E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035





**中图法分类号**:TP391.41;TP18　　**文献标识码**: A　　**文章编号**: 1006-8961(2026)06-1795-28

**论文引用格式**:Han Q, Li L F and Min W D. 2026. Research progress on object re-identification in special scenarios. Journal of Image and Graphics, 31(6):1795-1822(韩清, 李龙飞, 闵卫东. 2026. 特殊场景下的目标重识别研究进展. 中国图象图形学报, 31(6):1795-1822)[DOI:10.11834/jig.260117]

# 特殊场景下的目标重识别研究进展

韩清[1,2,3],李龙飞[1],闵卫东[1,2,3*]

1. 南昌大学数学与计算机学院,南昌　330031; 2. 南昌大学元宇宙研究院,南昌　330031;
3. 虚拟现实江西省重点实验室,南昌　330031

**摘　要**: 目标重识别旨在跨摄像头、跨时间的条件下实现对特定目标的检索与匹配,是支撑智能安防、智慧交通和智慧城市等领域的关键技术。经过十多年的持续关注与研究,目标重识别领域发展迅速,技术体系不断完善。本文系统性介绍了目标重识别背景,包括目标重识别的发展历程、数据集与评价指标。面向现实世界的落地应用需求与挑战,总结并分析了无监督目标重识别、多光谱目标重识别、跨模态行人重识别、遮挡行人重识别、换装行人重识别以及小股行人重识别等特殊场景下的研究,归纳其发展现状,并对每个研究方向的前沿方法进行梳理与性能对比。本文也对动物重识别的发展现状进行了简介。最后,对目标重识别的发展趋势进行分析与展望。

**关键词**: 特殊场景下的目标重识别;深度学习;无监督目标重识别;多光谱目标重识别;跨模态行人重识别;遮挡行人重识别;换装行人重识别;小股行人重识别;动物重识别

# Research progress on object re-identification in special scenarios

Han Qing[1,2,3], Li Longfei[1], Min Weidong[1,2,3*]

*1. School of Mathematics and Computer Science, Nanchang University, Nanchang　330031, China; 2. Institute of Metaverse, Nanchang University, Nanchang　330031, China; 3. Jiangxi Provincial Key Laboratory of Virtual Reality, Nanchang　330031, China*

**Abstract**: Object re-identification (Re-ID) is a pivotal technology for retrieving and associating specific targets across non-overlapping cameras and different time spans, and it serves as a core component of intelligent surveillance, smart transportation, and smart-city sensing systems. Over the past decade, Re-ID research has evolved dramatically, moving from handcrafted features in its early stage to deep-learning-based automatic representation learning. With the recent emergence of large-scale vision-language models, the field is now advancing toward more generalizable and versatile solutions. This paper provides a systematic review of the foundations of object Re-ID, including its developmental trajectory, mainstream datasets, and evaluation metrics. It then focuses on six specialized directions that address challenges arising in complex real-world deployment: unsupervised object Re-ID, multi-spectral object Re-ID, cross-modality person Re-ID, occluded person Re-ID, clothes-changing person Re-ID, and group person Re-ID. For each direction, we summarize research progress and development trends, analyze representative state-of-the-art methods, and compare their performance. We also briefly review the current status of animal Re-ID, which is of substantial ecological importance. First, to address the high cost of manual annotation in practice, unsupervised object Re-ID aims to reduce or eliminate dependence on labeled data.

**收稿日期**:2026-03-04;**修回日期**:2026-03-18;**预印本日期**:2026-03-25
**所有作者为共同一作 * 通信作者**:闵卫东　minweidong@ncu.edu.cn
**基金项目**:国家自然科学基金项目(62166026, 62076117);虚拟现实江西省重点实验室(2024SSY03151)

**Supported by**: National Natural Science Foundation of China (62166026, 62076117); Jiangxi Provincial Key Laboratory of Virtual Reality (2024SSY03151)

This survey reviews two major streams in this area: domain-adaptation-based methods and fully unsupervised methods. Both typically rely on iterative optimization strategies in which clustering is used to generate pseudo-labels for unlabeled data. Current research focuses on mitigating label noise and learning more discriminative feature representations within unsupervised frameworks. Second, to overcome the limitations of visible-light imagery in adverse conditions such as night-time or heavy fog, multi-spectral object Re-ID integrates complementary cues from visible, near-infrared, and thermal infrared spectra. Current work mainly addresses the distribution gaps among spectral bands in order to exploit their complementarity and achieve effective cross-spectral alignment and aggregation. Third, we present an in-depth review of cross-modality person Re-ID, which performs retrieval using heterogeneous data beyond visible-light images. Based on the modalities involved, related studies can be divided into three sub-tasks. Visible-thermal Re-ID addresses the modality gap between day and night surveillance; mainstream methods focus on either style alignment through image generation or modality-invariant feature learning. Text-to-image Re-ID uses natural-language descriptions as queries when query images are unavailable, and current work emphasizes fine-grained semantic-visual interaction mechanisms. Sketch-to-photo Re-ID enables retrieval from intuitive visual sketches; its key challenge is the large stylistic variation inherent in hand-drawn sketches, and existing approaches therefore focus on improving style generalization. Fourth, because occlusion is pervasive in real-world scenarios, occluded person Re-ID has become an important research topic. We categorize mainstream methods into two groups. Visibility-aware methods improve matching by exploiting information from unoccluded regions, typically using human structural priors to locate visible parts or enabling the model to focus adaptively on visible discriminative cues. Feature-reconstruction-based methods instead seek to restore or hallucinate missing pedestrian information in occluded regions so that the resulting features are more discriminative. Fifth, to address the significant intra-class variation caused by clothing changes in long-term Re-ID, this paper reviews clothes-changing person Re-ID. Existing methods fall into two broad categories: those that extract clothing-agnostic features, such as facial characteristics, gait, and body shape, and those that explicitly disentangle clothing information from identity information. Sixth, because pedestrians often appear in groups in real scenes, we review group person Re-ID, which seeks to associate group images across cameras. In addition to the challenges found in single-person Re-ID, this task must handle variations in group membership, mutual occlusion, and changes in group layout. Current methods mainly use graph neural networks to model topological relationships among members or transformer-based architectures to learn higher-order statistical features and uncertainty representations for robust group matching. In addition, we provide a brief introduction to animal Re-ID, a field of growing importance for ecological conservation. This section outlines the technological evolution from traditional marker-based methods to modern deep-learning approaches and highlights the key shift toward non-invasive monitoring. Finally, the paper summarizes the development and trends of object Re-ID and discusses future directions, including the construction of large-scale datasets, research on large foundation models for general object Re-ID, and lightweight deployment strategies. These directions are essential for improving practical applicability and accelerating broad real-world adoption.

**Key words**: object Re-ID in special scenarios; deep learning; unsupervised object Re-ID; multi-spectral object Re-ID; cross-modal person Re-ID; occluded person Re-ID; clothes-changing person Re-ID; group person Re-ID; animal Re-ID

## 0 引 言

目标重识别(object re-identification,Re-ID)是计算机视觉中细粒度识别的关键分支,其核心目标是在非重叠的摄像头检索与查询目标相同的特定个体(Ye等,2022;Han等,2022;Wang等,2020b)。目标重识别可与目标检测、追踪等技术紧密结合,广泛应用于智能安防、智慧交通、智慧城市乃至野生动物监测等领域。通过实现跨摄像头、跨时空的个体关联,该技术为构建连续、智能的视觉感知系统提供核心支撑,使得对行人、车辆等目标的长期追踪与行为分析成为可能,具有重要广阔的应用前景。

作为人工智能领域计算机视觉的热门方向,目标重识别在过去十多年间吸引了学术界和工业界广泛而持续的关注,其技术体系日益成熟。该领域的研究范围也不断拓宽,其作用对象也从早期以行人、车辆为核心,逐步延伸并涵盖了动物、建筑等更为多样化的类别(Ye等,2025;孙伟 等,2025),展现出持续拓展的应用潜力。目标重识别技术的一般流程可

分为3个核心步骤：1)构建数据集，在监控场景下采集视频数据，并利用目标检测算法(Redmon等，2016；Ren等，2017)检测与截取包含目标的区域并进行人工或自动的身份标注，从而构建有标注的目标数据集；2)训练模型，利用步骤1)获取的有标注数据集，训练出具有判别性特征提取与相似性度量能力的重识别模型，从而能够区分不同身份的目标；3)检索推理，给定查询图像，基于步骤2)已训练的重识别模型，将其与来自不同摄像头的大量候选图像进行特征提取与相似度计算，返回最匹配的若干结果，以此实现跨摄像头的目标重识别。目标重识别的发展大致可分为早期基于手工特征和当前基于深度学习两大阶段。早期基于传统手工特征的阶段，研究主要依赖设计颜色、纹理和形状等特征，并结合度量学习进行匹配。而后进入基于深度学习的阶段，研究利用深度神经网络端到端地从目标图像中学习判别性特征，并计算特征相似度，从而直接完成重识别任务。与此同时，一系列大规模、高质量数据集的构建，也为深度学习模型的训练与评估提供了关键支撑，共同推动了重识别性能的显著提升。而大模型技术也为目标重识别领域带来了重要进展。因其强大的视觉表征能力与语义理解能力，模型在复杂场景与跨域任务中的泛化性能得到显著提升，从而推动该领域迈向更加通用、灵活的方向发展。

然而，目标重识别在现实应用的场景中仍面临诸多挑战，如标注数据稀缺、光照变化、目标遮挡以及外观改变等，导致现有方法在应对这些复杂情况时往往难以满足高鲁棒性、高精度的实际应用需求。因此，越来越多的研究人员致力于解决实际应用场景中的挑战，由此催生了一系列面向特殊场景的目标重识别研究，包括无监督目标重识别、多光谱目标重识别、跨模态行人重识别、遮挡行人重识别、换装行人重识别以及小股行人重识别。此外，重识别技术的研究对象也从行人、车辆扩展至动物等更广泛的类别，进而衍生出动物重识别这一研究领域。

本文首先系统性介绍目标重识别背景，包括目标重识别的发展历程、数据集与评价指标。然后总结并分析无监督目标重识别、多光谱目标重识别、跨模态行人重识别、遮挡行人重识别、换装行人重识别以及小股行人重识别等特殊现实场景下的研究，归纳其发展现状。同时，本文也对动物重识别的发展现状进行简介。最后对目标重识别的发展趋势进行分析与展望。

# 1 目标重识别背景

## 1.1 目标重识别发展历程

目标重识别主要研究对象为行人和车辆，其发展历程大致可分为早期基于手工特征方法的阶段与当前基于深度学习方法的阶段。在上世纪90年代，行人和车辆重识别已受到关注(Cai和Aggarwal，1996；Huang和Russell，1997)。行人重识别的研究起步相对较早，Gheissari等人(2006)在2006年首次提出行人重识别的概念。随后，Gray和Tao(2008)提出VIPeR(viewpoint invariant pedestrian recognition)行人重识别数据集，为后续研究提供了重要的参考和基准。在早期手工特征阶段，研究主要利用颜色与纹理特征(Farenzena等，2010；Zhao等，2013；Yang等，2014)、形状特征(Oreifej等，2010)、局部不变性特征(Jüngling等，2011)以及人工定义的语义属性(Layne等，2012)来构建行人的表征信息，并结合度量学习(Köstinger等，2012；Xiong等，2014)实现行人重识别任务(张永飞 等，2023)。相比之下，车辆重识别在同期虽然有一些基于手工特征(Feris等，2012；Zheng等，2015b)的初步探索，但由于缺乏清晰的任务定义与统一的评估体系，导致系统性研究开展得相对较晚。

然而，基于手工特征的方法存在主观性强、泛化能力受限等固有缺陷，难以应对复杂多变的真实场景。而随着深度学习在计算机视觉中展现出的强大表征学习能力，早期代表性工作如DeepReID(deep person re-identification)(Li等，2014)将深度学习应用于行人重识别，并取得比手工特征方法更为优异的性能，从而使深度学习方法逐渐成为该研究领域的主流。随后，深度学习的方法也扩展至车辆重识别领域，其发展的关键转折点是2016年VeRi(vehicle re-identification)(Liu等，2016b)和VehicleID(Liu等，2016a)专用数据集的发布。这些数据集为任务提供了明确的定义与评测基准，从而推动了基于深度学习的车辆重识别研究进入快速发展阶段。在基于深度学习阶段，研究人员借助卷积神经网络(convolutional neural network，CNN)(He等，2016)及视觉Transformer(Dosovitskiy等，2021)等模型，通过端到端的方式学习判别性特征，并在此基础上围绕

网络结构、损失函数等方面不断优化,大量相关算法不断涌现。与此同时,一系列大规模、高质量数据集的构建与公开,也显著推进了目标重识别的发展。

随着大模型技术的快速发展,目标重识别领域的研究方法也发生了显著变革(冯展祥 等,2025)。一方面,LUPerson(Fu 等,2021)等大规模无标记行人数据集的构建,为基于无监督与自监督预训练的研究提供了重要基础。先前,使用在 ImageNet(Deng 等,2009)上预训练的权重来初始化模型是目标重识别领域的标准方案。然而,ImageNet 作为通用计算机视觉数据集,其数据分布与以行人为主的重识别任务之间存在显著的域差异。相比之下,LUPerson 等专用数据集在行人身份密度、场景多样性和数据规模上更具针对性,大幅提升重识别模型在下游多个基准数据集的泛化能力。另一方面,CLIP(contrastive language-image pre-training)(Radford 等,2021)等视觉—语言大模型的兴起,其强大的通用视觉表征能力为目标重识别模型提供高质量的预训练权重,或作为现成的视觉编码器直接提取鲁棒的特征。同时,这类模型也推动了利用文本信息来增强模型表征学习的重识别方法研究(Li 等,2023)。此外,ChatGPT(chat generative pre-trained transformer)(Radford 等,2018)等大语言模型在语义理解与推理上的能力,也开始被探索用于解析复杂查询、生成判别性描述,以构建更智能的交互式重识别系统(Zhai 等,2024)。

随着目标重识别算法在主流数据集上的深入研究,其在现实世界的复杂特殊场景中的鲁棒性与泛化能力受到广泛关注。研究者围绕实际应用中的核心难题,拓展了多个针对性研究领域。无监督目标重识别旨在缓解对大规模标注数据的依赖;多光谱与跨模态目标重识别致力于异构数据间的可靠匹配;遮挡与换装行人重识别聚焦于目标外观显著变化下的鲁棒识别;小股行人重识别则面向多人场景中的群体关联与身份推断。同时,研究对象也从行人、车辆扩展至动物等更广泛的类别。在此基础之上,构建一个能自适应多场景、统一多任务的通用重识别框架正成为研究热点,已有工作(He 等,2024d;Zheng 等,2025)对此进行了探索。总体而言,目标重识别正朝着利用通用先验知识、融合多模态信息、具备开放式感知与推理能力的通用、灵活方向持续发展。

### 1.2 目标重识别数据集

为推进目标重识别技术的发展,研究者们构建并公开了一系列数据集以提供标准评估基准。常用的行人重识别数据集有 Market-1501(Zheng 等,2015a)、DukeMTMC-reID(Duke multi-target, multi-camera for person re-identification)(Ristani 等,2016)和 MSMT17(multi-scene multi-time 17)(Wei 等,2018)。常用的车辆重识别数据集有 VeRi-776(Liu 等,2018)、VehicleID(Liu 等,2016a)和 VERI-Wild(vehicle re-identification in the wild)(Lou 等,2019)。

Market-1501 数据集包含 1 501 个行人的 32 668 幅图像。按身份划分,751 个行人对应的 12 936 幅图像构成训练集,750 个行人对应的 19 732 幅图像构成测试集。在测试集中,3 368 幅图像被设定为查询集。

DukeMTMC-reID 数据集包含 1 404 个行人的 36 411 幅图像,其中 702 个行人对应的 16 522 幅图像构成训练集。测试集则包含其余 702 个行人中的 2 228 幅图像构成的查询集和 17 661 幅图像构成的图库集。需要说明的是,DukeMTMC-reID 因伦理问题已被官方下架,在后续研究中已不再适合使用。

MSMT17 数据集由室内外多个摄像机拍摄到的 4 101 个行人对应的 126 441 幅图像组成,其数据规模相对较大且场景多样化,是最具挑战性的基准之一。训练集包含 1 041 个行人对应的 32 621 幅图像。测试集包含其余 3 060 个行人对应的 93 820 幅图像,其中查询图像有 11 659 幅。

VeRi-776 数据集是从 24 h 内覆盖 1 $km^2$ 实际交通监控区域的 20 个摄像头收集而来。该数据集总共包含了 776 辆车的 50 117 幅图像。其中训练集有 37 778 幅图像,查询集有 1 678 幅图像,图库集有 10 661 幅图像。

VehicleID 数据集包含 26 267 辆车的 221 763 幅图像,其中 13 134 辆车的 110 178 幅图像用于训练。为了在不同数据规模下评估性能,它包括 3 个不同规模的测试集:Test800、Test1600 和 Test2400。具体而言,Test800、Test1600 和 Test2400 分别包含 800、1 600 和 2 400 幅图像作为查询集,以及 6 532、11 395 和 17 638 幅图像作为图库集。

VERI-Wild 是一个大规模且目前最具挑战的车辆重识别数据集。它包含 40 671 辆车的 416 314 幅图像,这些图像由 174 个摄像头在一个月内收集而

来，涵盖背景、光照、视角和天气等各种环境因素。训练集包含30 671辆车的277 794幅图像。测试集分为3个子集：小子集包括3 000辆车的41 816幅图像，中子集包括5 000辆车的69 389幅图像，大子集包括10 000辆车的138 517幅图像。

