

可见的红外人重新识别的隐性歧视性知识学习

Lei Zhang*的Kaijie Ren*微电子与传播工程学院，中

国重庆大学kaijieren@cqu.edu.cn, leizhang@cqu.edu.cn

抽象的

Visible-Infrared Person Re-identification (VI-ReID) is a challenging cross-modal pedestrian retrieval task, due to significant intra-class variations and cross-modal discrepancies among different cameras. Existing works mainly focus on embedding images of different modalities into a unified space to mine modality-shared features. They only seek distinctive information within these shared features, while ignoring the identity-aware useful information that is implicit in the modality-specific features. To address this issue, we propose a novel Implicit Discriminative Knowledge Learning (IDKL) network to uncover and leverage the implicit discriminative information contained within the modality-specific. First, we extract modality-specific and modality-shared features using a novel dual-stream network. Then, the modality-specific features undergo purification to reduce their modality style discrepancies while preserving identity-aware discriminative knowledge. Subsequently, this kind of implicit knowledge is distilled into the modality-shared feature to enhance its distinctiveness. Finally, an alignment loss is proposed to minimize modality discrepancy on enhanced modality-shared features. Extensive experiments on multiple public datasets demonstrate the superiority of IDKL network over the state-of-the-art methods. Code is available at <https://github.com/1KK077/idkl>.

1. 简介

可见的红外人员重新识别（VI-REID）旨在匹配多个非重叠的相机视图和不同方式的行人图像。当今的广告监视系统能够在夜幕降临时自动从可见的红外模式切换，从而确保大量可训练的数据供应。但是，红外图像的独特方式产生了重大的差异和更复杂的环境。这使得以前的单模式方法是基于的

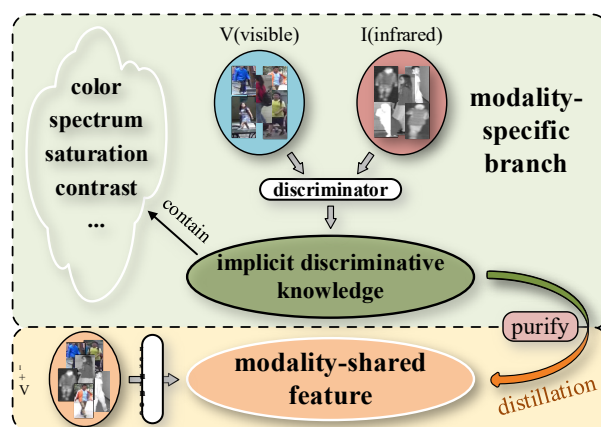


图1。以前的方法着重于在模态共享特征中寻求歧视性化，从而忽略了一个歧视性线索所隐含在模态特定特征中的事实。值得考虑使用隐性歧视信息来增强共享不变功能的利用。

仅在可见图像上，对VI-REID任务的效率较低。

最近，VI-Reid领域已经出现了许多高级方法。总而言之，这些方法通常可以分为两种类型：第一个目的是直接从原始模态数据中学习模态共享的特征，第二种方法旨在合并其他模态信息以完善特征空间或弥合模态差距，从而促进搜索模态共享的features。基于原始数据的方法[2, 4, 7, 7, 8, 17, 19, 33, 49]是将不同的模态图像嵌入相同的空间中，并在功能级上对齐，以便模型学习模态不变特征。尽管基于原始图像的方法取得了理想的结果，但它们仍然存在特征空间不同方式之间的巨大差距。为了弥合可见态度和循环方式之间的差距，并构建一个连续的空间，以更好地学习模式共享特征，基于引入额外信息的各种方法[14、29、30、34、38、42、45]不断出现。

尽管已经取得了重大改进

当前的VI-REID方法，这些模型不可避免地会丢弃一些依赖于模态特定特征的歧视性信息，这些信息并未完全利用和使用。这些以模态特定特征存在的歧视性提示可以称为隐式歧视天然信息，如图1所示的绿色区域所示，例如颜色，灰度谱，对比度，饱和度等。因此，仅依靠模态共享的提示可以限制特征表示的歧视能力的上限。有效地利用隐式模态特定特征对于增强模态不变特征的独特性至关重要。但是，由于其包含的模态差异，我们无法直接使用特定于模态的知识。不胜替，我们需要减少这些差异，同时保留隐性知识中固有的歧视性信息。同时，涉及蒸馏，对齐和相互学习的传统VI-REID方法通常依赖于逻辑[35, 44, 47]。但是，在测试阶段没有分类器的参与，仅在特征级别执行匹配。因此，在特征级别进行区分信息蒸馏也是必不可少的。

为了解决上述局限性，在本文中，我们提出了一个隐式歧视性知识学习（IDKL）框架，该框架从特定于模态特定的特征中捕获了隐式不变的信息，并将其提炼成模态共享的特征，以增强其歧视性可容纳性。我们首先使用模态鉴别器和模态混淆者提取特定于模态和模态共享特征。模态鉴别器的效率可以区分不同的模态特征，从而赋予了特定特征。而模式混淆者无法区分模态特征，从而赋予了共享特征。由于上一个阶段之后的特定于模式特定功能具有实质性的方式差异，因此它不适合直接蒸馏到共享特征中。我们最初采用实例归一化来减少域差异。但是，至关重要是要承认，可以看出，可以导致某些歧视特征的丧失。因此，我们旨在减少其方式风格差异，同时保留身份感知的歧视性知识。随后，我们将这种隐式知识提炼成特征图结构的特征级别的模态共享特征，以及通过logit矢量的语义级别来提高其独特性。最后，提出了对齐损失，以最大程度地减少增强模态共享特征上的模态差异。

本文的主要贡献可以总结如下：

- 我们建议隐性歧视性知识学习（IDKL）网络以利用歧视性

在特定于模态特征中隐含的知识，以增强模态共享特征的判别能力上限。

- 为了减少模式风格的差异而不会失去特定于模式的信息的歧视性信息，我们提出了一个指导的信息净化器（IP），该信息通过歧视增强损失和减少损失的差异来监督。
- 新颖的TGSA损失是为了将特定于犯罪模式的信息提炼成模态性特征，并减轻模态共享特征的模式间差异。实质性的实验结果证明了我们方法的优势。

2. 相关工作

可见的模态人里德。Reid人近年来获得了越来越多的成功，该公司旨在实施可见图像之间的行人检索。它面临着巨大的挑战，包括各种观点，照明，姿势等。为了解决这些概率，大多数方法[24, 25, 39, 48]旨在通过训练CNN网络获得统一的阶层内和歧视性级别表示。为了进一步增强特征的区分性，Sun *et al.* [27]直接提出对齐零件功能，而不是使用外部提示。他 *et al.* [9]抓住了变压器的热点，并首先提出了适用于单态人Reid的变压器变体。同时，Reid人建立了一些重要的分支机构，例如不透视的领域自适应人REID（UDA-REID）和域的概括人REID（DG-REID）[41]。尽管现有的方法在可见的模态人里德（Reid）取得了广泛的进步，但由于跨度差异差异，当将其应用于可见的红外人员里德（Reid）时，它们会遭受性能退化。

