

参见本出版物的讨论、统计数据和作者简介：<https://www.researchgate.net/publication/340021578>

# 高阶信息问题：闭塞者再识别的学习关系和拓扑

预印-2020 年 3 月

引文

阅读

269

9 名作者，包括：



关南王

自动化研究所

10 份出版物，53 份引文

[见简介](#)



杨洋

中国科学院

28 份出版物，923 份引文

[见简介](#)

本出版物的一些作者也在从事这些相关项目：

项目

人再识别视图项目



# 高阶信息：学习关系和拓扑 用于闭塞的人重新识别

关安王<sup>1\*</sup>, 杨朔<sup>3\*</sup>刘环宇<sup>2</sup>, 王志成<sup>2</sup>杨阳<sup>1</sup>, 王淑良<sup>3</sup>, 帮派  
玉<sup>2</sup>, 周尔金<sup>2</sup> 还有孙健<sup>2</sup>

自动化研究所, CAS

2MEGVII 技术

<sup>3</sup>北京理工大学

<sup>1</sup>王瓜南 2015@ia.ac.cn<sup>2</sup>{liuhuanyu, wangzhicheng, yugang, zej, sunjian}@megvii.com

<sup>3</sup>{Shuoyang, slwang2011}@bit.edu.cn<sup>1</sup>{yang.yang}@nlpr.ia.ac.cn

## 摘要

遮挡人再识别 (ReID) 旨在将遮挡人图像与不相交摄像机上的整体图像相匹配。在本文中, 我们提出了一个新的框架, 通过学习高阶关系和拓扑信息的判别特征和鲁棒对齐。首先, 我们使用 CNN 骨干和关键点估计模型来提取语义局部特征。即使如此, 被遮挡的图像仍然存在遮挡和离群点。然后, 我们将图像的局部特征看作是图的节点, 并提出了一种自适应方向图卷积层来传递节点之间的关系信息。所提出的 ADGC 层可以通过动态学习方向和联动程度自动抑制无意义特征的消息传递。当两组局部特征从两幅图像中对齐时, 我们将其看作是一个图形匹配问题, 并提出了一个交叉图嵌入对齐 (CGEA) 层, 以联合学习和嵌入拓扑信息到局部特征, 并直接预测相似性评分。所提出的 CGEA 层不仅充分利用了图匹配所学的对准, 而且用鲁棒的软匹配代替了敏感的一对一匹配。最后, 在遮挡、部分和整体 ReID 任务上的广泛实验表明了我们提出的方法的有效性。具体来说, 我们的框架显著优于最先进的 6.5% 的 AP 分数在闭塞的杜克数据集。守则可在 <https://github.com/wanguan/HoreID> 查阅。

## 1. 引言

人的重新识别 (ReID) [6, 43] 旨在匹配一个人的图像跨越不相交的相机, 这是广泛的

用于视频监控、安全和智能城市。最近, 人们提出了 [25, 39, 18, 44, 16, 19, 43, 11, 35] 人身份识别的各种方法。然而, 它们大多集中在整体图像上, 而忽略了被遮挡的图像, 这可能更实用和更具挑战性。如图 1(a) 所示, 人很容易被一些障碍 (例如, 行李、柜台、拥挤的公共场所、汽车、树木) 或走出相机场, 导致遮挡图像。因此, 有必要将人与闭塞观察相匹配, 这就是所谓的闭塞人 Re-ID 问题 [48, 26]。

与用整体图像匹配的人相比, 由于以下原因, 遮挡 ReID

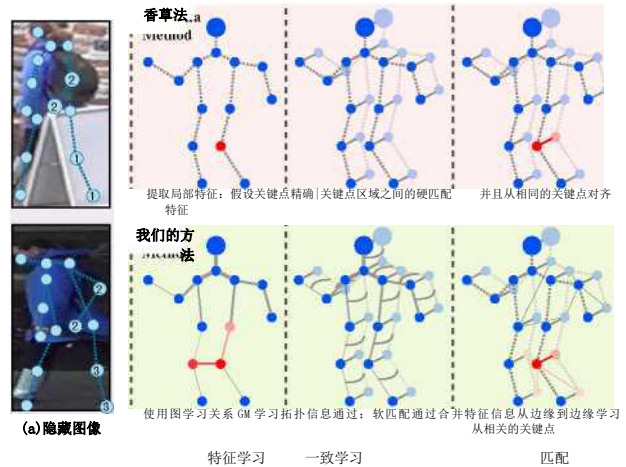


图 1. 说明高阶关系和拓扑信息。(a) 在闭塞的 ReID 中, 关键点受到闭塞 (①②) 和异常值 (③) 的影响。(b) 香草方法在所有三个阶段都依赖一阶关键点信息, 这一点不可靠。(c) 我们的方法通过图学习特征以建模关系信息, 并通过学习节点到节点和边缘到边缘的对应关系, 将对齐视为模型拓扑信息的图形匹配问题。

更具挑战性 [45, 48]: (1) 对于遮挡区域, 图像包含的鉴别信息较少, 更有可能匹配错误的人。(2) 通过部分对部分的匹配, 证明了基于部分的特征是有效的 [35]。但它们需要事先严格的人对齐, 因此在严重闭塞的情况下不能很好地工作。最近, 提出了许多闭塞/部分人 ReID 方法 [48, 49, 26, 10, 8, 34, 23], 其中大多数只考虑一阶信息进行特征学习和对齐。例如, 预定义的区域 [35]、姿态 [26] 或人类解析 [10] 用于特征学习和对齐。我们认为, 除了一阶信息外, 还应该导入高阶信息, 并且可能对阻塞的 ReID 更有效。

在图 1(A) 中, 我们可以看到关键点信息存在遮挡 (①②) 和异常值 (③)。例如, 关键点 ① 和 ② 被遮挡, 导致无意

义的特征。关键点③是异常值，导致失调。常见的解决方案如图 1(B)所示)。它提取关键点区域的局部特征，假设所有关键点都是准确的，并且局部特征很好地对齐。在这个解决方案中，所有三个阶段都依赖于二阶关键点信息，这并不是很健壮。在本文中，如图 1(C)所示，我们提出了一种新的判别特征和鲁棒对齐框架。在特征学习阶段，我们将图像的局部特征作为图的节点来学习关系信息。通过在图中传递消息，封闭关键点引起的无意义特征可以通过其邻居有意义的特征来改善。在对齐阶段，我们使用图匹配算法[40]学习鲁棒对齐。除了与节点对节点的对应关系保持一致外，它还模拟了额外的 edgeto 边缘对应关系。然后，我们通过构造交叉图像图将对齐信息嵌入到特征中，其中图像的节点消息可以传递给其他图像的节点。因此，离群点的特征可以通过它在另一幅图像上的相应特征来修复。最后，我们使用网络来学习由验证损失监督的相似性，而不是用预定义的距离来计算相似性。

