CVPR.2018.00016>
- [5] Wu, J., Yang, Y., Liu, H., Liao, S., Lei, Z. and Li, S.Z. (2019) Unsupervised Graph Association for Person Re-Identification. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 8321-8330. <https://doi.org/10.1109/ICCV.2019.00841>
- [6] Wu, Z., Xiong, Y., Yu, S.X. and Lin, D. (2018) Unsupervised Feature Learning via Non-Parametric Instance Discrimination. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 3733-3742. <https://doi.org/10.1109/CVPR.2018.00393>
- [7] Fan, H., Zheng, L., Yan, C. and Yang, Y. (2018) Unsupervised Person Re-Identification: Clustering and Fine-Tuning. *ACM Transactions on Multimedia Computing, Communications, and Applications*, **14**, Article No. 83. <https://doi.org/10.1145/3243316>
- [8] Fu, Y., Wei, Y., Wang, G., Zhou, Y., Shi, H. and Huang, T.S. (2019) Self-Similarity Grouping: A Simple Unsupervised Cross Domain Adaptation Approach for Person Re-Identification. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 6112-6121. <https://doi.org/10.1109/ICCV.2019.00621>
- [9] Lin, Y., Dong, X., Zheng, L., Yan, Y. and Yang, Y. (2019) A Bottom-Up Clustering Approach to Unsupervised Person Re-identification. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 8738-8745. <https://doi.org/10.1609/aaai.v33i01.33018738>
- [10] Wang, D. and Zhang, S. (2020) Unsupervised Person Re-Identification via Multi-Label Classification. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 10981-10990. <https://doi.org/10.1109/CVPR42600.2020.01099>
- [11] Ge, Y., Zhu, F., Chen, D., Zhao, R. and Li, H.S. (2020) Self-Paced Contrastive Learning with Hybrid Memory for

- Domain Adaptive Object Re-ID. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Vancouver, 6-12 December 2020, 11309-11321.
- [12] Ester, M., Kriegel, H.-P., Sander, J., Xu, X., *et al.* (1996) A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland Oregon, 2-4 August 1996, 226-231.
- [13] MacQueen, J., *et al.* (1967) Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Oakland, 27 December 1965-7 January 1966, 281-297.
- [14] Rosvall, M. and Bergstrom, C.T. (2008) Maps of Random Walks on Complex Networks Reveal Community Structure. *Proceedings of the National Academy of Sciences*, **105**, 1118-1123. <https://doi.org/10.1073/pnas.0706851105>
- [15] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [16] Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y. *et al.* (2023) A Survey on Visual Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 87-110. <https://doi.org/10.1109/TPAMI.2022.3152247>
- [17] Khan, S., Naseer, M., Hayat, M., *et al.* (2022) Transformers in Vision: A Survey. *ACM Computing Surveys*, **54**, Article No. 200. <https://doi.org/10.1145/3505244>
- [18] Zheng, L., Huang, Y.J., Lu, H.C., *et al.* (2019) Pose-Invariant Embedding for Deep Person Re-Identification. *IEEE Transactions on Image Processing*, **28**, 4500-4509. <https://doi.org/10.1109/TIP.2019.2910414>
- [19] Chen, H., Lagadec, B. and Bremond, F. (2021) ICE: Inter-Instance Contrastive Encoding for Unsupervised Person Re-Identification. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 14940-14949. <https://doi.org/10.1109/ICCV48922.2021.01469>
- [20] Isobe, T., Li, D., Tian, L., Chen, W., Shan Y. and Wang, S. (2021) Towards Discriminative Representation Learning for Unsupervised Person Re-Identification. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 8506-8516. <https://doi.org/10.1109/ICCV48922.2021.00841>
- [21] Cho, Y., Kim, W.J., Hong, S., *et al.* (2022) Part-Based Pseudo Label Refinement for Unsupervised Person Re-Identification. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 7298-7308. <https://doi.org/10.1109/CVPR52688.2022.00716>
- [22] Yu, H.X., Zheng, W.S., Wu, A., Guo, X., Gong, S. and Lai, J.H. (2019) Unsupervised Person Re-Identification by Soft Multilabel Learning. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 2148-2157. <https://doi.org/10.1109/CVPR.2019.00225>
- [23] Kheradpisheh, S.R., Ganjtabesh, M., Thorpe, S.J., *et al.*, (2018) STDP-Based Spiking Deep Convolutional Neural Networks for Object Recognition. *Neural Networks*, **99**, 56-67. <https://doi.org/10.1016/j.neunet.2017.12.005>
- [24] Wang, P.C., Wang, X., Luo, H., *et al.* (2022) Scaled ReLU Matters for Training Vision Transformers. *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**, 2495-2503. <https://doi.org/10.1609/aaai.v36i3.20150>
- [25] Xiao, T., Singh, M., Mintun, E., *et al.* (2021) Early Convolutions Help Transformers See Better. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, Beijing, 4 November 2021.
- [26] Tong, H.Y. and Ng, K.W. (2021) A Bottom-Up Clustering Approach to Identify Bus Driving Patterns and to Develop Bus Driving Cycles for Hong Kong. *Environmental Science and Pollution Research International*, **28**, 14343-14357. <https://doi.org/10.1007/s11356-020-11554-w>
- [27] Lin, Y., Wu, Y., Yan, C., *et al.* (2020) Unsupervised Person Re-Identification via Cross-Camera Similarity Exploration. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 3387-3396. <https://doi.org/10.1109/CVPR42600.2020.00345>
- [28] Zeng, K., Ning, M., Wang, Y. and Guo, Y. (2020) Hierarchical Clustering with Hard-Batch Triplet Loss for Person Re-Identification. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 13654-13662. <https://doi.org/10.1109/CVPR42600.2020.01367>
- [29] Han, X.M., Yu, X.H., Li, G.R., *et al.* (2023) Rethinking Sampling Strategies for Unsupervised Person Re-Identification. *IEEE Transactions on Image Processing*, **32**, 29-42. <https://doi.org/10.1109/TIP.2022.3224325>
- [30] Dai, Z., Wang, G., Yuan, W., Zhu, S. and Tan, P. (2022) Cluster Contrast for Unsupervised Person Re-Identification. *Computer Vision-ACCV 2022*, Macao, 4-8 December 2022, 319-337. https://doi.org/10.1007/978-3-031-26351-4_20
- [31] Li, M., Zhu, X. and Gong, S. (2018) Unsupervised Person Re-Identification by Deep Learning Tracklet Association. *Computer Vision-ECCV 2018*, Munich, 8-14 September 2018, 772-788. https://doi.org/10.1007/978-3-030-01225-0_45
- [32] Li, M., Zhu, X. and Gong, S. (2019) Unsupervised Tracklet Person Re-Identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 1770-1782. <https://doi.org/10.1109/TPAMI.2019.2903058>