

用于人在环行人重新识别的深度强化主动学习

刘子墨^{†,?}, 王静雅^{‡,?}, 宫少刚[§], 湖川路^{†*}, 陶大成[‡]

[†]大连理工大学, [‡]悉尼大学 UBTECH 悉尼人工智能中心, [§]伦敦玛丽女王大学

lzm920316@gmail.com、jingya.wang@sydney.edu.au、s.gong@qmul.ac.uk、lhchuan@dlut.edu.cn、da.Cheng.tao@sydney.edu.au

抽象的

大多数现有的行人重新识别 (Re-ID) 方法取得了优异的结果, 其基础是假设大量的预标记数据通常可用并且可以一次性放入训练短语中。然而, 这种假设并不适用于大多数现实世界的 Re-ID 任务部署。在这项工作中, 我们提出了一种基于强化学习的替代人机循环模型, 该模型释放了预标记的限制, 并通过逐步收集的数据保持模型升级。目标是最大限度地减少人工注释工作, 同时最大限度地提高 Re-ID 性能。它通过交替细化 RL 策略和 CNN 参数, 在迭代更新框架中工作。尤其, 我们制定了深度强化主动学习 (DRAL) 方法来指导代理 (强化学习过程中的模型) 由人类用户/注释者即时选择训练样本。强化学习的奖励是每个人类选择的样本的不确定性值。由人类注释者标记的二元反馈 (正或负) 用于选择样本, 这些样本用于微调预训练的 CNN Re-ID 模型。大量实验证明, 与现有的无监督和迁移学习模型以及主动学习模型相比, 我们的 DRAL 方法在基于深度强化学习的人机循环人员重新识别方面具有优越性。由人类注释者标记的二元反馈 (正或负) 用于选择样本, 这些样本用于微调预训练的 CNN Re-ID 模型。大量实验证明, 与现有的无监督和迁移学习模型以及主动学习模型相比, 我们的 DRAL 方法在基于深度强化学习的人机循环人员重新识别方面具有优越性。由人类注释者标记的二元反馈 (正或负) 用于选择样本, 这些样本用于微调预训练的 CNN Re-ID 模型。大量实验证明, 与现有的无监督和迁移学习模型以及主动学习模型相比, 我们的 DRAL 方法在基于深度强化学习的人机循环人员重新识别方面具有优越性。

一、简介

人员重新识别 (Re-ID) 是在分布于不同位置的非重叠摄像机视图中匹配人员的问题。大多数现有的受监督人员重新识别方法采用一次训练和部署方案, 即手动收集和注释成对的训练数据

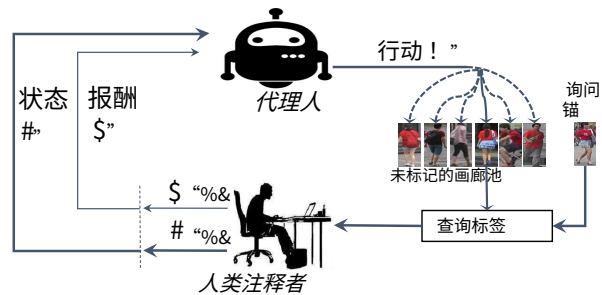


图1: 深度强化主动图解

学习 (DRAL)。对于每个查询锚点 (探针), 代理 (强化主动学习器) 将从图库池中选择顺序实例, 以在主动学习过程中使用二进制反馈 (正/负) 进行人工注释。

在学习模型之前对每对相机进行检查。基于这一假设, 有监督的 Re-ID 方法近年来在多个基准上取得了进展 [21,56,35,52,25]。

然而, 在实践中, 由于一些重新考虑, 这一假设并不容易适应。asons: 首先, pairwise 行人数据是禁止收集, 因为大量行人不太可能再次出现在其他摄像机视图中。其次, 摄像机视图数量的增加加大了在多个摄像机视图中搜索同一个人的难度。为了解决这些困难, 一种解决方案是设计无监督学习算法。一些工作开始关注无监督 Re-ID 的迁移学习或领域适应技术 [11,44,28]。然而, 与基于监督学习的模型相比, 基于无监督学习的 Re-ID 模型本质上较弱, 从而损害了任何实际部署中的 Re-ID 有效性。

另一种可能的解决方案是遵循半监督学习方案, 减少数据注释的要求。字典学习已经取得了成功的研究 [27 号] 或自定进度学习 [14] 为基础的方法。这些模型仍然基于一个强有力的假设, 即部分身份 (例如三分之一的身份)

* 通讯作者

? 平等贡献

训练集)针对每个摄像机视图进行了完全标记。对于城市应用中典型的具有数百个摄像头和 24/7 操作的 Re-ID 任务来说,这仍然不切实际。

在注释预算成本有限的情况下,为了实现有效的 Re-ID,我们专注于通过人类即时反馈进行选择标记的人机循环人员 Re-ID。[43]。这种方法不同于常见的一次性模型学习方法。相反,通过在较小的样本池上探索人类选择性注释来进行模型学习,采用逐步顺序主动学习过程。这些通过人工二进制验证累积标记的数据用于更新模型训练,以提高 Re-ID 性能。这种模型学习方法自然适合强化学习和主动学习,这是这项工作的重点。

主动学习是一种动态人类数据注释技术,旨在主动采样信息更丰富的训练数据,以优化模型学习,而无需详尽的数据标记。形式上,从未标记的集合中选择一些实例,然后由人类预言机进行注释,并且标签信息可以用于模型训练。这些操作将重复多次,直到满足终止标准,例如注释预算耗尽。这个过程中最关键的是样本选择策略。通过更少的人工注释成本获得更多信息的样本可以极大地提高性能。我们提出基于强化学习的标准,而不是手工设计策略。如图1说明了我们的深度强化主动学习 (DRAL) 模型的设计。具体来说,我们开发了一个模型,在单个人机循环模型学习框架中引入主动学习 (AL) 和强化学习 (RL)。通过将模型的 AL 部分表示为序列生成过程,由于每个动作都会影响未标记数据池之间的样本相关性(每一步都会重新计算相似性),因此它将影响下一步的决策。通过将所选样本带来的不确定性视为客观目标,我们模型的强化学习部分旨在给定人类反馈注释的情况下学习强大的样本选择策略。因此,从 RL 策略中选择的信息丰富的样本可以显著提高 Re-ID 的性能,从而增强样本选择策略的能力。