可见的红外人里德（Vi-Reid）。Vi-Reid是一个跨模式的人检索问题，旨在匹配白天可见和夜间红外图像。它不仅面临着传统的单明一年人REID任务遇到的困难，而且还面临着由不同的频谱引起的巨大模态差异的主要挑战。为了解决这些问题，提出了许多方法来搜索共享功能[1、5、18、20、32、36、39、46]。Wu *et al.* [32]首先制定了Vi-Reid Is-Sue，并贡献了新的跨模式数据集Sysu-MM01，这对于以下重新搜索至关重要。是的 *et al.* [36]提出了零件级特征的反应和图形结构的注意，以增强不变特征的损害。同时，一些工作开始引入其他模态信息以共同搜索不变特征[14、21、28、30、45]。Wang *et al.* [30] RGB组合三个通道功能和IR

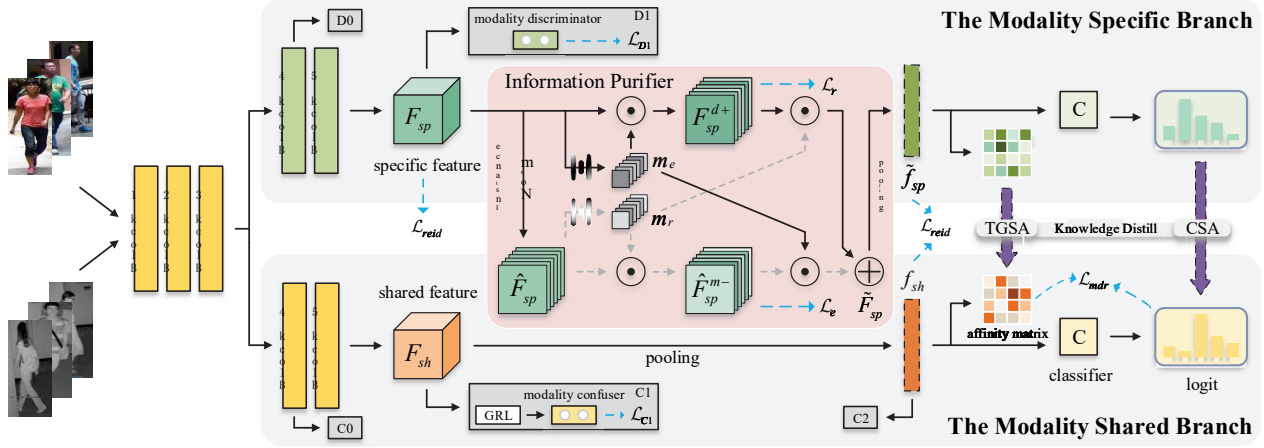


图2. 拟议的隐式判别知识学习 (IDKL) 模型的框架。由RESNET块构建的双重单流网络首先提取模式特异性 F_{sp} 和模态共享特征 F_{sh} 在模态歧视器和模态混淆器的约束下，而常见的REID损失则用于基于优化网络。然后，将特定于模式的特征馈入信息净化器中，以调节模态样式差异，同时保留隐式歧视性信息并获得纯化的特定于模式特异性功能 \hat{F}_{sp} 。随后，这种隐式知识通过TGSA和CSA提炼成模态共享的特征。最后，进一步提出了 L_{mdr} ，以最大程度地减少增强模态共享特征中的模态差异。

通道功能到United四个通道功能，以减少模态差异。江 *et al.* [14]采用全局影响，通过变压器生成缺失的模态。但是，这些基于一代的补偿方法不可避免地引入了很多噪声，并且大多数当前方法旨在使模态差异和提取模态不变特征。他们忽略了特定功能中包含的有益样式信息以提高独特性。

互动学习。在知识蒸馏中，使用教师模型将较大的复杂教师模型学到的知识转移到较小的模拟学生模型中。与教师和学生之间的单向转移不同，深度相互学习[44]是一组学生，他们在整个培训过程中进行了协作并互相教学。一些工作将此想法应用于Vi-Reid，以不同方式进行互动学习。Zheng *et al.* [47]通过使用两种模式之间的互动学习来增强不变的学习。是的 *et al.* [35]采用多分类器来弥补逻辑级别的模式差异。wu *et al.* [33]由四个分类器蒸馏了RGB和IR知识，以实现共享表示。但是，他们只学习语义级别中的知识。

3. 方法论

在本节中，我们详细介绍了所提出的隐性歧视知识学习 (IDKL) 框架，如图2所示。IDKL首先通过两个网络分支区分了模态共享的特定方式与模态共享的特征，这受到两个网络分支的约束，这受到模态歧视的限制。

和模式混淆者。同时，使用REID损失来赋予代表性的特征。然后，开发了信息净化器 (IP)，以赋予样式差异的影响，同时保留特定于模态特征的身份意识和歧视性知识。最后，我们将两个分支的隐式判别知识提炼成模态共享的特征，通过在logit级别的特征级别和类语义对齐 (CSA) 处的三重态图结构对齐 (TGSA)。此外，为了减少共享特征的模态差异，我们提出了模态分支分支中的模态减少 (MDR) 损失。随着酌处能信息的增加，模态差异不断减少，可以获得增强的模态共享特征。

正式地，我们将数据集中的可见图像分别表示为 $V = \{\mathbf{x}_i^V\}_{i=1}^{N_V}$ 和 $I = \{\mathbf{x}_i^I\}_{i=1}^{N_I}$ 。通常，在两个模态中，在迷你批次中采样的图像数量相等，即 $N_V = N_I = N = P \times K$ ，其中 N_V 和 N_I 分别表示从可见的和红外模态采样的图像数。在这里， N 表示来自单个模态的图像数量， $2N$ 是迷你批次中的图像总数， P 是不同的人类类的数量， K 是单个模式中每个类图像的数量。因此，我们可以分别将图像及其相应的标签表示为 $X = \{\mathbf{x}_i | \mathbf{x}_i \in V \cup I\}_{i=1}^{2N}$ 和 $Y = \{\mathbf{y}_i\}_{i=1}^{N_P=P}$ 。这些图像 X 被送入两个单独的网络分支，以提取包含 $F_{sp,V}$, $F_{sp,I}$ 和共享功能 F_{sh} 的模态特定功能 F_{sp} ，其中包含 $F_{sh,V}$, $F_{sh,I}$ 作为fol--

lows:

$$\mathbf{F}_{sp} = E_{sp}(\mathbf{x} \mid \Theta, \Psi), \mathbf{F}_{sh} = E_{sh}(\mathbf{x} \mid \Theta, \Phi), \quad (1)$$

其中 E_{sp} 和 E_{sh} 分别通过使用Resnet-50来表示特定和共享特征提取器。在Resnet-50的坚固中，全球平均平均集合取代了GEM POLING [24, 39]，这是最大池和平均池之间的合并操作。 Θ 是具有Resnet-50， Ψ 的前三个块和 Φ 的浅层参数，是不同分支的深层参数，其中最后两个resnet-50块。

