DOI: 10.13382/j.jemi.B2407721

多粒度共享-解离相关网络支持下的跨模态行人重识别算法*

宋婉茹[1] 郝川艳[1] 郑洁莹[2] 刘 峰[1,2]
(1.南京邮电大学教育科学与技术学院 南京 210046;2.南京邮电大学通信与信息工程学院 南京 210023)

摘 要:随着智能安防系统的不断升级,面向全天候监控实现行人检索成为了相关领域热点之一,可见光-红外跨模态行人重识别的研究应运而生。该研究面临的主要挑战是同一行人在不同模态的图像间展现出巨大的差异。现有方法通过探索不同模态之间共享信息,来减少同一行人在两种模态下的特征差异。为了进一步提升跨模态行人重识别的准确率,提出了一种多粒度共享-解离相关网络,通过共享-解离模块的嵌入,对主干网络中参数共享分支进行复制和分解,打破了原有基准模型在多粒度特征提取上的局限;通过多粒度相关特征学习模块的设计,充分挖掘了行人跨模态不变的身体结构关联信息,优化了全局-局部特征的对齐方案;通过多层次的损失函数构建,为模型的训练提供了有效的监督,提升了模型的判别力和鲁棒性。该算法在公开数据集 SYSU-MM01 和 RegDB 上均获得优秀的性能,其中,SYSU-MM01 全搜索模式下 Rank-1 和平均精度均值(mAP)分别达到 74.70%和 71.79%;在 RegDB 的两种检索模式下,Rank-1 和 mAP 均高于 90%,准确率优于多种先进方法。实验显示该网络在跨模态特征对齐和复杂场景适应性方面具有一定优势。

关键词:行人重识别;可见光-红外;共享-解离;多粒度;特征学习;关系网络
中图分类号:TP 391.41;TN911.7 **文献标识码**:A **国家标准学科分类代码**:510.40

Multi-granularity shared-disentangling relation network for cross-modality person re-identification

Song Wanru[1] Hao Chuanyan[1] Zheng Jieying[2] Liu Feng[1,2]
(1. School of Educational Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing 210046, China; 2. School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: With the continuous development of intelligent security systems, pedestrian retrieval for all-day surveillance has become one of the research hotspots. Thus, the research of visible-infrared cross-modality person re-identification has emerged. The main challenge faced in this task is the huge discrepancy between visible and infrared images of the same pedestrian. Existing methods focus on exploring the shared information and reducing the feature variances of the same pedestrian in the two modalities. To further improve the accuracy of the task, this paper proposes multi-granularity shared-disentangling relation network for re-identification. By embedding the shared-disentangling module, the parameter-sharing branch of the backbone is replicated and decomposed, thus breaking limitations of the original benchmark model in multi-granularity feature extraction. By designing the multi-granularity relation feature learning module, the modality-invariant correlation information of the pedestrian body is fully explored, enhancing the learning of the shared features. And through constructing a loss function in multiple levels, effective supervision is available for the training of the model, and the global-local feature alignment scheme is optimized. The proposed algorithm obtains superior performance on both public datasets named SYSU-MM01 and RegDB. The Rank-1 and mAP in All-search mode on the SYSU-MM01 dataset can reach 74.70% and 71.79% respectively. In both retrieval modes of RegDB, Rank-1 and mAP are higher than 90%, and the accuracy is superior to many state-of-the-art methods. Experiments demonstrate the advantages of this network in cross-modality feature alignment and complex scene adaptation.

Keywords: person re-identification; visible-infrared; shared-disentangling; multi-granularity; feature learning; relation network

收稿日期:2024-07-07 Received Date: 2024-07-07
*基金项目:国家自然科学基金(62307025)、江苏省高校自然科学基金面上项目(22KJB520025)资助

0 引 言

视频监控的快速普及促进着视频、图像识别与分析技术的蓬勃发展。其中,行人重识别(person re-identification, Re-ID)作为一项关键目标于 2006 年被提了出来,近些年来受到了学术界的广泛关注。Re-ID 旨在视域不重叠的摄像头下识别同一行人,实现行人跨镜头追踪。传统 Re-ID 聚焦在可见光域的行人检索,即 RGB 图像之间的检索问题。然而刑事案件、交通事故等情形常高发于夜晚等一些光线不足的环境,这时可见光监控摄像机几乎无法发挥作用。幸运的是现在大多数监控摄像机都可以在夜间捕获红外(IR)图像,实现 24 h 监控。因此,研究可见光域与红外域间的行人跨模态检索是实现智能安防的重要一环,对公共安全、刑事侦查等方面都有着非常重要的现实意义,可见光-红外跨模态行人重识别(visible-infrared person re-identification, VI Re-ID)应运而生[1-2]。

行人重识别研究面临着视角与姿态变化、复杂背景以及遮挡等问题。然而,困扰跨模态任务的不仅仅是模态内差异。在 VI Re-ID 中,由于成像原理不同,IR 图像相较于 RGB 图像缺少着重要的颜色信息,两个模态之间存在着巨大差异,这种差异被称之为跨模态差异。综上所述,在 VI Re-ID 的研究中除了解决单模态研究所遇到的问题外,探索两域之间的模态不变的共享信息,减少同一行人在两种模态下的特征差异,构建两者之间联系的桥梁,是该研究所面临的主要挑战。现有的 VI Re-ID 研究可以划分为 3 大类,分别是基于表征学习的研究、基于距离度量损失的研究以及基于中间模态辅助的研究,其目的都在于减小同一行人在不同模态下的差异并增大不同行人之间的距离。

1)基于表征学习的研究。通过共享特征学习将不同模态的特征映射到相同的特征空间,从而减小同一行人不同模态情形下的特征差异。该任务最常用的网络框架是双流网络模型[3-7]。其中,文献[6]提出了第一个大规模跨模态行人重识别数据集 SYSU-MM01,为其后的研究奠定了数据层面的基础。双流网络模型通常是将 RGB 图像和 IR 图像输入到单独的分支,分别学习行人在每个模态内特有的特征,再通过特征嵌入将不同模态的特征映射到同一特征空间。文献[7]首次将生成对抗网络(generative adversarial network, GAN)应用到 VI Re-ID 中,提出跨模态生成对抗网络(cross-modality generative adversarial network, cmGAN),先利用生成器学习不同模态的特有特征,再采用判别器判断特征属于哪个类别,通过训练这两个目标不同的网络,实现跨模态共享特征的学习。文献[8-9]提出了双流参数共享网络,探索了双流网络在以 ResNet50 为主干网络的跨模态任务中需要共享多少参数以实现最好的识别结果。通过双流 ResNet50 部分阶段的参数共享,将可见光域和红外域图像映射到共享空间,为后续的研究提供了基准。文献[10]通过全局-局部特征提取模块的设计,学习更具判别力的行人特征表达。事实上,在 VI Re-ID 任务中,全局-局部特征学习一直是一个关键研究点[5,11-13],且近些年来多粒度特征表达构建研究也受到广泛关注[14-15]。但现有的算法多为单模态算法的扩展,在跨模态任务上的表现受限于双流网络模型对共享特征的学习能力。

