

基于噪声标签自适应的行人再识别方法

唐 轲, 郎丛妍

(北京交通大学计算机信息与技术学院, 北京 100043)

摘要: 行人再识别技术目前逐步被应用于视频监控、智能安防等领域。监控设备与日俱增, 给研究工作提供了海量数据支持, 但人工标注或检测器识别难以避免地引入带有噪声的数据标签。在进行大规模深度神经网络训练时, 伴随数据量增加, 标签的噪声给模型训练带来不可忽视的损害。为解决行人再识别的噪声标签问题, 本文结合噪声、非噪声数据训练差异化特征, 提出一种噪声标签自适应的行人再识别方法, 不需要使用额外的验证集以及噪声比例、类型等先验信息, 完成对噪声数据的筛选过滤。此外, 本文方法自适应地学习噪声样本权重, 进一步降低噪声影响。在含噪声的 Market1501、DukeMTMC-reID 两个数据集上, 主流模型受噪声影响严重, 本文提出的方法可以在此基础上提高约 10% 的平均精度。

关键词: 行人再识别; 噪声标签; 深度学习; 噪声过滤; 深度神经网络

中图分类号: TP391.41 **文献标志码:** A

Noise Label Based Self-adaptive Person Re-identification

TANG Ke, LANG Congyan

(School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100043, China)

Abstract: As security issues have received widespread attention, the research on person re-identification has become more realistic, which is gradually being applied to video surveillance, intelligent security and other fields. The increasing of the number of monitoring equipments provides massive data support for research, but manual labeling or detector recognition inevitably introduces noisy labels. When training large-scale deep neural networks, as the amount of data increases, the noise of the label brings non-negligible damage to model training. In order to solve the noise label problem of person re-identification, this paper combines noise and non-noise data to train differentiated features, and proposes a noise-label adaptive pedestrian re-identification method without using additional verification sets, noise ratio, types and other priors. In addition, the method adaptively learns the weight of noise data to further reduce the influence. On the noisy Market1501 and DukeMTMC-reID data sets, the state of the art is severely affected by noise. The proposed method can improve the evaluation index by about 10% on this basis.

Key words: person re-identification; noise label; deep learning; noise filtering; deep neural network

引 言

行人再识别(Person re-identification, Re-ID)是指给定一张目标行人图像,在跨摄像头的图像库中重定位该行人的技术。随着近几年深度学习技术突飞猛进,安防问题受到广泛关注,行人再识别技术也受到越来越多的重视。同时,国内外安防监控设备的增多,给研究提供了海量数据,促使其在短时期内得到快速发展,逐步成熟。目前正在进入实际应用阶段,被应用于视频监控、智能安防等领域。近几年有许多标志性方法^[1-3]先后被提出,并取得了突破性进展。但在实际应用场景中,许多有效的方法无法得到预期结果和验证,这是因为:当用于训练的数据量激增时,人工标注或检测器识别均难以避免地引入大量噪声标签,例如将一个行人错误标注为数据集中的其他行人,即均匀噪声^[4],如图1,2所示。而目前主流使用的深度学习方法,均基于训练数据标签正确的基本假设,对数据标签有较强的依赖性。文献[7]阐述了神经网络对样本具有充足的拟合能力,可以记忆包括噪声数据在内的所有训练样本。文献[8]阐述了噪声数据给深度神经网络的泛化能力造成极大损害:随着训练进行,非噪声训练数据回传梯度数值减小,对模型参数优化产生的更新也随之减少。而训练迭代过程中,在噪声数据上计算的回传梯度仍然很大,模型参数开始向噪声数据进行拟合。



图1 DukeMTMC-reID^[5]数据集中遮挡严重时标签噪声示例(行人ID:0013)

Fig.1 Noise label example on DukeMTMC-reID^[5] when severe occlusion(ID 0013)



图2 Market1501^[6]数据集中标签噪声示例(行人ID:1365)

Fig.2 Noise label example on Market1501^[6] when severe occlusion(ID 1365)