3.1. 模式混淆者和歧视者

模式混淆者。对于每个示例图像 \mathbf{x}_i ，都有一个模态标签 $t_i \in \{0, 1\}$ 。要学习与[6]类似的模态共享信息，我们的目标是混淆不同的域，以使域分类器无法区分样本的原始域。我们采用基于梯度反向层（GRL）作为模态混淆者的广告模态分类器。模式混淆者的约束损失由以下方式给出：

$$\mathcal{L}_{Cj} = -\frac{1}{2N} \sum_{i=1}^{2N} t_i \cdot \log p(C_j(GRL(\mathbf{F}_{sh}^i))), \quad (2)$$

其中 C_j 代表模态共享分支中的j-th模式混淆者， $p(\cdot)$ 是通过softmax函数获得的预测概率， t_i 是模态标签。

模态歧视者。为了充分学习模式与特定信息相关，我们采用一种模态分类为模式歧视者。该分类器不使用GRL，在特定分支上应用于特定于特定的特定功能。分类损失的表达如下：

$$\mathcal{L}_{Dj} = -\frac{1}{2N} \sum_{i=1}^{2N} t_i \cdot \log p(D_j(\mathbf{F}_{sp}^i)), \quad (3)$$

其中 D_j 表示在模式特异性分支中的j-th模式歧视器。

模式混淆者和犯罪者的综合损失由以下方式给出：

$$\mathcal{L}_C = \sum_{j=1}^K \mathcal{L}_{Cj}, \quad \mathcal{L}_D = \sum_{j=1}^K \mathcal{L}_{Dj}. \quad (4)$$

为了有效提取模式特异性和模态共享的特征，我们将这些模态分类器损失与标准的REID损失 \mathcal{L}_{reid} 结合在一起，其中包括跨渗透性和硬三重态损失。这些都应用于特定于模态和模态共享的分支：如下：

$$\mathcal{L}_{sp} = \mathcal{L}_{reid}(\mathbf{f}_{sp}) + \mathcal{L}_D, \mathcal{L}_{sh} = \mathcal{L}_{reid}(\mathbf{f}_{sh}) + \mathcal{L}_C, \quad (5)$$

其中 $\mathbf{f} \in \mathbb{R}^{B \times C}$ 是对应于 $\mathbf{F} \in \mathbb{R}^{B \times C \times H \times W}$ 的池功能。

因此，我们的模型的基本损失被表达为：

$$\mathcal{L}_b = \mathcal{L}_{sh} + \mathcal{L}_{sp}. \quad (6)$$

3.2. 信息净化器

信息净化器（IP）旨在最大程度地减少样式差异的影响，同时在特定于模态特征中保留身份感知和歧视性知识。IP集成了实例归一化（in），已知可以减少域差异[13, 23, 48]。但是，重要的是要认识到，不可避免地会导致某些歧视特征的丧失[11, 15]，这可能会阻碍里德的高性能。

为了克服上述问题，我们签署了一个以正常方式指导（in）的双面掩盖网络，以减轻模式风格差异，同时保留隐性的歧视性知识。首先，我们在特定于模态的功能上申请获得Nor-Malized功能 $\hat{\mathbf{F}}_{sp}$ by：

$$\hat{\mathbf{F}}_{sp} = \text{IN}(\mathbf{F}_{sp}) = \frac{\mathbf{F}_{sp} - \mathbb{E}[\mathbf{F}_{sp}]}{\sqrt{\text{Var}[\mathbf{F}_{sp}] + \epsilon}}, \quad (7)$$

其中 ϵ 代表确保分辨率不是零的安全系数。沿每个通道计算平均 $\mathbb{E}[\cdot]$ 和方差 $\text{var}[\cdot]$ 。

遵循SE-NET [10]的方法，我们生成两个频道掩码 \mathbf{m}_e 和 \mathbf{m}_r by：

$$\mathbf{m}_e = \sigma(W_2 \delta(W_1 g(\mathbf{F}_{sp}))), \mathbf{m}_r = \sigma(W_4 \delta(W_3 g(\hat{\mathbf{F}}_{sp}))), \quad (8)$$

其中 $g(\cdot)$ 表示合并操作， $W_1, W_3 \in \mathbb{R}^{\frac{c}{r} \times c}$ 和 $W_2, W_4 \in \mathbb{R}^{c \times \frac{c}{r}}$ 是四个完全连接（FC）层中可学习的参数，其次是relu激活函数 $\delta(\cdot)$ 和Sigmoid Activation功能 $\sigma(\cdot)$ 。为了平衡计算消耗，降低比率 r 设置为16。

频道掩盖 \mathbf{m}_e 和 \mathbf{m}_r 分别表示歧视性特征和降低疾病的注意力掩码。因此，我们可以获得更强的独特性 \mathbf{F}_{sp}^{d+} 和较小的模态差异 $\hat{\mathbf{F}}_{sp}^{m-}$ by：

$$\mathbf{F}_{sp}^{d+} = \mathbf{m}_e \odot \mathbf{F}_{sp}, \quad \hat{\mathbf{F}}_{sp}^{m-} = \mathbf{m}_r \odot \hat{\mathbf{F}}_{sp}. \quad (9)$$

随后，计算歧视增强损失 \mathcal{L}_e 和减少损失 \mathcal{L}_r 的差异分别为：监督 \mathbf{m}_e 和 \mathbf{m}_r 为：

$$\mathcal{L}_e = \text{Softplus}\left(h\left(C_{sp}\left(\mathbf{f}_{sp}^{d+}\right)\right) - h\left(C_{sp}\left(\mathbf{f}_{sp}\right)\right)\right), \quad (10)$$

$$\mathcal{L}_r = \text{Softplus}\left(d\left(\hat{\mathbf{f}}_{sp,V}^{m-}, \hat{\mathbf{f}}_{sp,I}^{m-}\right) - d\left(\hat{\mathbf{f}}_{sp,V}, \hat{\mathbf{f}}_{sp,I}\right)\right). \quad (11)$$

在这里, \mathcal{L}_e 的目的是将生成的 \mathbf{F}_{sp}^{d+} 具有更大的语义独特性, 而 \mathbf{F}_{sp} 则 \mathcal{L}_r 试图确保 $\hat{\mathbf{F}}_{sp}^{m-}$ 表现出比 $\hat{\mathbf{F}}_{sp}$ 的模式差异。和 $\text{SoftPlus}(\cdot) = \ln(1 + \text{EXP}(\cdot))$ 是一个具有单调增加的函数, 旨在通过规避损耗中的负值来减轻优化挑战。