具体而言，我们提出了一种新的框架，联合建模高阶关系和人-拓扑信息的闭塞人重新识别。如图 2 所示，我们的框架包括三个模块，即。一阶语义模块(S)、高阶关系模块(R)和高阶人类拓扑模块(T)。在 S 中，我们利用 CNN 骨干来学习特征映射和人类关键点估计模型来学习关键点。然后，我们可以提取相应关键点的语义特征。在 R 中，我们将图像的学习语义特征看作是图的节点，并提出了一个自适应方向图卷积层(ADGC)来学习和传递边缘特征的消息。ADGC 层可以自动决定每个边缘的方向和程度。因此，它可以促进语义特征的信息传递，抑制无意义和嘈杂的信息传递。最后，学习的节点包含语义和相关信息。在 T 中，我们提出了一个交叉图嵌入对齐(CGEA)层。它以两个图作为输入，利用图匹配策略学习两个图之间节点的对应关系，并将学习到的对应关系作为邻接矩阵传递消息。因此，可以增强相关特征，并将对齐信息嵌入到特征中。最后，为了避免硬一对一的对齐，我们通过将两个图映射到 logit 中来预测它们的相似性，并以验证损失进行监督。

本文的主要贡献总结如下：(1)提出了一种新的高阶关系和人-拓扑信息联合建模框架，用于学习遮挡 ReID 的良好和鲁棒对齐特征。据我们所知，这是第一个将这种高阶信息引入到阻塞的 ReID 的工作。(2)提出了一种自适应有向图卷积层来动态学习图的有向链接，它可以促进语义区域的消息传递，抑制无意义区域的消息传递，如遮挡或离群点。利用该模型，可以更好地模拟遮挡 ReID 的关系信息。(3)提出了一种与验证损失相结合的交叉图嵌入对准(CGEA)层，用于学习特征对准和预测相似性评分。它们可以避免敏感的硬一对一的人匹配，并执行一个健壮的软匹配。(4)对闭塞、部分和整体 ReID 数据集的广泛实验结果表明，所提出的模型对最先进的方法具有良好的效果。特别是在闭塞的 Duke 数据集，我们的方法在 Rank-1 和 mAP 评分方面明显优于最先进的至少 3.7%和 6.5%。

## 2. 相关工程

人重新识别。人重新识别解决了在不相交的摄像机[6]上匹配行人图像的问题。关键的挑战在于，由于不同的观点、姿势、光照和遮挡而导致的大类内和小类间变化。现有的方法可以分为手工制作的描述符[25, 39, 18]、度量学习方法[44, 16, 19]和深度学习算法[43, 11, 35, 36, 37, 22]。所有这些 ReID 方法都集中在匹配整体人图像上，但对遮挡图像的性能不佳，这限制了在实际监控场景中的适用性。

闭塞的人重新识别。给定闭塞的探针图像，闭塞的人重新识别[48]旨在发现同一人的全身外观在不相交

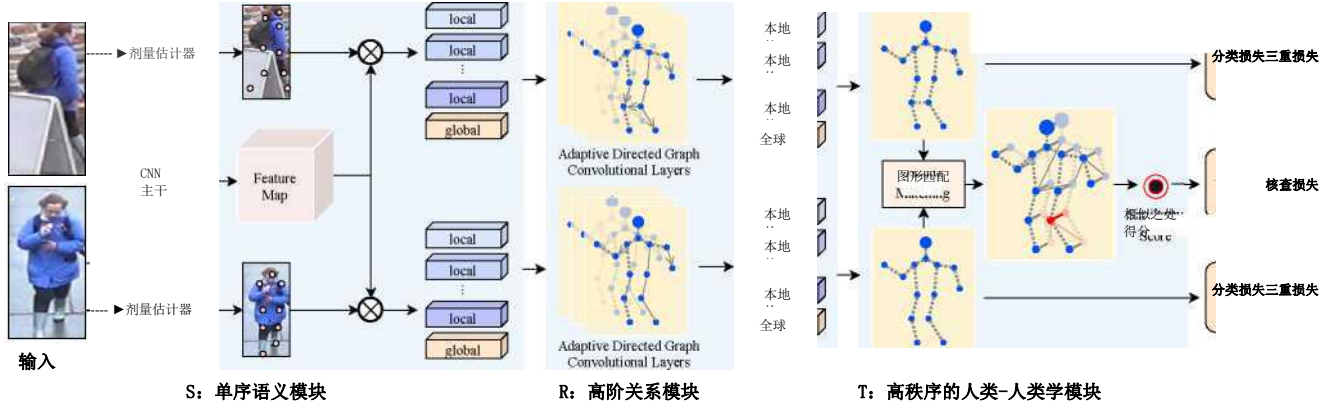


图 2。说明我们提议的框架。它由一阶语义模块 S、高阶关系模块 R 和高阶拓扑模块 T 组成。模块 S 学习关键点区域的语义局部特征。在 R 中，我们将图像的局部特征看作是图的节点，并提出了一种自适应方向图卷积层来传递节点之间的关系信息。在 T 中，我们将对齐问题看作是一个图形匹配问题，并提出了一个交叉图嵌入对齐 (CGEA) 层，将拓扑信息联合学习并嵌入到局部特征中，并直接预测相似性分数。

照相机。由于信息不完整和空间失调，这项任务更具挑战性。卓等人。[48]使用遮挡/非遮挡二值分类 (OBC) 损失来区分遮挡图像和整体图像。在他们接下来的工作中，预测了一个显着性地图来突出区分部分，师生学习方案进一步改进了学习特征。Miao 等人。[26]提出了一种基于人的语义关键点的姿态引导特征对齐方法来匹配探针和图库图像的局部补丁。它们使用预先定义的关键点置信度阈值来确定零件是否被遮挡。Fan 等人。[3]使用空间信道并行网络 (SCPNet) 将零件特征编码到特定的信道，并将整体特征和零件特征融合起来，得到判别特征。罗等人。[23]使用空间变换模块对整体图像进行变换，使其与部分图像对齐，然后计算对齐的距离。此外，还对部分 Re-ID 任务的空间对齐进行了一些努力。

部分人重新识别。伴随着遮挡的图像，局部图像往往发生由于不完善的检测和离群点的相机视图。与被遮挡的人 ReID 一样，部分人 ReID[45]的目的是将部分探针图像与画廊整体图像相匹配。郑等人。[45]提出了一种全局到局部匹配模型来捕获空间布局信息。他等人。[7]从整体行人中重建部分查询的特征映射，并通过前景-背景掩码进一步改进，以避免背景杂波对[10]的影响。Sun 等人。在[34]中提出了一种可见性感知部件模型 (VPM)，它通过自我监督学习感知区域的可见性。