行人再识别任务中的噪声类型包括标签噪声和样本噪声。为了解决样本噪声问题,文献[9]提出一个配合注意力机制的网络结构,以过滤行人图像中的噪声;行人再识别任务中的标签噪声问题此前少有人关注,本文专注于解决标签噪声问题。传统标签噪声属于图像分类任务中的一个子问题,常见的思路是去除疑似错误标注的训练数据或将错误标签进行修正。文献[10]提出一种修正噪声标签的方法,但需要使用额外一个确定无噪声的验证,在行人再识别的应用场景下,获取如此庞大的准确标注数据集必然带来极高的标注成本;在无需额外验证集的方法中,文献[11]提出一种修正噪声标签的方法,基于模型优先学习非噪声数据的基本假设,在训练过程中将模型的预测作为噪声标签修正的辅助信息,但修正噪声标签不可避免地污染非噪声数据;此外,一种合理的方法是使用无监督学习,文献[12]提出使用平滑多标签的无监督训练方法,根据训练数据特征表示之间的相似度,使模型学习输入数据的标签,在一定程度上缓解了训练对于标签的依赖性,但需要一个额外的训练集和训练过程,没有充分地利用本域内人工给出的准确标注,损失了部分潜在的有效信息。文献[13]提出一种改进的三元组损失函数,可应用于大规模带噪声行人再识别数据集,但主要面向的问题是改进三元组损失函数缺陷,文献[14]提出一种新的网络结构,不同于从输入图像中提取出一个特定维度的特征向量,而是与文献[15]相同,将图像的特征表示建模为一个高斯分布。

与上述不同,本文针对行人再识别任务具体应用场景及噪声类型,结合噪声、非噪声样本在训练过程中表现的差异化特征和任务特点,不需要难以获得的无噪声验证数据集、噪声比例先验知识,避免了

修复标签时污染非噪声数据的问题,自适应地对非噪声难样本和噪声样本进行权重学习,提出一种基于噪声标签自适应的行人再识别方法。

1 问题定义

在无噪声数据的条件下,对于一个 C 类别的分类问题,有 N 个样本量的训练数据集 $D = \{x_i, y_i\}_{i=1}^N$, x_i 表示数据集中第 i 个训练样本, y_i 表示第 i 个样本的标记,此标记为样本的真实无噪声类别标记。

在含有噪声数据的条件下,对于一个 C 类别的分类问题,有 N 个样本量的训练数据集 $\tilde{D} = \{x_i, \tilde{y}_i\}_{i=1}^N$, x_i 表示数据集中第 i 个训练样本, \tilde{y}_i 表示第 i 个样本的标记,此标记为样本可能带噪声的类别标记。

训练过程中,随机地从训练数据集 \tilde{D} 中采样一个批次的训练数据 $\mathcal{B} = \{x_i, \tilde{y}_i\}_{i=1}^b$, 其包含 b 个样本。假设学习率为 $\dot{\alpha}$, 损失函数为 l , 则模型参数 θ 的更新公式可表示为

$$\theta = \theta - \dot{\alpha} \nabla \left(\frac{1}{|\mathcal{B}|} \sum_{x \in \mathcal{B}} l(x, \tilde{y}; \theta) \right) \quad (1)$$

训练数据集中出现的标签噪声可通过均匀噪声转移矩阵计算得到

$$T_u = \begin{bmatrix} 1-p & \frac{p}{n-1} & \dots & \frac{p}{n-1} \\ \frac{p}{n-1} & 1-p & \frac{p}{n-1} & \dots & \frac{p}{n-1} \\ \vdots & & \ddots & & \vdots \\ \frac{p}{n-1} & \dots & \frac{p}{n-1} & 1-p \end{bmatrix} \quad (2)$$

式中: $T_{u[i][j]}$ 表示类别 i 被错误标注为类别 j 的概率; p 为训练数据集噪声比例; n 为训练集类别数。

2 本文方法

根据拟合特性,本文提出的方法分为两个部分:筛选学习和修正学习。在起始阶段,筛选学习模块减少噪声标签对模型训练的影响,加速模型对非噪声样本的充分学习。为了避免模型对非噪声训练数据过拟合,修正学习模块补充非噪声的难样本,增强模型泛化性能。

2.1 训练拟合特性

文献[7,15]提出神经网络首先倾向于学习简单的图像特征,之后逐步记住全部训练样本,根据实验观察发现噪声、非噪声样本的训练期间损失值大小的区别性特点。训练前期,由于非噪声数据一致性,其回传梯度也保持一致性特点,模型参数较快地向一致的梯度方向进行更新,以优化目标函数。另一方面,回传梯度由于受噪声影响,损失了其于训练数据特征的相关性,故在噪声数据上难以完成拟合,损失值相对较大,此阶段对噪声相对鲁棒。训练后期,模型完成对非噪声数据的拟合,在这些训练样本上计算的回传梯度较小,而在噪声训练数据上计算得到的回传梯度较大,在这一阶段对噪声数据进行拟合。