这项工作的主要贡献是:(1)我们引入了深度强化主动学习 (DRAL) 模型,该模型旨在单个 CNN 深度学习框架中共同探索强化学习和主动学习原理。(2)我们为人机循环人员重新识别设计了一种有效的 DRAL 模型,以便深度强化主动学习器(代理)可以直接在 CNN 深度网络上促进人机循环主动学习策略。大量的对比实验清楚地表明

所提出的 DRAL 模型在可扩展性和注释成本方面比现有的监督和迁移学习方法具有优势,比现有的半监督、无监督和主动学习方法具有显著的性能增益,同时使用更少的注释。

2.相关工作

人员重新识别。行人重识别任务旨在多个摄像机视图中搜索相同的人。最近,大多数人重新识别[50,45,8,10,33,38,7,53,19,5,51,9,39,36]尝试在监督学习框架下解决这个问题,其中训练数据是完全注释的。尽管这些方法实现了高性能,但它们巨大的注释成本也不容忽视。为了解决高标记成本问题,一些研究人员提出仅使用少量标记样本或不使用任何标记信息来学习模型。代表性算法[32,48,2,55,23,44,28,46]包括域转移方案、组关联方法和一些标签估计方法。

除了上述方法之外,一些研究人员还致力于降低人机交互 (HITL) 模型学习过程中的注释成本。当只有少量带注释的图像样本时,HITL模型学习可以通过在模型训练、调整或测试的循环中直接涉及人类交互来提高模型性能。随着人类纠正机器学习预测中出现的错误,模型可以得到有效纠正,从而获得更好的结果。这种情况听起来类似于行人 Re-ID 任务的情况,当图库候选大小远远超出查询锚点大小时,很难获得预标记信息。受此启发,王等人。[43]制定了人类验证增量学习 (HVIL) 模型,旨在通过灵活的人类反馈持续实时优化距离度量。该模型中采用的灵活的人类反馈(真、假、假但相似)能够包含更多信息并以渐进的方式提高性能。

AL 和 RL。主动学习在过去几十年中引起了人们的广泛关注,并在自然语言处理 (NLP)、数据注释和图像分类任务中得到了利用。[41,6,4,31]。它的过程可以被认为是人机循环设置,它允许算法交互式地查询人类注释者,并将实例识别为整个未标记数据池中信息最丰富的样本。这项工作通常是通过使用一些效果有限的启发式选择方法完成的。因此,一些研究人员旨在通过将主动学习视为强化学习问题来明确优化选择策略,以解决启发式选择方法的缺点。在 [15],而不是采用-

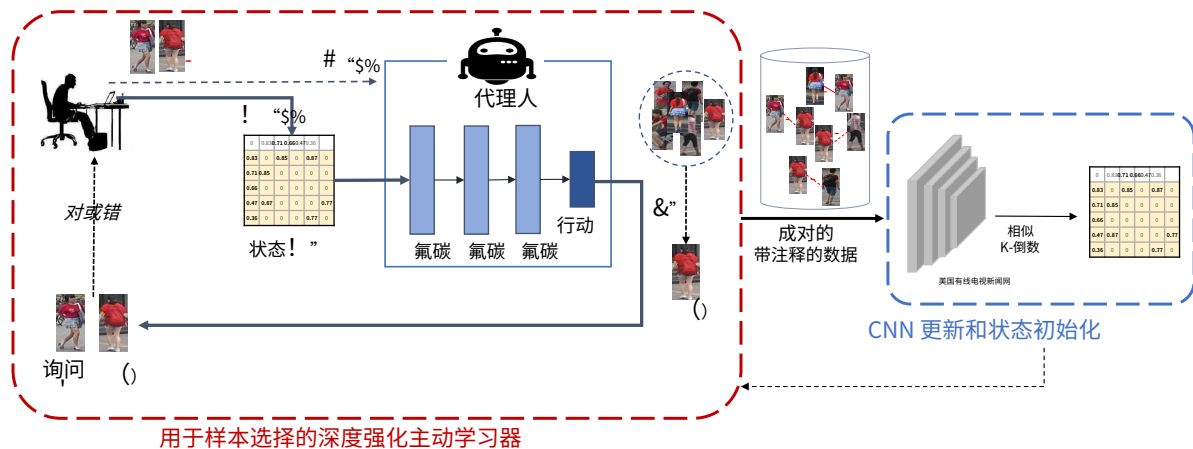


图 2: 深度强化主动学习 (DRAL) 框架: 状态衡量所有实例之间的相似关系。操作确定将发送哪个候选库供人工注释者查询。奖励是根据不同的人类反馈来计算的。采用 CNN 进行状态初始化, 并在部署模型时通过人工注释器动态注释的成对数据进行更新。当达到注释预算时, 该迭代过程停止。

Fang 采用固定的启发式选择策略等人。执行学习深度 Q 网络作为自适应策略来选择用于标记的数据实例。伍德沃德等人。[47] 尝试通过制定一种结合元学习的主动学习方法来解决一次性分类任务。通过深度强化学习。通过这种方法学习的代理能够决定如何以及何时请求 la。这些成功的应用表明强化学习非常适合主动学习。

3. 方法论

3.1. 基础 CNN 网络

我们采用 Resnet-50 [20] 作为骨干网络。在 ImageNet 预训练。为了有效地学习 ID 犯罪特征输入, 我们同步采用交叉输入损失进行分类, 并采用三元组损失进行相似性学习。

softmax 交叉熵损失函数定义为:

$$L_{\text{交叉}} = - \frac{1}{n} \sum_{i=1}^n \log(p(y_i)) \quad (1)$$

在哪里 n 表示批量大小, $p(y)$ 是真实类别的预测概率 y 输入图像的。

给定三元组样本 X_A , X_P , X_N , X_A 是一个锚点。 X_P 是同类中最难的正样本 X_A 和 X_N 是不同类别的最难负样本 X_A 。最后我们将三元组损失定义如下:

$$L_{\text{三元组}} = \sum_{X_A, X_P, X_N} [D(X_A, X_P) - D(X_A, X_N) + \text{margin}] \quad (2)$$

在哪里 margin 是正负对的裕度参数。

最后, 总损失可以通过下式计算:

$$L_{\text{全部}} = L_{\text{交叉}} + L_{\text{三元组}} \quad (3)$$

3.2. 深度强化主动学习者 - 代理

图的框架 2, 我们提出的 DRAL 是在 “一个时代通常选择型” 中提出的, 旨在动态地实例。对查询来说信息最丰富实例。正如每个查询 n_s 当实例到达时, 我们发现它是一最近邻作为 t 离散时间 t 个未标记的画廊池。在每一个步长 t , 状态 S_t 哪个雷夫环境提供实例关系的观察, 并通过选择操作来管理代理