与现有的只使用一阶信息进行特征学习和对齐的闭塞和部分 ReID 方法不同，我们使用高阶关系和体液学信息进行特征学习和对齐，从而获得更好的性能。

### 3. 拟议方法

本节介绍了我们提出的框架，包括一个一阶语义模块 (S) 来提取人类关键点区域的语义特征，一个高阶关系模块 (R) 来建模不同语义局部特征之间的关系信息，以及一个高阶体液学模块 (T) 来学习两幅图像之间的鲁棒对齐和预测相

似性。这三个模块是以端到端的方式联合培训的。所提出的方法的概述如图 2 所示。

语义特征提取。该模块的目标是提取关键点区域的一阶语义特征，这是由两个线索启发的。首先，基于部分的特性被证明是有效的人 ReID[35]。其次，在闭塞/部分 ReID[8, 34, 10]中，局部特征的精确对齐是必要的。根据上述想法，并在人 ReID[43, 35, 24, 4]和人类关键点预测[2, 33]的最新发展的启发下，我们利用 CNN 骨干提取不同关键点的局部特征。请注意，虽然人类关键点预测已经达到了很高的精度，但在遮挡/部分图像[17]下，它们仍然具有不满意的性能。这些因素导致关键点立场和信心不准确。因此，需要以下关系和人类主题信息，并将在下一节中讨论。

具体来说，给定一个行人图像  $x$ ，我们可以得到它的特征映射  $m_{CNN}$  和关键点热图  $m_{KP}$  通过 CNN 模型和关键点模型。通过外部产品 ( $\odot$ ) 和全局平均池操作 ( $G(\cdot)$ )。

我们可以得到一组关键点区域  $V$  的语义局部特征  $i$  以及一个全局特征 VGS。这些程序可以在 Eq 中制定。(1)，其中  $K$  是关键点数  $k \in \mathbb{R}^c$  而  $c$  是信道号。注意我  $KP$  通过用 Softmax 函数对原始关键点热图进行归一化，以防止噪声和异常值。这种简单的操作在实验部分是有效的。

$$v_i^{*s} = \{ \bar{v}_k \}_{k=1}^K = \text{argmax}_{k \in \{1, \dots, K\}} \{ \text{softmax}(\bar{v}_k) \} \quad (1)$$

培训损失。在[43, 11]之后，我们利用分类和三重态损失作为我们的目标，就像在 Eq 中一样。(2)。在这里， $k = \text{最大值}(MKP[k]) \in [0, 1]$  是  $k^{\text{th}}$  关键点置信度，和  $k+1=1$  的全局特征， $p_v$  是由分类器预测的特征  $v_k$  属于其地面真相标识的概率， $a$  是裕度， $d_{v^+}$ ， $v^-$  是正对 (VG) 之间的距离  $k$ ， $d_{v^-}$  来自同一身份  $k$  来自不同的身份。不同局部特征的分类器不共享。

$$\begin{aligned} & \text{是} = \frac{1}{K+1} \sum_{k=1}^{K+1} \left[ \text{Ltri}(v_k) + \text{Ltri}(v^+) \right] \\ & = \frac{1}{K+1} \sum_{k=1}^{K+1} \left[ \text{Ltri}(v_k) + \text{Ltri}(v^+) \right] \end{aligned}$$

### 3.1. 高阶关系学习

虽然我们有不同关键点区域的一阶语义信息，但由于行人图像不完整，遮挡 ReID 更具挑战性。因此，有必要利用更多的鉴别特征。我们转向图卷积网络 (GCN) 方法[1]并尝试对高阶关系信息进行建模。在 GCN 中，不同关键点区域的语义特征被看作是节点。通过在节点之间传递消息，不仅可以联合考虑一阶语义信息（节点特征），还可以联合考虑高阶关系信息（边缘特征）。

然而，对于遮挡 ReID 仍然存在挑战。闭塞区域的特征往往是毫无意义的，甚至是嘈杂的。当在图中传递这些特征时，它会带来更多的噪声，并对阻塞的 ReID 产生副作用。因此，我们提出了一种新的自适应方向图卷积层 (ADGC) 来动态学习消息传递的方向和程度。利用它，我们可以自动抑制无意义特征的消息传递，促进语义特征的传递。

自适应定向图卷积层。一个简单的图卷积层[15]有两个输入，一个图的相邻矩阵  $A$  和所有节点的特征  $X$ ，

输出可以通过：其中  $A$  是归一化版本， $A$  和  $W$  是指参数。

我们基于输入特征自适应地学习相邻矩阵（节点的链接）来改进简单的图卷积层。我们假设给定两个局部特征，有意义的特征比无意义的特征更类似于全局特征。因此，我们提出了一种自适应有向图卷积 (ADGC) 层，其输入是全局特征  $V_g$  和  $K$  局部特征  $V_l$ ，以及预定义的图（相邻矩阵为

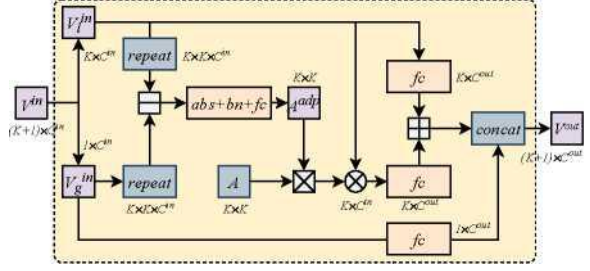


图 3。图示了所提出的自适应有向图卷积层 (ADGC)。甲是预先定义的相邻矩阵日，田，图是按元素进行减法，加法和乘法。腹肌、BN 和 FC 是绝对化、批归一化和完全连接层，反式是转置的。详情请参阅正文。

A)。我们使用局部特征  $V_l$  和全局特征  $V_g$  之间的差异  $g$  动态更新图中所有节点的边权重，得到  $A$  冲。然后，通过  $V_l$  和  $A$  之间的乘法可以得到一个简单的图卷积  $\odot$ 。为了稳定训练，我们将输入的本地特性  $V_l$  融合到我们的 ADGC 层的输出中，就像在 ResNet[7] 中一样。详情如图 3 所示。我们的自适应有向图卷积 (ADGC) 层可以用方程表示。(3)，其中  $f_i$  和  $f$  是两个未共享的完全连接层。

$$V^{\text{out}} = [F1(A^{\text{out}} \odot V_l^{\text{in}}) + F2(V_g^{\text{in}}), V_l^{\text{in}}] \quad (3)$$