2)基于度量损失函数的研究。在跨模态行人重识别任务中,损失函数不断拉近模态内与交叉模态下的同一行人特征距离[3-4,16-19],使得同一身份行人可以分类到同一类别中。文献[4]提出了异质中心损失,首先计算不同每个类别下不同模态的中心向量,再通过减小同一类别不同中心向量之间的距离来减小模态差异,提升 VI Re-ID 的准确率。文献[8]提出了异质中心三元组损失,将传统三元组损失中点与其他所有样本之间距离的比较替换为锚点与所有中心向量之间距离的比较。除此之外,文献[18]将困难三元组损失引入到跨模态任务中,同时考虑了全局、模态内和模态间的三元组损失。上述这些针对跨模态任务的度量学习损失函数都是在中心损失的基础上设计的,并在公开数据集上展现了优秀的性能。

3)基于中间模态生成的研究。GAN 在跨模态行人重识别中常用来实现可见光图像和红外图像的相互转换[20-22]。然而多数基于 GAN 的方法准确率并不高。这是因为跨模态的行人图像并不是成对出现的,图像的转换可能存在不同的合成结果,网络很难判断所生成的伪样本是否为正确的目标图像。最为关键的是基于 GAN 的模态转换可能会破坏原有的行人结构。因而更多的研究倾向于采用轻量级的网络或是灰度变换、通道选择等方式来对可见光模态图像进行处理。基于上述方案的辅助模态生成策略通常能够有效地提升 VI Re-ID 的准确率[14,23-27],缺点在于多数算法会在一定程度上增大模型计算量。

综上所述,VI Re-ID 研究多是采用双流网络或是双流参数共享网络[8]作为基准网络来提取特征,在异质图像共享信息的挖掘方面具有一定的局限性;在全局-局部特征学习方面多采用的是基于部件的卷积基准方法(part-based convolutional baseline, PCB)[28],针对多粒度特征构建,倾向于采用多尺度划分的方案,它们都忽略了多粒度信息和行人身体结构间的相关信息在 VI Re-ID 任务中发挥的重要作用;引入中间模态通常会增大计算量。受文献[29-30]对人体部件间相关关系探索的启发,本文面向可见光-红外跨模态行人重识别任务提出了一

种改进的基准网络即多粒度共享-解离相关网络,在不额外增加计算量的基础上实现不同模态下的行人特征对齐。

本文在双流参数共享网络框架的基础上嵌入了共享-解离模块,通过“特有分支-共享分支-解离分支”的架构实现对输入图像的特征学习;设计了一种多粒度相关特征学习模块,利用行人身体部件间的相关性对于模态间和模态内变化的鲁棒性,在粗细粒度层面充分挖掘了行人多层次模态不变信息;设计了多粒度全局-局部损失计算模块,通过联合异质三元组中心损失和 ID 分类损失进一步拉近模态内和模态间同一行人特征间的距离;在现有的大规模公开数据集 SYSU-MM01[6] 和 RegDB[31] 进行了大量的消融实验和对比试验,验证了本文方法的有效性和优越性。

1　基于多粒度共享-解离相关网络的跨模态行人重识别方法

1.1　网络结构

本文提出的网络整体框架如图 1 所示,采用在 ImageNet 数据集上预训练的 ResNet50 作为主干网络,输入为可见光图像集合 $V = \{I_{vis}^i, y_{vis}^i\}_{i=1}^N$ 和红外图像集合 $I = \{I_{inf}^i, y_{inf}^i\}_{i=1}^N$,其中 I_{vis}^i 和 I_{inf}^i 分别表示可见光和红外图像,y_{vis}^i 和 y_{inf}^i 分别表示图像中行人的 ID 标签,N 表示的行人 ID 的总数。从图 1 可以看出,本文模型包含了两个核心模块,分别是共享-解离模块和多粒度相关特征学习模块。

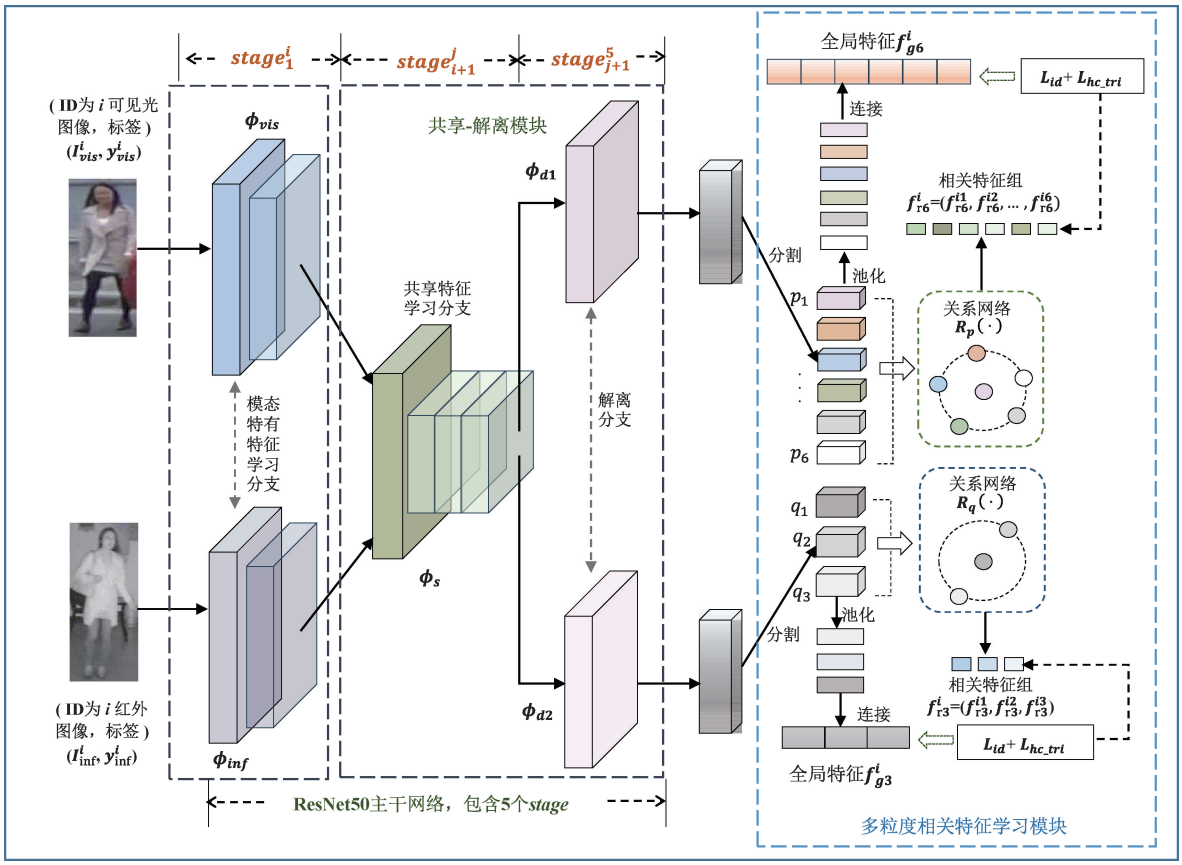

图 1　本文方法总体框图

Fig. 1　Overall framework of the proposed method

双流 ResNet50 通过参数共享的全连接层将模态特有特征映射到共享空间,这种方法忽略了行人身体结构的空间特征。研究表明在卷积层进行参数共享可以获得更好的识别结果[8]。本文所提出的方法为了提升双流参数共享网络在多粒度特征学习方面的有效性,在卷积层参数共享后又进行了解离,即在网络中嵌入共享-解离模块。主干网络 ResNet50 模型包含 5 个卷积模块,被标识为 5 个阶段(stage),分别是 $stage_1$、$stage_2$、$stage_3$、$stage_4$ 和 $stage_5$。通过嵌入共享-解离模块,将双流特征学习模型划分为:1)特有特征学习分支 $stage_1$ ~ $stage_i$,用于捕捉模态特有信息;2)共享特征学习分支 $stage_{i+1}$ ~ $stage_j$,用于挖掘跨模态不变信息;3)解离分支 $stage_{j+1}$ ~