本文对噪声给模型带来的损害和训练拟合特性进行了实验验证。值得注意的是,噪声、非噪声样本在训练阶段产生了差异化的特征,分别在噪声比例1%、3%、5%和10%的条件下进行了测试,可以得到具体的差异情况分别如图3、4所示。此外,随着噪声比例升高,模型精度受损愈加严重,为了观察到高比例噪声对模型精度的影响,本文测试了最高30%的均匀噪声,实验结果分别如图5、6所示。图

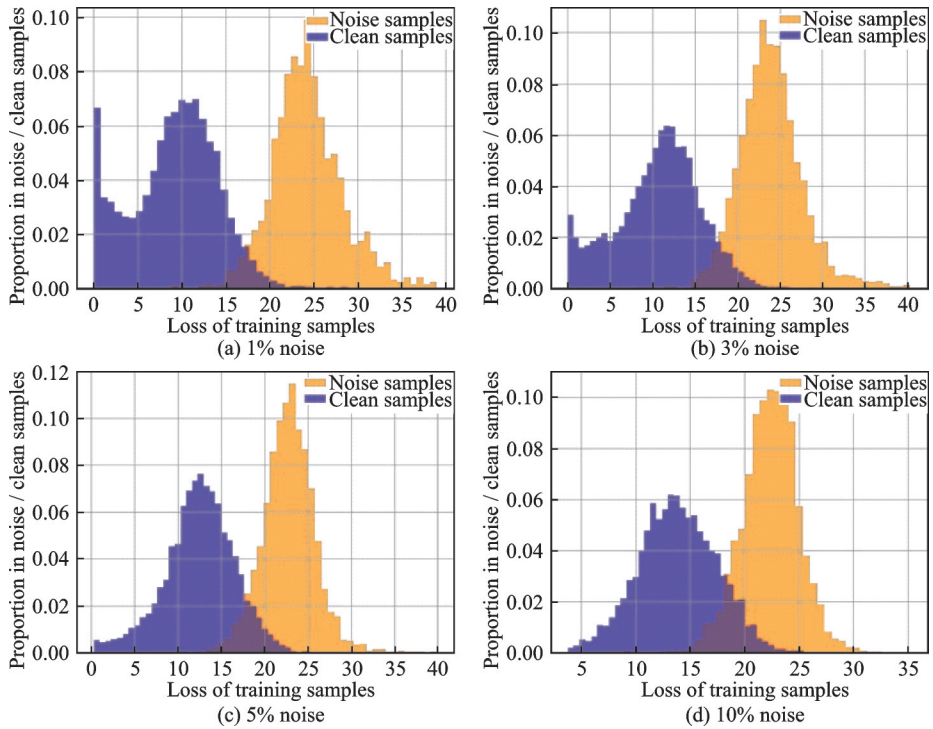


图3 Market1501^[6]数据集训练样本损失值分布

Fig.3 Training loss distribution on Market1501^[6]