收到回复 A_t 。为了行动 $A_t \in G_k$, 它要求 k -未标记画廊池中的第一个实例由 h_u 注释。man oracle, 针对查询回复二进制反馈 true 或 false。重复此操作, 直到 max_iteration 。每个查询的妈妈注释量已耗尽。什么时候获得足够多的成对标记数据, CNN 参数可以通过三元组损失函数进行更新。化, 这反过来会为传入数据生成一个新的初始状态。通过迭代执行样本选择和 CNN 网络刷新, 所提出的算法可以快速升级。当所有查询实例都被浏览一次时, 此进程终止。关于所提议的主动学习器的更多细节将在下文中披露。为了澄清我们模型的表述, 表 1 和算法 1 分别给出符号的定义和该方法的整个过程。

表 1：符号定义。

符号	描述
A_t, S	时间的动作、状态和奖励 t
Tr, R_t	火车组及其尺寸
TP	成对注释数据集
$\text{辛}(我, j)$	样本之间的相似度 $我, j$
$d_{我j}$	马哈拉诺比斯距离 $我, j$
q, g^k	查询, 将 k 第-个画廊候选人
y_k	二元反馈 G_k 在某个时间 t
X_{tp}, X_h	正/负样本批次直至时间 t
$K_{\text{最大限度}}$	为每个查询注释样本编号
n_s	动作大小
κ	倒数运算参数
线程	阈值参数

算法1德拉尔

输入：代理人 π , CNN 权重 w , Tr (尺寸 n), $TP = \emptyset$
 为了 $我 = 1 : n$ 做
 示例查询 q 和画廊池 G 从 Tr
 尽管 $t < K_{\text{最大限度}}$ 做
 $S_t \leftarrow (\text{西姆}, R(n_{我}, k))$ 通过方程式 4-8
 $A_t: G_k \leftarrow \pi(S_t)$, 请求标签对 (q, g^k)
 $TP \leftarrow TP \cup (q, g^k)$
 $(右_t, \text{辛}) \leftarrow (S_t, A_t)$ 通过方程式 9
 结束同时
 优化 $\pi \leftarrow \text{最大精量} E[右_t + \gamma 右_{t+1} + \dots]$
 优化 w 经过 TP 经过几步之后 结束于

3.2.1 行动

操作集定义从未标记的图库池中选择一个实例, 因此其大小与池相同。在每个时间步 t , 当遇到当前状态时 S_t , 代理根据其策略决定采取的操作 $\pi(A_t/S_t)$ 。因此 A_t 未标记图库池的实例将由人类预言机查询选择。一次 $A_t = G_k$ 执行后, 代理无法在后续步骤中再次选择。该过程的终止标准取决于预先定义的 $K_{\text{最大限度}}$ 它限制每个查询锚点的最大注释量。

3.2.2 状态

图相似性已广泛应用于主动学习框架中的数据选择[16,30]通过挖掘数据点之间的结构关系。通常, 采用稀疏图, 仅将数据点连接到其几个最相似的邻居以利用它们的上下文信息。在这项工作中, 我们还在查询样本和图库样本之间构建了一个稀疏相似图, 并将其作为状态值。带有查询的锚点 q 及其对应的

响应画廊候选集 $G = \{G_1, G_2, \dots, G_{n_s}\}$, 可以通过 CNN 网络提取 Re-ID 特征, 其中 n_s 是预定义数量的图库候选者。相似度值 $\text{辛}(我, j)$ 每两个样本之间

$我, j$ ($我 \neq j$) 然后计算为

$$\text{辛}(我, j) = 1 - \frac{d_{我j}}{\text{最大限度 } d_{我j}} \quad (4)$$

在哪里 $d_{我j}$ 是马哈拉诺比斯距离 $我, j$, 否则设置为 0。 k 倒数运算 [57] 被执行来构建稀疏相似度矩阵。对于任意节点 $n_{我} \in (q, g)$ 相似度矩阵 辛 , 其顶部 κ -最近邻定义为 $氮(n_{我}, \kappa)$ 。然后 κ -互惠邻居右 $(n_{我}, \kappa)$ 的 $n_{我}$

是通过获得

$$右(n_{我}, \kappa) = \{X_j | (n_{我} \in 氮(X_j, \kappa)) \wedge (X_j \in 氮(n_{我}, \kappa))\} \quad (5)$$

与之前的描述相比, κ -相互最近邻居与节点更相关 $n_{我}$, 其相似度值保持不变, 否则分配为零。然后将该稀疏相似矩阵作为初始状态并导入策略网络中以进行动作选择。一旦采取行动, 状态值将相应调整, 以更好地揭示样本关系。

为了更好地理解状态值的更新, 我们在图 1 中举例说明 3。对于一个州 S_t 在某个时间 t , 最优行动 $A_t = G_k$ 通过策略网络选择, 该网络表示画廊候选者 G_k 将被人工注释者选择进行查询。二进制反馈是

给出为 $y_t, \kappa \in \{1, -1\}$, 这表明 G_k 成为 pos-查询实例的正对或负对。因此相似度 $\text{辛}(q, g^k)$ 之间 q 和 G_k 将被设置为

$$\text{辛}(q, g^k) = \begin{cases} 1, & y_t, \kappa = 1 \\ 0, & y_t, \kappa = -1 \end{cases} \quad (6)$$

其余画廊样本的相似之处 $G_{我}, 我 \neq k$ 查询样本也将被重新计算, 其目的是放大正例之间的距离并推出负例之间的距离。因此, 在正反馈的情况下, 相似度 $\text{辛}(q, g_{我})$ 是之间的平均分数 $G_{我}$

和 $(q, g_{我})$, 在哪里

$$\text{辛}(q, g_{我}) = \frac{\text{辛}(q, g_{我}) + \text{辛}(q, g^k)}{2} \quad (7)$$

否则相似度 $\text{辛}(q, g_{我})$ 仅当之间的相似性时才会更新 G_k 和 $G_{我}$ 大于阈值 线程, 在哪里

$$\text{辛}(q, g_{我}) = \text{最大值}(\text{辛}(q, g_{我}) - \text{辛}(G_k, G_{我}), 0) \quad (8)$$