### 1.3　目标重识别评价指标

目标重识别模型的性能通常通过以下几项指标进行综合评估：CMC（cumulative match characteristic）曲线、mAP（mean average precision）以及 mINP（mean inverse negative penalty）。

1）CMC 曲线。CMC 曲线基于一系列 Rank-$n$ 的值构建而成。Rank-$n$ 为前 $n$ 个检索结果存在正确样本的准确率，若存在，则为1，否则为0。最后，对所有查询图像的 Rank-$n$ 值求平均。

2）mAP。mAP 为所有查询图像的平均精度（average precision，AP）的均值，AP 计算为

$$f_{\mathrm{AP}}=\frac{1}{c}\sum_{i=1}^{c}\frac{i}{s_i}\tag{1}$$

式中，$c$ 为检索的正确样本总数，$s_i$ 为第 $i$ 个正确样本在检索中的位次。

3）mINP。mINP 是 Ye 等人（2022）提出的新评估指标，旨在衡量模型检索最难匹配正样本的能力，从而弥补 CMC 和 mAP 在此类性能衡量上的不足，计算为

$$f_{\mathrm{mINP}}=\frac{1}{N}\sum_{i=1}^{N}\frac{\left|G_i\right|}{R_i^{\mathrm{hard}}}\tag{2}$$

式中，$N$ 为查询图像总数，$\left|G_i\right|$ 为第 $i$ 幅查询图像的正确样本总数，$R_i^{\mathrm{hard}}$ 为最靠后的正确样本在检索中的位次。

## 2　特殊场景下的目标重识别

随着目标重识别技术向大规模复杂现实的特殊场景应用，其面临的挑战催生了多个关键研究方向。由于行人目标的复杂性与社会应用中的核心地位，面向行人的各类特殊场景重识别研究也更为丰富。本节旨在归纳无监督目标重识别、多光谱目标重识别、跨模态行人重识别、遮挡行人重识别、换装行人重识别以及小股行人重识别的研究进展与发展趋势，并对各研究方向的部分前沿方法进行性能对比。

### 2.1　无监督目标重识别

在现实应用场景中，为目标重识别任务获取精确的样本标签不仅需要投入大量时间进行数据收集与整理，更依赖于专业人员进行反复核对与标注，耗费大量的人力与物力。这种对标注数据的强依赖，成为有监督目标重识别方法在实际落地时的主要障碍。因此，能够显著降低甚至摒弃对人工标注依赖的无监督目标重识别得到广泛研究，其旨在让模型直接从无标签的数据中学习有效的特征表示。根据在训练过程中是否使用带标签的数据，无监督目标重识别主要分为两类：无监督域自适应目标重识别与完全无监督目标重识别。

1）无监督域自适应目标重识别。此任务通常给定一个带标签的源域和一个无标签的目标域，通过将源域学习到的知识迁移至目标域，从而训练出能良好适应目标域的模型。早期的一些方法（Wang 等，2022b，2023b）主要利用生成对抗网络（generative adversarial networks，GAN）（Goodfellow 等，2014）进行图像风格迁移，旨在缓解源域与目标域之间的分布差异。SPGAN（similarity preserving GAN）（Deng 等，2018）基于 CycleGAN（cycle-consistent GAN）（Zhu 等，2017）与孪生网络结合的架构，通过在风格迁移过程中显式约束图像的自相似性与域间不相似性，从而缓解风格迁移导致源域身份信息丢失的问题。CR-GAN（context rendering GAN）（Chen 等，2019）利用目标域中大量未标注的真实图像实例作为上下文参考，为每个源域身份生成视觉合理、风格多样的合成训练图像，从而增强生成数据的丰富性与真实性，进而有效提升跨域重识别模型的性能。Li 等人（2019）利用基于 PatchGAN（Isola 等，2017）的姿态判别器，对生成图像与条件姿态图进行局部匹配判别以有效解耦并迁移姿态信息，从而在无目标域身份标注的情况下，学习跨域且姿态不变的身份特征。为了充分利用源域有标记的数据信息，PAL（progressive adaptation learning）（Peng 等，2020）采用基于 CycleGAN 的数据适应模块生成与目标域数据分布相似的图像，以减少源域与目标域之间的域偏差。这些生成样本与通过动态采样策略挑选的未标记样本相结合，从而加快训练速度。同样地，PLM（progressive learning method with a multi-scale fusion network）（Wang 等，2022d）也采用 CycleGAN 生成与目标域数据分布相似的图像减少域偏差。与 PAL 不同的是，PLM 以多尺度注意力网络将低层纹理特征与高层语义特征相结合，从而学习更具鉴别性的特征。

然而,这种风格迁移的方法通常仅能缓解图像表观的低级差异(如颜色、纹理等),难以克服由数据分布偏移所导致的深层语义鸿沟。当源域与目标域差异显著时,此类方法的效果往往受限。为此,研究重心逐渐转向直接在目标域内部挖掘其固有的判别性结构。其中,通过聚类生成标签以迭代优化模型成为主流范式。

基于聚类的无监督域自适应方法通常分两步进行:首先利用带标签的源域对模型进行有监督预训练;随后在无标签的目标域上,通过聚类生成标签与基于标签优化模型两个步骤的迭代循环以自适应地微调预训练模型。然而,由于目标在不同条件下的特征变化以及聚类算法的有限性,聚类结果无法保证每个类别内的实例具有相同身份,从而降低标签可靠性,进而制约模型的学习能力与性能表现。因此,现有方法提出多种策略(Zhang 等,2024d;Wei 等,2023;Isobe 等,2021;Wang 等,2024a)以缓解上述问题。SpCL(self-paced contrastive learning)(Ge 等,2020b)通过混合内存生成多级监督信号,并以自步学习的方式迭代优化聚类结果,从而获得更为可靠的标签。MMT(mutual mean-teaching)(Ge 等,2020a)通过两个并行网络协同训练,并利用彼此的辅助网络预测生成的软标签进行相互教学,以此迭代地细化目标域的标签质量。Wang 等人(2023a)提出不确定性感知聚类方法,该方法通过分层聚类提升标签质量,并引入不确定性协同实例选择机制,从而抑制标签噪声。汪琦等人(2024)通过交叉置信软聚类建模域间相关性并提供监督信号,同时利用注意力机制提取的显著性特征与原始特征生成的度量因子来平滑标签,从而提高标签准确性。为缓解标签噪声问题,Sun 等人(2024)利用相似性排序细化标签,并通过对抗学习缩小域差异,进一步降低因域偏移导致的标签误差。

2)完全无监督目标重识别。此任务的核心挑战在于如何直接从无标签数据中自动学习判别性特征,而无需额外带标签的数据。与无监督域自适应类似,主流方法同样采用基于聚类的策略为样本生成标签来监督模型训练,也因此同样面临着标签噪声的问题。一些方法(Wang 等,2024b;Han 等,2024)利用标签平滑(Szegedy 等,2016)的思想来细化标签以缓解标签噪声。Zhang 等人(2021a)提出基于聚类共识与时间集成的标签精炼方法,该方法在模型迭代训练过程中,估计连续轮次间标签的相似性,并对其进行跨代传播与集成,从而实现标签的动态优化,并学习更具判别性的特征表示。Cho 等人(2022)通过挖掘全局特征与局部特征之间的互补关系来相互优化全局和局部的标签,从而有效降低标签噪声并使模型学习包含丰富局部上下文信息的判别性表征。大多数方法(Qiu 等,2024a;Zhang 等,2022a;Ji 等,2024,2025a;Han 等,2023;Liu 等,2024c)通过对比学习的策略强化表征以提升聚类质量,从而提高标签的准确性。Liu 等人(2022b)利用随机更新内存的策略,在对比学习中使用从集群中随机选取的非质心实例来代表该集群,从而学习随机图像对之间的关系,有效避免因不可靠标签带来的训练偏差积累。Yin 等人(2023)提出一种实时内存更新策略,通过随机采样实例特征更新聚类中心,并引入样本到实例和样本到聚类的对比损失,从而保持每个聚类特征的最新特性。He 等人(2024a)提出时空双重注意力机制方法,该方法包含基于类内邻居的空间注意力模块和基于类间序列的时间注意力模块,通过这两个模块与对比学习的结合进行训练减轻类间不一致的更新状态。Xiong 等人(2025)提出自适应聚类与加权正则化对比学习方法,该方法通过动态调整聚类阈值和基于相似度的加权损失函数,从而优化聚类质量并抑制标签噪声。此外,还有一些方法(Xuan 和 Zhang,2021;Wang 等,2021)通过利用相机标签信息增强模型的表征学习能力。随后,一系列利用相机标签信息以进一步提升无监督目标重识别模型性能的方法(Wang 等,2022a;Zhang 等,2023b;Qiu 等,2024b;Tao 等,2024;Li 等,2026)相继涌现。

综上所述,尽管无监督域自适应和完全无监督方法在不同的数据条件下进行重识别任务,但它们的主流方法均采用基于聚类的学习策略,并且取得了更好的效果,因此,这两种范式已逐渐统一。目前,无监督目标重识别常用的数据集为 1.2 小节中介绍的 6 个基准数据集。部分无监督目标重识别方法,包括 DiDAL(discriminative identity-feature exploring and differential aware learning)(Liu 等,2024c)、DSGL(disentangled sample guidance learning)(Ji 等,2024)、PISL(person intrinsic semantic learning)(Tao 等,2024)、STDA(spatial and temporal dual-attention)(He 等,2024a)、MPRD(meta pairwise relationship distillation)(Ji 等,2025a)和 ACWRCL(adaptive clustering and weighted regularization contrastive learning)(Xiong等,2025),在Market-1501和VeRi-776上的性能对比如表1所示。

表1　部分无监督目标重识别方法的性能表现
Table 1　Performance of some unsupervised object Re-ID methods

/%

| 方法 | 年份 | 数据集 | | | |
|---|---|---|---|---|---|
| | | Market-1501 | | VeRi-776 | |
| | | Rank-1 | mAP | Rank-1 | mAP |
| DiDAL* | 2024 | 94.2 | 84.8 | 89.0 | 43.5 |
| DSGL | 2024 | 94.4 | 85.8 | 85.6 | 43.2 |
| PISL | 2024 | 92.4 | 81.1 | 84.2 | 40.2 |
| PISL* | 2024 | 93.8 | 84.1 | 86.8 | 42.2 |
| STDA | 2024 | 93.1 | 82.7 | 87.4 | 42.3 |
| MPRD | 2025 | 94.6 | 85.8 | 88.1 | 43.3 |
| ACWRCL | 2025 | **95.8** | **89.6** | **90.7** | **46.5** |

注:加粗字体为每列最优值,*表示使用相机标签。

## 2.2　多光谱目标重识别

现有目标重识别研究主要基于可见光图像,在夜间或恶劣天气条件下,其性能因环境光照变化而显著下降。在此背景下,多光谱目标重识别技术应运而生,其核心在于融合可见光、近红外和热红外数据,利用不同光谱的互补性提升重识别性能。Li等人(2020)首次构建了多光谱车辆重识别基准数据集RGBN300(RGB-NIR 300)和RGBNT100(RGB-NIR-TIR 100),并提出异构协作感知多流卷积网络以融合多光谱特征。Zheng等人(2021)提出渐进式融合网络以序列化提取高效的多光谱特征,从而有效地利用多光谱数据的互补特性。同时,该研究首次构建了多光谱行人重识别基准数据集RGBNT201(RGB-NIR-TIR 201)。然而,简单的融合多光谱数据的特征,往往忽略光谱间显著差异及其关联性。

针对光谱间的异质性问题,Zheng等人(2023)提出跨方向一致性方法,该方法通过拉近同一身份在不同光谱下的特征中心缓解光谱间的差异,从而生成具有判别性的多光谱特征。此外,为建立完整的评估体系,该研究构建了多光谱车辆重识别基准数据集MSVR310(multi-spectral vehicle re-identification 310)。Wang等人(2024d)利用标记置换模块实现光谱间的循环特征聚合,并通过互补重建模块减少光谱间的分布差异,从而生成更具判别性的多光谱特征。EDITOR(select diverse tokens for multi-modal object re-identification)(Zhang等,2024b)利用共享的视觉Transformer提取不同光谱的标记化特征,并提出结合空间与频率信息的选择器以自适应地筛选出与目标最相关的标记化特征,进而有效抑制背景干扰并缓解光谱间的差异。为了筛选光谱特异性的关键信息,Yu等人(2025b)提出表征选择性耦合网络。该网络设计了注意力—傅里叶令牌稀疏化模块来保留光谱特异性关键信息,以减小光谱间差异。同时利用信息统一约束以进一步对齐多光谱信息并引导模型学习更具判别力的表征。

近期一些方法(Xi,2025;Wang等,2025b)均利用CLIP视觉—语言预训练模型作为骨干网络。Feng等人(2025)将光谱信息分解为共享特征与特定特征,从而使模型能够同时利用光谱间的共性信息与各光谱的独有信息,以学习对光谱变化鲁棒的判别性表征。Wang等人(2025c)提出基于解耦特征的专家混合方法,该方法通过层次化解耦模块将多光谱特征显式地分离并保留光谱特异性信息,并利用注意力触发的专家混合机制以动态融合,从而有效提升模型对多光谱成像中动态质量变化与光谱缺失的鲁棒性。Wan等人(2025)提出不确定性引导的图模型,该方法通过高斯补丁图表示来估计局部与全局的不确定性以建模其图结构关系,并设计不确定性引导的专家混合策略抑制噪声干扰并增强多光谱的信息融合。此外,一些方法利用文本语义或可学习提示缓解光谱间的视觉差异。Wang等人(2025d)则通过多模态大语言模型为图像生成结构化的文本描述,进而以此引导视觉特征的融合。Zhang等人(2025c)提出基于可学习提示方法,该方法通过可学习的多光谱提示并设计提示交换机制,以有效地实现光谱间信息的交互与融合。

目前,多光谱目标重识别常用的行人数据集有RGBNT201(Zheng等,2021),车辆数据集有RGBNT100(Li等,2020)和MSVR310(Zheng等,2023)。表2总结了部分多光谱目标重识别方法在RGBNT201和RGBNT100数据集上的性能结果,包括TOP-ReID(token permutation re-identification)(Wang等,2024d)、EDITOR(Zhang等,2024b)、RSC-

Net(representation selective coupling network)(Yu等,2025b)、MDReID(modality-decoupled re-identification)(Feng等,2025)、IDEA(inverted text with cooperative deformable aggregation)(Wang 等,2025d)、DeMo(decoupled feature-based mixture of experts)(Wang等,2025c)、PromptMA(prompt-based modality alignment)(Zhang 等,2025c)和 UGG-ReID(uncertainty-guided graph re-identification)(Wan等,2025)。

表2 部分多光谱目标重识别方法的性能表现
Table 2 Performance of some multi-spectral object Re-ID methods

/%

| 方法 | 年份 | 数据集 | | | |
|---|---|---|---|---|---|
| | | RGBNT201 | | RGBNT100 | |
| | | Rank-1 | mAP | Rank-1 | mAP |
| TOP-ReID | 2024 | 76.6 | 72.3 | 96.4 | 81.2 |
| EDITOR | 2024 | 68.3 | 66.5 | 96.4 | 82.1 |
| RSCNet | 2025 | 72.5 | 68.2 | 96.6 | 82.3 |
| MDReID | 2025 | 85.2 | **82.1** | 95.6 | 85.3 |
| IDEA | 2025 | 82.1 | 80.2 | 96.5 | 87.2 |
| DeMo | 2025 | 82.3 | 79.0 | 97.6 | 86.2 |
| PromptMA | 2025 | 80.9 | 78.4 | 97.4 | 85.3 |
| UGG-ReID | 2025 | **86.8** | 81.2 | **98.1** | **88.0** |

注:加粗字体为每列最优值。

### 2.3 跨模态行人重识别

跨模态行人重识别根据模态的不同可分为可见光—红外行人重识别、文本—图像行人重识别和素描—照片行人重识别3种类型的任务。随着多模态大模型的兴起,跨模态行人重识别逐渐转向通用多模态行人重识别,相关代表性工作有UNIReID(unified person re-identification)(Chen 等,2023)和 AIO(all-in-one)(Li等,2024a)。

#### 2.3.1 可见光—红外行人重识别

可见光—红外行人重识别旨在实现可见光图像与红外图像之间的身份匹配,其核心挑战在于可见光与红外之间成像的显著模态差异。现有的可见光—红外行人重识别主要可划分为图像对齐方法和特征学习方法。

1) 基于图像对齐的可见光—红外行人重识别。此类方法旨在统一两种模态图像的风格特征。一些方法主要利用生成模型(Goodfellow等,2014;Ho等,2020)生成跨模态图像缩小模态差异。Kniaz等人(2019)通过模拟行人身体热力特征将彩色图像转换为热力图,以利用热一致性进行身份匹配。Wang等人(2019a)则利用生成对抗过程对齐异构样本,从而在两个层级上同步缩减模态波动。为了缓解生成图像中的身份失真问题,Wang等人(2019b)提出双路重构学习方法,利用真实与伪图像对的循环重构,实现对身份信息的有效保留。Choi等人(2020)则提出分层跨模态解耦模型,将图像分解为身份相关和模态相关的因子,以在生成过程中保留身份特征。Liu等人(2022a)提出两阶段模态增强网络,通过深度跳跃连接显著提升生成图像的质量。Pan等人(2024)提出统一条件图像生成方法,利用自适应条件扩散模型生成语义对齐图像。此外,一些方法通过构建中间模态辅助跨模态检索。Ye等人(2021)利用灰度图像作为辅助模态,在不增加复杂生成的条件下缩小模态差异。Alehdaghi等人(2023)利用随机线性组合RGB通道生成无偏见图像,实现对色彩偏差的鲁棒特征提取。Lu等人(2023)提出通道交互生成器,通过产生模糊模态图像从而缓解了特征偏差。Kim等人(2023)和Hua等人(2023)分别提出跨模态部位描述符混合和跨模态通道混合技术,这些方法利用空间或通道层面的部件混合,实现模型对行人局部对齐能力的增强。Zhang等人(2024a)利用灰色直方图均衡化生成的辅助模态进行两阶段学习,从而缓解颜色和光照差异的影响。Alehdaghi等人(2025)提出自适应生成中间信息的方法,利用虚拟中间域实现在不改变推理模型的前提下利用中间模态信息辅助训练。