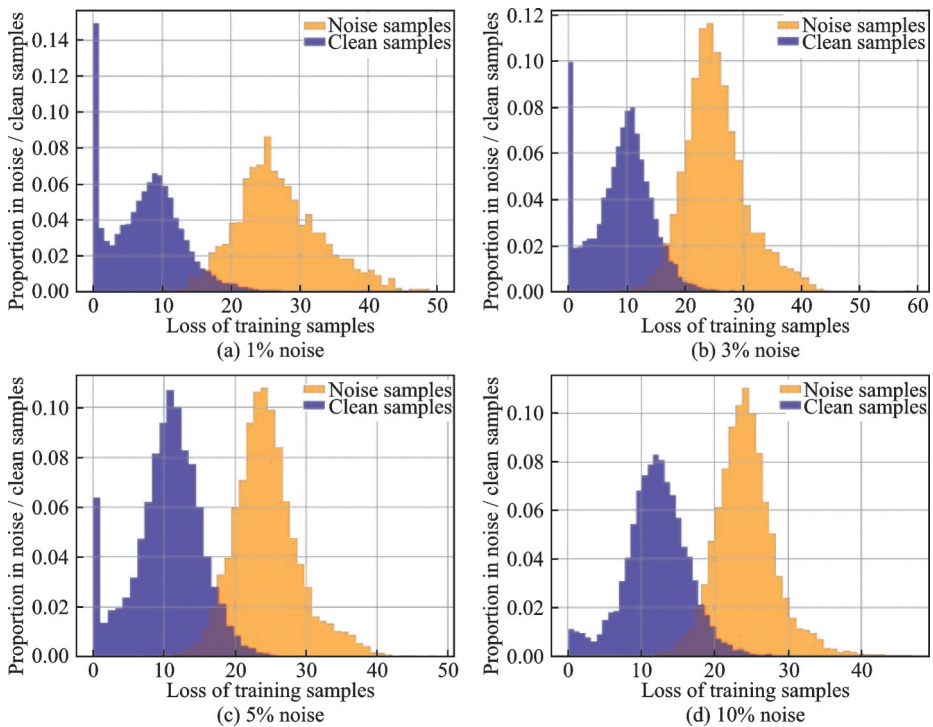


图4 DukeMTMC-reID^[5]数据集训练样本损失值分布

Fig.4 Train loss distribution on DukeMTMC-reID^[5]

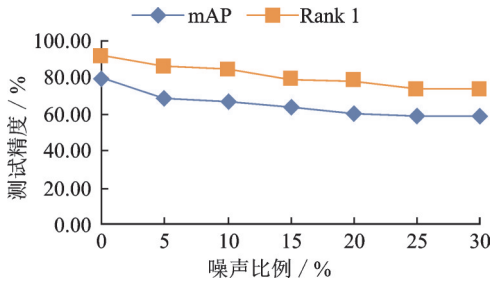


图5 标签噪声对模型精度的影响(Market1501^[6]数据集实验结果)

Fig.5 Influence of label noise on model accuracy (experiment on Market1501^[6])

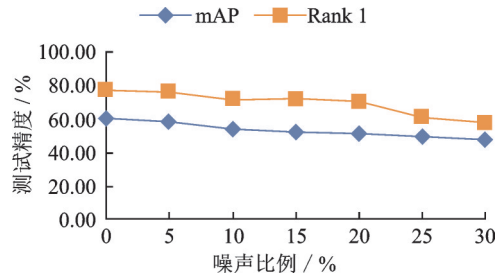


图6 标签噪声对模型精度的影响(DukeMTMC-reID^[5]数据集实验结果)

Fig.6 Influence of label noise on model accuracy (experiment on DukeMTMC-reID^[5])

中, mAP表示平均精度均值,计算方法为:对于一个查找图,求检索库中该查找图的检索精度的均值; Rank $N(N=1, 5, 10)$ 表示匹配正确率,计算方法为:对于一张查找图像,将全部的检索结果图像按相似度降序排列,计算相似度前 N 名的检索图像中匹配正确的概率。

2.2 样本记忆

神经网络对训练样本的记忆最早于文献[15]提出, $\phi(x_i|\theta_t)$ 为模型在时刻 t 对样本 (x_i, y_i) 进行的预测,记为 $\hat{y}_i = \phi(x_i|\theta_t)$,其中 y 为样本 x 可能含噪声的真实类别标签。 $F_x(m) = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m\}$ 是训练阶段,样本 (x_i, y_i) 在近 m 轮被预测的类别标签,指示函数 $I(\cdot)$,根据式(3)可以计算模型参数对样本记忆的准确率 $H(x; m)$ 。

$$H(x; m) = \frac{\sum_{\hat{y} \in F_x(m)} I(\hat{y} = y)}{|F_x(m)|} \quad (3)$$

2.3 筛选学习阶段

筛选学习网络结构和模块分别如图7,8所示。筛选学习模块统计数据集 \tilde{D} 的损失值分布。根据图3,4,本文采用一种广泛使用的二元高斯混合模型对训练数据的损失值进行参数估计。样本 x_i 属于噪声样本的概率可以根据最大似然函数计算得到,根据高斯混合模型,可以得到训练集非噪声比例的估计值 $\hat{\tau}$ 。当训练过程中全部样本记忆比例达到 $\hat{\tau}$ 时,完成这一阶段的筛选学习,伪码由算法1给出,其中记忆长度、记忆准确度和筛选参数的影响在后面进行讨论。筛选学习表明模型在减少受到噪声训练数据干扰前提下,完成对非噪声样本的学习。值得注意的是,筛选学习模块会引入两个需要关注的问题:训练数据量减小和难样本过滤,它们将在修正学习模块中被关注。