最后, 通过应用 \mathbf{m}_e 的应用 \mathbf{m}_e 和从 \mathbf{F}_{sp}^{d+} 应用 \mathbf{F}_{sp}^{d+} 来提取 $\hat{\mathbf{F}}_{sp}^{m-}$ 的独特信息来提取独特信息, 从 \mathbf{F}_{sp}^{d+} 中提取 \mathbf{F}_{sp}^{d+} 。因此, 以这种方式提取的两组特征都表明了较小的方式差异和更强的隐式歧视信息。通过集成它们, 我们将纯化的模式特异性特征 $\tilde{\mathbf{F}}_{sp}$ 得出:

$$\tilde{\mathbf{F}}_{sp} = \mathbf{m}_e \odot \hat{\mathbf{F}}_{sp}^{m-} + \mathbf{m}_r \odot \mathbf{F}_{sp}^{d+}. \quad (12)$$

直觉上, 隐式功能的池功能 $\tilde{\mathbf{f}}_{sp}$ $\tilde{\mathbf{F}}_{sp}$ 也是REID损失的约束。因此, 本节的信息净化损失 \mathcal{L}_{ip} 的夏季为:

$$\mathcal{L}_{ip} = \mathcal{L}_e + \mathcal{L}_r + \mathcal{L}_{reid}(\tilde{\mathbf{f}}_{sp}). \quad (13)$$

3.3. 隐性知识蒸馏

为了确保模态共享的特征可以全面地学习和集成隐式信息, 我们会通过TGSA和通过CSA从功能级和Logit级别进行蒸馏。

3.3.1 三重态图结构对齐 (TGSA)

为了赋予共享特征具有判别性信息, 并减少了特征级别的模态差异, 我们会开发三胞胎特征图结构损失。这种方法是由于特征图结构包含有关关系和特征之间的分布的大量信息, 例如类间独特性和类内部多样性的动机。这些特征 - 用于发掘潜在特征关系并增强[17, 36]中的特征表示。图表结构亲和力矩阵, 该矩阵指示特征之间的关系 - 由以下方式计算。

$$\alpha_{ij} = \frac{\exp(L([l(\mathbf{f}_i) \| l(\mathbf{f}_j)] \cdot \mathbf{w}))}{\sum_{k \in \mathcal{N}_i} \exp(L([l(\mathbf{f}_i) \| l(\mathbf{f}_k)] \cdot \mathbf{w}))}, \quad (14)$$

其中 L 表示LeakyRelu激活函数, $[\cdot \| \cdot]$ 表示concatenate操作, \mathcal{N}_i 表示用于 i 的邻居样本。 $l(\cdot)$ 是特征维度转换层, \mathbf{w} 是在配对特征中计算得分的完整连接层。获得亲和力矩阵通过软最大函数归一化使这些得分。

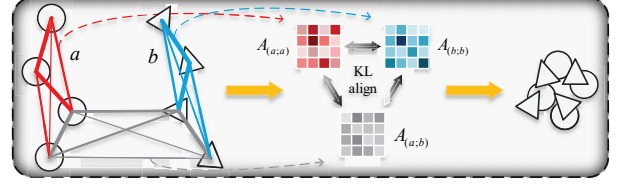


图3。拟议的TGSA损失的例证: ‘A’ 和 ‘B’ 表示两种不同类型的功能。 ‘A’ 表示图结构亲和力矩阵。对齐三个亲和力矩阵后, 将消除特征 ‘b’ 之间的图形结构分布的差异。

由于我们利用图形结构来对齐和提炼知识, 而不是增强特征, 而Eu-Clidean空间分布对于特征更有趣, 因此我们用欧几里得距离替换线性转换以计算注意力分数并重新定义两组特征的图形结构表达式:

$$\mathbf{A}_{(a;b)} = \{\alpha_{ij}\}_{i,j \in N} = \frac{\exp(D(\mathbf{f}_a^i, \mathbf{f}_b^j))}{\sum_{k \in N} \exp(D(\mathbf{f}_a^i, \mathbf{f}_b^k))}, \quad (15)$$

其中 α_{ij} 表示亲和力矩阵的元素, N 响应一个模态内的整个样本, a, b 代表两个模态, $D(\cdot)$ 表示euclidean distance $D(\cdot)$ 。

具体而言, 为交叉模态REID开发了三重态图结构对齐损失 \mathcal{L}_{tgsa} , 以使两种不同的模态类型对齐, 从而使它们能够符合相同的图形结构分布并减少模态差异。这种损失包括两个自模式亲和力矩阵和一个跨模式亲和力矩阵, 并带有图表结构分布的总体一致性, 如图3所示。这三个矩阵通过Kullback-Leibler-Leibler (kl) 差异通过利用来对齐。因此, 两种不同模态类型的对齐损失 \mathcal{L}_{tgsa} 定义为:

$$\mathcal{L}_{tgsa}^{(a;b)} = \sum_{p=1}^P \sum_{k=1}^K \left(\text{KL}(\mathbf{A}_{(a;a)}^{p_k}, \mathbf{A}_{(b;b)}^{p_k}) + \text{KL}(\mathbf{A}_{(a;a)}^{p_k}, \mathbf{A}_{(a;b)}^{p_k}) + \text{KL}(\mathbf{A}_{(b;b)}^{p_k}, \mathbf{A}_{(a;b)}^{p_k}) \right). \quad (16)$$

在这里, \mathbf{A}^{p_k} 表示 k th 示例的图形结构分布在 p - th 类中, 其中 P 是人类类的数量, K 表示单个模式中每个类的图像的数量。

为了将歧视性隐式模态 - 特定于特征级别的特定特征传达给特征级别的特定特征, 可以将通过TGSA的两个分支上的蒸馏损失损失为:

$$\mathcal{L}_{tgsa} = \mathcal{L}_{tgsa}^{(sp,V;sh,V)} + \mathcal{L}_{tgsa}^{(sp,I;sh,I)}. \quad (17)$$

3.3.2类静态对准 (CSA)

CSA用于将特定于模式的知识的语义信息提炼成模态共享的分支，以增强共享特征的特征表示。CSA以logit级别的两个分支之间的均匀特征运行。分类器背后的logit矩阵可以被配制为：

$$\mathbf{Z}_{sp} = C_{sp}(\mathbf{f}_{sp}), \mathbf{Z}_{sh} = C_{sh}(\mathbf{f}_{sh}), \quad (18)$$

其中 C_{sp} 和 C_{sh} 是单独的模态特定和模态分类器。特定和共享分支的logit \mathbf{Z}_{sp} 和 $\mathbf{Z}_{sh} \in \mathbb{R}^{2N \times C}$ 都包含可见的模态和红外模式， C 是火车数据集身份的总数。

为了学习关于语义级别的隐式歧视方式 - 特定于语义级别的知识，CSA损失是在两个分支之间的相同模态logit上的影响，该分支的表述为：

$$\mathcal{L}_{csa} = \sum_{i=1}^N (\text{KL}(\mathbf{Z}_{sh,V}^i, \mathbf{Z}_{sp,V}^i) + \text{KL}(\mathbf{Z}_{sh,I}^i, \mathbf{Z}_{sp,I}^i)). \quad (19)$$