最后，我们实现了我们的高阶关系模块  $f_n$  作为 ADGC 层的级联。因此，给定一个图像  $x$ ，我们可以得到它的语义特征  $V = \{v_k\}_{k=1}^K$  通过 Eq。(1)。则其关系特征  $V = \{v_R\}_{k=1}^K$  可表述如下：

$$v^r = fR(V^s) \quad (4)$$

损失和相似性。我们使用分类和三重态损失作为我们的目标，就像在 Eq 中一样。(5)  $L_{cls}(-)$  和  $L_{tri}(j)$  的定义可在 Eq 中找到。(2)。注意，仇是  $k^{\text{th}}$  关键点信心。

$$\begin{aligned} & \frac{1}{K+1} \sum_{k=1}^{K+1} \left[ \text{Lcls}(v^r) + \text{Ltri}(v^r) \right] \\ & \text{输出} = \frac{1}{K+1} \sum_{k=1}^{K+1} \left[ \text{Lcls}(v^r) + \text{Ltri}(v^r) \right] \end{aligned} \quad (5)$$



给定两幅图像和  $X_2$ ，我们可以得到它们的关系特征  $v_i$  特征

——  $1^k$  还有  $v_i$  ——  $Jv^T 1^k$  我通过

勾  $1^k$  还有  $v_i$  ——  $Jv^T 1^k$  通过

1  $k+1$

——  $Jv^T 1^k$  余弦  $(V_k, v)$  (6)

情商。（4）计算它们与方程中余弦距离的相似性。（6）。

### 3.2. 高层次的人学学习

基于部件的特性已被证明是非常有效的人 ReID[35, 34]。一个简单的对齐策略是直接匹配相同关键点的特征。然而，这种一阶对齐策略不能处理一些坏的情况，如异常值，特别是在严重闭塞的情况下[17]。图形匹配[40, 38]自然可以考虑高阶人类拓扑信息。但只能学会一一对应。这种硬对齐仍然对异常值敏感，并对性能产生副作用。在该模块中，我们提出了一种新的交叉图嵌入对齐层，它不仅可以利用图匹配算法所学的人类拓扑信息，而且可以避免敏感的一对一对齐。

**图形匹配的修订。** 从图像中给出两个图——和  $G_2$ ——（电）和  $X_2$ ，the 图匹配的目标是学习匹配矩阵  $U \in [0, 1]^{n_1 \times n_2}$  在  $V_1$  和  $V_2$  之间。让  $U \in [0, 1]$  一个指示向量，使  $U$  内部是  $v_1$  和  $v_2$  之间的匹配度。建立一个平方对称正矩阵  $M \in \mathbb{R}^{n_1 \times n_2}$ ，使  $M$  内部  $j$  测量每一对  $(i, j) \in E_1$  与  $(a, b) \in E_2$  的匹配程度。对于不形成边的对，它们在矩阵中的对应项被设置为 0。对角线条目包含节点-节点分数，而非对角线条目包含  $edge$  边缘分数。因此，最优匹配  $u^*$  可以表述如下：

$$U^* = \arg \max_U U^T M U, \text{ s.t. } ||U|| = 1 \quad (7)$$

在[40]之后，我们用一元和成对点特征来参数矩阵  $M$ 。优化过程是通过功率迭代和双随机操作来制定的。因此，我们可以在具有随机梯度下降的深度学习框架中优化  $U$ 。受页面限制，我们没有显示更多的图形匹配细节，请参阅论文[38, 40]。

**具有相似性预测的交叉图嵌入对齐层。** 我们提出了一种新的交叉图嵌入对齐层 (CGEA)，它既考虑了  $GM$  学习的高阶人类拓扑信息，又避免了敏感的一对一对齐。所提出的 CGEA 层以两幅图像中的两个子图作为输入，输出嵌入的特征，包括两者

语义特征和人类拓扑引导的对齐特征。

我们提出的 CGEA 层的结构如图 4 所示。它需要两组特征，并输出两组特征。首先，有两组节点  $V_i \in \mathbb{R}^{l \times 1} \times \mathbb{C}^d$  和  $V_j \in \mathbb{R}^{k \times 1} \times \mathbb{C}^d$  我们将它们嵌入到具有完全连接层和 ReLU 层的隐藏空间中，得到两组隐藏特征  $V^h \in \mathbb{R}^{(l+1) \times 1} \times \mathbb{C}^d$  和  $V^h \in \mathbb{R}^{(k+1) \times 1} \times \mathbb{C}^d$ 。其次，我们通过  $E_q$  执行“和”之间的图匹配。（7），得到一个亲和矩阵  $U$  介于  $V^h$  和  $V^h$  之间。在这里， $U(i, j)$

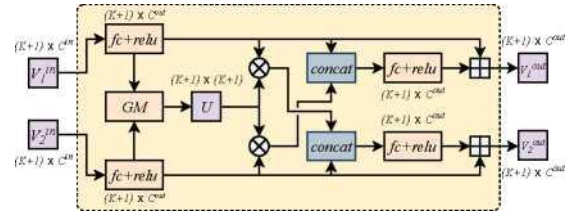


图 4。交叉图嵌入对齐层的说明。这里，区是矩阵乘法，FC+relu 表示完全连接层和恢复线性单元，GM 表示图形匹配操作，U 是学习的亲和矩阵。详情请参阅正文。

表示  $v_i$  和  $v_j$  之间的对应关系。最后，输出可以用  $E_q$  表示。

（8）其中  $[\cdot, \cdot]$  表示沿信道维度的级联操作， $f$  是一个完全连接的层。

$$\begin{aligned} & \text{咋呀} \text{——} f([V^h, U^h V^h]) + \\ & \text{——} f([V^h, U^h V^h]) + V^h \end{aligned}$$

我们实现了我们的高阶拓扑模块 (T) 与级联的 CGEA 层  $f$  和相似性预测层  $f$ 。给定一对图像  $(x_1, x_2)$ ，我们可以得到它们的关系特征  $(V)$  通过  $E_q$ 。（4），然后通过方程（“，”）的拓扑特征。（9）。在得到拓扑特征对  $(V, V^h)$  后，我们可以使用  $E_q$  计算它们的相似性。（10），其中  $|\cdot|$  是按元素进行的绝对化操作， $f_s$  是从  $\mathbb{C}^d$  到 1 的完全连接层， $a$  是乙状结肠激活函数。

$$(V^T, V^T) \text{——} (V^T V^T) \quad (9)$$

$$\text{——} a(f_s(-|V^T - V^T|)) \quad (10) \quad (8)$$

核查损失。我们的高阶体液学模块的损失可以用方程来表示。（11），其中  $y$  是他们的基本真理， $y \text{——} 1$ ，如果  $(x_1, x_2)$  来自同一个人，否则  $y \text{——} 0$ 。

$$L \text{——} y / \log \frac{1}{1 - y} / \log \frac{1}{1 - y} \text{——} s^T, x_2 \quad (11)$$