之后也会采用 k -倒数运算, 更新状态 S_{t+1} 然后得到。

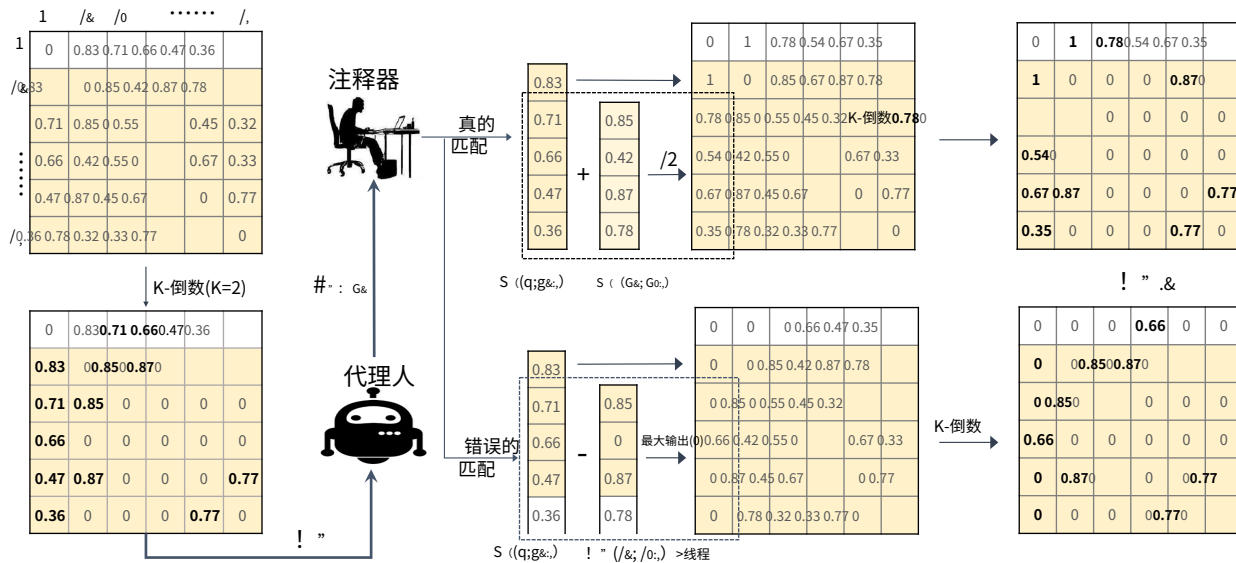


图 3：使用不同人类反馈进行状态更新的示例，旨在缩小相似度

环h 高c 联系离子w第n伽蒂模本，并放大背景较低的国家进口即代理实例之间的相似位置动性放大器。针评估低线网都佳。

礼数 阿莫n杜松埋麵s

重新嘶里欧拉例 e 到第

3.2.3R 埃瓦尔

斯坦达德活跃的学习英米方法闲话dt确定性平均信息当然蒙特,炒作这些迪萨 格力米恩奥密度作为塞行动功能离子f类 s菲卡椅子 [4,18,58,49] 和雷特瓦可 [17,3]。她e,我翻过的a不确定性作为NT对象活跃功能的离子的r e信息C埃梅学习策略。

对于数据不确定度测量,不确定度越高表明样本越难区分。遵循同样的原则[42]扩展了三重态损失

为来乌拉提 米奥德异类欧塞德三重态损失中的弹雷特瓦任务,我们p罗福k阿斯伊拉不确定性[21]至米易苏re的翁塞污点达的塔。乐的 $X_t p, X_n$ 表明这正面维安d内加主动燃普莱批直到时间获得 t ,马哈 d_{δ_k} 成为梅特里功能的米易苏英拉诺比斯距离之间恩任何w欧萨普莱斯GA NDX。那么奖励就是con普特d作为

$$右 = [米 + y_k (米斧头 \frac{y_k}{X_{j \in X_t p} X_{j \in X_n t}} - \frac{分钟 d_{X_{G_t}}}{X_{j \in X_n t}})] + \quad (9)$$

何关于 $[\cdot]_+$ 是s 易英瑞金翁蒂至少一个火星(右 $t+1$ 杜松在酒像雷塞部福真正病房, R_{t+2}, \dots) dis-按一个因素计算 γ 在某个时间 t 可以计算为

$$\text{问} = \text{最大限度} E[\gamma_{右t+1} + \gamma_2 \gamma_{右t+2} \dots / \pi, S_t, A_t] \quad (10)$$

一次问学习到最优策略 π^* 可以通过选择具有最大的动作来直接推断问价值。

3.3. CNN 网络更新

对于每个查询锚点,通过建议的DRAL代理主动选择几个样本,并手动选择

记为th 恶呼曼奥拉克莱和这些鹅牌明智数据w生病了添加到up 日期火车 英德阿塔池。这中国w兽人我然后更新 毕业终于使用 微调 ing。瓦我们的三倍t损失为o 目标 玩得很开心1行动,和w 母鸡莫相对美女数据A参与了, 这模型 贝科梅斯莫 雷罗乙乌斯和叫动脉更新后的 网络奥克是 雇员忠诚的佛重新D 有隳元法rac-化,其中 在ret盆h 埃尔普格德奥这s泰特初始的化。当训练数据池中的每个图像时,我们用固定的注释预算停止这种迭代训练方案我们的DRAL代理已浏览过一次。

4. 实验 耳鼻喉科

4.1. 数据集 谭 d 设置

数据集F 或前佩里娜塔评价 纳秒是的是奥尔结果在两个大 -尺度和小 l尺度p森 r恩蒂批基准 对于粗壮的分析姐姐:

(1) 中号 方愈-1501 [54]是宽度伊利广易于大的-规模重新识别数据集 那孔塔插入第501章乳头 吧太埃德y德可成型帕 钼 德尔德斯特尔 德泰克 托尔。包括德斯3,668图像从校园内6个不重叠的摄像机视图中获取,包含751个身份的12936张图像用于训练。在测试阶段,使用3368个查询作为查询集在剩下的候选者中搜索真正的匹配。

(2) 中大01 [24]是非凡的小规模之一re-id数据集,由来自两个摄像机视图的971个身份组成,每个身份每个摄像机视图有两个图像,因此总共包含3884个手动裁剪的图像。整个数据集分为两部分:485个用于训练的身份和486个用于测试的身份。

(3) DukeMTMC-ReID(杜克) [34] 是最流行的大规模重新识别数据集，包含从 8 个不同摄像机视图捕获的 36411 个行人图像。其中，采用16522张图像（702个身份）进行训练，将2228张（702个身份）图像作为查询从剩余的17661张图像中检索。