2) 基于特征学习的可见光—红外行人重识别。此类方法侧重直接提取与模态无关的身份表征。大多数方法(Ye等,2022;Chen等,2022b;Liang等,2023;Jiang等,2025;Wu等,2025a)通过强化模型的表征学习实现鲁棒的可见光—红外行人重识别,Wu等人(2017)提出深度零填充方法以学习模态共享特征。Ye等人(2020)提出动态双注意力聚合学习框架,以学习行人判别性区域特征。Lu等人(2020)提出共享—特定特征转移算法,同时利用模态共享与模态特定特征进行重识别。Wu等人(2021)通过分离模态特异性信息并挖掘跨模态细粒度信息,从而提升特征判别性与跨模态一致性。FMCNet(feature-

level modality compensation network）（Zhang 等，2022c）及 FMCNet+（Xi 等，2025）提出特征级模态补偿网络，以优化特征空间的模态特异性表征。Zhang 和 Wang（2023）提出多样性嵌入扩展网络对模态共享信息进行深度挖掘。Li 等人（2024b）则提出跨模态特定—共享协作网络以实现模态对齐。Chen 等人（2025）通过层次化特征交互与跨模态重建，从多尺度视角挖掘模态共享特征，从而缓解可见光与红外模态间的显著差异。为了挖掘隐式判别信息，Ren 和 Zhang（2024）提出隐式判别知识学习网络，该方法提取模态特定特征中的隐式判别信息并蒸馏至共享特征中，显著增强判别性表征。Cheng 等人（2025）提出模态感知领域对齐网络，该方法通过模态感知全局—局部上下文注意力捕获多粒度特征，从而对齐跨模态信息。Lu 等人（2025）提出基于三流网络与正交分解的模态—姿态解耦模型，通过记忆注意力机制融合跨模态特征与骨架信息，实现模态差异消除与姿态无关特征提取。Zhang 等人（2025f）提出频域细节挖掘方法，利用相位增强和幅度细节挖掘建模频域视角的物理属性。针对细粒度对齐问题，Liu 等人（2025a）提出跨模态语义一致性学习网络，利用人体解析信息实现行人部件的精确对齐。魏思和杨文璐（2025）利用姿态估计构建双流网络以提取骨骼和视觉特征，然后通过交互注意力融合并引入结构内聚损失优化特征一致性，从而学习判别性的跨模态行人表示。此外，Qi 等人（2025）提出辅助表示引导网络，利用模态共享空间内的辅助表征引导学习。在度量学习方面，Zhu 等人（2020）和 Liu 等人（2021）分别提出异构中心损失和异构中心三元组损失，以优化类内紧凑与类间分离。

一些方法（Wu 等，2025b）通过大模型（如 CLIP、ChatGPT 等）的语义理解能力，增强视觉特征与文本语义的交互，从而缓解模态差异对身份判别的干扰。Yu 等人（2025a）提出 CLIP 驱动的语义挖掘网络，利用高层文本语言描述引导视觉特征与语义信息的融合。Hu 等人（2025a）提出基于 CLIP 的模态补偿方法，利用文本提示生成恢复缺失信息，从而生成包含更多判别信息的完整模态特征。Dong 等人（2025）提出多样性语义引导的网络，利用 ChatGPT 生成的多样化句式引导视觉特征的跨模态对齐。

目前，可见光—红外行人重识别常用的数据集有 SYSU-MM01（SYSU multiple modality 01）（Wu 等，2017）和 RegDB（Nguyen 等，2017）。表 3 总结了部分可见光—红外行人重识别方法在 SYSU-MM01 和 RegDB 数据集上的性能结果，包括 MSALNet（multi-stage auxiliary learning network）（Zhang 等，2024a）、IDKL（implicit discriminative knowledge learning）（Ren 和 Zhang，2024）、HTCR（hierarchical token-aware cross-modality reconstruction）（Chen 等，2025）、DMPF（disentangling modality and posture factors）（Lu 等，2025）、MDANet（modality-aware domain alignment network）（Cheng 等，2025）、CSCL（cross-modality semantic consistency learning）（Liu 等，2025a）、DSFA（diverse semantics-guided feature alignment）（Dong 等，2025）和 ARGN（auxiliary representation guided network）（Qi 等，2025）。需要说明的是，上述讨论集中于有监督学习设置下的方法。此外，该领域也涌现出一系列无监督可见光—红外行人重识别方法（Yang 等，2024；Ji 等，2025b；He 等，2025；Liu 等，2025c）。

**表 3 部分可见光—红外行人重识别方法的性能表现**

**Table 3 Performance of some visible-infrared person Re-ID methods**

/%

| 方法 | 年份 | 数据集 | | | |
|---|---|---|---|---|---|
| | | SYSU-MM01 (All-Search) | | RegDB (VIS-to-IR) | |
| | | Rank-1 | mAP | Rank-1 | mAP |
| MSALNet | 2024 | 71.2 | 68.0 | 91.1 | 85.2 |
| IDKL | 2024 | **81.4** | **79.9** | 94.7 | **90.2** |
| HTCR | 2025 | 74.6 | 72.1 | 92.6 | 87.9 |
| DMPF | 2025 | 76.4 | 71.6 | 88.8 | 81.0 |
| MDANet | 2025 | 79.5 | 77.0 | 93.6 | 84.6 |
| CSCL | 2025 | 75.7 | 72.1 | 92.2 | 84.3 |
| DSFA | 2025 | 78.4 | 74.7 | 95.8 | **90.2** |
| ARGN | 2025 | 77.0 | 72.7 | **96.2** | 90.0 |

注：加粗字体为每列最优值。

2.3.2 文本—图像行人重识别

文本—图像行人重识别旨在通过文本描述实现行人检索，以应对待查询行人的视觉图像可能缺失的现实困境。Li 等人（2017b）首次提出文本—图像行人重识别任务，并构建首个大规模基准数据集 CUHK-PEDES（CUHK person description），同时通过

门控神经注意力机制捕捉文本与图像的关联性。在此基础上,Li等人(2017a)进一步提出身份感知的两阶段网络以优化文本与图像的匹配结果。Zhang和Lu(2018)则提出跨模态投影匹配和分类两种损失,优化文本与图像在共享特征空间中的分布一致性。Sarafianos等人(2019)则利用对抗学习,旨在通过学习与模态无关的特征表示来缩小文本与图像的分布差异。随后,一些方法通过局部细节的显式对齐与多粒度特征的挖掘来提升模型性能。Niu等人(2020)提出包含全局与局部匹配的多粒度对齐模型,以增强多层级文本与视觉特征的关联性。Wang等人(2020c)通过辅助分割网络获得行人的多个语义区域,并将其与文本中的属性短语细粒度对齐。Zhu等人(2021a)将人物与背景分离,并采用多样化的细粒度对齐范式,从而有效提取和匹配人物信息。Shu等人(2023)通过单词、短语和句子3个层级的对齐并结合双向掩码建模,强制全局特征包含丰富的多层级语义信息。Farooq等人(2022)则提取多尺度视觉特征并与文本的语义特征隐式对齐。Wang等人(2022e)提出相互学习的多分支架构,从而实现全方位信息的捕捉,以避免模型过度依赖颜色信息。

CLIP等视觉—语言大模型的兴起,为文本—图像行人重识别提供了强大的跨模态基础。模型能够直接将文本描述与行人图像映射至统一的语义空间,实现更高效的端到端检索。因此,一些方法采用视觉—语言预训练大模型作为基础骨干网络以执行文本—图像行人重识别任务。Jiang和Ye(2023)提出隐式关系推理与对齐方法,该方法通过学习局部视觉信息与文本信息之间的关系,以增强全局的图文匹配能力。Yan等人(2023b)提出细粒度信息挖掘框架,该方法通过隐式地发掘模态内共享的判别性细节并增强全局特征,以实现高效的跨模态全局对齐。为了缓解CLIP模型在细粒度行人检索上的不足,Cao等人(2024)系统性地探索数据增强、损失函数等训练技巧,仅微调CLIP模型而无需引入任何复杂模块,便取得出色的性能。He等人(2024c)提出视觉引导的语义分组网络,该网络在通道维度的语义分组上提取文本局部特征,并借助视觉知识转移模块,实现高效的跨模态局部特征隐式对齐。Lin等人(2024)构建文本—图像与图像—文本的双向局部关联,从而实现更精准地跨模态细粒度对齐。Wei等人(2024)则将文本解构为关键短语引导视觉区域的定位,以实现无需外部检测器的端到端细粒度图文对齐。此外,为克服显式对齐方法的局限,Sun等人(2025)提出基于隐式对齐的跨模态共生网络降低图文的模态异质性。Zhang等人(2025d)利用图神经网络构建渐进式关系挖掘框架,通过建模模态内与模态间的特征语义关联,从而显著增强行人特征的判别性与跨模态语义对应能力。

目前,文本—图像行人重识别常用的数据集有CUHK-PEDES(Li等,2017b)、ICFG-PEDES(identity-centric and fine-grained person description)(Ding等,2021)和RSTPReid(real scenario text-based person re-identification)(Zhu等,2021a)。表4总结了部分文本—图像行人重识别方法在CUHK-PEDES和ICFG-PEDES数据集上的性能结果,包括CADA(cross-modal adaptive dual association)(Lin等,2024)、SCVD(semantics-centric visual division)(Wei等,2024)、TBPS-CLIP(text-based person search CLIP)(Cao等,2024)、VGSG(vision-guided semantic-group)(He等,2024c)、RMGNet(relationship-mining graph neural network)(Zhang等,2025d)和CMSN(cross-modal symbiotic network)(Sun等,2025)。

**表4 部分文本—图像行人重识别方法的性能表现**

**Table 4 Performance of some text-to-image person Re-ID methods**

/%

| 方法 | 年份 | 数据集 | | | |
|---|---|---|---|---|---|
| | | CUHK-PEDES | | ICFG-PEDES | |
| | | Rank-1 | mAP | Rank-1 | mAP |
| CADA | 2024 | **78.4** | 68.9 | 67.8 | 39.9 |
| SCVD | 2024 | 76.7 | – | **69.8** | – |
| TBPS-CLIP | 2024 | 73.5 | 65.4 | 65.1 | 39.8 |
| VGSG | 2024 | 71.4 | 67.9 | 63.1 | – |
| RMGNet | 2025 | 77.2 | **70.6** | 68.4 | **41.6** |
| CMSN | 2025 | 76.5 | 70.3 | 64.5 | 41.2 |

注:加粗字体为每列最优值,“–”表示该指标暂无相关数据。

#### 2.3.3 素描—照片行人重识别

素描—照片行人重识别旨在通过素描图像实现行人检索。相较于文本,素描能提供更直观地细粒度视觉线索,以弥补文本描述在细节和视觉结构信息上的不足。Pang等人(2018)首次提出基于素描

的行人重识别任务，并构建首个基准数据集PKU-Sketch（Peking University sketch）。此外，为应对素描与照片间的模态差异，该工作进一步提出跨域对抗特征学习框架以学习模态不变的判别性表征。Yang等人（2021a）提出实例级异构域适应框架来利用源域属性构建共享语义空间并运用对抗性域适应，从而有效增强模型在数据稀缺条件下的跨模态检索性能。Chen等人（2022a）提出非对称解耦与动态合成学习方法，该方法利用辅助素描将照片中缺失的风格信息迁移至素描模态，从而缩小跨模态差异。在此基础上，Chen等人（2024）进一步通过模态感知原型对比学习以更高效地实现跨模态对齐。Zhang等人（2022e）提出跨兼容性嵌入与语义一致特征构建方法，该方法通过特征层的交互移植与语义分组对齐，从而增强特征的判别性。为了克服素描中因目击者主观性与艺术家绘画风格差异导致的检索难题，Lin等人（2023）提出包含非局部融合与属性对齐模块的框架，其通过融合多视角草图并利用CLIP驱动的属性掩码以对齐跨模态特征，同时该工作还构建大规模多风格数据集Market-Sketch-1K。Liu等人（2024a）则动态生成一个去噪且风格化的辅助模态，并利用模态交互注意力来对齐特征并通过辅助模态学习跨模态的不变模式。针对素描主观风格多变和标注数据稀缺的问题，Hu等人（2025b）提出利用风格多样的非行人素描数据提升模型泛化能力。该方法的核心是自适应增量提示调优框架，通过从外部数据学习可共享的风格，并动态适配新输入，从而有效实现跨类别的风格知识迁移，并显著提升性能。

目前，素描—照片行人重识别常用的数据集有PKU-Sketch（Pang等，2018）。表5总结了部分素描—照片行人重识别方法在PKU-Sketch数据集上的性能结果，包括SketchTrans（sketch transformer）（Chen等，2022a）、CCSC（cross-compatible embedding and semantic consistent feature construction）（Zhang等，2022e）、AttrAlign（attribute alignment）（Lin等，2023）、SketchTrans+（Chen等，2024）、DALNet（differentiable auxiliary learning network）（Liu等，2024a）和AIP（adaptive incremental prompt-tuning）（Hu等，2025b）。

表5　部分素描—照片行人重识别方法的性能表现

**Table 5　Performance of some sketch-to-photo person Re-ID methods**

/%

| 方法 | 年份 | PKU-Sketch数据集 | |
|---|---|---|---|
| | | Rank-1 | mAP |
| SketchTrans | 2022 | 84.6 | – |
| CCSC | 2022 | 86.0 | 83.7 |
| AttrAlign | 2023 | 78.0 | 78.8 |
| SketchTrans+ | 2024 | 85.8 | – |
| DALNet | 2024 | 90.0 | 86.2 |
| AIP | 2025 | **96.0** | **94.3** |

注：加粗字体为每列最优值，“–”表示该指标暂无相关数据。

### 2.4　遮挡行人重识别

真实世界中的行人和物体随机移动导致行人被遮挡的可能性很高，从而显著降低视觉信息的质量，严重影响现有行人重识别方法的性能。现有的遮挡行人重识别主要有基于可见性感知与基于特征重建两大类方法。

1）基于可见性感知的遮挡行人重识别。此类方法主要通过未遮挡的可见区域信息以提升模型在遮挡场景下的鉴别力。一些方法利用人体结构先验获取可见区域信息。Miao等人（2019）提出姿态引导的特征对齐方法，该方法利用姿态关键点生成注意力图指示遮挡部位，并在特征构建与匹配阶段引导模型仅关注和利用非遮挡的可见身体区域。Gao等人（2020）提出姿态引导的可见区域匹配框架，该方法通过姿态引导注意力提取判别性局部特征，并利用姿态引导可见性预测器在无遮挡标注的情况下训练行人区域的可见性。Wang等人（2020a）将图像的局部特征视为图节点，并构建自适应方向图卷积与跨图像嵌入对齐模块，从而能够同时建模特征间的高阶关系与跨图像的拓扑结构信息，以提升遮挡场景下特征匹配的准确性。Zhang等人（2021b）提出语义感知的遮挡鲁棒网络，该网络通过局部、全局和语义3个分支的联合训练，并利用语义分支生成的前景—背景掩码标识行人的非遮挡区域，从而克服遮挡的干扰。Yang等人（2021b）利用离散化后的姿态可见性标签界定身体部位的遮挡状态，从而促使模型在特征学习过程中聚焦于可见区域。Tan等人（2022）提出CBDB-Net（consecutive batch drop-block network）方法，通过在特征图上连续丢弃局部区域生成多个不完整特征，促使模型在遮挡条件下

也能学习判别性表征。Miao等人(2022)融合人体姿态估计所提供的身体位置与部位可见性信息抑制遮挡等噪声信息的干扰。Wang等人(2022c)提出基于Transformer的姿态引导特征解耦方法，该方法利用姿态信息分离出与身体部位相关的语义特征，并以此有选择地匹配可见部位，从而有效降低遮挡干扰。Dou等人(2023)通过弱监督的协同解析网络提取语义信息来对齐未被遮挡的身体区域。Gao等人(2024b)利用人体语义辅助模块来提取多粒度人体区域特征以应对遮挡。Geng等人(2025)提出姿态—骨架引导的交叉注意力表示融合方法，通过可见外观区域注意力模块抑制遮挡干扰，并结合骨骼区域建模生成局部掩码，从而在网络中间层有效区分遮挡信息与行人特征，实现对行人特定语义区域的定位。Liu等人(2025b)则基于关键点特征和注意力机制的协同引导，以自适应地优化多粒度特征表示，从而增强行人可见区域特征的判别性并抑制遮挡部分。此外，还有部分方法通过文本提示来引导可见区域定位。He等人(2024b)提出区域生成与评估网络，该方法通过区域生成模块利用CLIP提取的语义原型定位身体区域，并通过区域评估模块为不同区域分配置信度分数以降低遮挡影响。为了克服通用文本提示无法捕获样本细粒度语义的问题，Lin等人(2025)通过语义分割与文本生成模块提供细粒度语义描述，从而引导模型学习判别性细粒度特征，同时有效抑制无关区域的干扰。Wang等人(2025a)则提出文本引导的层次化上下文混合网络以充分利用文本信息对行人外观属性的描述能力，从而有效引导视觉模型学习更具判别性的身份特征。