### 4. 训练和推理

在训练阶段，我们的框架的总体目标函数是在方程中制定的。（12）其中  $\lambda$  为相应项的权重。我们通过最小化  $L$  来训练我们的框架端到端。

$$L = L_s + \lambda L_r + T \quad (12)$$

对于相似性，给定一对图像  $(x_1, x_2)$ ，我们可以得到基于相似性  $s^1$  的关系信息，来自情商。（六）基于拓扑信息的相似性  $s^1$ ，来自情商。（10）。最后的相似性可以通过梳理这两种相似性来计算。

$$s = Y_s X_i X^T (1 - y)^{s^T} \text{我}, x_2 \quad (13)$$

推断时，给定查询图像  $x_q$  我们首先计算它的相似性  $s^1$  与所有画廊图像，并得到它的顶部  $n$  最近的邻居。然后计算方程中的最终相似度  $s$ 。（13）改进最高的  $n$ 。

5. 实验

5.1. 实施细节

模型架构。对于 CNN 骨干，与[43]一样，我们通过删除其全局平均池 (GAP) 层和完全连接层，将 ResNet50[7]作为 CNN 骨干。对于分类器，在[24]之后，我们使用批处理归一化层[13]和完全连接层，然后是 Softmax 函数。对于人类关键点模型，我们使用 HR-Net[33]预先训练的 COCO 数据集[20]，一个最先进的关键点模型。该模型预测了 17 个关键点，我们融合了头部区域的所有关键点，得到了最终的 K=14 个关键点，包括头部、肩膀、肘部、手腕、臀部、膝盖和脚踝。

培训细节。我们用 Pytorch 实现了我们的框架。图像被调整到 256×128，并通过随机水平翻转、填充 10 像素、随机裁剪和随机擦除[47]增强。当对遮挡/部分数据集进行测试时，我们使用额外的颜色抖动增强来避免域方差。批量大小设置为 64，每人 4 张图片。在训练阶段，所有三个模块都以端到端的方式联合训练 120 个时代，初始化学率为 3.5e-4，衰减到 0.1at30 和 70 个时代。请参考我们的代码<sup>4</sup>有关实现细节。

评价指标。我们使用标准指标，如大多数人 ReID 文献，即累积匹配特性 (CMC) 曲线和平均精度 (MAP)，以评估不同的人重新识别模型的质量。所有实验都是在单个查询设置中执行的。

5.2. 实验结果

封闭数据集的结果。我们在两个被遮挡的数据集上评估我们提出的框架，即。阻塞杜克[26]和阻塞-重新 ID[48]。阻塞-杜克是

数据集	训练妈妈 (ID/Image)	测试 Nums (ID/图像)	
		画廊	查询
市场-1501	751/12, 93	750/19, 732	750/3, 36
杜克 MTMC-是 ID	702/16, 52	1, 110/17, 6	702/2, 22
闭塞-杜克	702/15, 61	1, 110/17, 6	519/2, 21
闭塞-重新 ID	-	200/1, 000	200/1, 00
部分恢复	-	60/300	60/300
部分-我的	-	119/119	119/119

表 1。数据集的详细信息。我们在 6 个公共数据集上对我们提出的方法进行了广泛的评价，包括 2 个整体数据集、2 个闭塞数据集和 2 个部分数据集。

方法	闭塞-杜克		闭塞-里德	
	排名第	m AP	排名第	m AP
部分不结盟[41]	28. 8	20. 2	-	-
PCB[35]	42. 6	33. 7	41. 3	38.
部分苯胺[32]	36. 9	-	-	-
FD-gan[5]	40. 8	-	-	-
空管+SWM[45]	-	-	31. 2	27.
DSR[8]	40. 8	30. 4	72. 8	62.

SFR[9]	42. 3	32	-	-
加入了[12]	44. 5	32. 2	-	-
[49]	-	-	73. 7	77.
FPR[10]	-	-	78. 3	68.
PGFA[26]	51. 4	37. 3	-	-
<b>HORE ID (Ours)</b>	<b>55. 1</b>	<b>43. 8</b>	<b>80. 3</b>	<b>70. 2</b>

表 2。与两个闭塞数据集的最先进技术比较，即。闭塞-杜克[26]和闭塞-里德[48]。

从杜克 MTMC-reID 中选择，留下被遮挡的图像，并过滤掉一些重叠的图像。它包含 15, 618 个训练图像、17, 661 个图库图像和 2, 210 个遮挡查询图像。遮挡-ReID 由移动摄像机捕获，由 200 名被遮挡者的 2000 幅图像组成。每个身份有五个全身人图像和五个被遮挡的人图像，具有不同类型的严重遮挡。

比较了四种方法，分别是香草整体 ReID 方法[41, 35]、具有关键点信息[32, 5]的整体 ReID 方法、部分 ReID 方法[45, 8, 9]和闭塞 ReID 方法[12, 49, 10, 26]。实验结果见表 2。正如我们所看到的，香草整体 ReID 方法和具有关键点信息的整体方法之间没有显著的差距。例如，PCB[34]和 FD-GAN[5]都在闭塞的 Duke 数据集上获得了大约 40% 的 Rank-1 分，这表明简单地使用关键点信息可能不会显著地有利于闭塞的 ReID 任务。对于部分 ReID 和阻塞 ReID 方法，它们都在阻塞数据集上取得了明显的改进。例如，DSR[8]获得 72. 8%，FPR[10]在闭塞的 REID 数据集上获得 78. 3% 的 Rank-1 分数。这表明闭塞和 PAR-

<sup>4</sup><https://github.com/wanguan/HOREID>



方法	部分恢复		部分-我的 LIDS		方法	市场-1501		杜克地铁公司	
	排名第	排名第	排名第	排名第三		排名第	m AP	排名第	m AP
DSR[8]	50.	70.	58.	67.2	PCB[35]	92.3	77.4	81.	66.1
SFR[9]	56.	78.	63.	74.8	VPM[34]	93.0	80.8	83.	72.6
VPM[34]	67.	81.	65.	74.8	机器人[24]	94.1	85.7	86.	76.4
PGFA[26]	68.	80.0	69.	80.9	SPRE ID[14]	92.5	81.3	-	-
法新社[48]	78.	-	-	-	mgcam[30]	83.8	74.3	46.	46.0
FPR[10]	81.	-	68.	-	面具 ReID[27]	90.0	75.3	-	-
[49]	82.	-	-	-	FPR[10]	95.4	86.6	88.	78.4
<i>HORE</i>	<b>85.</b>	<b>91.0</b>	<b>72.6</b>	<b>86.4</b>	PDC[31]	84.2	63.4	-	-
					剂量转移[21]	87.7	68.9	30.	28.2
					PSE[29]	87.7	69.0	27.	30.2
					PGFA[26]	91.2	76.8	82.	65.5
					<b>HORE ID(Ours)</b>	<b>94.2</b>	<b>84.9</b>	<b>86.</b>	<b>75.6</b>