评估协议该方法采用两个评估指标来评估 Re-ID 性能。第一个是累积匹配特征（CMC），第二个是将行人重识别任务作为对象检索问题的平均精度（mAP）。实施细节。我们在 Pytorch 框架中实现了所提出的 DRAL 方法。我们通过三元组损失和交叉熵损失的组合预训练了 resnet-50 多类身份识别网络 60 个 epoch（在 Duke 上预训练 Market1501 和 CUHK01，在 Market1501 上预训练 Duke），在一次学习中使用 Adam 优化器，率达到 5E-4。通过将所有训练图像的大小调整为 256，最终的 FC 层输出特征向量 (2,048-D) 被提取为我们模型中的重新识别特征向量 $\times 128$ 。所提出方法中的策略网络由三个 FC 层组成，设置为 256。所提出的 DRAL 模型随机初始化，然后以 2E-2 的学习率进行优化，并且 ($K_{\text{最大限度}}, n_s, k$) 默认设置为 (10, 30, 15)。平衡参数 α 和 β 分别设置为 0.4 和 0.2。每 25% 的训练需求都经过人类注释者的审核，我们开始以 5E-6 的学习率微调 CNN 网络。

4.2. 与无监督/迁移学习/半监督方法的比较

人机循环的行人重识别不需要预先标记数据，而是一点一点地接收用户对输入查询的反馈。标记许多图库实例是可行的，但为了降低人工注释成本，我们使用主动学习技术进行样本选择。因此，我们将所提出的 DRAL 方法与一些基于主动学习的方法和基于无监督/转移/半监督的方法进行比较，在表中我们使用“uns/trans/semi”、“active”来指示训练风格。此外，报告的基线结果是通过直接采用预训练的 CNN 模型计算得出的，上限结果表明该模型在具有完全监督训练数据的数据集上进行了微调。

对于无监督/迁移学习和半监督环境，选择了 16 种最先进的方法进行比较，包括 UMDL [32], 拉 [14], SPGAN [11], Tfusion [28], TL-AIDL [44], ARN [26], TAUDL [23]、骆驼[48], SSDAL [40]、斯帕科[29]，一个示例 [13] 和 DML [52]。表中2,3和4，我们展示了 Market1501 上的 1、5、10 级匹配准确率和 mAP(%) 性能 [54]，杜克[34] 和 CUHK01 [24] 数据集，其中我们的应用程序的结果

表 2: Market1501 数据集上一些无监督、半监督和自适应方法的 Rank-1、5、10 准确度和 mAP (%)。

风格	方法	市场1501			
		地图	R-1	R-5	R-10
uns/trans/semi	UMDL [32]	22.4	34.5	52.6	59.6
	拉 [14]	20.7	45.5	60.7	66.7
	SPGAN [11]	26.9	58.1	76.0	82.7
	Tfusion [28]	-	60.75	74.4	79.25
	TL-AIDL [44]	26.5	58.2	74.8	81.1
	ARN [26]	39.4	70.3	80.4	86.3
	TAUDL [23]	41.2	63.7	77.7	82.8
	骆驼[48]	26.3	54.5	-	-
	SSDAL [40]	19.6	36.4	-	-
	斯帕科[29]	-	68.3	-	-
	一个例子[13]	26.2	55.8	72.3	78.4
积极的	数据标记语言 [52]	46.57	-	-	-
	随机的	35.15	58.02	79.07	85.78
	邱[22 号]	44.99	67.84	85.69	91.12
	QBC [1]	46.32	68.35	86.07	91.15
	GD [12]	49.3	71.44	87.05	91.42
我们的	高压线 [43]	-	78.0	-	-
	基线	20.04	42.79	62.32	70.04
	上限	73.25	87.95	95.25	96.79
	德拉尔	66.26	84.2	94.27	96.59

表 3: 在杜克数据集上使用一些无监督、半监督和自适应方法的 Rank-1、5、10 准确度和 mAP (%)。

风格	方法	公爵			
		地图	R-1	R-5	R-10
uns/trans/semi	UMDL [32]	7.3	17.1	28.8	34.9
	拉 [14]	16.4	30.0	43.4	48.5
	SPGAN [11]	26.2	46.4	62.3	68.0
	TL-AIDL [44]	23.0	44.3	-	-
	ARN [26]	33.4	60.2	73.9	79.5
	TAUDL [23]	43.5	61.7	-	-
	骆驼[48]	-	57.3	-	-
	一个例子[13]	28.5	48.8	63.4	68.4
积极的	随机的	25.68	44.7	63.64	70.65
	邱[22 号]	36.78	56.78	74.15	79.31
	QBC [1]	40.77	61.13	77.42	82.36
	GD [12]	33.58	53.5	69.97	75.81
我们的	基线	14.87	28.32	43.27	50.94
	上限	60.93	77.96	88.69	91.61
	德拉尔	56	74.28	84.83	88.42

方法均以粗体显示。该方法在rank-1和mAP上达到了 84.2%和66.26%，优于第二好的无监督/转移/半监督方法

表 4：在 CUHK01 数据集上使用一些无监督和自适应方法的 Rank-1、5、10 准确率和 mAP (%)。

风格	方法	香港中文大学01			
		地图	R-1	R-5	R-10
联合国/区式	股东总回报率[37]	-	22.4	35.9	47.9
	UCTL [32]	-	32.1	-	-
	骆驼[48]	61.9	57.3	-	-
	TRSTP [28]	-	60.75	74.44	79.25
积极的	随机的	52.46	51.03	71.09	81.28
	邱[22 号]	56.95	54.84	76.85	85.29
	QBC [1]	58.88	57.1	80.04	86.83
	GD [12]	54.79	52.37	75.21	83.44
我们的	基线	45.55	43.21	65.74	73.46
	上限	79.96	79.22	93.00	95.37
	德拉尔	71.52	74.07	88.99	93.93

Market1501 上分别上涨 13.9% 和 19.69% [54] 基准。对于杜克大学[34] 和 CUHK01 [24] 数据集上，DRAL 也取得了相当不错的性能，Rank-1 匹配率分别为 74.28% 和 74.07%。这些结果清楚地证明了我们通过 DRAL 方法实现的主动样本选择策略的有效性，并表明在不注释大量训练数据的情况下，DRAL 可以有效地构建良好的重新识别模型。