另一些方法(Yan等，2023a；Li等，2024c)则让模型自适应地聚焦于行人的可见判别性区域。Chen等人(2021b)提出注意力引导的掩码模块以捕捉身体可见部位。Li等人(2021)利用部件多样性与部件可判别性两种弱监督学习机制，使用身份标签来引导模型挖掘具有鉴别力的行人部件特征。Tan等人(2023)则通过多头自注意力自适应地捕获行人的关键局部信息。Jia等人(2023)提出解耦表示学习网络，该方法通过对局部特征来全局推理以实现无需对齐的重识别。为了克服遮挡与整体图像间存在信息不对称的问题，Gao等人(2024a)利用对比学习机制来引导其注意力模块捕捉行人前景。

2）基于特征重建的遮挡行人重识别。此类方法旨在通过对遮挡区域进行特征恢复或生成，以构建更加完整和判别性的行人表征。Hou等人(2022)通过空间区域特征补全模块利用空间上下文信息预测遮挡区域特征，并引入时序区域特征补全模块利用时序上下文信息进行优化，从而在特征空间中恢复被遮挡区域的语义。FED(feature erasing and diffusion network)(Wang等，2022f)利用遮挡擦除与特征扩散两个模块协同训练，在特征空间中分别处理非行人遮挡的噪声与非目标行人的特征污染，从而提升模型对目标行人的感知能力。FRT(feature recovery transformer)(Xu等，2022)利用图相似度筛选出的近邻样本特征集来恢复查询图像的完整特征。Huang等人(2023)通过推理身体部位的可见性并补偿遮挡部位的语义损失，从而学习遮挡图像的完整行人表示。Ye等人(2024)提出动态特征剪枝与整合框架，该框架通过自注意力机制自动剪枝噪声特征，并利用检索到的近邻特征恢复鲁棒的身份表示。Wang等人(2024c)提出特征补全Transformer，通过遮挡实例增强技术生成对齐的遮挡—整体图像对，利用特征补全解码器以自监督方式聚合信息来补全遮挡区域特征，并引入特征补全一致性损失使补全特征与身份信息相关联。Bian等人(2024)对检索的图像中未遮挡的局部特征加权平均，以重建查询图像中被遮挡的特征，从而恢复更全面的特征表示以优化检索。张红颖等人(2024)利用邻近信息恢复被遮挡语义，并通过姿态来引导特征对齐，同时聚焦局部鉴别性特征以抑制背景干扰。Li等人(2025)通过全局和局部双流结构分别捕获全局上下文信息和局部语义信息，并提出基于距离的特征恢复模块以动态恢复被遮挡的行人区域特征。Zhang等人(2025a)则通过掩码信息嵌入模块引导模型关注非遮挡身体部位，并结合层次特征聚合模块挖掘细粒度局部特征，接着利用多样化特征补全化模块整合多路径特征以补全全局特征信息。

目前，遮挡行人重识别常用的数据集有Occluded-DukeMTMC(Miao等，2019)、Occluded-ReID(occluded person re-identification)(Zhuo等，2018)、Partial-REID(partial person re-identification)(Zheng等，2015c)和Partial-iLIDS(partial international logistic identification)(Zheng等，2011)。部分遮挡行人重识别方法，包括FPC(feature pruning and consolidation)(Ye等，2024)、SPH(semantic perception and CNN-transformer hybrid network)(Gao 等，2024b)、RGANet(region generation and assessment network)(He 等，2024b)、FCFormer(feature completion transformer)(Wang等，2024c)、HGTDR(heterogeneous generative tokens and distance-aware recovery network)(Li 等，2025)、AOANet(adaptive occlusion-aware network)(Liu 等，2025b)、MAHATMA(mask-aware hierarchical aggregation transformer)(Zhang等，2025a)、PSCR(pose-skeleton guided cross-attention representation fusion)(Geng 等，2025)、THCB-Net(text-guided hierarchical context blending network)(Wang等，2025a)和SGFNet(semantically guided and focused network)(Lin 等，2025)，在 Occluded-DukeMTMC 和 Occluded-ReID 上的性能对比如表 6 所示。

表6　部分遮挡行人重识别方法的性能表现

Table 6　Performance of some occluded person Re-ID methods

/%

| 方法 | 年份 | Occluded-DukeMTMC | | Occluded-ReID | |
|---|---|---|---|---|---|
| | | Rank-1 | mAP | Rank-1 | mAP |
| FPC | 2024 | 76.7 | **72.8** | 86.3 | 84.6 |
| SPH | 2024 | 71.9 | 63.7 | 82.6 | 82.1 |
| RGANet | 2024 | 71.6 | 62.4 | 86.4 | 80.0 |
| FCFormer | 2024 | 73.0 | 63.1 | 83.6 | 85.7 |
| HGTDR | 2025 | 71.8 | 62.3 | 87.2 | 83.3 |
| AOANet | 2025 | **81.2** | 70.6 | 91.7 | 88.8 |
| MAHATMA | 2025 | 73.3 | 62.3 | 85.8 | 79.5 |
| PSCR | 2025 | 73.4 | 63.1 | 87.7 | 83.1 |
| THCB-Net | 2025 | 80.0 | 70.7 | 88.9 | 84.8 |
| SGFNet | 2025 | 76.5 | 67.2 | **93.0** | **90.3** |

注：加粗字体为每列最优值。

### 2.5　换装行人重识别

由于行人在不同时间可能穿着不同服装，导致依赖短期表观一致性的模型判别力显著下降。因此，换装行人重识别这一重要任务被提出，其目标是在行人着装发生变化的条件下，仍能跨摄像头实现准确的身份关联与检索。

一类方法致力于挖掘与服装无关的生物特征，以构建对服装变化鲁棒的身份表示。Xue 等人(2018)提出衣物变化感知的识别网络，通过协同学习衣物变化检测与身份识别，并自适应地调整对脸部与身体特征的关注权重，从而在衣物变化与不变两种情形下都能判定行人身份。Yang 等人(2021c)利用人体轮廓草图作为输入，并借助空间极坐标变换将其映射至极坐标空间，以过滤颜色和纹理等与服装相关的干扰，从而提取身份判别特征。Hong 等人(2021)提出细粒度形状—外观互学习方法，该方法通过包含形状流与外观流的双流架构，在形状流中以身份为指导学习细粒度掩码并提取姿态特定的身体形状特征，并通过交互式学习将此知识迁移至外观流，从而使模型获得对服装变化鲁棒的识别能力。Jin 等人(2022)利用步态识别作为辅助监督任务，驱动模型在训练过程中捕获不受服装变化影响的生物运动信息。Chen 等人(2022c)则引入轮廓特征作为额外的监督信号，通过最大化其与外观特征的互信息，从而约束模型从彩色图像中挖掘与形状强相关的判别特征。Zhang 等人(2023a)提出多生物线索挖掘框架，该方法通过协同提取头部、颈部和肩部等受衣物影响较小的生物特征，从而实现换装场景下鲁棒的身份表征，并设计轻量级领域自适应模块以提升跨场景的泛化性。Yang 等人(2023b)利用分层竞争策略在全局、通道和像素维度挖掘细粒度的身份线索，从而实现在不依赖额外辅助信息情况下的鲁棒识别。为了克服 CNN 感受野受限导致难以挖掘细粒度特征的问题，Wu 等人(2024)提出 CNN 与 Transformer 双流混合网络，以捕捉如发型、饰品等与服装无关的判别性细节。Liu 等人(2024b)利用姿态估计生成的关键点对齐细粒度特征，引导模型聚焦于头部和四肢等受服装变化影响较小的稳定区域。Nguyen 等人(2024)利用关系图注意力网络捕获衣物不变的体型信息，并实现在有限标签数据下的高质量特征提取。

另一类方法则在特征表示中分离身份信息与服装信息，旨在获得对服装变化不变的身份特征。Yu 等人(2020)通过双分支网络将行人的生物特征与服装信息解耦。Huang 等人(2021a)提出服装状态感知网络，该方法通过显式建模行人换装状态，并以此为元信号来正则化特征学习过程，从而使模型自适应地平衡对身份的判别性与对服装变化的稳健性。Chen 等人(2021a)将三维人体重建作为辅助任务，

并在二维图像中直接提取对纹理不敏感的三维形状特征,实现身份与衣物纹理的解耦。Gu等人(2022)提出基于服装的对抗性损失,该方法通过在图像上引入对抗性学习框架,惩罚模型对服装属性的预测能力,从而驱动模型从图像中挖掘与服装无关的判别性特征。Yang等人(2023a)基于因果推理框架,构建双分支架构以模拟对服装变量的干预,从而自动解耦特征中纠缠的服装与身份信息,引导模型聚焦于服装不变的身份特征。Liu等人(2023b)则采用双层自适应加权方案,利用人类解析技术从图像和特征双重维度量化换装程度,通过自动调整各语义部位的权重消除服装模式带来的干扰,从而识别并隔离服装变动信息。Liu等人(2023a)利用神经隐式函数分别建模裸体形状、服装几何与纹理,旨在从单幅图像中解耦出对衣服和姿态高度不变的三维裸体形状特征。Cui等人(2023)则提出组件重建解耦框架,该方法通过重建人体的各组成部分(如躯干、四肢),实现对衣物相关与无关特征的可控分离。Gao等人(2024c)提出多流协作学习架构,该方法利用身份不变性引导特征提取,从而获得在服装变化与姿态变化双重干扰下表现出与服装弱相关、与身份强相关的稳健特征。Zhang等人(2025b)则提出基于CLIP的多模态特征学习框架,通过联合优化视觉—文本对齐、全局—局部特征融合与多尺度上下文感知,从而抑制服装属性对身份特征的干扰。Gao等人(2025)提出语义感知注意力与视觉屏蔽网络,通过屏蔽服装区域并对人体语义部位进行加权,使模型在姿态和视角变化下仍聚焦于与服装无关的本质属性。此外,一些方法利用文本提示或掩码机制引导模型学习与服装无关的恒定属性。Wei等人(2025)提出多信息提示学习方案,通过服装信息剥离模块,在文本提示引导下将服装信息从图像特征中解耦。Liang和Rawat(2025)则提出利用文本描述引导的对抗学习,利用梯度反转层强制使身份表征与服装等非生物因素在独立子空间中解耦。Peng等人(2025)与Hao等人(2025)均采用文本掩码策略,前者通过遮蔽文本描述词中的颜色与服装类型等信息,后者通过掩码文本对抗训练,共同强制模型将视觉特征与性别、发型和体态等衣物无关的属性对齐。

目前,换装行人重识别常用的数据集有PRCC(person re-id under moderate clothing change)(Yang等,2021c)、LTCC(long-term cloth-changing)(Qian等,2021)、Celeb-reID(Huang等,2020)、VC-Clothes(virtually changing-clothes)(Wan等,2020)和Deep-Change(Xu和Zhu,2023)。表7总结了部分换装行人重识别方法在PRCC和LTCC数据集上的性能结果,包括CCPG(contrastive clothing and pose generation)(Nguyen等,2024)、IGCL(identity-guided collaborative learning)(Gao等,2024c)、PGAL(pose-guided attention learning)(Liu等,2024b)、MADE(masked attribute description embedding)(Peng等,2025)、DIFFER(disentangle identity features from entangled representations)(Liang和Rawat,2025)、SAVS(semantic-aware attention and visual shielding network)(Gao等,2025)和MIPL(multiple information prompt learning)(Wei等,2025)。

表7 部分换装行人重识别方法的性能表现
Table 7 Performance of some clothes-changing person Re-ID methods

/%

| 方法 | 年份 | 数据集 | | | |
|---|---|---|---|---|---|
| | | PRCC | | LTCC | |
| | | Rank-1 | mAP | Rank-1 | mAP |
| CCPG | 2024 | 61.8 | 58.3 | 46.2 | 22.9 |
| IGCL | 2024 | 64.4 | 63.0 | **77.8** | **47.1** |
| PGAL | 2024 | 59.5 | 58.7 | 62.5 | 27.7 |
| MADE | 2025 | 64.3 | 59.1 | 47.4 | 24.4 |
| DIFFER | 2025 | 68.5 | 64.7 | 58.2 | 31.6 |
| SAVS | 2025 | 69.4 | 57.6 | 71.2 | 32.5 |
| MIPL | 2025 | **71.0** | **67.0** | 75.1 | 38.8 |

注:加粗字体为每列最优值。

## 2.6 小股行人重识别

现实中行人常以群体形式活动,导致针对单人的重识别方法难以适用。小股行人重识别旨在解决这一实际问题,其目标是通过匹配不同摄像头拍摄的群体图像之间的视觉相似性,实现对同一群体的跨摄像头关联。除单人重识别存在的视觉挑战外,小股行人重识别任务还面临着群体成员变化、群体成员间相互遮挡和群体布局变化等挑战。

早期的小股行人重识别方法主要利用特征描述符(Cai等,2010)和直方图特征(Lisanti等,2017)等

手工设计的传统特征捕捉群体的表观信息。然而，这些手工设计的方法难以应对群体人数与布局的动态变化，泛化能力有限。随着深度学习技术的井喷式发展，研究人员逐渐将重心转向基于深度学习的小股行人重识别方法。一些方法主要利用图神经网络（graph neural network，GNN）建模群体内成员间的复杂关系。为应对群体成员变化与布局变化，Xiao 等人（2018）和 Lin 等人（2021）分别提出多粒度匹配框架 LIMI（leveraging and integrating multi-grain information）和 MGR（multigrained），其将群体图像分解为个体、两人子组、三人子组和全局对象等多个粒度，通过动态权重评估各对象的判别性，并利用多阶匹配算法优化跨视图关联，从而有效处理布局和成员变化。GCGNN（group context GNN）（Zhu 等，2021b）显式地将群体建模为空间 K 近邻图，并设计群体上下文图神经网络捕捉群体局部结构。MACG（multi-attention context graph）（Yan 等，2023c）基于图神经网络将群体构建为上下文图结构，并设计多级注意力机制建模群体内与群体间的上下文依赖。DoT-GNN（domain-transferred GNN）（Huang 等，2019）引入 CycleGAN 将行人重识别数据风格迁移至小股行人重识别领域，并利用图神经网络对群体建模，从而学习鲁棒表示。DotSCN（domain-transferred single and couple representation learning network）（Huang 等，2021b）在此基础上结合个体与成对表征学习，在保持性能的同时大幅降低计算成本。随后，一些方法尝试将 Transformer 模型应用于小股行人重识别任务，并展现出显著提升。3DT（3D Transformer）（Zhang 等，2022b）从三维视角对成员间的空间布局进行建模，通过重建相对的三维布局关系提取布局特征，以克服传统二维方法存在的布局模糊性问题。Zhang 等人（2022d，2024c）提出基于二阶 Transformer 的 SOT（second-order Transformer）（Zhang 等，2022d）和 UMSOT（uncertainty modeling SOT）（Zhang 等，2024c）方法，这些方法引入不确定性建模将群体图像表征为概率分布，并通过成员变化与布局变化模块动态挖掘潜在群体结构，从而学习对未知成员与布局变化更具鲁棒性的群体表示。为提升复杂遮挡场景下的模型判别能力，Lu 等人（2024）提出基于孪生 Transformer 的方法，其通过多尺度特征变换与联合学习机制有效增强特征鲁棒性。Zhang 等人（2025e）提出基于二阶 Transformer 的 PBSOT（parallel branches-based SOT）方法，该方法通过并行的全局和局部分支分别学习组级特征，并设计布局引导的局部特征采样模块对最密集区域的局部特征采样，以更有效地缓解因布局变化引起的遮挡问题。

目前，小股行人重识别常用的数据集有 CSG（CUHK-SYSU-Group）（Yan 等，2023c）、RoadGroup（Xiao 等，2018）和 DukeMTMCGroup（Xiao 等，2018）。表 8 总结了部分小股行人重识别方法在 CSG 和 RoadGroup 数据集上的性能结果，包括 SOT（Zhang 等，2022d）、3DT（Zhang 等，2022b）、ST（siamese Transformer）（Lu 等，2024）、UMSOT（Zhang 等，2024c）和 PB-SOT（Zhang 等，2025e）。

**表 8　部分小股行人重识别方法的性能表现**
**Table 8　Performance of some group person Re-ID methods**

/%

| 方法 | 年份 | 数据集 | | | |
|---|---|---|---|---|---|
| | | CSG | | RoadGroup | |
| | | Rank-1 | mAP | Rank-1 | mAP |
| SOT | 2022 | 91.7 | 90.7 | 86.4 | 91.3 |
| 3DT | 2022 | 92.9 | 92.1 | 91.4 | 94.3 |
| ST | 2024 | 96.3 | **96.6** | **95.5** | 96.7 |
| UMSOT | 2024 | 93.6 | 92.6 | 88.9 | 91.7 |
| PB-SOT | 2025 | **96.4** | 95.1 | 95.1 | **97.0** |

注：加粗字体为每列最优值。

## 3　动物重识别

目标重识别的研究对象已由经典的行人和车辆，延伸至野生动物领域。动物重识别在生态保护、环境监测及农业管理等场景中展现出广阔的应用前景，正成为保护生物多样性、维持生态平衡的有力工具。

在计算机视觉技术之前，动物重识别主要依赖于生物学家的定性观察与人工干预手段。具体方法包括：标记法，即为动物佩戴特定标识物进行个体区分（Wolcott，1995；Block 等，2001；Cooke 等，2004）；疤痕识别法，依据动物体表自然或受伤形成的疤痕特征进行辨识（Bertulli 等，2016）；环志法，主要通过为鸟类等动物佩戴编号脚环实现识别（Carroll 等，