3.4. 模差减少 (MDR)

在这一部分，为了确保模态共享特征的不变表示形式，TGSA和CSA用于减少模态分支中的模态差异 - 如下：

$$\mathcal{L}_{mdr} = \mathcal{L}_{tgsa}^{(sh,V;sh,I)} + \sum_{i=1}^N \text{KL}(\mathbf{Z}_{sh,V}^i, \mathbf{Z}_{sh,I}^i). \quad (20)$$

这样，模态共享分支中的可见特征和红外功能就可以从功能级别和语义级别获得相互学习。它使两种模态具有相互对齐的信息，同时减轻模态差距并保持模态共享功能的不变性。

3.5. 优化

最终，通过不断从特定于模态特征的隐式歧视知识中不断提炼，并始终减少模式共享功能中的模态差异，我们可以实现更具歧视和不变的模态共享特征。

模型IDKL的总损失定义为：

$$\mathcal{L}_{total} = \mathcal{L}_b + \lambda_1 \mathcal{L}_{ip} + \lambda_2 \mathcal{L}_{tgsa} + \lambda_3 \mathcal{L}_{csa} + \mathcal{L}_{mdr}, \quad (21)$$

其中 λ_1 , λ_2 和 λ_3 是炒作参数，可以平衡单个损失项的贡献。

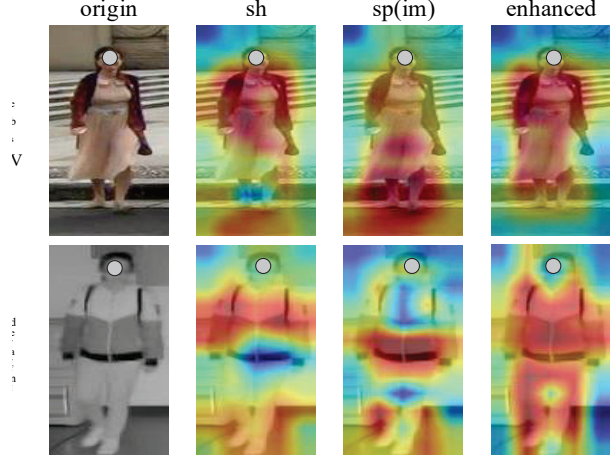


图4. 通过Grad-CAM观察隐式判别信息。和“sh”和“sp(im)”分别呈现训练有素的IDKL的模态特征和特定于知识蒸馏的模式特征；“增强”表示带有知识蒸馏的IDKL的模态共享特征。

4. 实验

4.1. 数据集和实验设置

数据集。使用三个公共VI-REID数据集Sysu-MM01 [28], LLCM [43]和REGDB [22]用于评估我们的模型。SYSU-MM01是一个受欢迎的大规模数据集，该数据集由四台可见摄像机和两个近红外摄像机收集，包括室内和室外环境。测试协议由全搜索和室内搜索组成。LLCM数据集是一个大规模且低光的跨模式数据集，该数据集以2: 1的比例分为训练和测试集。使用双相机系统收集REGDB，可见和红外图像成对捕获。LLCM和REGDB均包含侵权下可见的，并且可以看到两种搜索模式。

评估指标。标准级别 k 匹配的准确性和平均平均精度 (MAP) 被采用为评估指标。所有报告的结果都是10个试验的平均值。

实施详细信息。所提出的方法和所有实验均在具有Pytorch框架的单个NVIDIA GEFORCE 3090 GPU上实现。基线模型采用 \mathcal{L}_b 的ImageNet上预先训练的RESNET-50。输入图像的大小为 $3 \times 384 \times 128$ 。火车迷你批量大小设置为120，每个身份都包含12个随机身份和10张图像。使用初始学习率 3×10^{-5} 的Adam Optimizer被利用，在60次和100个时期内，衰减系数为0.1。Hyper-Parameters λ_1 , λ_2 和 λ_3 设置为0.1, 0.6和0.8。在测试阶段，仅使用模态共享的功能来评估性能。

表1. CMC (%) 和MAP (%) 性能与SYSU-MM01数据集中的最新方法的比较。

Methods	All-search								Indoor-search							
	Single-shot				Multi-shot				Single-shot				Multi-shot			
	r=1	r=10	r=20	map	r=1	r=10	r=20	map	r=1	r=10	r=20	map	r=1	r=10	r=20	map
Zero-Padding [32]	14.80	54.12	71.33	15.95	19.13	61.40	78.41	10.89	20.58	68.38	85.79	26.92	24.43	75.86	91.32	18.86
D-HSME [8]	20.68	62.74	77.95	23.12	-	-	-	-	-	-	-	-	-	-	-	-
AlignGAN [28]	42.40	85.00	93.70	40.70	51.50	89.40	95.70	33.90	45.90	87.60	94.40	54.30	57.10	92.70	97.40	45.30
DDAG [36]	54.75	90.39	95.81	53.02	-	-	-	-	61.02	94.06	98.41	67.98	-	-	-	-
NFS [2]	56.91	91.34	96.52	55.45	63.51	94.42	97.81	48.56	62.79	96.53	99.07	69.79	70.03	97.70	99.51	61.45
PIC [47]	57.51	89.35	95.03	55.14	-	-	-	-	60.40	-	-	67.70	-	-	-	-
MID [12]	60.27	92.90	-	59.40	-	-	-	-	64.86	96.12	-	70.12	-	-	-	-
cm-SSFT [21]	61.60	89.20	93.90	63.20	63.40	91.20	95.70	62.00	70.50	94.90	97.70	72.60	73.00	96.30	99.10	72.40
MCLNet [7]	65.40	93.33	97.14	61.98	-	-	-	-	72.56	96.98	99.20	76.58	-	-	-	-
FMCNet [42]	66.34	-	-	62.51	73.44	-	-	56.06	68.15	-	-	74.09	78.86	-	-	63.82
SMCL [31]	67.39	92.87	96.76	61.78	72.15	90.66	94.32	54.93	68.84	96.55	98.77	75.56	79.57	95.33	98.00	66.57
CAJ [37]	69.88	95.71	98.4	66.89	-	-	-	-	76.26	97.88	99.49	80.37	-	-	-	-
MPANet [33]	70.58	96.21	98.80	68.24	75.58	97.91	99.43	62.91	76.74	98.21	99.57	80.95	84.22	99.66	99.96	75.11
CMT [14]	71.88	96.45	98.87	68.57	80.23	97.91	99.53	63.13	76.9	97.68	99.64	79.91	84.87	99.41	99.97	74.11
DEEN [43]	74.7	97.6	99.2	71.8	-	-	-	-	80.3	99.0	99.8	83.3	-	-	-	-
SAAI [3]	75.90	-	-	77.03	82.8	-	-	82.39	83.20	-	-	88.01	90.73	-	-	91.30
MUN [40]	76.24	97.84	-	73.81	-	-	-	-	79.42	98.09	-	82.06	-	-	-	-
MSCLNet [45]	76.99	97.63	99.18	71.64	-	-	-	-	78.49	99.32	99.91	81.17	-	-	-	-
PartMix [16]	77.78	-	-	74.62	80.54	-	-	69.84	81.52	-	-	84.38	87.99	-	-	79.95
IDKL(Ours)	81.42	97.38	98.89	79.85	84.34	98.89	99.73	78.22	87.14	98.28	99.26	89.37	94.30	99.71	99.93	88.75