$stage_s$,用于提取更具有判别力的多粒度共享信息。

多粒度相关特征学习模块替代基于 PCB 的局部特征学习模块,囊括了从 3 等分到 6 等分再到全局的特征表征方式,将多模态数据的多粒度信息深层次融入到共享特征学习中,构建更具有表征性的全局-局部行人特征表达模型。为了更好地捕捉身体各个部件与其他部件之间的联系,提升模型判别力和鲁棒性,本文方法在特征图切分后采用了两个关系网络(relation network)[31]去计算相关局部特征,替代原始只经过切分和池化操作所获得的局部特征。图 1 中采用 $R_p(\cdot)$ 和 $R_q(\cdot)$ 描述上述两个关系网络。此外,在跨模态行人重识别模型的训练过程中构建了多粒度损失函数,优化模型的训练过程,使得训练好的模型尽可能提取到跨模态不变的行人表征。

1.2 共享-解离模块

双流网络分为特征学习和特征嵌入两部分,前者采用两个离散分支分别从红外和可见光行人图像中提取中浅层的模态特有特征;后者通过共享主干网络部分卷积层或全连接层,将异质特征映射到同一特征空间,挖掘出深层模态共享信息。事实上,从卷积层开始进行参数共享,考虑了行人的身体结构信息,进而可以获得更多行人跨模态不变的中级特征和空间信息。为了获得更高判别力的多粒度信息,共享-解离模块被引入到了双流参数共享的 ResNet50 网络中。该模块将共享后的卷积层再一次进行解离。解离操作的实质是在阶段共享后又进行了一次复制操作,复制后的网络不再共享参数,被用来学习两种精细度不同的局部信息。不同于特有特征学习分支,解离后的网络虽然不再共享参数,但仍是在两个不同的共享空间中学习模态不的特征。研究表明,在 ResNet50 的不同 *stage* 进行参数共享可产生不同的效果[8],因此本文提出若在不同阶段进行共享-解离模块的嵌入,也可以获得不同的识别效果。共享-解离模块嵌入方案如表 1 所示。

表 1 共享-解离模块嵌入方案

Table 1 Embedding Schemes of the shared-disentangling module

方案	特有分支 ϕ_{vis} 和 ϕ_{inf}	共享分支 ϕ_s	解离分支 ϕ_{d1} 和 ϕ_{d2}
1	$stage_1$	$stage_2$	$stage_{3-5}$
2	$stage_1$	$stage_{2-3}$	$stage_{4-5}$
3	$stage_1$	$stage_{2-4}$	$stage_5$
4	$stage_{1-2}$	$stage_3$	$stage_{4-5}$
5	$stage_{1-2}$	$stage_{3-4}$	$stage_5$
6	$stage_{1-3}$	$stage_4$	$stage_5$

表 1 中,ϕ_{vis} 和 ϕ_{inf} 分别表示的是可见光图像和红外图像特有特征学习分支,ϕ_s 表示的是跨模态共享特征学习分支,而 ϕ_{d1} 和 ϕ_{d2} 表征的是经过解离后的参数不共享分支。对于可见光输入图像 I_{vis}^i 来说,经过有共享-剥离模块嵌入后的网络后可以学习到两个 3D 行人特征 $\boldsymbol{f}_{v1}^i$ 和 $\boldsymbol{f}_{v2}^i$,定义如下:

$$\begin{cases}\boldsymbol{f}_{v1}^i = \phi_{d1}[\phi_s(\phi_{vis}(I_{vis}^i))] \\ \boldsymbol{f}_{v2}^i = \phi_{d2}[\phi_s(\phi_{vis}(I_{vis}^i))]\end{cases} \tag{1}$$

同理可得,输入的红外图像经网络处理后得到特征 $\boldsymbol{f}_{t1}^i$ 和 $\boldsymbol{f}_{t2}^i$,定义如下:

$$\begin{cases}\boldsymbol{f}_{t1}^i = \phi_{d1}[\phi_s(\phi_{inf}(I_{inf}^i))] \\ \boldsymbol{f}_{t2}^i = \phi_{d2}[\phi_s(\phi_{inf}(I_{inf}^i))]\end{cases} \tag{2}$$

通过在主干网络中嵌入共享-解离模块,整个网络被划分为 3 个部分,只有在跨模态共享特征学习分支中所有网络参数共享。

1.3 多粒度相关特征学习模块

通过对 ResNet50 残差块的解离,对多粒度特征的学习不仅是对输出特征图划分层次的不同,还存在一定的特征差异性。模块将 3D 特征 $\boldsymbol{f}_{v1}^i$、$\boldsymbol{f}_{v2}^i$、$\boldsymbol{f}_{t1}^i$ 和 $\boldsymbol{f}_{t2}^i$ 进行纵向的等比切分。为了获得多粒度的特征,第 1 个解离分支和第 2 个解离分支输出的特征划分层次不同,即采用 3 等分切分和 6 等分切分两种方案。不同的切分程度对细节的表征程度不同。数据集中的行人图片包含着背景信息,因此采用不超过 6 等分的纵向划分方式更符合人体的身体结构分布,且可以避免某些局部特征块中干扰信息集中的情况出现。针对输入图片,本模块的输出为两个全局特征和两组局部特征。

1)在局部特征学习方面,本文受关系网络中的 One-vs. -rest 模块启发,将人体局部块间的相关联系嵌入到局部特征的学习过程中。如图 2 所示,经过关系网络 $R_p(\cdot)$ 和 $R_q(\cdot)$ 的处理,对于划分大小不同的局部块,可以获取到两个不同粒度的相关局部特征组 $(\boldsymbol{f}_{r6}^{i1}, \boldsymbol{f}_{r6}^{i2}, \cdots, \boldsymbol{f}_{r6}^{i6})$ 和 $(\boldsymbol{f}_{r3}^{i1}, \boldsymbol{f}_{r3}^{i2}, \boldsymbol{f}_{r3}^{i3})$。

图 2(a)选择 6 等分分支进行分析,对于局部特征图 $\boldsymbol{p}_1$,目标是经过一个由 1×1 卷积、批量归一化(batch normalization, BN)以及激活函数 ReLU 共同构成的关系网络 $R_p(\cdot)$ 后,可以获得局部相关特征 $\boldsymbol{f}_{r6}^{i1}$,其定义如下:

$$\boldsymbol{f}_{r6}^{i1} = \overline{\boldsymbol{p}_1} + R_p(concat(\overline{\boldsymbol{p}_1}, \overline{\boldsymbol{r}_{p1}})) \tag{3}$$

式中:$\overline{\boldsymbol{p}_1}$ 是将特征 $\boldsymbol{p}_1$ 经过 1×1 卷积后得到的 1×1×512 维特征。$\boldsymbol{r}_{p1}$ 定义如下:

$$\boldsymbol{r}_{p1} = \frac{1}{5}\sum_{i=2}^{6}\boldsymbol{p}_i \tag{4}$$

对 $\boldsymbol{r}_{p1}$ 做相似处理,可以得到 1×1×512 维特征 $\overline{\boldsymbol{r}_{p1}}$;$concat(\cdot)$ 表示特征间的串联。相似地,$\boldsymbol{f}_{r6}^{i2}, \boldsymbol{f}_{r6}^{i3}, \cdots, \boldsymbol{f}_{r6}^{i6}$ 也通过上述方式学习到。此外,由图 2(b)可以观察到,$R_q(\cdot)$ 具有和 $R_p(\cdot)$ 相同的网络结构。通过挖掘 3 等分支局部特征间的相互关系学习到相关特征 $\boldsymbol{f}_{r3}^{i1}$、$\boldsymbol{f}_{r3}^{i2}$ 和 $\boldsymbol{f}_{r3}^{i3}$。