2017)。然而,这些传统方法普遍存在劳动强度大、实施成本高的问题(Schneider等,2019),并且可能对动物造成生理不适,甚至干扰其自然行为(Walker等,2012)。之后,在计算机视觉技术应用于动物重识别领域的早期,主要基于手工设计的特征进行个体识别(Duyck等,2015;Bolger等,2012;Crall等,2013)。2014年后,受行人重识别技术的启发,动物重识别迈入深度学习时代。根据Ye等人(2025)综述的总结与分析,基于深度学习的动物重识别方法可分为基于全局图像学习的方法、基于关键局部区域学习的方法以及基于辅助信息学习的方法。同时,Ye等人(2025)通过制定统一标准和测试多种先进的通用重识别算法,为未来开发更鲁棒、可扩展的动物重识别技术指明方向。

研究人员针对野生动物重识别领域提出了一系列开创性基础模型,如:MegaDescriptor(Čermák等,2024)和WildFusion(Cermak等,2025)。近期的一项重要进展是大型基准数据集WildlifeReID-10k(wildlife re-identification dataset with 10k individual animals)(Adam等,2025)的发布。该数据集系统性地整合了全球37个现有公开资源,包含超过14万幅图像和约1万个动物个体,涵盖包括野生肉食动物、非洲食草动物、海洋生物及灵长类在内的33个物种,构成了当前规模最大、最多样化的公开基准。同时,该工作针对传统评估中普遍存在的训练—测试数据泄露问题,提出创新的数据划分协议,为动物重识别模型提供更严谨、更贴近真实场景的评估基准。因此,WildlifeReID-10k通过其提出的数据划分协议和公开基准,为未来面向动态野生种群的开放重识别研究确立了更真实、可靠的科学评估标准。

## 4 结语

目标重识别旨在跨摄像头、跨时间的条件下实现对特定目标的检索与匹配,在智能安防、智慧交通和智慧城市等领域发挥了重要作用。经过十多年的快速发展与进步,其研究对象已从早期的行人、车辆逐步拓展到动物等多样化类别,研究方法也经历了从传统手工特征提取到深度学习自动特征学习的转变。随着大模型等技术的兴起,目标重识别进一步向着通用化、鲁棒化和强泛化的方向推进。与此同时,面向现实世界中复杂多变的特殊场景,相关研究也取得了一系列进展。下面将对目标重识别未来的发展趋势进行分析与展望。

1)大规模数据集构建。当前,无论是行人还是车辆重识别,可用于大模型预训练的高质量数据在规模上仍显不足,且日益严格的隐私保护要求使得大规模收集与公开真实数据面临挑战。为解决数据瓶颈,未来的大规模数据集构建主要有两个值得探索的方向。其一,发展基于虚拟引擎(如Unity3D、UE5等)的合成数据生成技术,该方法能低成本构建可控的虚拟场景、人物、车辆甚至其他目标。当前虽已有虚拟合成的行人和车辆数据集,但由于其与真实数据集的差异较大,应用价值主要体现在辅助训练层面。随着虚拟现实、增强现实以及虚拟生成技术的持续发展,为提供海量、多样、安全且接近现实世界风格的数据带来可能。其二,利用生成式模型合成高保真图像。随着扩散模型等先进生成技术的快速持续发展,有望实现生成具有高度真实感且细节丰富的目标图像,并支持对视角、光照和天气等条件的精准控制,从而为构建覆盖广泛场景、具备高度多样性的大规模数据提供了新方法。此外,如何充分有效地利用生成数据用于重识别模型训练也是未来值得研究的方向。

2)通用目标重识别大模型研究。当前基于大规模预训练技术的重识别方法受限于可用数据规模,普遍采用参数量较小的模型作为骨干网络。尽管这些方法在多个下游任务上的基准数据集上已展现出优越的性能和较强的泛化能力,但随着目标类别数据的持续积累,其学习能力终将面临瓶颈。此外,一些方法直接采用通用视觉任务设计的大模型作为基准架构,并不完全适合作为细粒度目标重识别任务的基准模型。因此,如何设计适合目标重识别任务特性的基准大模型架构,如何训练多场景、多任务统一的强泛化通用目标重识别大模型是未来的重要研究方向及主要发展趋势。

3)高效轻量化模型设计。当前智能交通和安防监控对实时性要求日益提高,而海量摄像头网络产生的数据规模不断扩大,这使得模型计算效率和响应速度成为制约实际应用的主要瓶颈。高效化模型设计是实现技术落地的关键所在,未来目标重识别研究也应聚焦于模型轻量化与检索优化。在模型

轻量化方面,需突破网络架构的参数冗余问题,开发兼顾识别精度与计算效率的精简架构,重点解决边缘设备部署中的资源约束问题。在检索优化方面,需构建智能化的快速匹配系统,通过优化特征比对、建立高效索引等技术创新,实现检索精度与响应速度的动态平衡。这两方面的突破将共同推动目标重识别技术的规模化应用。

## 参考文献(References)


Adam L, Čermák V, Papafitsoros K and Picek L. 2025. WildlifeReID-10k: wildlife re-identification dataset with 10k individual animals//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville, USA: IEEE: 2090-2100 [DOI: 10.1109/CVPRW67362.2025.00197]

Alehdaghi M, Josi A, Cruz R M O and Granger E. 2023. Visible-infrared person re-identification using privileged intermediate information//Proceedings of the European Conference on Computer Vision. Tel Aviv, Israel: Springer: 720-737 [DOI: 10.1007/978-3-031-25072-9_48]

Alehdaghi M, Josi A, Cruz R M O, Shamsolmoali P and Granger E. 2025. Adaptive generation of privileged intermediate information for visible-infrared person re-identification. IEEE Transactions on Information Forensics and Security, 20: 3400-3413 [DOI: 10.1109/TIFS.2025.3541969]

Bertulli C G, Rasmussen M H and Rosso M. 2016. An assessment of the natural marking patterns used for photo-identification of common minke whales and white-beaked dolphins in Icelandic waters. Journal of the Marine Biological Association of the United Kingdom, 96(4): 807-819 [DOI: 10.1017/S0025315415000284]

Bian Y, Liu M, Wang X P, Tang Y and Wang Y N. 2024. Occlusion-aware feature recover model for occluded person re-identification. IEEE Transactions on Multimedia, 26: 5284-5295 [DOI: 10.1109/TMM.2023.3331192]

Block B A, Heidi D, Blackwell S B, Williams T, Prince E, Boustany A M, et al. 2001. Archival and pop-up satellite tagging of Atlantic Bluefin tuna//Sibert J R and Nielsen J L, eds. Electronic Tagging and Tracking in Marine Fisheries. Dordrecht: Springer: 65-88 [DOI: 10.1007/978-94-017-1402-0_3]

Bolger D T, Morrison T A, Vance B, Lee D and Farid H. 2012. A computer-assisted system for photographic mark-recapture analysis. Methods in Ecology and Evolution, 3(5): 813-822 [DOI: 10.1111/j.2041-210X.2012.00212.x]

Cai Q and Aggarwal J K. 1996. Tracking human motion using multiple cameras//Proceedings of the 13th International Conference on Pattern Recognition. Vienna, Austria: IEEE: 68-72 [DOI: 10.1109/ICPR.1996.546796]

Cai Y H, Takala V and Pietikainen M. 2010. Matching groups of people by covariance descriptor//Proceedings of the 20th International Conference on Pattern Recognition. Istanbul, Turkey: IEEE: 2744-2747 [DOI: 10.1109/ICPR.2010.672]

Cao M, Bai Y, Zeng Z Y, Ye M and Zhang M. 2024. An empirical study of clip for text-based person search//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI: 465-473 [DOI: 10.1609/aaai.v38i1.27801]

Carroll J M, Hamm R L, Hagen J M, Davis C A and Guthery F S. 2017. Evaluation of leg banding and attachment of radio-transmitters on ring-necked pheasant chicks. Wildlife Biology, 2017(1): 1-6 [DOI: 10.2981/wlb.00263]

Čermák V, Picek L, Adam L, Neumann L and Matas J. 2025. WildFusion: individual animal identification with calibrated similarity fusion//Proceedings of the Computer Vision. Milan, Italy: Springer: 18-36 [DOI: 10.1007/978-3-031-92387-6_2]

Čermák V, Picek L, Adam L and Papafitsoros K. 2024. WildlifeDatasets: an open-source toolkit for animal re-identification//Proceedings of 2024 IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE: 5941-5951 [DOI: 10.1109/WACV57701.2024.00585]

Chen C Q, Ye M and Jiang D. 2023. Towards modality-agnostic person re-identification with descriptive query//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 15128-15137 [DOI: 10.1109/CVPR52729.2023.01452]

Chen C Q, Ye M, Qi M B and Du B. 2022a. Sketch transformer: asymmetrical disentanglement learning from dynamic synthesis//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal: ACM: 4012-4020 [DOI: 10.1145/3503161.3547993]

Chen C Q, Ye M, Qi M B and Du B. 2024. SketchTrans: disentangled prototype learning with transformer for sketch-photo recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(5): 2950-2964 [DOI: 10.1109/TPAMI.2023.3337005]

Chen C Q, Ye M, Qi M B, Wu J J, Jiang J G and Lin C W. 2022b. Structure-aware positional transformer for visible-infrared person re-Identification. IEEE Transactions on Image Processing, 31: 2352-2364. [DOI: 10.1109/TIP.2022.3141868]

Chen J X, Jiang X Y, Wang F D, Zhang J, Zheng F, Sun X, et al. 2021a. Learning 3D shape feature for texture-insensitive person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 8142-8151 [DOI: 10.1109/CVPR46437.2021.00805]

Chen J X, Zheng W S, Yang Q Z, Meng J K, Hong R C and Tian Q. 2022c. Deep shape-aware person re-identification for overcoming moderate clothing changes. IEEE Transactions on Multimedia, 24: 4285-4300 [DOI: 10.1109/TMM.2021.3114539]

Chen P X, Liu W F, Dai P Y, Liu J Z, Ye Q X, Xu M L, et al. 2021b. Occlude them all: occlusion-aware attention network for occluded person Re-ID//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 11813-11822 [DOI: 10.1109/ICCV48922.2021.01162]

Chen S, Qiu L X, Wang D H, Zhu W T, Hua Y and Yan Y. 2025. Hierarchical token-aware Cross-modality reconstruction for visible-infrared person re-identification. IEEE Transactions on Multimedia, 27: 7012-7027 [DOI: 10.1109/TMM.2025.3590933]

Chen Y B, Zhu X T and Gong S G. 2019. Instance-guided context rendering for cross-domain person re-identification//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 232-242 [DOI: 10.1109/ICCV.2019.00032]

Cheng X, Yu H, Cheng K H M, Yu Z T and Zhao G Y. 2025. MDANet: modality-aware domain alignment network for visible-infrared person re-identification. IEEE Transactions on Multimedia, 27: 2015-2027 [DOI: 10.1109/TMM.2024.3521822]

Cho Y, Kim W J, Hong S and Yoon S E. 2022. Part-based pseudo label refinement for unsupervised person re-identification//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 7308-7318 [DOI: 10.1109/CVPR52688.2022.00716]

Choi S, Lee S, Kim Y, Kim T and Kim C. 2020. Hi-CMD: hierarchical cross-modality disentanglement for visible-infrared person re-identification//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 10257-10266 [DOI: 10.1109/CVPR42600.2020.01027]

Cooke S J, Hinch S G, Wikelski M, Andrews R D, Kuchel L J, Wolcott T G, et al. 2004. Biotelemetry: a mechanistic approach to ecology. Trends in Ecology and Evolution, 19(6): 334-343 [DOI: 10.1016/j.tree.2004.04.003]

Crall J P, Stewart C V, Berger-Wolf T Y, Rubenstein D I and Sundaresan S R. 2013. HotSpotter—Patterned species instance recognition//Proceedings of 2013 IEEE Workshop on Applications of Computer Vision. Clearwater Beach, USA: IEEE: 230-237 [DOI: 10.1109/WACV.2013.6475023]

Cui Z Y, Zhou J H, Peng Y X, Zhang S L and Wang Y W. 2023. DCR-ReID: deep component reconstruction for cloth-changing person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 33(8): 4415-4428 [DOI: 10.1109/TCSVT.2023.3241988]

Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]

Deng W J, Zheng L, Ye Q X, Kang G L, Yang Y and Jiao J B. 2018. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 994-1003 [DOI: 10.1109/CVPR.2018.00110]

Ding Z F, Ding C X, Shao Z Y and Tao D C. 2021. Semantically self-aligned network for text-to-image part-aware person re-identification [EB/OL]. [2026-03-01]. https://arxiv.arg/pdf/2107.12666.pdf

Dong N, Yan S L, Zhang L Y and Tang J H. 2025. Diverse semantics-guided feature alignment and decoupling for visible-infrared person re-identification. IEEE Transactions on Information Forensics and Security, 20: 12245-12259 [DOI: 10.1109/TIFS.2025.3629581]

Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, et al. 2021. An image is worth 16 × 16 words: transformers for image recognition at scale//Proceedings of the 9th International Conference on Learning Representations. [s.l.]: OpenReview.net

Dou S G, Zhao C R, Jiang X Y, Zhang S S, Zheng W S and Zuo W M. 2023. Human co-parsing guided alignment for occluded person re-identification. IEEE Transactions on Image Processing, 32: 458-470 [DOI: 10.1109/TIP.2022.3229639]

Duyck J, Finn C, Hutcheon A, Vera P, Salas J and Ravela S. 2015. Sloop: a pattern retrieval engine for individual animal identification. Pattern Recognition, 48(4): 1059-1073 [DOI: 10.1016/j.patcog.2014.07.017]

Farenzena M, Bazzani L, Perina A, Murino V and Cristani M. 2010. Person re-identification by symmetry-driven accumulation of local features//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE: 2360-2367 [DOI: 10.1109/CVPR.2010.5539926]

Farooq A, Awais M, Kittler J and Khalid S S. 2022. AXM-Net: implicit cross-modal feature alignment for person re-identification//Proceedings of the 36th AAAI Conference on Artificial Intelligence. Palo Alto, USA: AAAI: 4477-4485 [DOI: 10.1609/aaai.v36i4.20370]

Feng Y Y, Li J, Hu J, Zhang Y K, Tan L and Ji J Y. 2025. MDReID: modality-decoupled learning for any-to-any multi-modal object re-identification//Proceedings of the 39th Conference on Neural Information Processing Systems. San Diego, USA: [s.n.]

Feng Z X, Lai J H, Yuan Z, Huang Y L and Lai P J. 2025. Advancing universal person reidentification: a survey on the applications of large-scale, pretraining models for identifying individuals. Journal of Image and Graphics, 30(6): 1638-1660 (冯展祥, 赖剑煌, 袁藏, 黄宇立, 赖培杰. 2025. 走向通用行人重识别: 预训练大模型技术在行人重识别的应用综述. 中国图象图形学报, 30(6): 1638-1660) [DOI: 10.11834/jig.240426]

Feris R S, Siddiquie B, Petterson J, Zhai Y, Datta A, Brown L M, et al. 2012. Large-scale vehicle detection, indexing, and search in urban surveillance videos. IEEE Transactions on Multimedia, 14(1): 28-42 [DOI: 10.1109/TMM.2011.2170666]

Fu D P, Chen D D, Bao J M, Yang H, Yuan L, Zhang L, et al. 2021. Unsupervised pre-training for person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 14745-14754 [DOI: 10.1109/CVPR46437.2021.01451]

Gao L Y, Jiao B L, Long Y Z, Niu K, Huang H, Wang P, et al. 2024a.

Contrastive pedestrian attentive and correlation learning network for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 34(9): 8862-8880 [DOI: 10.1109/TCSVT.2024.3379577]

Gao S, Wang J Y, Lu H C and Liu Z M. 2020. Pose-guided visible part matching for occluded person ReID//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 11741-11749 [DOI: 10.1109/CVPR42600.2020.01176]

Gao Z, Chen P, Zhuo T, Liu M, Zhu L, Wang M, et al. 2024b. A semantic perception and CNN-transformer hybrid network for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 34(4): 2010-2025 [DOI: 10.1109/TCSVT.2023.3296680]

Gao Z, Wei H W, Guan W L, Nie J, Wang M and Chen S Y. 2025. A semantic-aware attention and visual shielding network for cloth-changing person re-identification. IEEE Transactions on Neural Networks and Learning Systems, 36(1): 1243-1257 [DOI: 10.1109/TNNLS.2023.3329384]

Gao Z, Wei S X, Guan W L, Zhu L, Wang M and Chen S Y. 2024c. Identity-guided collaborative learning for cloth-changing person reidentification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(5): 2819-2837 [DOI: 10.1109/TPAMI.2023.3334741]

Ge Y X, Chen D P and Li H S. 2020a. Mutual mean-teaching: pseudo label refinery for unsupervised domain adaptation on person re-identification//Proceedings of the 8th International Conference on Learning Representations. Addis Ababa, Ethiopia: OpenReview.net

Ge Y X, Zhu F, Chen D P, Zhao R and Li H S. 2020b. 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, Canada: Curran Associates Inc.