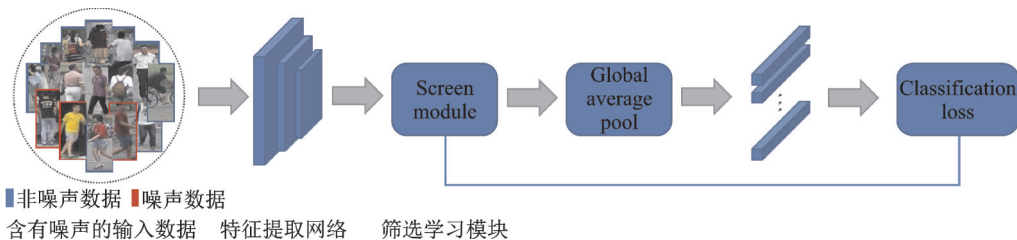


图7 筛选学习网络结构

Fig.7 Network architecture of screen learning module

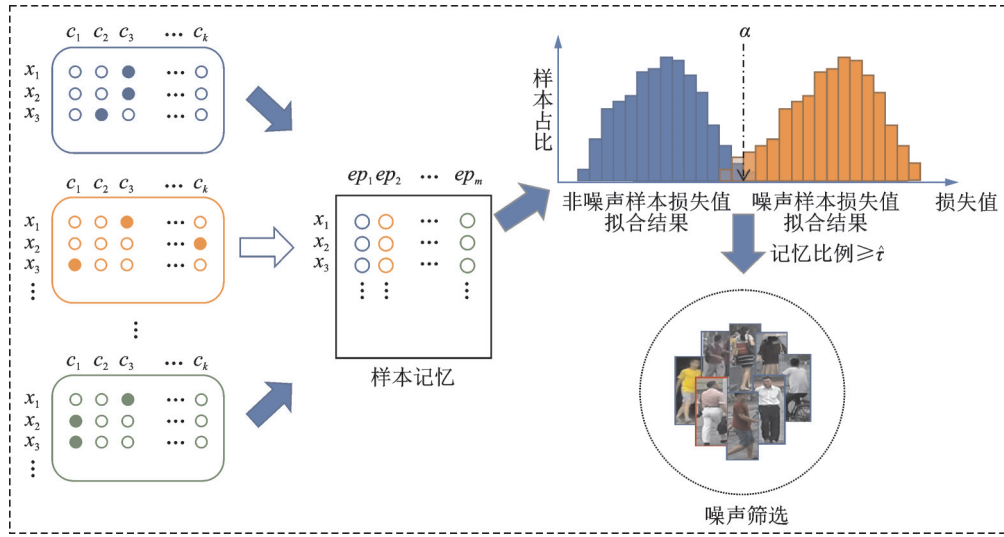


图8 筛选学习模块

Fig.8 Screen learning module

算法 1 筛选学习

输入

样本 x_i , 噪声标签 $\tilde{y}_i \in \tilde{D}$

样本 x_i , 干净标签 $y_i \in D$

模型参数 θ

记忆长度参数 m

记忆准确度 ϵ

从 $t=1$ 到 N , 执行

取批数据 $\forall (x_i, y_i) \in \tilde{D} \cup \forall (x_i, \tilde{y}_i) \in D$

当 $\frac{\sum x_i}{|\tilde{D}| + |D|} \geq \hat{\tau}$ 时停止筛选学习阶段

取得噪声筛选结果 \tilde{D} , 其中 $\tilde{x}_i \in \{\tilde{D} \mid L_{\tilde{x}_i} < \alpha + \hat{\tau} \times L_{x_i}\}$ // L_{x_i} 表示样本 \tilde{x}_i 的损失值

结束

2.4 修正学习阶段

修正学习阶段的网络结构如图9所示。修正学习在被过滤的估计噪声样本中挖掘非噪声难样本并补充学习。训练初期对于难样本的学习不足, 存在难样本被筛选学习模块过滤的风险, 例如图3、4中非噪声、噪声损失值分布交叉部分。为此, 对可能噪声样本进行权重学习, 自适应地补充学习难样本, 同时衰减对真实错误标签样本学习。全局权重 W 包含当前训练轮次的局部权重 w 和历史轮次全局权重的指数衰减, 局部、全局权重计算方式如式(3,4)所示, 并引入损失函数 L_{revise} , 其中 $n = |\tilde{D}|$, 即

$$W_{\text{order}_i} = \frac{n - \text{order}_i}{n} \tag{4}$$

$$W_{order_i^t} = W_{order_i^{t-1}} \times \beta + (1 - \beta) \times w_{order_i} \quad (5)$$

$$L_{revise} = -\frac{1}{C} \sum_{i=1}^C W_{order_i} \times \log p(\tilde{y}_i | \tilde{x}_i) \quad (6)$$