4.3. 与主动学习的比较

除了上述方法之外，我们还进一步与一些基于主动学习的方法进行比较，这些方法在训练过程中涉及人机交互。我们选择四种主动学习策略作为比较，其中模型是通过与我们的方法相同的框架进行训练的，其中执行这些主动样本选择策略和 CNN 参数更新的迭代过程，直到达到注释预算。这里，通过报告的主动学习方法选择了整个训练样本的 20%（大约 4% 对），这表明 388、2588、3304 被设置为 CUHK01 上终止的注释预算[24]、市场 1501[54]，和杜克[34]数据集，分别除了这些主动学习方法之外，我们还与另一种主动学习方法HVIL进行了性能比较[43]，它在人机交互设置下运行实验。这些方法的细节描述如下：（1）随机，作为基线主动学习方法，我们随机挑选一些样本进行查询；（2）查询实例不确定性【22 号】（QIU），QIU策略选择不确定性最高的样本进行查询；（3）委员会查询[1]（QBC），QBC 是一种非常有效的主动学习方法，它学习假设的集合并询问在委员会中引起最大分歧的实例；

（4）图密度[12]（GD），GD主动学习是一种构建图结构来识别高度连接的节点并确定最具代表性的数据进行查询的算法。（5）人工验证增量学习【12】（HVIL），HVIL 使用人机循环设置进行训练，该设置接收软用户反馈（真、假、

false 但类似）在模型训练期间，要求注释器标记每个查询实例的前 50 个候选者。

桌子2,3,4将主动学习模型的 Rank-1、5、10 和 mAP 率与 DRAL 进行比较，其中基线模型结果是直接采用预训练好的 CNN 模型。从这些结果我们可以观察到（1）所有的主动学习方法都比主动学习方法表现得更好随机挑选策略，验证了主动SAM-选择确实有利于人员重新识别性能。2）DRAL 优于所有其他主动学习方法，在 CUHK01 上，Rank-1 匹配率超过第二好的模型 QBC、HVIL 和 GD 16.97%、6.2% 和 13.15% [24]、市场1501[54]和杜克[34] 数据集，分别具有低得多的注释成本。这表明，通过引入策略作为样本选择策略，DRAL 比其他主动学习方法更有效地进行人员重新识别。

4.4. 不同标注成本下的比较

在这项工作中，成本是通过图像对之间的注释数量来衡量的。具有训练集大小 n ，完全监督设置的成本将为 $n * (n-1) / 2$ ，和 $10 * n$ 获取报告的 DRAL 结果。因此我们的 DRAL 注释大约为 0.12% (Duke [34])，0.15% (市场 1501 [54]) 和 1% (CUHK01 [24]) 对。我们进一步比较了所提出的 DRAL 方法在不同数量的标记数据中的性能（由 $K_{\text{最大限度}}$ 对三个报告的数据集进行完全监督学习（UpperBound）。随着训练数据规模的扩大，标注所有数据的成本呈指数级增长。其中，基线是直接使用预训练的 CNN 进行测试得到的。对于完全监督的设置，所有训练数据都已注释，它可以使用三元组损失和交叉熵损失来微调 CNN 参数，以寻求更好的性能。对于 DRAL 方法，我们用以下方式呈现性能 $K_{\text{最大限度}}$ 设置为表中 3、5、10。可以看出，1）随着注释数据的增加，模型随着注释成本的增加而变得更强大。随着每个查询的注释数量从 3 个增加到 10 个，Duke 上的 1 级匹配率分别提高了 13.37%、8.72% 和 15.43%。34]、市场 1501[54] 和 CUHK01 [24] 基准。2）与完全监督的设置相比，所提出的主动学习方法在每个数据集上仅显示大约 4% 的 1 级准确率。然而，DRAL 的标注成本远低于有监督的。

表 5: 直接采用 (Baseline)、完全监督学习 (UpperBound) 和 DRAL 的 Rank-1 准确度和 mAP (%) 结果 $K_{\text{最大限度}}$ 在三个报告的数据集上, 其中 n 表示每个基准的训练实例数。标注成本是通过每两个样本的标注行为次数来计算的。

方法	公爵				市场1501				香港中文大学01				成本
	地图	R-1	R-5	R-10	地图	R-1	R-5	R-10	地图	R-1	R-5	R-10	
基线	14.05	28.32	43.27	50.94	18.96	42.79	62.32	70.04	41.59	43.21	65.74	73.46	0
德拉尔	43.82	64.77	78.19	82.81	52.5	75.48	89.9	93.26	55.71	58.64	77.78	85.6	n^*3
	51.48	70.51	84.16	87.43	59.71	81.06	93.21	95.1	66.98	70.06	87.14	91.87	n^*5
	56	74.28	83.43	88.42	66.26	84.2	94.27	96.59	71.52	74.07	88.99	93.93	n^*10
上限	60.93	77.96	88.69	91.61	73.25	87.95	95.25	96.79	79.96	79.22	93.00	95.37	$n^* (n-1) / 2$

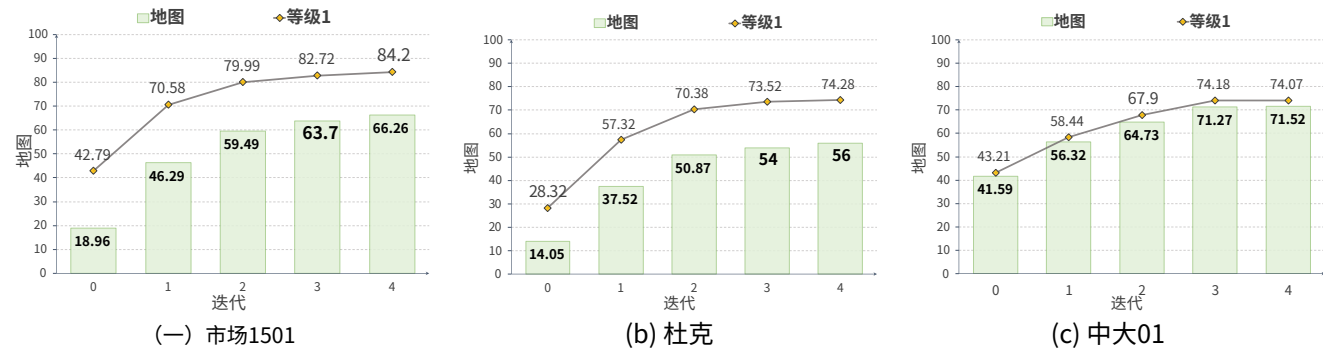


图 4: 相对于 (a) Market1501、(b) Duke 和 (c) CUHK01 数据集上的迭代, Rank-1 准确度和 mAP(%) 改进。灰线和绿条 (粗体数字) 分别表示Rank-1精度和mAP。