表 3。与两个部分数据集的最先进技术比较，i.e. 部分 REID[45] 和部分 ILIDS[8] 数据集。我们的方法在两个部分数据集上取得最佳性能。

TiAl ReID 任务也有类似的困难，即。学习鉴别特征和特征对齐。最后，我们提出的框架在闭塞 Duke 和闭塞-REID 数据集上达到了最好的性能，在 Rank-1 评分方面分别为 55.1% 和 80.4%，显示了有效性。

部分数据集的结果。伴随着被遮挡的图像，由于检测不完善、摄像机视图异常等原因，往往会出现部分图像。为了进一步评估我们提出的框架，在表 3 中，我们还报告了两个部分数据集，部分 REID[45] 和部分 ILIDS[8] 的结果。部分 REID 包括 60 人的 600 张图像，每人 5 张全身图像和 5 张部分图像，仅用于测试。Partiali LIDS 基于 i LIDS[8] 数据集，共包含机场多个不重叠摄像机捕获的 119 人的 238 幅图像，并手动裁剪其遮挡区域。接下来的 [34, 10, 49]，由于两个部分数据集太小，我们使用 Market1501 作为训练集，两个部分数据集作为测试集。正如我们所看到的，我们提出的框架在两个数据集上的 Rank-1 评分方面明显优于其他方法至少 2.6% 和 4.4。

整体数据集的结果。虽然最近的闭塞/部分 ReID 方法在闭塞/部分数据集上得到了改进，但它们往往无法在整体数据集上获得令人满意的性能。这是由特征学习和对齐过程中的噪声引起的。在这一部分中，我们证明了我们提出的框架也可以在包括 Market-1501 和 DuekMTMC-reID 在内的整体 ReID 数据集上获得满意的性能。市场-1501[42] 包含从 6 个摄像机视点观察到的 1,501 个身份，19,732 个画廊图像和 12,936 个训练图像，所有数据集都包含很少被遮挡或部分人图像。杜克 MTMC-reID[28, 46] 包含 1,404 个身份、16,522 个训练图像、2,228 个查询和 17,661 个画廊图像。

具体地，我们对两个常见的整体 ReID 数据集 Market-1501[42] 和 DukeMTMCre ID[28, 46] 进行了实验，并与 3 种香草 ReID 方法 [35, 34, 24]、3 种人工解析信息的 ReID 方法和 4 种关键点信息 [31, 21, 29, 26] 的整体 ReID 方法进行了比较。实验结果如表 4 所示。正如我们所看到的，3

表 4。与两个整体数据集上的最新技术比较，Market-1501[42] 和 DukeMTMC-reID[28, 46]。我们的方法在整体 ReID 上实现了类似的性能。香草整体 ReID 方法获得了非常有竞争力的性能。例如，BOT[24] 在两个数据集上获得 94.1% 和 86.4% 的 Rank-1 评分。然而，对于使用外部线索的整体 ReID 方法，例如人为解析和关键点信息的性能更差。例如，SPREID[14] 使用人工解析信息，在 Market-1501 数据集上只获得 92.5% 的 Rank-1 分。PFGA[26] 使用关键点信息，在 Duke MTMC-reID 数据集上只获得 82.6% 的 Rank-1 分数。这表明，简单地使用外部线索，如人为解析和关键点，可能不会带来整体 ReID 数据集的改进。这是由于大多数图像整体 ReID 数据集被很好地检测到，香草整体 ReID 方法足够强大，足以学习鉴别特征。最后，我们提出了一种自适应方向图卷积层 (ADGC)，它可以抑制噪声特征，并提出了一种避免硬一对一对齐的交叉图嵌入对齐 (CGEA) 层。在所提出的 ADGC 和 CGEA 层中，我们的框架在两个整体 ReID 数据集上也取得了类似的性能。具体来说，我们在 Market-1501 和 DukeMTMC-reID 数据集上获得了大约 94% 和 87% 的 Rank-1 分数。

### 5.3. 模型分析

拟议模块的分析。在这一部分中，我们分析了我们提出的一阶语义模块 (S)，高-

指	It' 是的	排名第	m AP
1	xxx	49.	39.5
2	/xx	52.	42.8
3	/ /x	53.	43.2
4	/ / /	<b>55.</b>	<b>43.8</b>

表 5。一阶语义模块 (S)、高阶关系模块 (R) 和高阶人类拓扑模块 (T) 的分析。实验结果表明了我们提出的三个模块的有效性。

秩序关系模块 (R) 和高阶人类拓扑模块 (T)。实验结果见表

5. 首先,在索引-1中,我们删除了所有三个模块,将我们的框架降级为 IDE 模型[43],其中只有一个全局特性  $V_g$  可用。其业绩不满足,仅达到 49.9%Rank-1 分。其次,在索引-2中,当使用一阶语义信息时,性能提高了 2.5%,达到 52.4%的 Rank-1 评分。这表明,关键点的语义信息对于学习和对齐特征是有用的。第三,在指数-3中,添加了额外的高阶关系信息,性能进一步提高了 1.5%,达到 53.9%。这证明了我们模块 R 的有效性。最后,在索引-4中,我们的完整框架在 55.1%的 Rank-1 评分下达到了最佳的精度,显示了我们模块 T 的有效性。

建议的层次分析。在这一部分中,我们进一步分析了语义模块(S)、关系模块(R)和拓扑模块(T)的关键组成部分——关键点保密(NORM)、自适应方向图卷积(ADGC)层和交叉图嵌入对齐(CGEA)层的归一化)。具体来说,在删除 NORM 时,直接使用原始的置信度分数。当移除 ADGC 时,在  $E_q$  中。我们替换  $A^{adj}$  用一个固定的邻接矩阵连接起来,就像人类拓扑一样。因此,关系模块(S)降解为香草 GCN,不能抑制噪声信息。当移除 CGEA 时,在  $E_q$  中。我们用一个完全连接的矩阵来代替  $U$ 。即图 1 的每个节点都连接到图 2 的所有节点。然后,拓扑模块(T)不包含用于特征对齐的高阶人类拓扑信息,并降级为香草验证模块。实验结果见表 6。正如我们所看到的,当删除 NORM、ADGC 或 CGEA 时,性能显著下降 2.6%、1.4%和 0.7%的秩-1 分数。实验结果表明,我们提出的 NORM、ADGC 和 CGEA 组分的有效性。