Geng S Z, Liu Y F, Wang Z J, Yan G, Yu Y and Guo Y C. 2025. Pose-skeleton guided cross-attention representation fusion for occluded pedestrian re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(9): 8598-8613 [DOI: 10.1109/TCSVT.2025.3556250]

Gheissari N, Sebastian T B and Hartley R. 2006. Person reidentification using spatiotemporal appearance//Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE: 1528-1535 [DOI: 10.1109/CVPR.2006.223]

Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. 2014. Generative adversarial nets//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press

Gray D and Tao H. 2008. Viewpoint invariant pedestrian recognition with an ensemble of localized features//Proceedings of the 10th European Conference on Computer Vision. Marseille, France: Springer: 262-275 [DOI: 10.1007/978-3-540-88682-2_21]

Gu X Q, Chang H, Ma B P, Bai S T, Shan S G and Chen X L. 2022. Clothes-changing person re-identification with RGB modality only//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 1050-1059 [DOI: 10.1109/CVPR52688.2022.00113]

Han D C, Liu W F, Zou M C and Liu B D. 2023. Non-contrastive nearest neighbor identity-guided method for unsupervised object re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 33(6): 2713-2723 [DOI: 10.1109/TCSVT.2022.3224994]

Han Q, Li L F, Min W D, Wang Q, Zeng Q P, Cui S M, et al. 2024. Joint training with local soft attention and dual cross-neighbor label smoothing for unsupervised person re-identification. Computational Visual Media, 10(3): 543-558 [DOI: 10.1007/s41095-023-0354-4]

Han Q, Liu H T, Min W D, Huang T M, Lin D Y and Wang Q. 2022. 3D Skeleton and two streams approach to person re-identification using optimized region matching. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 18(2s): #129 [DOI: 10.1145/3538490]

Hao C R, Wang B Y, Peng C L, Liu D C, Wang N N, Hu R M, et al. 2025. Masked text adversarial training for cloth-changing person re-identification. IEEE Transactions on Information Forensics and Security, 20: 11516-11527 [DOI: 10.1109/TIFS.2025.3625662]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

He Q L, Wang Z H, Zheng Z J and Hu H F. 2024a. Spatial and temporal dual-attention for unsupervised person re-identification. IEEE Transactions on Intelligent Transportation Systems, 25(2): 1953-1965 [DOI: 10.1109/TITS.2023.3314453]

He Q L, Yang Y M, Wang Z H and Hu H F. 2025. Part-based bi-directional enhancement learning for unsupervised visible-infrared re-identification. IEEE Transactions on Information Forensics and Security, 20: 12051-12065 [DOI: 10.1109/TIFS.2025.3629565]

He S T, Chen W H, Wang K, Luo H, Wang F, Jiang W, et al. 2024b. Region generation and assessment network for occluded person re-identification. IEEE Transactions on Information Forensics and Security, 19: 120-132 [DOI: 10.1109/TIFS.2023.3318956]

He S T, Luo H, Jiang W, Jiang X D and Ding H H. 2024c. VGSG: vision-guided semantic-group network for text-based person search. IEEE Transactions on Image Processing, 33: 163-176 [DOI: 10.1109/TIP.2023.3337653]

He W Z, Deng Y H, Tang S X, Chen Q H, Xie Q S, Wang Y Z, et al. 2024d. Instruct-ReID: a multi-purpose person re-identification task with instructions//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE:

17521-17531 [DOI: 10.1109/CVPR52733.2024.01659]

Ho J, Jain A and Abbeel P. 2020. Denoising diffusion probabilistic models//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.

Hong P X, Wu T, Wu A C, Han X T and Zheng W S. 2021. Fine-grained shape-appearance mutual learning for cloth-changing person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 10508-10517 [DOI: 10.1109/CVPR46437.2021.01037]

Hou R B, Ma B P, Chang H, Gu X Q, Shan S G and Chen X L. 2022. Feature completion for occluded person re-identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(9): 4894-4912 [DOI: 10.1109/TPAMI.2021.3079910]

Hu G, Lyu Y F, Zhang J T, Wu Q and Wen Z D. 2025a. Clip-based modality compensation for visible-infrared image re-identification. IEEE Transactions on Multimedia, 27: 2112-2126 [DOI: 10.1109/TMM.2024.3521764]

Hu Z C, Yang Z W, Li H, Wang Z and Zou Y X. 2025b. Cross-category subjectivity generalization for style-adaptive sketch Re-ID//Proceedings of 2025 IEEE/CVF International Conference on Computer Vision. Honolulu, USA: IEEE: 22644-22653

Hua B Y, Zhang J Y, Li Z Q and Ge Y X. 2023. Cross-modality channel mixup and modality decorrelation for RGB-infrared person re-Identification. IEEE Transactions on Biometrics, Behavior, and Identity Science, 5(4): 512-523. [DOI: 10.1109/TBIOM.2023.3287275]

Huang M Y, Hou C P, Yang Q Y and Wang Z P. 2023. Reasoning and tuning: graph attention network for occluded person re-identification. IEEE Transactions on Image Processing, 32: 1568-1582 [DOI: 10.1109/TIP.2023.3247159]

Huang T and Russell S. 1997. Object identification in a Bayesian context//Proceedings of the 15th International Joint Conference on Artificial Intelligence. Nagoya, Japan: Morgan Kaufmann Publishers Inc.: 1276-1282

Huang Y, Wu Q, Xu J S, Zhong Y and Zhang Z X. 2021a. Clothing status awareness for long-term person re-identification//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 11875-11884 [DOI: 10.1109/ICCV48922.2021.01168]

Huang Y, Xu J S, Wu Q, Zhong Y, Zhang P and Zhang Z X. 2020. Beyond scalar neuron: adopting vector-neuron capsules for long-term person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 30(10): 3459-3471 [DOI: 10.1109/TCSVT.2019.2948093]

Huang Z L, Wang Z, Hu W, Lin C W and Satoh S. 2019. DoT-GNN: domain-transferred graph neural network for group re-identification//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 1888-1896 [DOI: 10.1145/3343031.3351027]

Huang Z L, Wang Z, Tsai C C, Satoh S and Lin C W. 2021b. DotSCN: group re-identification via domain-transferred single and couple representation learning. IEEE Transactions on Circuits and Systems for Video Technology, 31(7): 2739-2750 [DOI: 10.1109/TCSVT.2020.3031303]

Isobe T, Li D, Tian L, Chen W H, Shan Y and Wang S J. 2021. Towards discriminative representation learning for unsupervised person re-identification//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 8526-8536 [DOI: 10.1109/ICCV48922.2021.00841]

Isola P, Zhu J Y, Zhou T H and Efros A A. 2017. Image-to-image translation with conditional adversarial networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5967-5976 [DOI: 10.1109/CVPR.2017.632]

Ji H X Y, Wang L, Zhou S P, Tang W and Hua G. 2024. Disentangled sample guidance learning for unsupervised person re-identification. IEEE Transactions on Image Processing, 33: 5144-5158 [DOI: 10.1109/TIP.2024.3456008]

Ji H X Y, Wang L, Zhou S P, Tang W, Zheng N N and Hua G. 2025a. Meta pairwise relationship distillation for unsupervised person re-identification. IEEE Transactions on Neural Networks and Learning Systems, 36 (5) : 9165-9179 [DOI: 10.1109/TNNLS. 2024.3409685]

Ji W H, Cheng X, Jiang Y, Sun Z D and Zhao G Y. 2025b. Learning from yourself to others for unsupervised visible-infrared re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(10): 9823-9836 [DOI: 10.1109/TCSVT.2025.3572697]

Jia M X, Cheng X H, Lu S J and Zhang J. 2023. Learning disentangled representation implicitly via transformer for occluded person re-identification. IEEE Transactions on Multimedia, 25: 1294-1305 [DOI: 10.1109/TMM.2022.3141267]

Jiang D and Ye M. 2023. Cross-modal implicit relation reasoning and aligning for text-to-image person retrieval//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 2787-2797 [DOI: 10.1109/CVPR52729.2023.00273]

Jiang Y, Cheng X, Yu H, Liu X Y, Chen H Y and Zhao G Y. 2025. DSAF: dual space alignment framework for visible-infrared person re-identification. IEEE Transactions on Multimedia, 27: 5591-5603 [DOI: 10.1109/TMM.2025.3542988]

Jin X, He T Y, Zheng K C, Yin Z H, Shen X, Huang Z, et al 2022. Cloth-changing person re-identification from a single image with gait prediction and regularization//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 14258-14267 [DOI: 10.1109/CVPR52688.2022.01388]

Jüngling K, Bodensteiner C and Arens M. 2011. Person re-identification in multi-camera networks//Proceedings of 2011 WORKSHOPS.

Colorado Springs, USA: IEEE: 55-61 [DOI: 10.1109/CVPRW.2011.5981771]

Kim M, Kim S, Park J, Park S and Sohn K. 2023. PartMix: regularization strategy to learn part discovery for visible-infrared person re-identification//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 18621-18632 [DOI: 10.1109/CVPR52729.2023.01786]

Kniaz V V, Knyaz V A, Hladůvka J, Kropatsch W G and Mizginov V. 2019. ThermalGAN: multimodal color-to-thermal image translation for person re-identification in multispectral dataset//Proceedings of the Computer Vision. Munich, Germany: Springer: 606-624 [DOI: 10.1007/978-3-030-11024-6_46]

Köstinger M, Hirzer M, Wohlhart P, Roth P M and Bischof H. 2012. Large scale metric learning from equivalence constraints//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE: 2288-2295 [DOI: 10.1109/CVPR.2012.6247939]

Layne R, Hospedales T and Gong S G. 2012. Person re-identification by attributes//Proceedings of 2012 British Machine Vision Conference. Surrey, UK: BMVA Press: #24 [DOI: 10.5244/C.26.24]

Li H, Ye M, Zhang M and Du B. 2024a. All in one framework for multimodal re-identification in the wild//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 17459-17469 [DOI: 10.1109/CVPR52733.2024.01653]

Li H C, Li C L, Zhu X P, Zheng A H and Luo B. 2020. Multi-spectral vehicle re-identification: a challenge//Proceedings of 2020 AAAI Conference on Artificial Intelligence. New York, USA: AAAI: 11345-11353 [DOI: 10.1609/aaai.v34i07.6796]

Li H J, Li M X, Peng Q J, Wang S J, Yu H and Wang Z H. 2024b. Correlation-guided semantic consistency network for visible-infrared person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 34(6): 4503-4515 [DOI: 10.1109/TCSVT.2023.3340225]

Li L F, Han Q, Min W D, Wang Q, Zhu M, Huang Z, et al. 2026. Outliers adaptation exploration and centroids matching label refinement for unsupervised person re-identification. IEEE Transactions on Multimedia: #3654461 [DOI: 10.1109/TMM.2026.3654461]

Li S, Xiao T, Li H S, Yang W and Wang X G. 2017a. Identity-aware textual-visual matching with latent co-attention//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: #209 [DOI: 10.1109/ICCV.2017.209]

Li S, Xiao T, Li H S, Zhou B L, Yue D Y and Wang X G. 2017b. Person search with natural language description//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5187-5196 [DOI: 10.1109/CVPR.2017.551]

Li S Y, Sun L and Li Q L. 2023. CLIP-ReID: exploiting vision-language model for image re-identification without concrete text labels//Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington, USA: AAAI: 1405-1413 [DOI: 10.1609/aaai.v37i1.25225]

Li W, Zhao R, Xiao T and Wang X G. 2014. DeepReID: deep filter pairing neural network for person re-identification//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 152-159 [DOI: 10.1109/CVPR.2014.27]

Li Y J, Lin C S, Lin Y B and Wang Y C F. 2019. Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 7918-7928 [DOI: 10.1109/ICCV.2019.00801]

Li Y L, He J F, Zhang T Z, Liu X, Zhang Y D and Wu F. 2021. Diverse part discovery: occluded person re-identification with part-aware transformer//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 2897-2906 [DOI: 10.1109/CVPR46437.2021.00292]

Li Y P, Liu Y Z, Zhang H Y, Zhao C R, Wei Z H and Miao D Q. 2024c. Occlusion-aware transformer with second-order attention for person re-identification. IEEE Transactions on Image Processing, 33: 3200-3211 [DOI: 10.1109/TIP.2024.3393360]

Li Z H, Zhang H X, Zhu L, Sun J D and Liu L. 2025. Heterogeneous generative tokens and distance-aware recovery network for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(5): 5022-5036 [DOI: 10.1109/TCSVT.2024.3519312]

Liang T F, Jin Y, Liu W and Li Y D. 2023. Cross-modality transformer with modality mining for visible-infrared person re-Identification. IEEE Transactions on Multimedia, 25: 8432-8444. [DOI: 10.1109/TMM.2023.3237155]

Liang X and Rawat Y S. 2025. DIFFER: disentangling identity features via semantic cues for clothes-changing person re-id//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 13980-13989 [DOI: 10.1109/CVPR52734.2025.01305]

Lin D X, Peng Y X, Meng J K and Zheng W S. 2024. Cross-modal adaptive dual association for text-to-image person retrieval. IEEE Transactions on Multimedia, 26: 6609-6620 [DOI: 10.1109/TMM.2024.3355644]

Lin G R, Yang S Z, Zheng W S, Li Z Y and Huang Z H. 2025. A semantically guided and focused network for occluded person re-identification. IEEE Transactions on Information Forensics and Security, 20: 9716-9731 [DOI: 10.1109/TIFS.2025.3608672]

Lin K J, Wang Z X, Wang Z, Zheng Y Q and Satoh S. 2023. Beyond domain gap: exploiting subjectivity in sketch-based person retrieval//Proceedings of the 31st ACM International Conference on Multimedia. Ottawa, Canada: ACM: 2078-2089 [DOI: 10.1145/3581783.3611732]

Lin W Y, Li Y X, Xiao H, See J, Zou J N, Xiong H K, et al. 2021. Group reidentification with multigrained matching and integration.

IEEE Transactions on Cybernetics, 51(3): 1478-1492 [DOI: 10.1109/TCYB.2019.2917713]

Lisanti G, Martinel N, Del Bimbo A and Luca Foresti G. 2017. Group re-identification via unsupervised transfer of sparse features encoding//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2468-2477 [DOI: 10.1109/ICCV.2017.268]

Liu F, Kim M, Gu Z A, Jain A and Liu X M. 2023a. Learning clothing and pose invariant 3D shape representation for long-term person re-identification//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 19560-19569 [DOI: 10.1109/ICCV51070.2023.01797]

Liu F Y, Ye M and Du B. 2023b. Dual level adaptive weighting for cloth-changing person re-identification. IEEE Transactions on Image Processing, 32: 5075-5086 [DOI: 10.1109/TIP.2023.3310307]

Liu H J, Tan X H and Zhou X C. 2021. Parameter sharing exploration and hetero-center triplet loss for visible-thermal person re-identification. IEEE Transactions on Multimedia, 23: 4414-4425 [DOI: 10.1109/TMM.2020.3042080]

Liu H Y, Tian Y H, Wang Y W, Pang L and Huang T J. 2016a. Deep relative distance learning: tell the difference between similar vehicles//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 2167-2175 [DOI: 10.1109/CVPR.2016.238]

Liu J N, Wang J L, Huang N C, Zhang Q and Han J G. 2022a. Revisiting modality-specific feature compensation for visible-infrared person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 32(10): 7226-7240 [DOI: 10.1109/TCSVT.2022.3168999]

Liu M, Zhang Z, Bian Y, Wang X P, Sun Y Q, Zhang B D, et al. 2025a. Cross-modality semantic consistency learning for visible-infrared person re-identification. IEEE Transactions on Multimedia, 27: 568-580 [DOI: 10.1109/TMM.2024.3521843]

Liu T Y, Lin Y T and Du B. 2022b. Unsupervised person re-identification with stochastic training strategy. IEEE Transactions on Image Processing, 31: 4240-4250 [DOI: 10.1109/TIP.2022.3181811]

Liu X C, Liu W, Ma H D and Fu H Y. 2016b. Large-scale vehicle re-identification in urban surveillance videos//Proceedings of 2016 IEEE International Conference on Multimedia and Expo. Seattle, USA: IEEE: 1-6 [DOI: 10.1109/ICME.2016.7553002]

Liu X C, Liu W, Mei T and Ma H D. 2018. PROVID: progressive and multimodal vehicle reidentification for large-scale urban surveillance. IEEE Transactions on Multimedia, 20(3): 645-658 [DOI: 10.1109/TMM.2017.2751966]

Liu X Y, Cheng X, Chen H Y, Yu H and Zhao G Y. 2024a. Differentiable auxiliary learning for sketch re-identification//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI: 3747-3755 [DOI: 10.1609/aaai.v38i4.28165]

Liu X Z, Guo J F, Chen H, Miao Q G, Xi Y and Liu R Y. 2025b. Adaptive occlusion-aware network for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(5): 5067-5077 [DOI: 10.1109/TCSVT.2024.3524555]

Liu X Z, Liu K P, Guo J F, Zhao P P, Quan Y N and Miao Q G. 2024b. Pose-guided attention learning for cloth-changing person re-identification. IEEE Transactions on Multimedia, 26: 5490-5498 [DOI: 10.1109/TMM.2023.3334975]

Liu Y X, Ge H W, Wang Z, Hou Y Q and Zhao M D. 2024c. Discriminative identity-feature exploring and differential aware learning for unsupervised person re-identification. IEEE Transactions on Multimedia, 26: 623-636 [DOI: 10.1109/TMM.2023.3268369]

Liu Y X, Zhang W M, Vasilakos A V and Wang L. 2025c. Unsupervised visible-infrared ReID via pseudo-label correction and modality-level alignment. IEEE Transactions on Neural Networks and Learning Systems, 36(11): 19631-19643 [DOI: 10.1109/TNNLS.2025.3591641]

Lou Y H, Bai Y, Liu J, Wang S Q and Duan L Y. 2019. VERI-wild: a large dataset and a new method for vehicle re-identification in the wild//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3230-3238 [DOI: 10.1109/CVPR.2019.00335]