在修正学习训练过程中,真实噪声样本权重随训练轮数逐渐衰减,局部权重如图10所示。

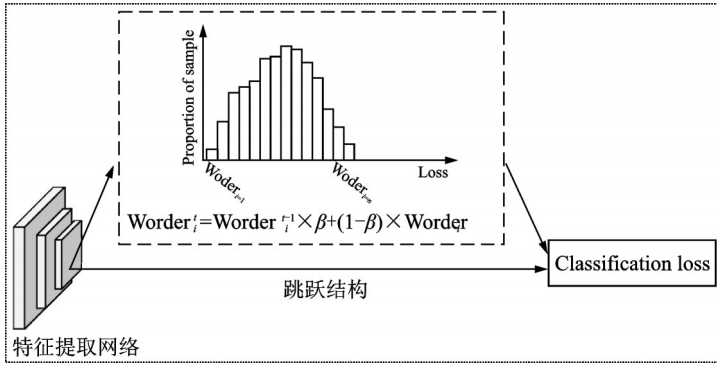


图9 修正学习模块

Fig.9 Revise learning module

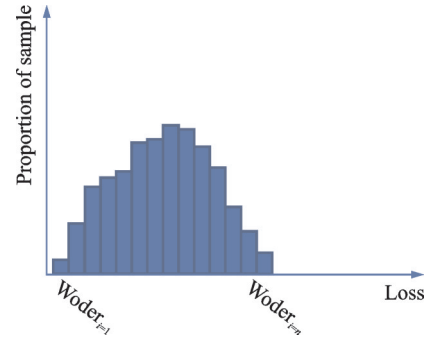


图10 自适应学习策略

Fig.10 Self-adaptive learning strategy

3 实验过程及结果

3.1 数据集

为了验证训练方法在行人重识别任务上的有效性,本文基于两个主流数据集进行实验,分别为Market1501^[6]和DukeMTMC-reID^[5]。Market1501^[6]数据集于2015年公开,由6个摄像头拍摄到1501个行人,每人至少由2个摄像头拍摄到。其中,训练集含有751个行人,共12936张图片,平均每人17张训练样本;测试集有750人,包含19732张图片,平均每人有26张图像数据。查找库(Query)有3368张图像,检测框由人工绘制,检索库(Gallery)中行人由DPM检测器检测得到。

DukeMTMC-reID^[5]数据集是DukeMTMC数据集的子集,该数据集包含8个85 min长度的高清晰视频,分别由8个摄像头采集,行人的检测框是由人工绘制的。本文使用了该数据集的行人重识别图像子集,数据集格式与Market1501^[6]相同,共1812个行人,有1404个行人出现在至少2个摄像头中,其他408个行人只出现在1个摄像头中。在测试集中从每个摄像头采集的图像集中,每个人选取1张图像作为查找图像,其他图像组成检索库。因此,训练集共16522张图像,由702个行人组成,测试集702个行人,查找库2228张图像,检索库17661张图像。

3.2 实验设置

3.2.1 噪声生成

为了最大程度模拟可能的标签噪声问题,实验在噪声比例5%和10%的数据集上进行,这是因为学术数据集样本量较少,一般经过人为标注,所以噪声比例较低,可以通过5%的噪声比例数据集实验进行模拟;而工业数据集由于样本量巨大,往往采用检测器进行非人工的数据标注,噪声比例一般接近10%。噪声类型方面,将选定样本的标签按照标签转移矩阵计算出噪声标签结果并替换该样本的原标签。具体来说,噪声包括两种情况:(1)一个行人标记的图像被划分成多个行人标记;(2)一个行人标记的图像可能由多个行人标记的图像组成。

3.2.2 模型结构

研究行人再识别问题普遍采用的网络结构是类别区分度嵌入(ID-discriminative embedding,

IDE^[16]),骨干网络为50层残差网络,输出的特征图经过全局平均池化层后输出一个特定维度对应输入的特征表示。分类器由带软最大化激活函数的全连接层组成,类别数为训练集的行人数,基本分类损失函数采用交叉熵损失函数。模型的训练与图像分类任务相同,被视为一个类别等于训练集行人数的图像分类任务。在模型测试阶段,经过训练的模型作为一个对输入图像提取特征表示的端到端网络,对查找库、检索库中的图像数据进行特征提取,提出得到的特征表示根据余弦相似度等其他度量方法,计算测试数据集中图像的相似程度。为了保证与最好方法(State of the art, SOTA)方法对比的公平性,本文图7、9中特征提取网络均使用IDE网络结构。