参数分析。我们评估了方程中参数的影响。(13),即。你和 N。结果如图 5 所示,最优设置为  $y=0.5, n=8$ 。在分析一个参数时,另一个参数固定在 OP-

规范 ADGCCGEA			排名第	m AP
x	/	/	52.5	40.4
/	x	/	53.7	42.2
/	/	x	54.4	43.5
/	/	/	<b>55.1</b>	<b>43.8</b>

表 6. 关键点保密(NORM)、自适应方向图卷积层(ADGC)和交叉 GRPA H 嵌入对准(CGEA)层归一化分析。实验结果表明了我们提出的层的有效性。

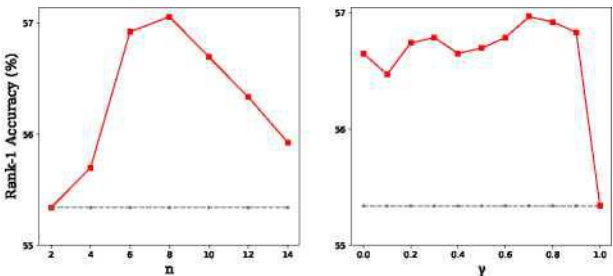


图 5. 方程中参数  $n$  和  $y$  的分析。(13). 最优值为  $y=0.5$  和  $n=8$ 。当分析其中一个时,另一个固定为其最优值。实验结果表明,该模型对不同参数具有较强的鲁棒性。

时间值。很明显,当使用不同的  $y$  和  $n$  时,我们的模型稳定地优于基线模型。实验结果表明,我们提出的框架对不同的权重是鲁棒的。请注意,这里的表现与表 2 不同,前者达到 57%,后者达到 55%。这是因为后者是使用平均 10 次计算的,以便进行公平比较。

## 6. 结论

在本文中,我们提出了一个新的框架来学习高阶关系信息的判别特征和拓扑信息的鲁棒对齐。对于学习关系信息,我们将图像的局部特征表示为图的节点,并提出了一种自适应方向图卷积层(ADGC),以促进语义特征的消息传递,抑制无意义和噪声特征的消息传递。对于学习拓扑信息,我们提出了一种具有验证损耗的交叉图嵌入对准(CGEA)层,它可以避免敏感的硬一对一对准,并执行健壮的软对准。最后,在闭塞、部分和整体数据集上的广泛实验证明了我们提出的框架的有效性。

## 承认

本研究得到了国家重点研发计划的支持。2017yfa0700800)。

## 参考资料

- [1] Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vincius Flores Zambaldi 和 Mateusz Malinowski. 关系归纳偏差、深度学习和图形网络。ar Xiv 预印 arXiv: 1806.01261, 2018 年。
- [2] 曹哲、海达尔戈、托马斯·西蒙、魏世恩和亚瑟·谢赫。开放姿态: 使用部分亲和场实时多人 2D 姿态估计。在 ARXIV 预印 ARXIV: 1812.08008, 2018 年。
- [3] 邢凡, 郝洛, 宣章, 凌霄和, 迟章和魏江。Scpnet: 空间通道并行网络, 用于联合整体和部分人的重新识别。亚洲计算机展望会议, 第 19-34 页。斯普林格, 2018 年。
- [4] 杨富, 魏云超, 周玉谦, 石洪辉, 高黄, 王新超, 姚志强, 黄托马斯。水平金字塔匹配的人重新识别。在第一^人工智能会议记录, 第 33 卷, 第 8295-8302 页, 2019 年。
- [5] 葛义晓, 李卓文, 赵海宇, 尹国军, 帅毅, 王晓刚, 等。FD-gan: POSE 引导的特征提取 GaN, 用于稳健的人重新识别。在神经信息处理系统的进展, 第 1222-1233 页, 2018 年。
- [6] 邵岗公, 马可克里斯坦尼, 水城燕, 陈变洛。人重新识别。2014.
- [7] 何开明, 张向宇, 任少卿, 孙健。深度残差学习进行图像识别。在 2016 年 IEEE 计算机视觉和模式识别会议 (CVPR), 第 770-778 页, 2016 年。
- [8] 凌霄河, 建良, 海青李, 振安孙。部分人重新识别的深度空间特征重建: 无对准方法。第 7073-7082 页, 2018 年。
- [9] 凌霄河, 孙振安, 朱玉浩, 王云波。识别部分生物特征。AR XIV 预印 AR XIV: 1810.07399, 2018 年。
- [10] 何凌霄, 王应刚, 吴柳, 廖星宇, 何昭, 孙振安, 冯嘉世。