Lu X Y, Li X D, Hu C F, Dunkin F, Li H Q and Ge S S. 2024. Siamese transformer for group re-identification via multiscale feature transform and joint learning. IEEE Transactions on Instrumentation and Measurement, 73: #5002910 [DOI: 10.1109/TIM.2023.3330216]

Lu Y, Wu Y, Liu B, Zhang T Z, Li B P, Chu Q, et al. 2020. Cross-modality person re-identification with shared-specific feature transfer//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 13379-13389 [DOI: 10.1109/CVPR42600.2020.01339]

Lu Z F, Lin R H and Hu H F. 2023. Modality and camera factors bi-disentanglement for NIR-VIS object re-identification. IEEE Transactions on Information Forensics and Security, 18: 1989-2004 [DOI: 10.1109/TIFS.2023.3262130]

Lu Z F, Lin R H and Hu H F. 2025. Disentangling modality and posture factors: memory-attention and orthogonal decomposition for visible-infrared person re-identification. IEEE Transactions on Neural Networks and Learning Systems, 36(3): 5494-5508 [DOI: 10.1109/TNNLS.2024.3384023]

Miao J X, Wu Y, Liu P, Ding Y H and Yang Y. 2019. Pose-guided feature alignment for occluded person re-identification//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 542-551 [DOI: 10.1109/ICCV.2019.00063]

Miao J X, Wu Y and Yang Y. 2022. Identifying visible parts via pose estimation for occluded person re-identification. IEEE Transactions on Neural Networks and Learning Systems, 33(9): 4624-4634 [DOI: 10.1109/TNNLS.2021.3059515]

Nguyen D T, Hong H G, Kim K W and Park K R. 2017. Person recognition system based on a combination of body images from visible light and thermal cameras. Sensors, 17(3): #605 [DOI: 10.3390/s17030605]

Nguyen V D, Mantini P and Shah S K. 2024. Contrastive clothing and pose generation for cloth-changing person re-identification//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA: IEEE: 7541-7549 [DOI: 10.1109/CVPRW63382.2024.00749]

Niu K, Huang Y, Ouyang W L and Wang L. 2020. Improving description-based person re-identification by multi-granularity image-text alignments. IEEE Transactions on Image Processing, 29: 5542-5556 [DOI: 10.1109/TIP.2020.2984883]

Oreifej O, Mehran R and Shah M. 2010. Human identity recognition in aerial images//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE: 709-716 [DOI: 10.1109/CVPR.2010.5540147]

Pan H H, Pei W J, Li X and He Z Y. 2024. Unified conditional image generation for visible-infrared person re-identification. IEEE Transactions on Information Forensics and Security, 19: 9026-9038 [DOI: 10.1109/TIFS.2024.3426335]

Pang L, Wang Y W, Song Y Z, Huang T J and Tian Y H. 2018. Cross-domain adversarial feature learning for sketch re-identification//Proceedings of the 26th ACM International Conference on Multimedia. Seoul, Korea (South): ACM: 609-617 [DOI: 10.1145/3240508.3240606]

Peng C L, Wang B Y, Liu D C, Wang N N, Hu R M and Gao X B. 2025. Masked attribute description embedding for cloth-changing person re-identification. IEEE Transactions on Multimedia, 27: 1475-1485 [DOI: 10.1109/TMM.2024.3521730]

Peng J J, Wang Y, Wang H B, Zhang Z, Fu X P and Wang M. 2020. Unsupervised vehicle re-identification with progressive adaptation//Proceedings of the 29th International Joint Conference on Artificial Intelligence. Yokohama, Japan: IJCAI: 913-919

Qi M Z, Chan S X, Hang C, Zhang G X, Zeng T Y and Li Z. 2025. Auxiliary representation guided network for visible-infrared person re-identification. IEEE Transactions on Multimedia, 27: 340-355 [DOI: 10.1109/TMM.2024.3521773]

Qian X L, Wang W X, Zhang L, Zhu F R, Fu Y W, Xiang T, et al. 2021. Long-term cloth-changing person re-identification//Proceedings of the 15th Asian Conference on Computer Vision. Kyoto, Japan: Springer: 71-88 [DOI: 10.1007/978-3-030-69535-4_5]

Qiu M K, Lu Y H, Li X Y and Lu Q. 2024a. Camera-aware differentiated clustering with focal contrastive learning for unsupervised vehicle re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 34(10): 10121-10134 [DOI: 10.1109/TCSVT.2024.3402109]

Qiu M K, Lu Y H, Li X Y and Lu Q. 2024b. Inter-intra cluster reorganization for unsupervised vehicle re-identification. IEEE Transactions on Intelligent Transportation Systems, 25(12): 20493-20507 [DOI: 10.1109/TITS.2024.3464585]

Radford A, Kim J W, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. 2021. Learning transferable visual models from natural language supervision//Proceedings of the 38th International Conference on Machine Learning. Vienna, Austria: ICML: 8748-8763

Radford A, Narasimhan K, Salimans T and Sutskever I. 2018. Improving language understanding by generative pre-training [EB/OL]. [2026-03-04].
https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]

Ren K J and Zhang L. 2024. Implicit discriminative knowledge learning for visible-infrared person re-identification//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE: 393-402 [DOI: 10.1109/CVPR52733.2024.00045]

Ren S Q, He K M, Girshick R S J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031]

Ristani E, Solera F, Zou R, Cucchiara R and Tomasi C. 2016. Performance measures and a data set for multi-target, multi-camera tracking//Proceedings of the Computer Vision. Amsterdam, the Netherlands: Springer: 17-35 [DOI: 10.1007/978-3-319-48881-3_2]

Sarafianos N, Xu X and Kakadiaris I. 2019. Adversarial representation learning for text-to-image matching//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: #591 [DOI: 10.1109/ICCV.2019.00591]

Schneider S, Taylor G W, Linquist S and Kremer S C. 2019. Past, present and future approaches using computer vision for animal re-identification from camera trap data. Methods in Ecology and Evolution, 10(4): 461-470 [DOI: 10.1111/2041-210X.13133]

Shu X J, Wen W, Wu H Q, Chen K Y, Song Y R, Qiao R Z, et al. 2023. See finer, see more: implicit modality alignment for text-based person retrieval//Proceedings of Computer Vision. Tel Aviv, Israel: Springer: 624-641 [DOI: 10.1007/978-3-031-25072-9_42]

Sun R, Du Y, Huang G X, Wang X B and Wu J J. 2025. Implicit alignment-based cross-modal symbiotic network for text-to-image person re-identification. IEEE Transactions on Information Forensics and Security, 20: 8069-8082 [DOI: 10.1109/TIFS.2025.3594558]

Sun W, Guan F, Zhang X R and Shen X Y. 2025. Foggy ship re-identification network based on multiple feature cascade enhancement and cross-layer adaptive fusion. Journal of Image and Graphics, 30(8): 2808-2821 (孙伟，管菲，张小瑞，沈欣怡. 2025. 多元特征级联增强和跨层自适应融合的雾天船舶重识别网络. 中国图象图形学报，30(8): 2808-2821) [DOI: 10.11834/jig.240646]

Sun W, Hu Y H, Zhang X R, Yao X and He X Z. 2024. Adversarial style-irrelevant feature learning with refined soft pseudo labels for domain-adaptive vehicle re-identification. IEEE Transactions on Intelligent Transportation Systems, 25(12): 20602-20615 [DOI: 10.1109/TITS.2024.3480156]

Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z. 2016. Rethinking the inception architecture for computer vision//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 2818-2826 [DOI: 10.1109/CVPR.2016.308]

Tan H C, Liu X P, Bian Y H, Wang H S and Yin B C. 2022. Incomplete descriptor mining with elastic loss for person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 32(1): 160-171 [DOI: 10.1109/TCSVT.2021.3061412]

Tan H C, Liu X P, Yin B C and Li X. 2023. MHSA-Net: multihead self-attention network for occluded person re-identification. IEEE Transactions on Neural Networks and Learning Systems, 34(11): 8210-8224 [DOI: 10.1109/TNNLS.2022.3144163]

Tao X F, Kong J, Jiang M, Lu M and Mian A. 2024. Unsupervised learning of intrinsic semantics with diffusion model for person re-identification. IEEE Transactions on Image Processing, 33: 6705-6719 [DOI: 10.1109/TIP.2024.3514360]

Walker K A, Trites A W, Haulena M and Weary D M. 2012. A review of the effects of different marking and tagging techniques on marine mammals. Wildlife Research, 39 (1): 15-30 [DOI: 10.1071/WR10177]

Wan F B, Wu Y, Qian X L, Chen Y X and Fu Y W. 2020. When person re-identification meets changing clothes//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA: IEEE: 3620-3628 [DOI: 10.1109/CVPRW50498.2020.00423]

Wan X X, Zheng A H, Jiang B, Wang B B, Li C L and Tang J. 2025. UGG-ReID: uncertainty-guided graph model for multi-modal object re-identification//Proceedings of the 39th Conference on Neural Information Processing Systems. San Diego, USA: [s.n.]

Wang C S, Gao X Y, Wu M Q, Lam S K, He S T and Tiwari P. 2025a. Looking clearer with text: a hierarchical context blending network for occluded person re-identification. IEEE Transactions on Information Forensics and Security, 20: 4296-4307 [DOI: 10.1109/TIFS.2025.3558586]

Wang D, Wang Q, Min W D, Gai D, Han Q, Li L F, et al. 2024a. SAM-driven MAE pre-training and background-aware meta-learning for unsupervised vehicle re-identification. Computational Visual Media, 10 (4): 771-789 [DOI: 10.1007/s41095-024-0424-2]

Wang D, Wang Q, Tu Z W, Min W D, Xiong X, Zhong Y L, et al. 2024b. Vision-language constraint graph representation learning for unsupervised vehicle re-identification. Expert Systems with Applications, 255: #124495 [DOI: 10.1016/j.eswa.2024.124495]

Wang G A, Yang S, Liu H Y, Wang Z C, Yang Y, Wang S L, et al. 2020a. High-order information matters: learning relation and topology for occluded person re-identification//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 6448-6457 [DOI: 10.1109/CVPR42600.2020.00648]

Wang G A, Zhang T Z, Cheng J, Liu S, Yang Y and Hou Z G. 2019a. RGB-infrared cross-modality person re-identification via joint pixel and feature alignment//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 3623-3632 [DOI: 10.1109/ICCV.2019.00372]

Wang M L, Lai B S, Huang J Q, Gong X J and Hua X S. 2021. Camera-aware proxies for unsupervised person re-identification//Proceedings of the 35th AAAI Conference on Artificial Intelligence. [s.l.]: AAAI Press: 2764-2772 [DOI: 10.1609/aaai.v35i4.16381]

Wang M L, Li J C, Lai B S, Gong X J and Hua X S. 2022a. Offline-online associated camera-aware proxies for unsupervised person re-identification. IEEE Transactions on Image Processing, 31: 6548-6561 [DOI: 10.1109/TIP.2022.3213193]

Wang P F, Ding C X, Tan W T, Gong M M, Jia K and Tao D C. 2023a. Uncertainty-aware clustering for unsupervised domain adaptive object re-identification. IEEE Transactions on Multimedia, 25: 2624-2635 [DOI: 10.1109/TMM.2022.3149629]

Wang Q, Min W D, Han Q, Liu Q, Zha C, Zhao H Y, et al. 2022b. Inter-domain adaptation label for data augmentation in vehicle re-identification. IEEE Transactions on Multimedia, 24: 1031-1041 [DOI: 10.1109/TMM.2021.3104141]

Wang Q, Min W D, He D J, Zou S, Huang T M, Zhang Y, et al. 2020b. Discriminative fine-grained network for vehicle re-identification using two-stage re-ranking. Science China Information Sciences, 63 (11): #212102 [DOI: 10.1007/s11432-019-2811-8]

Wang Q, Xue X Y, Min W D, Wang S, Gai D and Han Q. 2024. Cross-domain joint learning and shared subspace metric for vehicle re-identification. Journal of Image and Graphics, 29(5): 1364-1380 (汪琦，雪心远，闵卫东，汪晟，盖迪，韩清. 2024. 跨域联合学习与共享子空间度量的车辆重识别. 中国图象图形学报, 29(5): 1364-1380) [DOI: 10.11834/jig.230405]

Wang Q, Zhong Y L, Min W D, Zhao H Y, Gai D and Han Q. 2023b. Dual similarity pre-training and domain difference encouragement learning for vehicle re-identification in the wild. Pattern Recognition, 139: #109513 [DOI: 10.1016/j.patcog.2023.109513]

Wang T, Liu H, Song P H, Guo T Y and Shi W. 2022c. Pose-guided feature disentangling for occluded person re-identification based on transformer//Proceedings of the 36th AAAI Conference on Artificial Intelligence. Palo Alto, USA: AAAI: 2540-2549 [DOI: 10.1609/aaai.v36i3.20155]

Wang T, Liu M Y, Liu H, Li W H, Ban M J, Guo T Y, et al. 2024c. Feature completion transformer for occluded person re-identification. IEEE Transactions on Multimedia, 26: 8529-8542 [DOI: 10.1109/TMM.2024.3379908]

Wang Y, Peng J J, Wang H B and Wang M. 2022d. Progressive learning with multi-scale attention network for cross-domain vehicle re-identification. Science China Information Sciences，65（6）：#160103［DOI：10.1007/s11432-021-3383-y］

Wang Y H，Liu X H，Yan T Y，Liu Y，Zheng A H，Zhang P P，et al. 2025b. Mambapro：multi-modal object re-identification with mamba aggregation and synergistic prompt//Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia，USA：AAAI：8150-8158［DOI：10.1609/aaai.v39i8.32879］

Wang Y H，Liu X H，Zhang P P，Lu H，Tu Z Z and Lu H C. 2024d. TOP-ReID：multi-spectral object re-identification with token permutation//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver，Canada：AAAI：5758-5766［DOI：10.1609/aaai.v38i6.28388］

Wang Y H，Liu Y，Zheng A H and Zhang P P. 2025c. DeMo：decoupled feature-based mixture of experts for multi-modal object re-identification//Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia，USA：AAAI：8141-8149［DOI：10.1609/aaai.v39i8.32878］

Wang Y H，Lyu Y F，Zhang P P and Lu H C. 2025d. IDEA：inverted text with cooperative deformable aggregation for multi-modal object re-identification//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville，USA：IEEE：29701-29710［DOI：10.1109/CVPR52734.2025.02765］

Wang Z，Fang Z Y，Wang J and Yang Y Z. 2020c. ViTAA：visual-textual attributes alignment in person search by natural language//Proceedings of the 16th European Conference on Computer Vision. Glasgow，UK：Springer：402-420［DOI：10.1007/978-3-030-58610-2_24］

Wang Z J，Zhu A C，Xue J Y，Wan X L，Liu C，Wang T，et al. 2022e. CAIBC：capturing all-round information beyond color for text-based person retrieval//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa，Portugal：ACM：5314-5322［DOI：10.1145/3503161.3548057］

Wang Z K，Zhu F，Tang S X，Zhao R，He L H and Song J N. 2022f. Feature erasing and diffusion network for occluded person re-identification//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans，USA：IEEE：4744-4753［DOI：10.1109/CVPR52688.2022.00471］

Wang Z X，Wang Z，Zheng Y Q，Chuang Y Y and Satoh S. 2019b. Learning to reduce dual-level discrepancy for infrared-visible person re-identification//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach，USA：IEEE：618-626［DOI：10.1109/CVPR.2019.00071］

Wei L H，Zhang S L，Gao W and Tian Q. 2018. Person transfer GAN to bridge domain gap for person re-identification//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City，USA：IEEE：79-88［DOI：10.1109/CVPR.2018.00016］

Wei R，Gu J Y，He S T and Jiang W. 2023. Transformer-based domain-specific representation for unsupervised domain adaptive vehicle re-identification. IEEE Transactions on Intelligent Transportation Systems，24(3)：2935-2946［DOI：10.1109/TITS.2022.3225025］

Wei S and Yang W L. 2025. Visible-infrared person re-identification algorithm integrating structural and visual features. Journal of Image and Graphics，30(10)：3335-3345（魏思，杨文璐. 2025. 融合结构与视觉特征的可见光—红外行人重识别. 中国图象图形学报，30(10)：3335-3345）［DOI：10.11834/jig.240600］

Wei S X，Gao Z，Ma C J，Zhao Y B，Guan W L and Chen S Y. 2025. Multiple information prompt learning for cloth-changing person re-identification. IEEE Transactions on Image Processing，34：801-815［DOI：10.1109/TIP.2025.3531217］

Wei Z M，Zhang Z P，Wu P，Wang J，Wang P and Zhang Y N. 2024. Fine-granularity alignment for text-based person retrieval via semantics-centric visual division. IEEE Transactions on Circuits and Systems for Video Technology，34(9)：8242-8252［DOI：10.1109/TCSVT.2024.3392831］

Wolcott T G. 1995. New options in physiological and behavioural ecology through multichannel telemetry. Journal of Experimental Marine Biology and Ecology，193(1/2)：257-275［DOI：10.1016/0022-0981(95)00121-2］

Wu A C，Zheng W S，Yu H X，Gong S G and Lai J H. 2017. RGB-infrared cross-modality person re-identification//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice，Italy：IEEE：5380-5389［DOI：10.1109/ICCV.2017.575］

Wu J Y，Huang Y，Gao M，Gao Z P，Zhao J Q，Zhang H J，et al. 2024. A two-stream hybrid convolution-transformer network architecture for clothing-change person re-identification. IEEE Transactions on Multimedia，26：5326-5339［DOI：10.1109/TMM.2023.3331569］

Wu Q，Dai P Y，Chen J，Lin C W，Wu Y J，Huang F Y，et al. 2021. Discover cross-modality nuances for visible-infrared person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville，USA：IEEE：4330-4339［DOI：10.1109/CVPR46437.2021.00431］

Wu Q，Xia J E，Dai P Y，Zhou Y Y，Wu Y J and Ji R R. 2025a. CycleTrans：learning neutral yet discriminative features via cycle construction for visible-infrared person re-identification. IEEE Transactions on Neural Networks and Learning Systems，36(3)：5469-5479［DOI：10.1109/TNNLS.2024.3382937］

Wu R Q，Jiao B L，Liu M，Wang S N，Wang W X and Wang P. 2025b. Enhancing visible-infrared person re-identification with modality- and instance-aware adaptation learning. IEEE Transactions on Circuits and Systems for Video Technology，35(8)：8086-8103［DOI：10.1109/TCSVT.2025.3560118］

Xi R D. 2025. Miss-ReID：delivering robust multi-modality object re-identification despite missing modalities//Proceedings of the 39th Conference on Neural Information Processing Systems. San Diego，USA：［s.n.］

Xi R D，Huang N C，Lai C Z，Zhang Q and Han J G. 2025. FMCNet+：feature-Level modality compensation for visible-infrared person re-identification. IEEE Transactions on Neural Networks and Learning Systems，36(7)：13247-13261［DOI：10.1109/TNNLS.2024.