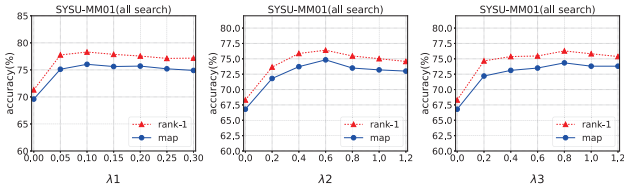


图5. 对于 \mathcal{L}_{ip} , \mathcal{L}_{tgsa} , \mathcal{L}_{tgsa} 和 \mathcal{L}_{csa} 的 λ_2 和 \mathcal{L}_{csa} 分别在sysu-mm01 dataset上分别为 \mathcal{L}_{ip} , λ_3 , λ_3 的消融分析。

4.2. 与最先进的方法进行比较

我们将我们的IDKL模型与近年来在公共VI-REID数据集SYSU-MM01, REGDB和LLCM上发布的最新VI-Reid方法进行了比较。

SYSU-MM01数据集的比较。比较实验结果显示在选项卡中。1显示所提出的IDKL方法的表现优于现有的剪边方法。具体而言, IDKL方法通过单拍的所有搜索协议实现了81.42%的排名1和79.85%的映射, 而单张室内搜索pro-tocol的精度为87.14%等级1和89.37%的映射。比较的SOTA包括各种基本甲基化*i.e.*。基于图形结构增强方法DDAG [36]和逻辑技能的相互学习 (PIC [47]和MPANET [33])。与其他几种基于模式的方法C M-SSFT [21], MUN [40]和MSCLNET [45]相比, 我们的结果比它们的表现优于它们。

REGDB数据集的比较。我们还在小型数据集REGDB上评估了IDKL, 如选项卡所示。2。表现出强大的性能IDKL表现并胜过现有解决方案。具体而言, 我们实现了排名-1 ac-

在可见的红外模式下可见的94.72%, 在红外线上的排名1精度为94.22%。

LLCM数据集的比较。IDKL模型在大型且复杂的LLCM数据集上实现了重大改进, 如选项卡所示。3, 在两种模式下, 表现出极好的排名1精度为72.2%和70.7%。这表明IDKL模型在复杂和多模式方案中表现出很强的鲁棒性。

表2. 在REGDB数据集中使用最新方法的CMC (%) 和MAP (%) 性能的比较。

Methods	Visible to infrared		Infrared to visible	
	rank-1	map	rank-1	map
Zero-Padding [32]	17.8	18.9	16.7	17.9
AlignGAN [28]	57.9	53.6	56.3	53.4
DDAG [36]	69.34	63.46	68.06	61.80
cm-SSFT [21]	72.3	72.9	71.0	71.7
MCLNet [7]	80.31	73.07	75.93	69.49
PIC [47]	83.6	79.6	79.5	77.4
MPANet [33]	83.7	80.9	82.8	80.7
SMCL [31]	83.93	79.83	83.05	78.57
MSCLNet [45]	84.17	80.99	83.86	78.31
CAJ [37]	85.03	77.82	84.75	77.82
MID [12]	87.45	84.85	84.29	81.41
FMCNet [42]	89.12	84.43	88.38	83.86
SAAI [3]	91.07	91.45	92.09	92.01
DEEN [43]	91.1	85.1	89.5	83.4
CMT [14]	95.17	87.3	91.97	84.46
MUN [40]	95.19	87.15	91.86	85.01
IDKL(Ours)	94.72	90.19	94.22	90.43

4.3. 消融研究

在本小节中, 我们进行了消融实验, 以详尽地评估我们提出的模型。

表3. 与LLCM数据集上最新方法的CMC (%) 和MAP (%) 性能的比较。

Methods	Visible to infrared		Infrared to visible	
	rank-1	map	rank-1	map
DDAG [36]	40.3	48.4	48.0	52.3
CAJ [37]	56.5	59.8	48.8	56.6
DEEN [43]	62.5	65.8	54.9	62.9
IDKL(Ours)	72.22	66.43	70.72	65.19

表4. 评估不同组件对级别1 (%) 和MAP (%) 对SYSU-MM01数据集的影响。

\mathcal{L}_b	\mathcal{L}_{ip}	\mathcal{L}_{tgsa}	\mathcal{L}_{csa}	\mathcal{L}_{mdr}	SYSU-MM01	
					rank1	map
✓	×	×	×	×	67.60	66.47
✓	×	✓	×	×	68.38	67.30
✓	✓	✓	×	×	76.40	74.83
✓	✓	×	✓	×	76.28	74.34
✓	✓	×	×	✓	77.04	75.66
✓	✓	✓	✓	✓	81.42	79.85

每个组件的有效性。我们在所有搜索单拍模式下评估了每个组件在SYSU-MM01数据集上的有效性。每个组合都可以独立添加，以显示其作为选项卡的性能。4。这表明每个组件都非常有用，并且 \mathcal{L}_{mdr} 可实现显著的性能提高。这进一步表明TGSA和CSA在减少模态差异中的有效性。将第二行与第三行进行比较，证明了净化隐式模式特定信息的必要性，并证明了我们信息净化器模块的有效性。

IP, TGSA和CSA的超参数分析。在这一部分中，我们提出了一个线图，以通过逐渐增加超参数的价值来检查IP, TGSA和CSA的细节。如图5所示，IP, TGSA和CSA的最大贡献分别在0.1、0.6和0.8达到。曲线的上升趋势表明了每个模块的有效性。

4.4. 可视化分析

注意地图可视化。为了进一步说明IDKL的有效性，将Grad-CAM [26]用于对不同特征热图进行视频检查。在图4中，隐式判别信息的重点领域通常与共享功能的重点区域不同，这表明有效，明智地利用此信息来增强共享特征可能是非常有益的。

特征分布可视化。我们利用T-SNE [22]特征图可视化来观察IDKL模型和TGSA对模型的影响。如图6所示，每种颜色代表不同的身份，而圆形和三角形的形状分别表示可见和循环的模态信息。从图6 (a) ，