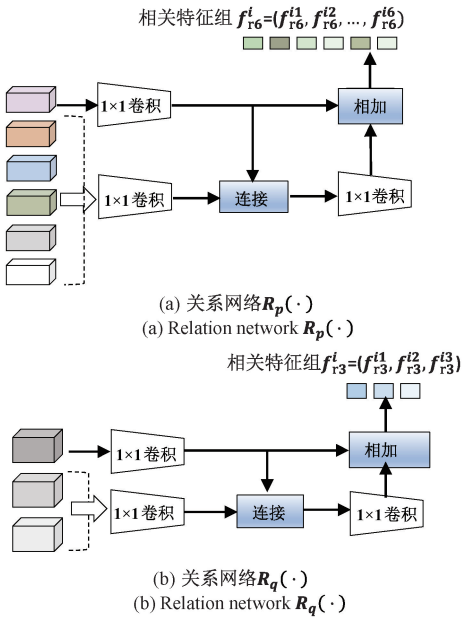

图 2　关系网络 $R_p(\cdot)$ 以及关系网络 $R_q(\cdot)$ 的网络结构

Fig. 2　The relation networks $R_p(\cdot)$ and relation network $R_q(\cdot)$, respectively

模块提出局部特征的学习不仅仅是依赖于当前的局部块,还可以通过挖掘它与其余局部块间的相互关系,构建出一种跨模态不变的局部相关表征。该策略避免了不同行人因局部相似所造成的识别错误问题,提升了局部特征在跨模态情境下的判别力和鲁棒性。鉴于在解离分支后网络参数不再共享,且模块在解离分支后采用了两种不同大小的切分方案,因此所获得的细粒度共享特征从不同层次表征着行人细节信息。

2)在全局特征获取方面,对于划分后的特征图采用广义平均池化(generalized mean pooling, GeM)方案[32],再通过串联操作对特征图拼接获得两个不同的全局特征 $\boldsymbol{f}_{g3}^{i}$ 和 $\boldsymbol{f}_{g6}^{i}$。

1.4　整体损失

模型选择 ID 分类损失和异质中心三元组损失来监督整个训练的过程。经过多粒度相关特征学习模块的处理后,模型可以学习到多种行人特征,它们表征着行人不同层次、不同方面的信息。因此,在损失函数的设计上,需要平衡细粒度损失与粗粒度损失之间的关系。

行人重识别可以视为一个分类任务,ID 分类损失主要用于监督经 FC 层预测的身份信息,实现更加准确有效的分类:

$$Loss_{id}=-\frac{1}{N}\sum_{i=1}^{N}q_i\log\left(\frac{e^{x_i}}{\sum_{j}^{N}e^{x_j}}\right)$$
$$q_i=\begin{cases}N-(N-1)\xi, & y=i\\ \xi, & y\neq i\end{cases}\tag{5}$$

式中:y 表示可见光或者红外图像中行人的真实身份标签;$\frac{e^{x_i}}{\sum_{j}^{N}e^{x_j}}$ 实现将原始输出 x_i 转换为概率分布,表示模型预测的第 i 个类别的概率;N 为一个训练 batch 中的总 ID 个数;ξ 为一个常数,表示分配给错误类别的概率,用于阻碍网络模型对数据集的完全依赖,在该类实验中常被设置为 0.1。针对不同的全局特征和局部特征,可以获得不同 ID 分类损失 L_{id}^{g1}、L_{id}^{g2}、$L_{id}^{l6i}(i=1,2,\cdots,6)$ 和 $L_{id}^{l3i}(i=1,2,3)$,因此用于任务的 ID 分类损失 L_{id} 可以表示为:

$$L_{id}=\frac{1}{4}\left(L_{id}^{g1}+L_{id}^{g2}+\frac{1}{6}\sum_{i=1}^{6}L_{id}^{l6i}+\frac{1}{3}\sum_{i=1}^{3}L_{id}^{l3i}\right)\tag{6}$$

式中:L_{id}^{g1} 和 L_{id}^{g2} 是全局特征产生的分类损失;L_{id}^{l6i} 和 L_{id}^{l3i} 分别是 6 等分和 3 等分后每个局部特征产生的分类损失。仅仅依赖于 ID 分类损失并不能有效地解决跨模态场景所带来的困难和挑战。本文方法将异质中心三元组损失与 ID 分类损失联合起来,引入同类模态和异类模态的中心来限制特征的分布。对于可见光模态和红外模态,首先计算每个模态下的特征中心为:

$$\boldsymbol{c}_i^{vis}=\frac{1}{N_i^{vis}}\sum_{j=1}^{N_i^{vis}}\boldsymbol{f}_{\boldsymbol{vis},i,j}\tag{7}$$

$$\boldsymbol{c}_i^{inf}=\frac{1}{N_i^{inf}}\sum_{j=1}^{N_i^{inf}}\boldsymbol{f}_{\boldsymbol{inf},i,j}\tag{8}$$

式中:i 表示身份 ID 的类别;$\boldsymbol{c}_i^{vis}$ 和 $\boldsymbol{c}_i^{inf}$ 分别表示可见光和红外模态下的特征中心;$\boldsymbol{f}_{\boldsymbol{vis},i,j}$ 示可见光模态下第 i 类样本中的第 j 个特征;$\boldsymbol{f}_{\boldsymbol{inf},i,j}$ 类似;N_i^{vis} 和 N_i^{inf} 分别表示可见光和红外模态下第 i 类的样本数量。异质中心三元组表示如下:

$$L_{hc\_tri}^{i}=\sum_{j=1}^{P}\left[\rho+\|\boldsymbol{c}_i^{vis}-\boldsymbol{c}_i^{inf}\|_2-\min_{\substack{n\in[vis,\inf]\\ j\neq i}}\|\boldsymbol{c}_i^{vis}-\boldsymbol{c}_j^{n}\|_2\right]_+ +\left[\rho+\|\boldsymbol{c}_i^{inf}-\boldsymbol{c}_i^{vis}\|_2-\min_{\substack{n\in[vis,\inf]\\ j\neq i}}\|\boldsymbol{c}_i^{inf}-\boldsymbol{c}_j^{n}\|_2\right]_+\tag{9}$$

式中:P 是样本数量;ρ 为边界;$\|\cdot\|_2$ 表示 L2 范数;$\|\boldsymbol{c}_i^{vis}-\boldsymbol{c}_i^{inf}\|_2$ 和 $\|\boldsymbol{c}_i^{inf}-\boldsymbol{c}_i^{vis}\|_2$ 表示同一行人异质特征中心间的距离;$\min_{\substack{n\in[vis,\inf]\\ j\neq i}}\|\boldsymbol{c}_i^{vis}-\boldsymbol{c}_j^{n}\|_2$ 和 $\min_{\substack{n\in[vis,\inf]\\ j\neq i}}\|\boldsymbol{c}_i^{inf}-\boldsymbol{c}_j^{n}\|_2$ 表示不同行人特征中心间的最小距离。该损失可以优化模型的训练过程,使得模型在跨模态任务中更具有判别力和鲁棒性。因此,用于任务的异质中心三元组分类损失可以表示为:

$$L_{hc_{tri}}=\frac{1}{4}\left(L_{hc_{tri}}^{g1}+L_{hc_{tri}}^{g2}+\frac{1}{6}\sum_{i=1}^{6}L_{hc\_tri}^{l6i}+\frac{1}{3}\sum_{i=1}^{3}L_{hc\_tri}^{l3i}\right)\tag{10}$$