3453292]

Xiao H, Lin W Y, Sheng B, Lu K, Yan J C, Wang J, et al. 2018. Group re-identification: leveraging and integrating multi-grain information//Proceedings of the 26th ACM International Conference on Multimedia. Seoul, Korea(South): ACM: 192-200 [DOI: 10.1145/3240508.3240539]

Xiong F, Gou M R, Camps O and Sznaier M. 2014. Person re-identification using kernel-based metric learning methods//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer: 1-16 [DOI: 10.1007/978-3-319-10584-0_1]

Xiong M F, Hu K K, Wang Z Y, Hu R M, Muhammad K, Del Ser J, et al. 2025. Adaptive clustering and weighted regularization contrastive learning framework for unsupervised person re-identification. IEEE Transactions on Multimedia, 27: 6896-6907 [DOI: 10.1109/TMM.2025.3590938]

Xu B Q, He L X, Liang J and Sun Z N. 2022. Learning feature recovery transformer for occluded person re-identification. IEEE Transactions on Image Processing, 31: 4651-4662 [DOI: 10.1109/TIP.2022.3186759]

Xu P and Zhu X T. 2023. DeepChange: a long-term person re-identification benchmark with clothes change//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 11162-11171 [DOI: 10.1109/ICCV51070.2023.01028]

Xuan S Y and Zhang S L. 2021. Intra-inter camera similarity for unsupervised person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 11926-11935 [DOI: 10.1109/CVPR46437.2021.01175]

Xue J, Meng Z B, Katipally K, Wang H B and van Zon K. 2018. Clothing change aware person identification//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, USA: IEEE: 2193-21938 [DOI: 10.1109/CVPRW.2018.00285]

Yan G, Wang Z J, Geng S Z, Yu Y and Guo Y C. 2023a. Part-based representation enhancement for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 33(8): 4217-4231 [DOI: 10.1109/TCSVT.2023.3241764]

Yan S L, Dong N, Zhang L Y and Tang J H. 2023b. CLIP-driven fine-grained text-image person re-identification. IEEE Transactions on Image Processing, 32: 6032-6046 [DOI: 10.1109/TIP. 2023.3327924]

Yan Y C, Qin J, Ni B B, Chen J X, Liu L, Zhu F, et al. 2023c. Learning multi-attention context graph for group-based re-identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(6): 7001-7018 [DOI: 10.1109/TPAMI.2020.3032542]

Yang B, Chen J and Ye M. 2024. Shallow-deep collaborative learning for unsupervised visible-infrared person re-identification//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: #1596 [DOI: 10.1109/CVPR52733.2024.01596]

Yang F, Wu Y, Wang Z, Li X, Sakti S and Nakamura S. 2021a. Instance-level heterogeneous domain adaptation for limited-labeled sketch-to-photo retrieval. IEEE Transactions on Multimedia, 23: 2347-2360 [DOI: 10.1109/TMM.2020.3009476]

Yang J R, Zhang J W, Yu F F, Jiang X Y, Zhang M D, Sun X, et al. 2021b. Learning to know where to see: a visibility-aware approach for occluded person re-identification//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 11865-11874 [DOI: 10.1109/ICCV48922.2021.01167]

Yang Q Z, Wu A C and Zheng W S. 2021c. Person re-identification by contour sketch under moderate clothing change. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(6): 2029-2046 [DOI: 10.1109/TPAMI.2019.2960509]

Yang Y, Yang J M, Yan J J, Liao S C, Yi D and Li S Z. 2014. Salient color names for person re-identification//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer: 536-551 [DOI: 10.1007/978-3-319-10590-1_35]

Yang Z W, Lin M, Zhong X, Wu Y and Wang Z. 2023a. Good is bad: causality inspired cloth-debiasing for cloth-changing person re-identification//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 1472-1481 [DOI: 10.1109/CVPR52729.2023.00148]

Yang Z W, Zhong X, Zhong Z, Liu H, Wang Z and Satoh S I. 2023b. Win-win by competition: auxiliary-free cloth-changing person re-identification. IEEE Transactions on Image Processing, 32: 2985-2999 [DOI: 10.1109/TIP.2023.3277389]

Ye M, Chen S Y, Li C Y, Zheng W S, Crandall D and Du B. 2025. Transformer for object re-identification: a survey. International Journal of Computer Vision, 133(5): 2410-2440 [DOI: 10.1007/s11263-024-02284-4]

Ye M, Shen J B, Crandall D J, Shao L and Luo J B. 2020. Dynamic dual-attentive aggregation learning for visible-infrared person re-identification//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer: 229-247 [DOI: 10.1007/978-3-030-58520-4_14]

Ye M, Shen J B, Lin G J, Xiang T, Shao L and Hoi S C H. 2022. Deep learning for person re-identification: a survey and outlook. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6): 2872-2893 [DOI: 10.1109/TPAMI.2021.3054775]

Ye M, Shen J B and Shao L. 2021. Visible-infrared person re-identification via homogeneous augmented tri-modal learning. IEEE Transactions on Information Forensics and Security, 16: 728-739 [DOI: 10.1109/TIFS.2020.3001665]

Ye Y T, Zhou H, Cai J L, Gao C X, Zhang Y J, Wang J L, et al. 2024. Dynamic feature pruning and consolidation for occluded person re-identification//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI: 6684-6692 [DOI:

10.1609/aaai.v38i7.28491]

Yin J H, Zhang X Y, Ma Z Y, Guo J and Liu Y F. 2023. A real-time memory updating strategy for unsupervised person re-identification. IEEE Transactions on Image Processing, 32: 2309-2321 [DOI: 10.1109/TIP.2023.3266166]

Yu S J, Li S H, Chen D P, Zhao R, Yan J J and Qiao Y. 2020. COCAS: a large-scale clothes changing person dataset for re-identification//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 3397-3406 [DOI: 10.1109/CVPR42600.2020.00346]

Yu X Y, Dong N, Zhu L H, Peng H and Tao D P. 2025a. CLIP-driven semantic discovery network for visible-infrared person re-identification. IEEE Transactions on Multimedia, 27: 4137-4150 [DOI: 10.1109/TMM.2025.3535353]

Yu Z, Huang Z Y, Hou M Y, Pei J M, Yan Y, Liu Y S, et al. 2025b. Representation selective coupling via token sparsification for multi-spectral object re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(4): 3633-3648 [DOI: 10.1109/TCSVT.2024.3509817]

Zhai Y J, Zeng Y W, Huang Z Y, Qin Z, Jin X and Cao D. 2024. Multi-prompts learning with cross-modal alignment for attribute-based person re-identification//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI: 6979-6987 [DOI: 10.1609/aaai.v38i7.28524]

Zhang G Q, Liu J, Chen Y H, Zheng Y H and Zhang H W. 2023a. Multi-biometric unified network for cloth-changing person re-identification. IEEE Transactions on Image Processing, 32: 4555-4566 [DOI: 10.1109/TIP.2023.3279673]

Zhang G Q, Yang Y, Zheng Y H, Martin G and Wang R L. 2025a. Mask-aware hierarchical aggregation transformer for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(6): 5821-5832 [DOI: 10.1109/TCSVT.2025.3531142]

Zhang G Q, Zhang H W, Lin W S, Chandran A K and Jing X. 2023b. Camera contrast learning for unsupervised person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 33(8): 4096-4107 [DOI: 10.1109/TCSVT.2023.3240001]

Zhang G Q, Zhou J Q, Jiang L, Zheng Y H and Lin W S. 2025b. Clip-based multi-modal feature learning for cloth-changing person re-identification. IEEE Transactions on Image Processing, 34: 5570-5583 [DOI: 10.1109/TIP.2025.3602641]

Zhang H D, Cheng S L and Du A Y. 2024a. Multi-stage auxiliary learning for visible-infrared person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 34(11): 12032-12047 [DOI: 10.1109/TCSVT.2024.3425536]

Zhang H W, Zhang G Q, Chen Y H and Zheng Y H. 2022a. Global relation-aware contrast learning for unsupervised person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 32(12): 8599-8610 [DOI: 10.1109/TCSVT.2022.3194084]

Zhang H Y, Liu T F, Luo Q and Zhang T. 2024. Pose guidance and multi-scale feature fusion for occluded person re-identification. Journal of Image and Graphics, 29(8): 2364-2376 (张红颖，刘腾飞，罗谦，张涛. 2024. 融合姿态引导和多尺度特征的遮挡行人重识别. 中国图象图形学报, 29(8): 2364-2376 [DOI: 10.11834/jig.230523]

Zhang P P, Wang Y H, Liu Y, Tu Z Z and Lu H C. 2024b. Magic tokens: select diverse tokens for multi-modal object re-identification//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 17117-17126 [DOI: 10.1109/CVPR52733.2024.01620]

Zhang Q, Dang K H, Lai J H, Feng Z X and Xie X H. 2022b. Modeling 3D layout for group re-identification//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 7502-7510 [DOI: 10.1109/CVPR52688.2022.00736]

Zhang Q, Lai C Z, Liu J N, Huang N C and Han J G. 2022c. FMCNet: feature-level modality compensation for visible-infrared person re-identification//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 7349-7358 [DOI: 10.1109/CVPR52688.2022.00720]

Zhang Q, Lai J H, Feng Z X and Xie X H. 2022d. Uncertainty modeling with second-order transformer for group re-identification//Proceedings of the 36th AAAI Conference on Artificial Intelligence. Palo Alto, USA: AAAI: 3318-3325 [DOI: 10.1609/aaai.v36i3.20241]

Zhang Q, Lai J H, Feng Z X and Xie X H. 2024c. Uncertainty modeling for group re-identification. International Journal of Computer Vision, 132(8): 3046-3066 [DOI: 10.1007/s11263-024-02013-x]

Zhang S Z, Luo W L, Cheng D, Xing Y H, Liang G Q, Wang P, et al. 2025c. Prompt-based modality alignment for effective multi-modal object re-identification. IEEE Transactions on Image Processing, 34: 2450-2462 [DOI: 10.1109/TIP.2025.3556531]

Zhang X, Ge Y X, Qiao Y and Li H S. 2021a. Refining pseudo labels with clustering consensus over generations for unsupervised object re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 3436-3445 [DOI: 10.1109/CVPR46437.2021.00344]

Zhang X, Ling Y N, Li K G, Shi W M and Zhou Z. 2024d. Multimodality adaptive transformer and mutual learning for unsupervised domain adaptation vehicle re-identification. IEEE Transactions on Intelligent Transportation Systems, 25(12): 20215-20226 [DOI: 10.1109/TITS.2024.3455330]

Zhang X, Liu K, Wang X W, Zhou Z and Chen H Y. 2025d. RMGNet: the progressive relationship-mining graph neural network for text-to-image person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 35(6): 5749-5761 [DOI: 10.1109/TCSVT.2025.3532685]

Zhang X, Wu Z G, Zhang Q H and Zhang Z Y. 2025e. Parallel branches-based second-order transformer for robust group re-identification with layout-guided occlusion mitigation. Expert Systems with Applications, 293: #128679 [DOI: 10.1016/j. eswa. 2025.128679]

Zhang X K, Yan Y, Xue J H, Hua Y and Wang H Z. 2021b. Semantic-aware occlusion-robust network for occluded person re-identification. IEEE Transactions on Circuits and Systems for Video Technology, 31 (7): 2764-2778 [DOI: 10.1109/TCSVT. 2020. 3033165]

Zhang Y and Lu H C. 2018. Deep cross-modal projection learning for image-text matching//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 686-701 [DOI: 10.1007/978-3-030-01246-5_42]

Zhang Y F, Wang Y Z, Li H F and Li S. 2022e. Cross-compatible embedding and semantic consistent feature construction for sketch re-identification//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal: ACM: 3347-3355 [DOI: 10.1145/3503161.3548224]

Zhang Y F, Yang H Y, Zhang Y J, Dou Z P, Liao S C, Zheng W S, et al. 2023. Recent progress in person re-ID. Journal of Image and Graphics, 28(6): 1829-1862 (张永飞, 杨航远, 张雨佳, 豆朝鹏, 廖胜才, 郑伟诗, 等. 2023. 行人再识别技术研究进展. 中国图象图形学报, 28(6): 1829-1862) [DOI: 10.11834/jig.230022]

Zhang Y K and Wang H Z. 2023. Diverse embedding expansion network and low-light cross-modality benchmark for visible-infrared person re-identification//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 2153-2162 [DOI: 10.1109/CVPR52729.2023.00214]

Zhang Y K, Wang H Z, Lu Y, Yan Y and Li X L. 2025f. Frequency domain nuances mining for visible-infrared person re-identification. IEEE Transactions on Information Forensics and Security, 20: 5411-5424 [DOI: 10.1109/TIFS.2025.3569176]

Zhao R, Ouyang W L and Wang X G. 2013. Unsupervised salience learning for person re-identification//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE: 3586-3593 [DOI: 10.1109/CVPR.2013.460]

Zheng A H, Wang Z, Chen Z H, Li C L and Tang J. 2021. Robust multi-modality person re-identification//Proceedings of the 35th AAAI Conference on Artificial Intelligence. [s. l.]: AAAI: 3529-3537 [DOI: 10.1609/aaai.v35i4.16467]

Zheng A H, Zhu X P, Ma Z Q, Li C L, Tang J and Ma J X. 2023. Cross-directional consistency network with adaptive layer normalization for multi-spectral vehicle re-identification and a high-quality benchmark. Information Fusion, 100: #101901 [DOI: 10.1016/j.inffus. 2023.101901]

Zheng L, Shen L Y, Tian L, Wang S J, Wang J D and Tian Q. 2015a. Scalable person re-identification: a benchmark//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 1116-1124 [DOI: 10.1109/ICCV.2015.133]

Zheng Q, Liang C, Fang W H, Xiang D, Zhao X, Ren C P, et al. 2015b. Car re-identification from large scale images using semantic attributes//Proceedings of the 17th IEEE International Workshop on Multimedia Signal Processing. Xiamen, China: IEEE: 1-5 [DOI: 10.1109/MMSP.2015.7340861]

Zheng W S, Gong S G and Xiang T. 2011. Person re-identification by probabilistic relative distance comparison//Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE: 649-656 [DOI: 10.1109/CVPR. 2011.5995598]

Zheng W S, Li X, Xiang T, Liao S C, Lai J H and Gong S G. 2015c. Partial person re-identification//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 4678-4686 [DOI: 10.1109/ICCV.2015.531]

Zheng W S, Yan J K and Peng Y X. 2025. A versatile framework for multi-scene person re-identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 47(3): 1362-1380 [DOI: 10. 1109/TPAMI.2024.3381184]

Zhu A C, Wang Z J, Li Y F, Wan X L, Jin J, Wang T, et al. 2021a. DSSL: deep surroundings-person separation learning for text-based person retrieval//Proceedings of the 29th ACM International Conference on Multimedia. [s. l.]: ACM: #3475369 [DOI: 10.1145/3474085.3475369]

Zhu J, Yang H, Lin W Y, Liu N, Wang J and Zhang W J. 2021b. Group re-identification with group context graph neural networks. IEEE Transactions on Multimedia, 23: 2614-2626 [DOI: 10.1109/TMM.2020.3013531]

Zhu J Y, Park T, Isola P and Efros A A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2242-2251. [DOI: 10.1109/ICCV. 2017.244]

Zhu Y X, Yang Z, Wang L, Zhao S, Hu X and Tao D P. 2020. Hetero-Center loss for cross-modality person re-identification. Neurocomputing, 386: 97-109 [DOI: 10.1016/j.neucom.2019.12.100]

Zhuo J X, Chen Z Y, Lai J H and Wang G C. 2018. Occluded person re-identification//Proceedings of 2018 IEEE International Conference on Multimedia and Expo. San Diego, USA: IEEE: [DOI: 10.1109/ICME.2018.8486568]

## 作者简介

韩清，女，副教授，主要研究方向为图像与视频处理、网络管理、虚拟现实及元宇宙。E-mail：hanqing@ncu.edu.cn

闵卫东，通信作者，男，教授，主要研究方向为虚拟现实、计算机视觉、人工智能、大数据和元宇宙。
E-mail：minweidong@ncu.edu.cn

李龙飞，男，博士研究生，主要研究方向为计算机视觉。
E-mail：lilongfei@email.ncu.edu.cn