4.5. 迭代次数的影响

主动学习的前景是, 通过迭代增加标记数据的大小, 模型性能正在逐渐增强。对于每个输入查询, 我们仅将标签与从 DRAL 派生的图库候选相关联, 并采用这些成对标记数据进行 CNN 参数更新。我们将迭代设置为固定的数字

在我们对所有数据集的实验中, ber 4。如图4节目通过三个数据集的迭代, Rank-1 精度和 mAP 尊重得到改善。从这些r 结果, 我们可以观察到 propos 主动学习器的性能提升很快, rank-1 ac 提高了 20% 左右 ~比前两项提高 40% 精确性 埃德·德拉尔 口粮上 运动单位: m 奥德尔佩尔 饮食。钷 是暗示 s可能不会 t 是埃森- 信息 已经 派生于 通过探索- 电子飞行。

5. 结论

在这项工作中, we 致辞 哦可的缺陷 w 减少 人类标签e 公约中的努力 AI 数据预置 为 ing。与我 仿- 人员重新识别 阳离子火车模型 p 的数量 重新贴上标签 符号成本或我 不可访问的大q

训练数据, 我们的模型设计旨在最大化效果- 少量选择性样本标记的 Re-ID 模型学习的有效性。模型设计的关键任务 变成如何在固定的时间内选择更多信息的样本 注释成本。具体来说, 我们制定了一种强制 深缰绳- 主动学习 (DRAL) 方法 具有灵活的 盈利政策 选举信息 样品 (排名表) fo ra 给定输入 qu 埃里那些 样品是 然后送入一个 人类注释者 哦那 米奥德尔可以重新 中欧二进制费用 dback (真或假 e) 作为强化 元素学习- 在g 对 D 的奖励 RAL型号更新 婷。莫雷奥 v呢, 迭代- 蒂大方案是 e为 up ve 比较 eva吃DRA L和重新识别 米奥德尔。延伸性 执行 我们的关系 re 进行 在 两个大s 规模和小规模 重新识别 基准点 德米论奥 r 模型稳健性 。

AC知识渊博的 G水泥

这部作品 我持 N hina 国家自然科学基金 修村科学 佛你的确定 No.6172520 2、61829102 ,61751212; 福n基本R 研究基金 中央 大学 联合国兰特·N哦 DUT19GJ201; 维 西曼 tcs有限公司; 这 中国学校 哦领导委员会; A兰图灵 我研究所; 在不v吃了英国工业 你试验挑战项目 德上的对象v私奔和 钴 商业 ng 智能视频 哦分析 S解决方案 普公共安全 (9 8111-571149) ; 一个 d 澳大利亚人 一项研究 钴 联合国委员会项目 S: FL-170100117, DP-1801034 24.

参考

- [1] 阿部直树和间冢博. 使用 boosting 和 bagging 查询学习策略. 在 *ICML*, 1998. 6,7
- [2] 斯瓦沃米尔·巴克、彼得·卡尔和让-弗朗索瓦·拉隆德. 通过综合进行领域适应, 以实现无人监督的人员重新识别. 在 *ECCV*, 2018. 2
- [3] Björn Barz、Christoph Käding 和 Joachim Denzler. 基于内容的图像检索的信息论主动学习. 在 *公关*, 2018. 5
- [4] William H. Beluch、Tim Genewein、Andreas Nürnberger 和 Jan M. Köhler. 图像分类中主动学习的集成的力量. 在 *CVPR*, 2018. 2,5
- [5] 张晓斌, Timothy M. Hospedales, 向涛. 用于人员重新识别的多级分解网络. 在 *CVPR*, 2018. 2
- [6] 莫伊托亚·查特吉和安东·洛伊斯基. 一种基于主动学习的有效视频注释和检索方法.
在国家信息处理计划, 2015. 2
- [7] 陈伟华, 陈晓棠, 张建国, 黄凯奇. 超越三元组损失: 用于人员重新识别的深度四元组网络. 在 *CVPR*, 2017. 2
- [8] 陈逸伦, 王志成, 彭宇翔, 张志强, 余刚, 孙健. 用于多人姿态估计的级联金字塔网络. 在 *CVPR*, 2018. 2
- [9] 成德, 龚一红, 周三平, 王进军, 郑南宁. 通过基于多通道零件的 cnn 进行人员重新识别, 并改进了三元组损失函数. 在 *CVPR*, 2016. 2
- [10] Dahjung Chung、Khalid Tahboub 和 Edward J. Delp. A 用于人员重新识别的两流连体卷积神经网络. 在 *ICCV*, 2017. 2
- [11] 邓伟建, 郑亮, 康国良, 杨易, 叶启翔, 焦建斌. 图像-图像域适应, 保留自相似性和域相异性, 用于人员重新识别. 在 *CVPR*, 2018. 1,6
- [12] 桑德拉·艾伯特、马里奥·弗里茨和伯恩特·席勒. RALF: 用于对象类识别的强化主动学习公式. 在 *CVPR*, 2012. 6,7
- [13] 于武等人. 以一个例子来进行人员重新识别的渐进式学习. *提示*, 2019. 6
- [14] 范禾禾, 郑亮, 颜成刚, 杨易. 无监督人员重新识别: 聚类和微调. *ACM*, 2018. 1,6
- [15] 孟芳, 李媛, 特雷弗·科恩. 学习如何主动学习: 一种深度强化学习方法. 在 *EMNLP*, 2017. 2
- [16] Eyal En Gad、Akshay Gadde、Amir Salman Avestimehr 和 Antonio Ortega. 使用自适应和非自适应方法对加权图进行主动学习. 在 *ICASSP*, 2016. 4
- [17] 菲利普·亨利·戈塞林和马蒂厄·科德. 交互式图像检索的主动学习方法. *提示*, 2008年. 5
- [18] 郭虎生, 王文建. 基于主动学习的SVM多类分类模型. *公关*, 2015. 5
- [19] 郭一奎, 张毅文. 使用多级相似性进行高效、深入的人员重新识别. 在 *CVPR*, 2018. 2
- [20] 何凯明, 张翔宇, 任少清, 孙健. 用于图像识别的深度残差学习. 在 *CVPR*, 2016. 3
- [21] 亚历山大·赫尔曼斯、卢卡斯·拜尔和巴斯蒂安·莱贝. 为人员重新识别的三重损失辩护. *结RR*, 2017. 1,5
- [22] 大卫·D·盖尔·刘易斯和威廉·A·盖尔. 通过不确定性采样训练文本分类器. 在 *西吉尔*, 1994. 6,7
- [23] 李敏贤, 朱夏天, 宫少刚. 通过深度学习轨迹关联进行无监督人员重新识别. 在 *ECCV*, 2018. 2,6
- [24] 李伟, 赵锐, 王小刚. 通过迁移度量学习进行人类重新识别. 在 *ACCV*, 2012. 5,6,7
- [25] 李伟, 朱夏天, 宫少刚. 用于人员重新识别的和谐注意力网络. 在 *CVPR*, 2018. 1
- [26] 李玉杰、杨福恩、刘彦成、叶玉英、杜晓飞、王玉强. 适应和重新识别网络: 一种用于人员重新识别的无监督深度迁移学习方法. 在 *CVPR*, 2018. 6
- [27] 刘晓, 宋明利, 陶大成, 周星辰, 陈春, 卜家军. 用于人员重新识别的半监督耦合字典学习. 在 *CVPR*, 2014年. 1
- [28] 吕建明, 陈伟航, 李庆, 杨灿. 通过时空模式的迁移学习进行无监督的跨数据集人员重新识别. 在 *CVPR*, 2018. 1,2,6,7
- [29] 马凡, 孟德宇, 谢琪, 李子娜, 董宣仪. 自定进度的联合培训. 在 *ICML*, 2017. 6
- [30] 马一飞, 黄子国, 杰夫·G·施奈德. 使用西格玛最优性在图上进行主动搜索和强盗. 在 *联合人工智能*, 2015. 4
- [31] Sujoy Paul、Jawadul H. Bappy 和 Amit K. Roy-Chowdhury. 用于结构化数据主动学习的非均匀子集选择. 在 *CVPR*, 2017. 2
- [32] 彭培熙, 向涛, 王耀伟, Massimiliano Pontil, 宫少刚, 黄铁军, 田永红. 用于人员重新识别的无监督跨数据集迁移学习. 在 *CVPR*, 2016. 2,6,7
- [33] 钱学林, 付艳伟, 蒋玉刚, 向涛, 薛向阳. 用于人员重新识别的多尺度深度学习架构. 在 *ICCV*, 2017. 2
- [34] Ergys Ristani、Francesco Solera、Roger S. Zou、Rita Cucchiara 和 Carlo Tomasi. 用于多目标、多摄像机跟踪的性能测量和数据集. 在 *ECCV研讨会*, 2016. 6,7
- [35] M. Saquib Sarfraz、Arne Schumann、Andreas Eberle 和 Rainer Stiefelhagen. 用于人员重新识别的姿势敏感嵌入, 具有扩展的跨邻域重新排名. *CVPR*, 2018. 1
- [36] 沉彦涛, 李红生, 易帅, 陈大鹏, 王小刚. 使用深度相似性引导图神经网络进行人员重新识别. 在 *ECCV*, 2018. 2
- [37] 施志远, Timothy M. Hospedales, 向涛. 传输语义表示以进行人员重新识别和搜索. 在 *CVPR*, 2015. 7