式中:$L_{hc\_tri}^{g1}$ 和 $L_{hc\_tri}^{g2}$ 是全局特征产生的分类损失;$L_{hc\_tri}^{l6i}$ 和

$L_{hc\_tri}^{[3i}$ 分别是 6 等分和 3 等分后每个局部特征产生的损失。因此,用于监督网络训练的整体损失 L_{total} 为:

$$L_{total} = L_{id} + \alpha L_{hc_{tri}} \tag{11}$$

式中:α 是用来平衡两种不同特征的参数。经过广泛的实验,对于 SYSU-MM01[6] 和 RegDB[31] 数据集,α 均设置为 2。

2 实验结果与分析

2.1 数据集和评价标准

本文方法在公开数据集上进行了实验,分别是 SYSU-MM01[6] 和 RegDB[31]。

SYSU-MM01[6] 是一个大规模的可见光-红外行人数据集,主要由 6 个分布在室内和室外的摄像机采集整理所得,包括 4 个可见光摄像机和 2 个红外摄像机。数据集已经划分出训练集和测试集,其中训练集包含 395 个行人 ID,共计 22 258 张可见光图像和 11 909 张红外图像;测试集包含 96 个行人 ID,提供了 3 803 张用于查询的红外图像和 301 张随机选择的可见图像。SYSU-MM01 涉及到拍摄环境的改变,行人的姿态、穿着等都发生了变化,是一个非常具有挑战性的数据集。

RegDB[31] 是另外一个广泛使用的跨模态数据集,由一个可见光摄像机和一个热红外摄像机共同拍摄完成。数据集包含 412 个行人 ID,每个 ID 下有 10 张可见图像和 10 张热红外图像。数据集被随机划分成两个部分,206 个行人 ID 所包含的图像用于训练,另一半用于测试。

为了更好描述模型的性能,采用了累积匹配特征(CMC)中的 Rank-k(k = 1, 5, 10, 20)和平均精度均值(mAP)作为评价指标。Rank-k 衡量了检索的前 k 个结果中出现相同 ID 的行人图像的概率;mAP 用于衡量方法的平均检索性能,在查询集中存在多个匹配图像的情况下尤为重要。

2.2 实验设置

本文所有实验均在深度学习框架 PyTorch 下完成,基本环境配置如表 2 所示。

表 2 实验环境设置

Table 2 Experimental environment settings

环境	版本
操作系统	Linux Mint 20.3
GPU	A800
Pytorch 版本	1.10.2
CUDA	12.2

算法采用经过 ImageNet 预训练的 ResNet50 作为骨干网络。为了捕获更多特征细节,在训练过程中将最后的卷积块步幅从 2 调整为 1,实现更好的细节捕捉。训练阶段,输入图像大小设置统一为 288×144,并在图像周围进行 10 pixels 的零填充。为了增加训练数据的多样性,训练时采用了随机左右翻转图片的方法,并将其裁剪到指定的大小,这样的数据增强策略有助于提高模型的泛化能力和稳定性。在网络优化器的选择方面,采用了随机梯度下降(SGD)优化器,动量参数设置为 0.9,并初始化学习率为 0.1。此外,实验采用了热身学习率策略,进而引导网络更快地收敛并获得更高的性能。

2.3 消融实验

1)共享-解离模块的有效性

共享-解离模块通过对共享卷积层的剥离,并复制成两个不再共享参数的卷积操作,实现两种不同精细度局部信息的学习。共享-解离模块的嵌入方案如表 1 所示。实验主要是为了探究在两个公开数据集 SYSU-MM01 和 RegDB 上采取何种嵌入方案获得最好的结果,并验证模块相较于双流参数共享网络更为有效,实验结果如图 3~5 所示。

如图 3 和 4 所示,对比 6 种共享-解离模块嵌入方案可以得出,在 SYSU-MM01 数据集上,对主干网络的 $stage_3$ 进行参数共享,并在 $stage_4$ 处进行复制,$stage_4$ ~ $stage_5$ 继续采用双流结构可以取得最好的结果(即采用方案 ④),其中 Rank-1 和 mAP 可以达到 74.70% 和 71.79%;在 RegDB 数据集上,综合两种检索模式可以得到 $stage_2$ ~ $stage_3$ 采用共享参数,$stage_4$ ~ $stage_5$ 采用双流结构可以获得更好的结果(即采用方案 ②)。

为了进一步验证本模块的有效性,将两个数据集上的结果与双流参数共享网络的结果进行了对比。如图 3 所示,当 $shared$=2 时,即 $stage_1$ ~ $stage_2$ 为特有特征提取分支,而 $stage_3$ ~ $stage_5$ 为共享特征学习分支时,Rank-1 和 mAP 在 SYSU-MM01 上取得了最高的准确率。共享-解离模块嵌入的最佳表现与之相比,在这两个指标上分别有着 1.55% 和 3.11% 的性能优势。从平均值来说,模块嵌入后 6 种方案的平均 Rank-1 和 mAP 分别是 73.23% 和 69.99%,相较于未嵌入该模块的平均结果 72.61% 和 68.54% 来说,也有一定程度的提升。在 RegDB 上的比较结果如图 5 所示,可以得到相似的结论。特别是在红外到可见光检索模式下,模块嵌入后准确率得到了显著的提升。

2)多粒度相关特征学习模块的有效性

多粒度相关特征模块通过关系网络来捕捉身体各个部分与其他部分之间的联系,进一步实现跨模态下的全局-局部特征对齐。通过一组实验来验证其有效性,实验结果如表 3 所示。表 3 中共享-解离模块均以最佳方案嵌入,若不采用 3 等分或 6 等分的局部特征,则将局部特

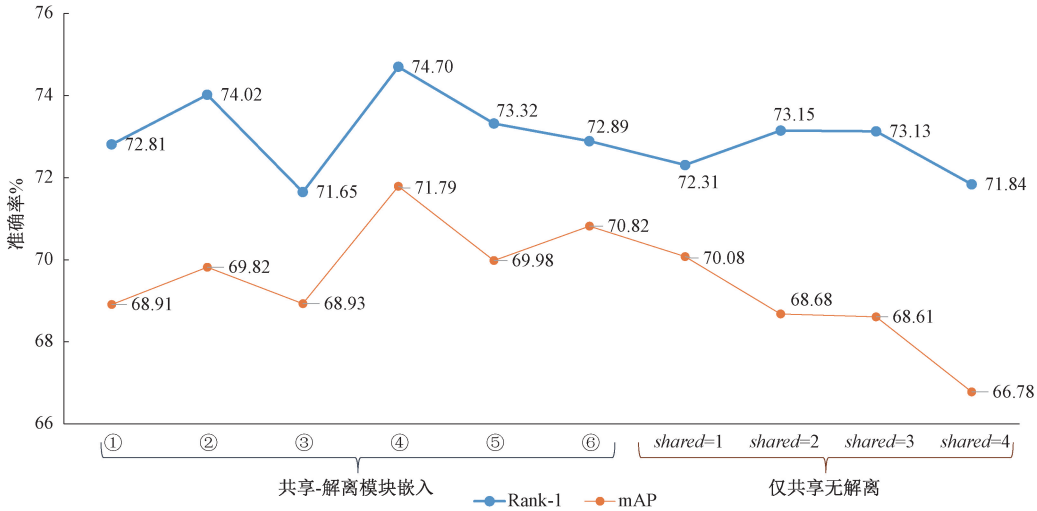

图 3　全搜索模式 SYSU-MM01 采用不同共享-解离模块嵌入方案和仅采用参数共享方案的 Rank-1 和 mAP 准确率对比

Fig. 3　Comparison of Rank-1 and mAP with different shared-disentangling module embedding schemes and parameter sharing-only scheme for SYSU-MM01 in the all-search mode