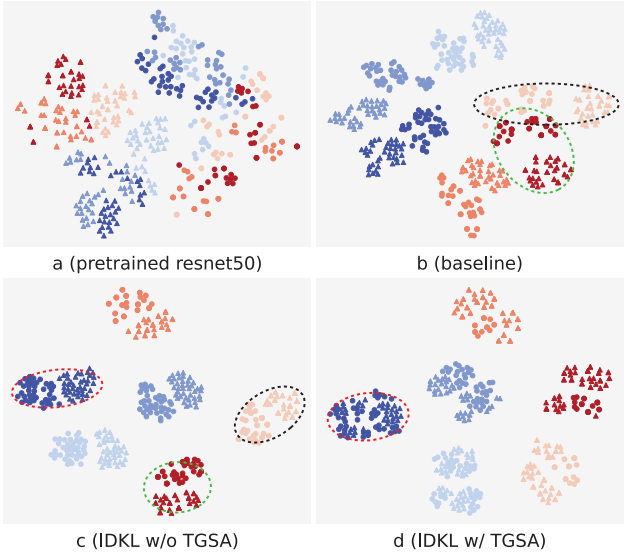


图6. t-sne对学习特征的可视化。其中“a”是指影像网上的预训练resnet50；“B”代表基线；“c”是w/o tgsa的IDKL模型；“d”是w/ tgsa的IDKL型号。

两种方式之间存在显著差异。图6 (b) 显示了模态差异的降低，模型表现出一些歧视能力。将图6 (b) 与图6 (c) 进行比较，可以观察到IDKL模型具有较小的类内部差异和阶层间歧视。在图6 (d) 中，暗蓝色类别比图6 (c) 更分散，并且两种方式的图形结构更相似且紧密地对齐，这表明了TGSA在降低模态差异中的有效性。

5. 结论

本文在特定于模态特征中利用了隐式歧视性信息，并引入了强大的模型IDKL，以利用异质相关特征的潜在区分并增强共享特征。IDKL模型包括双重单流净作品，一种新颖的内向信息净化器，三重态图结构对齐解决方案以及对数字的精制蒸馏。总的来说，这些组件表现出非凡的有效性，并有助于改善结果。

致谢

这项工作得到了中国国家密钥研发计划（2021YFB3100800），国家自然科学基金（62271090，61771079），重庆自然科学基金（CSTC2021JCYJCYJ-JQX0023）和国家青年人才项目的部分支持。这项工作还得到了重庆人工中心创新中心的华为计算能力的支持。

参考

[1] Cuiqun Chen, Mang Ye, Meibin Qi, Jingjing Wu, Jianguo Jiang和Chia-wen Lin. 可见的红外人员重新识别的结构意识到位置跨前期。 *IEEE Transactions on Image Processing*, 31: 2352–2364, 2022。2 [2] Yehansen Chen, Lin Wan, Zhihang Li, Qianyan Jing和Zongyuan Sun. 神经功能搜索RGB-Infrared人的重新识别。在2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*中, 第587-597、2021页。1, 7 [3] Xingye Fang, Yang Yang和Ying Fu. 通过语义对齐和亲和力重新识别可见的红外命名。在 *Proceedings of the IEEE/CVF International Conference on Computer Vision*中, 第11270–11279、2023页。7[4] Chaoyou Fu, Yibo Hu, Xiang Wu, Hailin Shi, Tao Mei, 然后跑了他。CM-NAS: 跨模式神经建筑搜索可见的红外人重新识别。在2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*中, 第11803–11812、2021页。1[5] xiaowei fu, fuxiang huang, yuhang zhou, huimin a, huimin ma, xin xu和lei zhang. 跨模式的跨模式跨域双对准网络, 用于重新识别RGB-Infrared人。 *IEEE Transactions on Circuits and Systems for Video Technology*, 32 (10): 6874–6887, 2022。2 [6] Yaroslav Ganin和Victor Lempitsky. 无监督的域通过反向传播适应。在 *International conference on machine learning*中, 第1180–1189页。PMLR, 2015。4[7] Xin Hao, Sanyuan Zhao, Mang Ye和Jianbing Shen. 通过模态融合和中心聚合重新识别跨模式的人。在2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*中, 第16383–16392、2021、2021、7 [8] Yi Hao, Nannan Wang, Jie Li和Xinbo Gao. HSME: 可见的热人识别的Hy-persphere歧管嵌入。在 *Proceedings of the AAAI conference on artificial intelligence*中, 第8385–8392、2019。1, 7 [9] 关闭He, Hao Luo, Pichao Wang, Fan Wang, Hao Li和Wei Jiang. TransReid: 基于变压器的对象重新标识。在2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*中, 第14993–15002、2021页。2[10] Jie Hu, Li Shen和Gang Sun. 挤压网络工作。在 *Proceedings of the IEEE conference on computer vision and pattern recognition*中, 第7132–7141、2018页。4[11] Xun Huang and Serge属于。自适应实例归一化的任意样式转移。在 *Proceedings of the IEEE international conference on computer vision*, 第1501–1510页, 2017年。4[12] Zhipeng Huang, Jiawei Liu, Liang Li, Kecheng Zheng和Zheng-Jun Zha. RGB-Infrared人重新识别的模态自适应混合和不变构成。在 *Proceedings of the AAAI Conference on Artificial Intelligence*中, 第1034–1042、2022页。7[13] Jieru Jia, Qiuqi Ruan和Timothy M Hospedales. 简单的人重新识别: 在实践中概括人重新识别。 *arXiv preprint arXiv:1905.03422*, 2019年。4[14] Kongzhu Jiang, Tianzhu Zhang, Xiang Liu, Bingqiao Qian, Yongdong Zhang和Feng Wu. 跨模式变压器

用于可见的红外人员重新识别。在 *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIV*中, 第480–496页。Springer, 2022。1, 2, 3, 7 [15] Xin Jin, Cuiling Lian, Wenjun Zeng, Zhibo Chen和Li Zhang. 风格的正常化和恢复原状, 以重新识别可推广的人。在 *proceedings of the IEEE/CVF conference on computer vision and pattern recognition*中, 第3143–3152、2020年。4[16] Minsu Kim, Seungryong Kim, Jungin Park, Seongheon Park和Kwanghoon Sohn. PartMix: 正式化策略要学习可见的红外人识别的部分发现。在 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*中, 第18621–18632年, 2023年。7[17] Xulin Li, Yan Lu, Yan Lu, bin Liu, Yating Liu, Yating Liu, Qi Chu, Qi Chu, Qi Chu, Jinyang Huang, Feng Zhu, Feng Zhu, Rui Zhao, Rui Zhao和Nenghai Yu. 可见红外人员重新识别的反事实干预特征转移。在 *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXVI*中, 第381–398页。施普林格, 2022年, 1, 5 [18] Wenqi Liang, Guangcong Wang, Jianhuang Lai和Xiao-Hua Xie. 均质到异质: RGB - 含红外人的无监督学习重新识别。 *IEEE Transactions on Image Processing*, 30: 6392–6407, 2021。2 [19] Haijun Liu, Xiaoheng Tan和Xichuan Zhou. 可见热人重新识别的参数共享探索和异核中心三重损失。 *IEEE Transactions on Multimedia*, 23: 4414–4425, 2021。1 [20] Jialun Liu, Yifan Sun, Feng Zhu, Hongbin Pei, Yi Yang和Wenhui Li. 学习记忆启动的跨模式人员重新识别的Unidirection指标。在2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*中, 第19344–19353、2022页。2[21] Yan Lu, Yue Wu, bin Liu, Tianzhu Zhang, Baopu Li, Qi Chu, Qi Chu和Nenghai Yu. 跨模式的人重新识别, 具有特定于共享的特征转移。在2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 第13376–13386、2020年。2, 7 [22] dat tien nguyen, Hyung Gil Hong, Ki Wan Kim和Kang Ryoung Park. 人识别系统基于可见光和热摄像机的身体图像的组。 *Sensors*, 17 (3): 605, 2017。6, 8 [23] Xingang Pan, Ping Luo, Jianping Shi和Xiaoou Tang. 同时两个: 通过IBN-NET增强学习和概括能力。在 *Proceedings of the European Conference on Computer Vision (ECCV)*中, 第464–479、2018页。4[24] Filip Radenović, Giorgos Tolias和Ondřej Chum. 精细调整CNN图像检索, 没有人类注释。 *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41 (7): 1655–1668, 2019。2, 4 [25] Min Ren, Lingxiao He, Xingyu Liao, Wu Liu, Yunlong Wang和Tieniu Tan. 学习实例级的空间 - 重新识别的时间模式。在2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, 第14910–14919、2021页。2