3.2.3 实验细节

实验配置方面,批尺寸为64,训练75轮,初始学习率设定为0.000 354,前5轮训练使用热启动(Warmup)学习率策略,并在第25、50轮对学习率进行折半衰减。

3.2.4 实验结果及分析

为了分析各参数在筛选学习模块中的作用,本文分别对记忆长度、记忆准确度和筛选参数进行了对比实验,实验结果分别如表1、2所示。其中筛选参数 α 的大小与筛选学习保留的难样本及可能噪声样本量相关,并且影响修正学习的效果,而记忆长度、记忆准确度参数相对不敏感。因此在实验中取记忆长度 $m=10$,记忆准确度 $\epsilon=0.85$,以有效地分析筛选参数的影响。同时为了确保泛化能力,本文使用IDE^[16]网络结构。IDE*为无噪声数据集,在表1所进行的实验中,IDE*实验结果为Rank 1 91.11, mAP 79.24;在表2所进行的实验中,IDE*实验结果为Rank 1 78.14, mAP 60.38。P1指筛选学习模块,P2指修正学习模块。由表1、2可见,在含有噪声的数据集中,使用筛选学习均可以提高模型在噪声条件下的精度, α 取0.3时取得最好效果,直观上是因为更大的 α 可以保留更多的难样本,这与筛选学习模块一致。值得注意的是,更大的 α 虽然会带入更多可能的噪声样本,但会在修正学习模块中被赋予更低的权重,降低噪声样本给模型带来的损耗,如果仅增大 α 而不使用修正学习对噪声、非噪声样本进行权重调整,精度会一定程度降低。此外,根据指数加权平均一般设置, β 取0.1。

最终本文将图7、9的特征提取网络更换为目前SOTA方法,包括MGN^[1]、ABD-NET^[17]、HORE-ID^[18]和PCB^[2],以验证本文方法在不同主干网络中的有效性、泛化和可移植性。使用 $\alpha=0.3$,噪声比例10%对方法的有效性进行验证,实验结果如表3所示。根据实验结果可以证明,本文提出的方法均可以降低SOTA受到噪声的影响,在含噪声的数据集上(实验噪声比例为10%)可以适当避免噪声损害。

表1 在Market1501^[6]上超参数 α 对比试验
Table 1 Ablation study of α on Market1501^[6]

Method	10% noise		5% noise	
	Rank 1	mAP	Rank 1	mAP
IDE	85.12	66.54	87.51	70.07
IDE+P1($\alpha=0.1$)	86.29	69.18	87.48	69.82
IDE+P1+P2($\alpha=0.1$)	86.95	71.33	87.96	70.36
IDE+P1($\alpha=0.2$)	86.47	69.54	87.88	70.17
IDE+P1+P2($\alpha=0.2$)	87.14	72.56	89.52	71.67
IDE+P1($\alpha=0.3$)	86.99	71.73	88.13	70.64
IDE+P1+P2($\alpha=0.3$)	88.91	72.27	89.81	73.66

表 2 在 DukeMTMC-reID^[5]上超参数 α 对比试验
 Table 2 Ablation study of α on DukeMTMC-reID^[5]

Method	10% noise		5% noise	
	Rank 1	mAP	Rank 1	mAP
IDE	76.39	58.69	77.42	60.03
IDE+P1($a=0.1$)	75.64	58.35	76.44	60.51
IDE+P1+P2($a=0.1$)	75.93	59.09	77.12	60.94
IDE+P1($a=0.2$)	75.85	57.13	76.31	60.14
IDE+P1+P2($a=0.2$)	75.97	58.41	77.36	61.01
IDE+P1($a=0.3$)	75.31	58.22	76.29	59.78
IDE+P1+P2($a=0.3$)	77.24	60.48	78.34	62.58

表 3 本文方法对比 SOTA 方法实验结果
 Table 3 Comparison of the proposed method with SOTA

Method	Market1501 ^[6]		DukeMTMC-ReID ^[5]	
	Rank 1	mAP	Rank 1	mAP
IDE*	91.11	79.24	78.14	60.38
IDE	85.12	66.54	76.39	58.69
MGN*	96.63	94.23	88.73	78.45
MGN	81.89	55.12	60.54	52.24
PCB*	92.38	77.43	81.75	66.13
PCB	85.17	62.14	67.16	42.23
HOREID*	79.66	70.35	69.47	62.12
HOREID	71.78	60.83	63.12	57.31
ABDNET*	95.65	88.36	89.01	78.61
ABDNET	79.66	70.37	69.48	62.19
P1+P2+MGN	83.57	62.94	64.93	59.64
P1+P2+PCB	88.96	66.63	73.82	57.15
P1+P2+HOREID	77.18	66.07	67.91	62.08
P1+P2+ABDNET	80.35	72.32	69.91	65.16