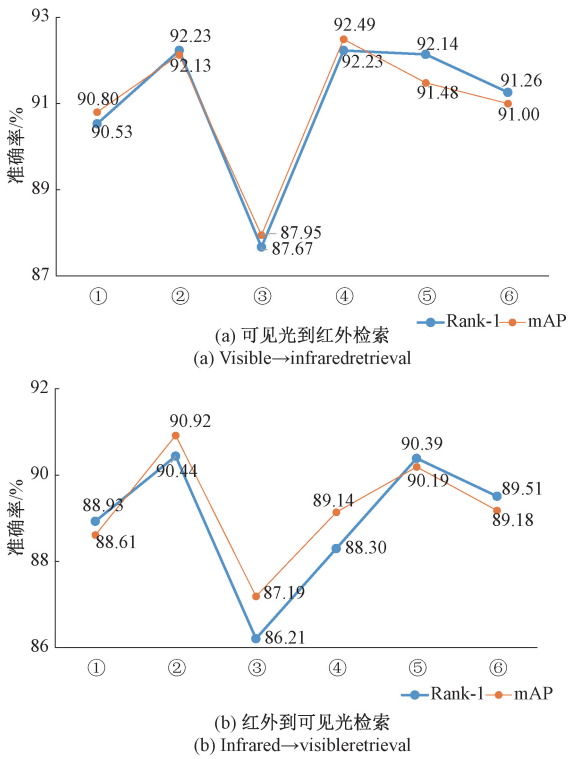

图 4　两种模式下 RegDB 上执行不同共享-解离模块嵌入方案的 Rank-1 和 mAP 准确率对比

Fig. 4　Illustrates the results of Rank-1 and mAP with different shared-disentangling embedding schemes in RegDB of the two modes

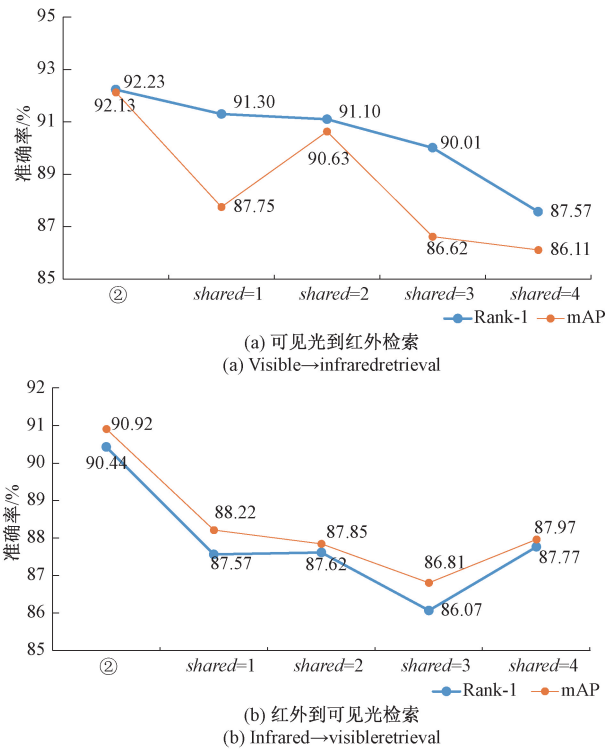

图 5　两种模式下 RegDB 上执行最优共享-解离模块嵌入方案和仅采用参数共享方案的 Rank-1 和 mAP 对比

Fig. 5　Illustrates the results of Rank-1 and mAP performing the optimal share-disentangling embedding scheme and the parameter sharing-only scheme in RegDB of two modes

征部分损失权重设置为 0,全局特征仍是由局部特征串联获得。

从表 3 可以观察到,若仅学习两个全局特征(即组合 1),Rank-1 和 mAP 分别是 64.77%和 61.48%;组合 2 和组合 3 相较于组合 1 来说,加入了不同粒度的局部特征,

准确率获得了一定的提升。组合 4 相较于组合 2,在特征学习阶段增加了一个关系网络 1,因此 Rank-1 和 mAP 分别提升 1.60% 和 1.82%。相似的,组合 5 相较于组合 3 准确率也有了显著的提升。这些结果都验证了引入两个粒度的局部相关特征的有效性。此外,本文方法和组合 6 进行对比,区别在于是否采用学习两个全局特征表达。从准确率可以看出,本文方法 Rank-1 和 mAP 均显著高于组合 6,因此可以验证全局相关特征在可见光-红外跨模态任务中的有效性。

表 3 多粒度相关特征学习模块在 SYSU-MM01 数据集上的准确率

Table 3 Effectiveness of the multi-granularity relation feature learning module in SYSU-MM01 (%)

方法	3 等分	6 等分	全局	关系网络 1	关系网络 2	Rank-1	mAP
组合 1			√			64.77	61.48
组合 2	√		√			66.96	64.02
组合 3		√	√			67.53	64.72
组合 4	√		√	√		68.56	65.84
组合 5		√	√		√	70.12	67.56
组合 6	√	√		√	√	72.08	69.10
本文	√	√	√	√	√	74.70	71.79

2.4 与其他方法进行比较

本文提出的算法与经典的以及先进的跨模态行人重识别方法在两个公开数据集上进行结果的比较,结果如表 4 和 5 所示。

1)SYSU-MM01 数据集上结果

SYSU-MM01 数据集上分别进行了全搜索(all-search)和室内搜索(indoor-search)搜索下的实验,与现有先进算法和经典算法对比结果如表 4 所示。从表 4 可以看出,本文算法在两种模式下 Rank-1 和 mAP 均获得了最高的准确率。其中,在全搜索模式下,本文算法 Rank-1 和 mAP 分别达到了 74.70% 和 71.79%;在室内搜索模式下,分别达到了 79.67% 和 83.58%。相较于经典的双流参数不共享网络如 Zero-padding、AGW、HC 等方法,本文算法准确率有了显著提升。HcTri 提出了参数共享的双流网络和异质中心三元组损失,但本方法在 HcTri 的基础上增加了共享-解离模块,并提取了局部特征间的关系信息,因此实现了准确率的提升。

PSFLNet 采用了参数共享的双流网络作为框架,并且考虑了多粒度特征。本文方法在两种搜索模式下,Rank-1 和 mAP 相较于它有一定程度的提升,分别是 0.70%、1.28%、0.17% 和 1.48%,但 Rank-10 和 Rank-20 均低于 PSFLNet。不仅如此,CAJ、AGMNet 等方法的 Rank-10 和 Rank-20 指标都高于本文算法。这可能是因为上述这些方法均采用了中间模态生成策略来提升准确率。当对输入的可见光图像进行灰度化处理后,本文方法的所有指标都获得了显著提升。