[26] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh和Dhruv Batra. Grad-CAM: 通过基于梯度的本地化, 深网的视觉解释。在 *Proceedings of the IEEE international conference on computer vision* 中, 第618–626页, 2017年。8 [27] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian和Shengjin Wang. 超越零件模型: 带有精制零件合并的人员检索 (以及强烈的卷积基线)。在 *Proceedings of the European conference on computer vision (ECCV)* 中, 第480–496, 2018页。2 [28] Guan'an Wang, Tianzhu Zhang, Jian Cheng, Si Liu, Yang Yang Yang and Zengguang Hou. 通过联合像素和特征对齐方式重新识别RGB - 信号交叉模式。在 *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* 中, 第3622–3631, 2019。2、6、7 [29] Guan-An Wang, Tianzhu Zhang, Yang Yang Yang, Jian Cheng, Jianlong Chang, Xu Liang和Zeng-Guang Hou. RGB-Infrared人重新识别的跨模态成对图像生成。在 *Proceedings of the AAAI conference on artificial intelligence* 中, 第12144–12151, 2020页。1 [30] Zhixiang Wang, Zheng Wang, Yinqiang Zheng, Yung-Yu Chuang和Shin'ichi Satoh. 学会减少可见红外可见人重新识别的双重差异。在 *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 中, 第618–626页, 2019年。1, 2 [31] Ziyu Wei, Xi Yang, Nannan Wang和Xinbo Gao. 可见红外人员重新识别的合成方式协作学习。在 *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* 中, 第225–234, 2021页。7 [32] Ancong Wu, Wei-Shi Zheng, Hong-Xing Yu, Shaogang Gong和Jianhuang Lai. RGB - 信号交叉模式重新识别。在 *2017 IEEE International Conference on Computer Vision (ICCV)* 中, 第5390–5399, 2017页。2, 7 [33] Qiong Wu, Pingyang Dai, Jie Chen, Chia-Wen Lin, Yongjian Wu, Feiyue Huang, Bineng Zhong Zhong和Rongongrong Ji. 可见的红外人员识别的跨模式细微差别。在 *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 中, 第4328–4337, 2021页。1、3、7 [34] Yang Yang, Tianzhu Zhang, Jian Cheng, Zengguang Hou, Prayag Tiwari, Hari Mohan Pandey等。跨模式成对图像的生成和RGB-Infrared人的增强性重新识别。 *Neural Networks*, 128: 294–304, 2020。1 [35] Mang Ye, Xiangyuan Lan, Qingming Leng和Jianbing Shen. 跨模式的人通过模态意识合作集合学习重新识别。 *IEEE Transactions on Image Processing*, 29: 9387–9399, 2020。2, 3 [36] Mang Ye, Jianbing Shen, David J. Crandall, Ling Shao和Jiebo Luo. 为可见的边红色人重新识别的动态双重聚合学习。在 *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVII 16*, 第229–247页。Springer, 2020。2, 5, 7, 8 [37] Mang Ye, Weijian Ruan, Bo Du和Mike Zheng Shou. 通道增强联合学习以进行可见的红外识别

本质。在 *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* 中, 第13547–13556, 2021页。7, 8 [38] Mang Ye, Jianbing Shen和Ling Shao. 可见的红外人通过均质增强三模态学习重新识别。 *IEEE Transactions on Information Forensics and Security*, 16: 728–739, 2021。1 [39] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao和Steven C. H. Hoi. 对人的深入学习: 调查和前景。 *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 (6): 2872–2893, 2022。2, 4 [40] Hao Yu, Xu Cheng, Wei Peng, Weihao Liu和Guoying Zhao. 可见的红外人员重新识别的模态统一网络。在 *Proceedings of the IEEE/CVF International Conference on Computer Vision* 中, 第11185–11195, 2023页。7 [41] Lei Zhang, Zhipu Liu, Wensheng Zhang和David Zhang. 基于风格的不确定性基于自定义的元学习, 可用于重新识别一般人。 *IEEE Transactions on Image Processing*, 32: 2107–2119, 2023。2 [42] Qiang Zhang, Changhou Lai, Jianan Liu, Nianchang Huang和Jungong Han. FMCNET: 可见的红外人员重新识别的特征级别的方式补偿。在 *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 中, 第7339–7348, 2022页。1, 7 [43] Yukang Zhang和Hanzi Wang. 可见的红外人员重新识别的多样化的嵌入式扩展网络和低光跨模式基准。在 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 第2153–2162, 2023, 2023页。6、7、8 [44] Ying Zhang, Tao Xiang, Timothy M Hospedales和Huchuan Lu. 深厚的相互学习。在 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 第4320–4328, 2018页。2, 3 [45] Yiyuan Zhang, Sanyuan Zhao, Yuhao Kang和Jianbing Shen. 模态协同补充学习与级联的聚合, 可为可见的红外人重新识别。在 *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIV*, 第462–479页。施普林格, 2022。1, 2, 7 [46] 空间通道增强的变压器, 用于可见的红外人员识别。 *IEEE Transactions on Multimedia*, 第1–1页, 2022年。2 [47] Xiangtao Zheng, Xiumei Chen和Xiaoqiang Lu. 可见的红外人员通过部分互动式进行重新识别。 *IEEE Transactions on Image Processing*, 31: 6951–6963, 2022。2, 3, 7 [48] Kaiyang Zhou, Yongxin Yang, Andrea Cavallaro和Tao Xiang. OMNI规模的人学习特征学习。在 *Proceedings of the IEEE/CVF international conference on computer vision* 中, 第3702–3712, 2019。2, 4 [49] Yuanxin Zhu, Zhao Yang, Li Wang, Sai Zhao, Xiao Hu和Dapeng Tao. 跨模式人员重新识别的异性中心损失。 *Neurocomputing*, 386: 97–109, 2020。