4 结束语

本文讨论了行人再识别任务中的标签噪声问题,利用噪声、非噪声数据在训练过程中的差异化特征表现,保留更多的非噪声数据用于模型训练而对噪声数据和难样本进行自适应权重学习。实验结果验证了本文方法的有效性,通过本文提出的训练方法可以有效地筛选非噪声数据,降低模型受损程度。

参考文献:

- [1] WANG G S, YUAN Y F, CHEN X, et al. Learning discriminative features with multiple granularities for person re-identification[C]//Proceedings of the 26th ACM International Conference on Multimedia. Seoul, Korea: ACM, 2018: 274-282.
- [2] SUN Y F, ZHENG L, YANG Y, et al. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline)[C]//Proceedings of the European Conference on Computer Vision. Munich, Germany: Springer, 2018:

480-496.

- [3] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification[EB/OL]. [2017-11-21][2020-11-15]. <https://arxiv.org/abs/1703.07737>.
- [4] ALGAN G, ULUSOY I. Image classification with deep learning in the presence of noisy labels: A survey [EB/OL]. [2019-01-11][2020-11-15]. <https://arxiv.org/abs/1912.05170>.
- [5] ZHENG Z D, ZHENG L, YANG Y. Unlabeled samples generated by gan improve the person re-identification baseline in vitro [C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 3754-3762.
- [6] ZHENG L, SHEN L, TIAN L, et al. Scalable person re-identification: A benchmark[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1116-1124.
- [7] ZHANG C S, BENGIO S, HARDT M, et al. Understanding deep learning requires rethinking generalization[EB/OL]. [2017-02-26][2020-11-15]. <https://arxiv.org/abs/1611.03530>.
- [8] NETTLETON D F, ORRIOLS-PUIG A, FORNELLS A. A study of the effect of different types of noise on the precision of supervised learning techniques[J]. *Artificial Intelligence Review*, 2010, 33(4): 275-306.
- [9] LAN X, WANG H X, GONG S G, et al. Deep reinforcement learning attention selection for person re-identification[EB/OL]. [2018-07-07][2020-11-15]. <https://arxiv.org/abs/1707.02785>.
- [10] JAEHWAN L, DONGGEUN Y, HYO-EUN K. Photometric transformer networks and label adjustment for breast density prediction[C]//Proceedings of the IEEE International Conference on Computer Vision Workshops. Seoul, Korea: IEEE, 2019.
- [11] MANG Y, PONG C. PurifyNet: A robust person re-identification model with noisy labels[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 2655-2666.
- [12] YU H X, ZHENG W S, WU A C, et al. Unsupervised person re-identification by soft multilabel learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019: 2148-2157.
- [13] YUAN Y, CHEN W Y, YANG Y, et al. In defense of the triplet loss again: Learning robust person re-identification with fast approximated triplet loss and label distillation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. [S.l.]: IEEE, 2020: 354-355.
- [14] YU T Y, LI D, YANG Y X, et al. Robust person re-identification by modelling feature uncertainty[C]//Proceedings of the IEEE International Conference on Computer Vision. Seoul, Korea: IEEE, 2019: 552-561.
- [15] ARPIT D, JASTRZEBSKI S, BALLAS N, et al. A closer look at memorization in deep networks[C]//Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia: JMLR.org, 2017: 233-242.
- [16] ZHENG L, YANG Y, HAUPTMANN A G. Person re-identification: Past, present and future[EB/OL]. [2016-10-10][2020-11-15]. <https://arxiv.org/abs/1610.02984>.
- [17] CHEN T, DING S, XIE J, et al. ABD-net: Attentive but diverse person re-identification[C]//Proceedings of the IEEE International Conference on Computer Vision. Seoul, Korea. IEEE, 2019: 8351-8361.
- [18] WANG G, YANG S, LIU H, et al. High-order information matters: Learning relation and topology for occluded person re-identification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE, 2020: 6449-6458.

作者简介:



唐轲(1992-),男,硕士研究生,研究方向:计算机视觉、行人重识别。



郎从妍(1978-),通信作者,女,教授,研究方向:多媒体信息检索与分析、计算机视觉、机器学习, E-mail: cylan@bjtu.edu.cn。

(编辑:刘彦东)