- 前景感知金字塔重建,用于无对齐遮挡人的再识别。阿希夫: *计算机视觉和模式识别*, 2019 年。
- [11] 亚历山大·赫曼斯, 卢卡斯·比尔和巴斯蒂安·莱贝。在防御三重态损失时对人进行重新识别。阿希夫预印阿希夫: 1703.07737, 2017 年。
- [12] 黄厚静, 李党伟, 张章, 陈晓棠, 黄凯琪。对抗性闭塞样本用于人的再识别。在 IEEE 计算机视觉和模式识别会议记录中, 第 5098-5107 页, 2018 年。
- [13] Sergey Ioffe 和 Christian Szegedy。批量归一化: 通过减少内部协变量移位来加速深度网络训练。*机器学习国际会议*, 第 448-456 页, 2015 年。
- [14] Mahdi M Kalayeh, Emrah Basaran, Muhittin Gokmen, Mustafa E Kamasak 和 Mubarak Shah。人的语义解析用于人的再识别。在 IEEE 计算机视觉和模式识别会议记录中, 第 1062-1071 页, 2018 年。
- [15] 托马斯·N·基普夫和麦克斯·韦林。图卷积网络的半监督分类。阿希夫预印阿希夫: 1609.02907, 2016 年。
- [16] Martin Koestinger, Martin Hirzer, Paul Wohlhart, Peter M Roth 和 Horst Bischof。从等价约束中学习大规模度量。在 2012 年 IEEE 计算机视觉和模式识别会议, 第 2288-2295 页。IEEE, 2012 年。
- [17] 李杰峰, 王灿, 郝柱, 毛一环, 方浩树, 吕彩武。拥挤: 有效的拥挤场景姿态估计和一个新的基准。在 IEEE 计算机视觉和模式识别会议记录中, 第 10863-10872 页, 2018 年。
- [18] 廖圣才、杨虎、朱湘玉、李斯坦。通过局部最大发生表示和度量学习重新识别人。在 IEEE 计算机视觉和模式识别会议记录中, 第 2197-2206 页, 2015 年。
- [19] 廖圣才和斯坦·兹利。有效的 PSD 约束非对称度量学习的人重新识别。在 IEEE 国际计算机视觉会议记录, 第 3685-3693 页, 2015 年。
- [20] 林宗义, 迈克尔·迈尔, 塞尔日·布利基, 詹姆斯·海斯, 皮埃特罗·佩罗纳, 德瓦·拉曼, 皮奥特·多尔德尔和 C·劳伦斯·齐特尼克。Microsoft coco: 上下文中的通用对象。欧洲计算机视觉会议, 第 740-755 页。斯普林格, 2014 年。
- [21] 刘金贤, 倪冰冰, 严义超, 彭州, 朔成, 胡建国。可转移的人重新识别。在 2018 年 IEEE/CVF 计算机视觉和模式识别会议上, 第 4099-4108 页, 2018 年。
- [22] 严璐, 岳武, 刘斌, 张天柱, 李宝浦, 齐楚, 余恩海。跨模态人员重新识别与共享特定特征转移, 2020 年。
- [23] 郝洛, 邢凡, 迟章, 魏江。深度卷积网络: 具有成对空间变压器网络的深度卷积网络, 用于部分人的重新识别。AR XIV 预印 AR XIV: 1903.07072, 2019 年。
- [24] 郝洛, 顾友志, 廖星宇, 赖沈琪, 魏江。一袋的技巧和一个强大的基线深入重新识别。在 IEEE 计算机视觉和模式识别研讨会会议记录, 第 0-0 页, 2019 年。
- [25] 马炳鹏, 余素, 弗雷德里克·朱莉。基于生物启发特征的协方差描述符, 用于人的重新识别和人脸验证。图像和视觉计算, 32 (6-7): 379-390, 2014 年。
- [26] 苗家旭, 余武, 刘平, 丁宇航, 易阳。姿势引导特征对齐用于闭塞者的再识别。在 ICCV, 2019 年。
- [27] 雷奇, 景火, 雷王, 应欢石, 杨高。面具: 一种基于面具的深度排序神经网络, 用于人的重新识别。阿希夫预印阿希夫: 1804.03864, 2018 年。
- [28] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara 和 Carlo Tomasi。性能度量和数据集

- 多目标,多摄像机跟踪。欧洲计算机视觉会议,第17-35页。斯普林格,2016年。
- [29] M Saquib Sarfraz, Arne Schumann, Andreas Eberle 和 Rainer Stiefelhagen。一种位姿敏感的嵌入,用于人的重新识别和扩展的交叉邻域重新排序。在 IEEE 计算机视觉和模式识别会议记录,第420-429页,2018年。
- [30] 宋春风,黄燕,欧阳万里,王梁。面具引导的人再识别对比注意模型。在 IEEE 计算机视觉和模式识别会议记录中,第11791188页,2018年。
- [31] 迟苏,嘉宁,张世良,邢俊良,文高,齐田。姿态驱动的深卷积模型用于人的再识别。在2017年 IEEE 国际计算机视觉会议 (ICCV),第3980-3989页,2017年。
- [32] 苏玉民、王景东、唐思玉、陶梅、李景明。部分对齐的人重新识别的双线性表示。欧洲计算机视觉会议,第418-437页,2018年。
- [33] 克孙,肖斌,刘东,王敬东。人类姿态估计的深度高分辨率表示学习。在 CVPR,2019年。
- [34] 孙一凡,秦旭,李亚丽,张驰,李一康,王胜金,孙健。感知要聚焦的地方:学习可见性-感知部分人重新识别的部分级别特征。第393-402页,2019年。
- [35] 孙一凡,梁征,易阳,齐天,王胜金。超越零件模型:人检索与精化零件池(和一个强大的卷积基线)。在欧洲计算机视觉会议记录 (ECCV),第480-496页,2018年。
- [36] 关安王,天柱张,建成,司柳,杨洋,增光侯。通过联合像素和特征对齐,对 RGB-红外交叉模型人进行再识别。在 IEEE 国际计算机视觉会议 (ICCV),2019年10月。
- [37] 王冠安,张天柱,杨阳,建成,张建龙,徐亮,增光侯。交叉模态配对图像生成的 RGB-红外人重新识别。在 AAAI-20 人工智能会议,2020年。
- [38] 王润忠,颜俊奇,杨晓康。学习组合嵌入网络进行深图匹配。arXiv preprint arXiv:1904.00597,2019年。
- [39] 杨洋,杨继美,严俊杰,廖胜才,董毅,斯坦兹利。高级颜色名称,用于人员重新识别。欧洲计算机视觉会议,第536-551页。斯普林格,2014年。
- [40] 安德烈·赞菲尔和克里斯蒂安·斯明奇塞斯库。图匹配的深度学习。在2018年 IEEE/CVF 计算机视觉和模式识别会议上,第2684-2693页,2018年。
- [41] 赵黎明,李熙,庄月亭,王敬东。深度学习的部分对齐表示人的重新识别。在2017年 IEEE 国际计算机视觉会议 (ICCV),第3239-3248页,2017年。
- [42] 梁征,李岳申,陆田,生金王,京东王,齐田。可缩放的人重新识别:一个基准。在 IEEE 计算机视觉国际会议记录,第1116-1124页,2015年。
- [43] 梁征,易阳,亚历山大·高普特曼。人的重新识别:过去、现在和未来。arXiv preprint arXiv:1610.02984,2016年。
- [44] 郑伟石、龚绍刚、陶翔。通过相对距离比较重新识别。IEEE 模式分析和机器智能事务,35(3):653-668,2013年。
- [45] 郑伟石、李祥、陶祥、廖胜才、赖建煌、龚绍刚。部分人重新识别。在2015年 IEEE 计算机视觉国际会议 (ICCV),第4678-4686页,2015年。
- [46] 浙东正,梁正,易阳。由 GaN 产生的未标记样本改善了人在体外重新识别基线。ARXIV预印ARXIV:1701.07717,2017年。
- [47] 淮忠,梁正,康国良,李少子,易阳。随机擦除数据增强。阿西夫预印阿西夫:1708.04896,2017年。
- [48] 稼轩卓,陈泽宇,赖建煌,王光聪。闭塞人员重新识别。在2018年 IEEE 多媒体和博览会国际会议 (ICME),第1-6页。IEEE,2018年。
- [49] 稼轩卓,建煌赖,培嘉陈。一种新颖的遮挡人再识别的师生学习框架。arXiv preprint arXiv:1907.03253,2019年。