- [38] 苏驰, 李佳宁, 张世良, 邢俊良, 高文, 田齐. 用于人员重新识别的姿势驱动深度卷积模型. 在 *ICCV*, 2017. 2
- [39] 苏驰, 杨范, 张世良, 田齐, Larry S. Davis, 高文. 具有低等级属性嵌入的多任务学习, 用于人员重新识别. 在 *ICCV*, 2015. 2
- [40] 苏驰, 张世良, 邢俊良, 高文, 田齐. 深层属性驱动多摄像头人员重新识别. 在 *ECCV*, 2016. 6
- [41] Hang Su, Zhaozheng Yin, Takeo Kanade, Seungil Huh. 在逐渐增强的图上进行主动样本选择和校正传播. 在 *CVPR*, 2015. 2
- [42] Ahmed Taha, Yi-Ting Chen, Teruhisa Misu, Abhinav Shrivastava 和 Larry Davis. 视觉检索系统中的无监督数据不确定性学习. 在 *RR*, 2019. 5
- [43] 王汉晓, 宫少刚, 朱夏天, 向涛. 人机交互的人员重新识别. 在 *ECCV*, 2016. 2,6,7
- [44] 王静亚, 朱夏天, 宫少刚, 李伟. 用于无监督人员重新识别的可转移联合属性身份深度学习. 在 *CVPR*, 2018. 1,2,6
- [45] 王一成, 陈振中, 吴峰, 王刚. 使用级联成对卷积进行人员重新识别. 在 *CVPR*, 2018. 2
- [46] 魏龙辉, 张世良, 高文, 田齐. 人员转移 gan 以弥合域差距以进行人员重新识别. 在 *CVPR*, 2018. 2
- [47] 马克·伍德沃德和切尔西·芬恩. 主动的一次性学习. 在 *RR*, 2017. 3
- [48] 于红星, 吴安聪, 郑伟时. 用于无监督人员重新识别的 Crossview 非对称度量学习. 在 *ICCV*, 2017. 2,6,7
- [49] 张弛诚, 卡玛莉卡·乔胡里. 超越基于分歧的不可知论主动学习. 在 *国家信息处理计划*, 2014年. 5
- [50] 张莉, 向涛, 宫少刚. 学习用于人员重新识别的判别零空间. 在 *CVPR*, 2016. 2
- [51] 张英, 李保华, 陆虎川, Atshushi Irie, 阮向. 用于人员重新识别的特定于样本的支持向量机器学习. 在 *CVPR*, 2016. 2
- [52] 张颖, 向涛, Timothy M Hospedales, 陆虎川. 深入互学. *CVPR*, 2018. 1,6
- [53] 赵海宇, 田茂清, 孙书阳, 邵静, 严俊杰, 易帅, 王晓刚, 唐晓鸥. 主轴网络: 通过人体区域引导特征分解和融合进行人员重新识别. 在 *CVPR*, 2017. 2
- [54] 郑梁, 沉丽月, 陆田, 王胜金, 王敬东, 田齐. 可扩展的人员重新识别: 一个基准. 在 *ICCV*, 2015. 5,6,7
- [55] 郑哲东, 郑亮, 杨易. gan 生成的未标记样本改善了体外人员重新识别基线. 在 *ICCV*, 2017. 2
- [56] 郑哲东, 郑亮, 杨易. 用于大规模行人重新识别的行人对齐网络. *TCSVT*, 2018. 1
- [57] 钟准, 郑梁, 曹东林, 李少子. 使用 k 倒数编码对人员重新识别进行重新排序. 在 *CVPR*, 2017. 4
- [58] 朱静波, 王惠珍, 邹凯, 马修Y.马. 通过数据注释的不确定性和密度进行主动学习. *TASLP*, 2010. 5