表 4 SYSU-MM01 数据集上的实验结果对比

Table 4 Comparison of our method and state-of-the-art methods on SYSU-MM01 (%)

检索模式	All-search				Indoor-search			
	Rank-1	Rank-10	Rank-20	mAP	Rank-1	Rank-10	Rank-20	mAP
Zero-padding[6]	14.80	47.99	65.50	12.85	15.60	61.18	81.02	21.49
eBDTR[3]	27.82	67.34	81.34	28.42	32.46	77.42	89.62	42.46
cmGAN[7]	26.97	67.51	80.56	31.49	31.63	77.23	89.18	42.19
D^2RL[20]	28.90	70.60	82.40	29.20	28.12	70.23	83.67	29.01
AlignGAN[21]	42.40	85.00	93.70	40.70	45.90	87.60	94.40	54.30
AGW[1]	47.50	84.39	92.14	47.65	54.17	91.14	95.98	62.97
HAT[23]	55.29	92.14	97.36	53.89	62.10	95.75	99.20	69.37
HC[4]	56.96	91.50	96.82	54.95	59.74	92.07	96.22	64.91
MCLNet[19]	65.40	93.33	97.14	61.98	72.56	96.98	99.20	78.30
HcTri[8]	61.68	93.10	97.17	57.51	63.41	91.69	95.28	68.17
SFANet[24]	65.74	92.98	97.05	60.83	71.60	96.60	99.45	80.05
CAJ[25]	69.88	95.71	98.46	66.89	76.26	97.88	99.49	80.37
CAJ+[25]	71.48	96.23	98.71	68.15	78.36	98.36	99.78	78.44
CMIT[33]	70.94	94.93	96.37	65.51	73.28	95.20	99.43	77.18
CMTR[34]	65.45	94.47	98.16	62.90	71.99	96.37	99.09	57.07
AGMNet[26]	69.63	96.27	98.82	66.11	74.68	97.51	99.14	78.30
PSFLNet[15]	74.00	96.50	99.00	70.51	79.50	97.50	99.24	82.10
本文算法	74.70	94.06	96.77	71.79	79.67	98.41	99.25	83.58

表 5　RegDB 数据集上的实验结果对比

Table 5　Comparison of our method and state-of-the-art methods on RegDB　　(%)

方法	Visible→infrared				Infrared→visible			
	Rank-1	Rank-10	Rank-20	mAP	Rank-1	Rank-10	Rank-20	mAP
Zero-padding[6]	17.75	34.21	44.35	18.90	16.63	34.68	44.25	17.82
eBDTR[3]	34.62	58.96	68.72	33.46	34.21	58.74	68.64	32.49
D^2RL[20]	43.40	66.10	76.30	44.10	–	–	–	–
AlignGAN[21]	57.90	–	–	53.60	56.30	–	–	53.40
AGW[1]	70.05	86.21	91.55	66.37	70.49	87.21	91.84	65.90
HAT[23]	71.83	87.16	92.16	67.56	70.02	86.45	91.61	66.30
MCLNet[19]	80.31	92.70	96.03	73.07	75.93	90.93	94.59	69.49
HcTri[8]	91.05	97.16	98.57	83.28	89.30	96.41	98.16	81.46
SFANet[24]	76.31	91.02	94.27	68.00	70.15	85.24	89.27	63.77
CAJ[25]	85.03	95.49	97.54	65.33	84.75	95.33	97.51	77.82
CAJ+[25]	85.69	95.45	97.54	79.70	84.88	95.66	97.74	78.55
CMIT[33]	88.78	94.76	97.04	88.49	84.55	93.72	95.83	83.64
CMTR[34]	88.11	–	–	81.66	84.92	–	–	80.79
AGMNet[26]	88.40	95.10	96.94	81.45	85.34	94.56	97.48	81.19
本文算法	92.23	95.78	97.62	92.13	90.44	96.21	98.30	90.92

2) RegDB 数据集上结果

RegDB 数据集上分别进行了可见光到红外(Visible→infrared)和红外到可见光(Infrared→visible)模式下的实验,与现有先进算法和经典算法对比结果如表 5 所示。从表 5 可以看出,在该数据集上的表现和 SYSU-MM01 相似,两种模式下 Rank-1 和 mAP 均获得了最高的准确率。而在可见光到红外模式下,Rank-10 和 Rank-20 略低于 HcTri;在红外到可见光模式下 Rank-10 也较 HcTri 低 0.20%。由于 RegDB 的数据量相较于 SYSU-MM01 更少,且拍摄环境更为简单,因此先进算法的准确率均较高,但是仍然可以看出本文算法在 mAP 上具有明显的优势。

3　结　论

本文针对可见光-红外行人重识别任务提出了一种基于多粒度共享-解离相关网络,为高判别力模态共享特征的学习设计了 3 个模块。首先,为了提升双流参数共享网络在多粒度特征学习方面的有效性,在基准网络中嵌入了共享-解离模块,并通过多组实验选取出最优嵌入方案;再者,为了提高行人特征的鲁棒性,将关系网络策略融入到全局-局部特征的构建中,从不同粒度上探索了行人身体结构间的关联信息;最后联合 ID 分类损失和异质中心三元组损失,通过对多粒度特征学习过程的多层次约束,优化了模型的训练过程。在两个公开数据集 SYSU-MM01 和 RegDB 上证明了该文算法较优秀。后续研究可以通过融合轻量级的辅助模态生成方法,在不过多增加计算量的同时,进一步提升模型的准确率。

参考文献

[1] YE M, SHEN J, LIN G, et al. Deep learning for person re-identification: A survey and outlook [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022,44(6):2872-2893.

[2] 罗浩,姜伟,范星,等. 基于深度学习的行人重识别研究进展[J]. 自动化学报,2019,45(11):2032-2049.
LUO H, JIANG W, FAN X, et al. A survey on deep learning based person re-identification [J]. Acta Automatica Sinica, 2019,45(11):2032-2049.

[3] YE M, LAN X, WANG Z, et al. Bi-directional center-constrained top-ranking for visible thermal person re-identification [J]. IEEE Transactions on Information Forensics and Security, 2020, 15:407-419.

[4] LIU H, CHENG J, WANG W, et al. Enhancing the discriminative feature learning for visible-thermal cross-modality person re-identification [J]. Neurocomputing, 2020, 398:11-19.

[5] 范馨月,张阔,张干,等. 细微特征增强的多级联合聚类跨模态行人重识别算法[J]. 电子测量与仪器学报,2024,38(3):94-103.
FAN X Y, ZHANG K, ZHANG G, et al. Cross-modal person re-identification algorithm based on multi-level join clustering with subtle feature enhancement [J]. Journal of Electronic Measurement and Instrumentation, 2024,38(3):94-103.

[6] WU A, ZHENG W, YU H, et al. RGB-infrared cross-modality person re-identification[C]. IEEE International Conference on Computer Vision, 2017: 5390-5399.

[7] DAI P, JI R, WANG H, WU Q, et al. Cross-modality

person re-identification with generative adversarial training[C]. International Joint Conference on Artificial Intelligence, 2018: 677-683.

[8] LIU H, TAN X, ZHOU X. Parameter sharing exploration and hetero-center triplet loss for visible-thermal person re-identification[J]. IEEE Transactions on Multimedia, 2020, 23:4414-4425.

[9] KONG J, HE Q, JIANG M, et al. Dynamic center aggregation loss with mixed modality for visible-infrared person re-identification [J]. IEEE Signal Processing Letters, 2021, 28:2003-2007.

[10] XUE C, DENG Z, WANG S, et al. GLSFF: Global-local specific feature fusion for cross-modality pedestrian re-identification[J]. Computer Communications, 2024, 215:157-168.

[11] KIM M, KIM S, PARK J, et al. PartMix: Regularization strategy to learn part discovery for visible-infrared person re-identification [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 18621-18632.

[12] 石林波,李华锋,张亚飞,等. 模态不变性特征学习和一致性细粒度信息挖掘的跨模态行人重识别[J]. 模式识别与人工智能,2022,35(12):1064-1077.
SHI L B, LI H F, ZHANG Y F, et al. Modal invariance feature learning and consistent fine-grained information mining based cross-modal person re-identification [J]. Pattern Recognition and Artificial Intelligence, 2022, 35(12):1064-1077.

[13] 张勃兴,马敬奇,张寿明, 等. 利用全局与局部关联特征的行人重识别方法[J]. 电子测量与仪器学报, 2022,36(6):205-212.
ZHANG B X, MA J Q, ZHANG SH M, et al. Person re-identification method based on global and local relation features [J]. Journal of Electronic Measurement and Instrumentation, 2022,36(6):205-212.

[14] 马潇峰,程文刚. 双粒度特征融合网络的跨模态行人再识别[J]. 中国图象图形学报, 2023, 28(5): 1422-1433.
MA X F, CHENG W G. Dual-grained feature fusion network-relevant cross-modality pedestrian re-identification [J]. Journal of Image and Graphics, 2023, 28(5): 1422-1433.

[15] CHAN S, DU F, TANG T, et al. Parameter sharing and multi-granularity feature learning for cross-modality person re-identification [J]. Complex Intelligence System,2024,10: 949-962.

[16] YE H, LIU H, MENG F, et al. Bi-directional exponential angular triplet loss for RGB-infrared person re-identification [J]. IEEE Transactions on Image Processing, 2021, 30:1583-1595.

[17] KONG J, HE Q, JIANG M, et al. Dynamic center aggregation loss with mixed modality for visible-infrared person re-identification [J]. IEEE Signal Processing Letters, 2021, 28: 2003-2007.

[18] 李灏, 唐敏, 林建武, 等. 基于改进困难三元组损失的跨模态行人重识别框架[J]. 计算机科学, 2020, 47(10): 180-186.
LI H, TANG M, LIN J W, et al. Cross-modality person re-identification framework based on improved hard triplet loss[J]. Computer Science, 2020, 47(10): 180-186.

[19] HAO X, ZHAO S, YE M, et al. Cross-modality person re-identification via modality confusion and center aggregation[C]. IEEE/CVF International Conference on Computer Vision, 2021:16383-16392.

[20] WANG Z, WANG Z, ZHENG Y, et al. Learning to reduce dual-level discrepancy for infrared-visible person re-identification[C]. IEEE International Conference on Computer Vision and Pattern Recognition, 2019: 618-626.

[21] WANG G, ZHANG T, CHENG J, et al. RGB-infrared cross-modality person re-identification via joint pixel and feature alignment[C]. IEEE International Conference on Computer Vision, 2019: 3622-3631.

[22] WEI Z, YANG X, WANG N, et al. Dual-adversarial representation disentanglement for visible infrared person re-identification[J]. IEEE Transactions on Information Forensics and Security, 2024,19: 2186-2200.

[23] ZHONG X, LU T, HUANG W, et al. Grayscale enhancement colorization network for visible-infrared person re-identification [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32 (3):1418-1430.

[24] LIU H, MA S, XIA D, et al. SFANet: A spectrum-aware feature augmentation network for visible-infrared person re-identification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 4:1958-1971.

[25] YE M, WU Z, CHEN C, et al. Channel augmentation for visible-infrared re-identification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(4): 2299-2315.

[26] LIU H, XIA D, JIANG W. Towards homogeneous modality learning and multi-granularity information exploration for visible-infrared person re-identification[J]. IEEE Journal of Selected Topics in Signal Processing, 2023, 17(3): 545-559.

[27] YE M, SHEN J, SHAO L. Visible-infrared person re-

identification via homogeneous augmented tri-modal learning[J]. IEEE Transactions on Information Forensics and Security, 2021,16:728-739.

[28] SUN Y, ZHENG L, YANG Y, et al. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline)[C]. European Conference on Computer Vision, 2018: 501-518.

[29] VARIOR R, SHUAI B, LU J, et al. A siamese long short-term memory architecture for human re-identification[C]. European Conference on Computer Vision, 2016: 135-153.

[30] PARK H, HAM B. Relation network for person re-identification[C]. AAAI Conference on Artificial Intelligence, 2019: 11839-11847.

[31] NGUYEN D, HONG H, KIM K,. Person recognition system based on a combination of body images from visible light and thermal cameras[J]. Sensors, 2017, 17(3):605.

[32] BERMAN M, HERVE J, VEDALDI A, et al. MultiGrain: A unified image embedding for classes and instances[J]. ArXiv preprint arXiv. 1902. 05509, 2019.

[33] FENG Y, YU J, CHEN F, et al. Visible-infrared person re-identification via cross-modality interaction transformer[J]. IEEE Transactions on Multimedia, 2023, 25:7647-7659.

[34] LIANG T, JIN Y, LIU W, et al. Cross-modality transformer with modality mining for visible-infrared person re-identification[J]. IEEE Transactions on Multimedia, 2023, 25:8432-8444.

作者简介

宋婉茹(通信作者),2020 年于南京邮电大学获得博士学位,现为南京邮电大学校聘副教授、硕士生导师,主要研究方向为模式识别及教育人工智能。

E-mail: songwanru@njupt.edu.cn

Song Wanru (Corresponding author) received her Ph. D. degree from Nanjing University of Posts and Telecommunications in 2020. Now she is a university-appointed associate professor and M. Sc. supervisor at Nanjing University of Posts and Telecommunications. Her main research interests include pattern recognition and AI for education.

郝川艳,2015 年于澳门大学获得博士学位,现为南京邮电大学教育科学与技术学院副教授,主要研究方向为图像处理、模式识别以及教育人工智能。

E-mail: hcy@njupt.edu.cn

Hao Chuanyan received her Ph. D. degree from University of Macau in 2015. Now she is an associate professor and M. Sc. supervisor at Nanjing University of Posts and Telecommunications. Her main research interests include image processing, pattern recognition and AI for education.

郑洁莹,2020 年于南京邮电大学获得博士学位,现为南京邮电大学讲师,主要研究方向为图像处理和模式识别。

E-mail: zhengjieying@njupt.edu.cn

Zheng Jieying received her Ph. D. degree from Nanjing University of Posts and Telecommunications in 2020. Now she is a lecturer at Nanjing University of Posts and Telecommunications. Her main research interests include image processing and pattern recognition.

刘峰,1997 年于南京理工大学获得博士学位,现为南京邮电大学教育科学与技术学院教授,主要研究方向为图像处理、模式识别以及教育人工智能。

E-mail: liuf@njupt.edu.cn

Liu Feng received his Ph. D. from Nanjing University of Science and Technology in 1997. Now he is a professor and Ph. D. supervisor at Nanjing University of Posts and Telecommunications. His main research interests include image processing, pattern recognition and AI for education.