



(12) 发明专利

(10) 授权公告号 CN 118506407 B

(45) 授权公告日 2024.09.13

(21) 申请号 202410950010.5

G06V 10/82 (2022.01)

(22) 申请日 2024.07.16

G06N 3/045 (2023.01)

(65) 同一申请的已公布的文献号

G06N 3/0464 (2023.01)

申请公布号 CN 118506407 A

G06N 3/084 (2023.01)

G06N 3/0985 (2023.01)

(43) 申请公布日 2024.08.16

(56) 对比文件

(73) 专利权人 南京信息工程大学

CN 117078967 A, 2023.11.17

地址 210044 江苏省南京市宁六路219号

CN 118115947 A, 2024.05.31

(72) 发明人 张艳艳 秦雨 丁紫郁

审查员 沈婷婷

(74) 专利代理机构 北京同辉知识产权代理事务
所(普通合伙) 11357

专利代理师 连朋册

(51) Int. Cl.

G06V 40/10 (2022.01)

G06V 10/44 (2022.01)

G06V 10/764 (2022.01)

权利要求书5页 说明书12页 附图6页

(54) 发明名称

基于随机颜色丢弃和注意力的轻量行人重识别方法及系统

(57) 摘要

本发明公开了基于随机颜色丢弃和注意力的轻量行人重识别方法及系统,涉及行人重识别技术领域,接收图像数据,对图像数据进行预处理,得到预处理图像数据;将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内,提取特征,得到图像特征;将图像特征通过全连接层进行分类,将图像特征映射到相应的类别标签上,得到分类后的图像特征;使用分类后的图像特征计算带有标签平滑的身份损失,通过反向传播更新梯度对预先建立的轻量行人重识别网络模型进行优化训练,得到优化后的轻量行人重识别网络模型;获取行人重识别数据集的测试集,将行人重识别数据集的测试集输入至优化后的轻量行人重识别网络模型内,得到轻量行人重识别结果。



1. 基于随机颜色丢弃和注意力的轻量行人重识别方法,其特征在于,方法包括以下步骤:

接收图像数据,对图像数据进行预处理,得到预处理图像数据;

所述对图像数据进行预处理的过程包括:

数据增强:采用随机翻转,随机擦除以及基于LAGT的随机颜色丢弃策略进行执行,最终得到预处理图像数据;

所述基于LAGT的随机颜色丢弃策略采用聚合灰度变换对图像灰度化,计算过程如下:

$$t_{AGT}(R, G, B) = \frac{1}{\varphi} [I_R(e, f) + I_G(e, f) + I_B(e, f)]$$

其中, R , G 和 B 表示红色、绿色和蓝色通道, $e \in \{1, \dots, H\}$, $f \in \{1, \dots, W\}$, $I_R(\cdot)$, $I_G(\cdot)$ 和 $I_B(\cdot)$ 分别表示红色通道、绿色通道和蓝色通道的特定位置处的像素值,权重系数表示为 φ , H 为图像的高, W 为图像的宽;

所述LAGT的实现过程如下:

在数据加载过程中,采用随机身份采样器,随机选取 K 种身份,每个身份选取 M 张行人,图片训练批次大小为 $B = K \times M$,集合表示为 $x_i^v = (x_i^v | i = 1, 2, \dots, K \times M)$,其中 $x_i^v = (x_i^v | y_i)$ 表示训练批次中的第 i 个图像, y_i 表示第 i 个图像的样本标签,LAGT以概率 P 将原始图像转变为灰度图像,从原始图像随机选择一个矩形区域,将灰度图像对应位置区域的灰度值替换到原始图像之中,给定一张原始行人图片 x^v ,以概率 P 进行聚合灰度变换,对应的AGT图像被定义为:

$$x^g = t_{AGT}(x^v)$$

原始图像 x^v 面积大小 S :

$$S = H \times W$$

其中, H 为图像的高, W 为图像的宽;

AGT矩形的面积 $S_{t_{AGT}}$:

$$S_{t_{AGT}} = Rand(s_l, s_h) \times S$$

其中, s_l 、 s_h 为AGT图像面积与原始图像面积比的最小值与最大值;

AGT矩形的高宽比 $r_{t_{AGT}}$ 及高 $H_{t_{AGT}}$ 和宽 $W_{t_{AGT}}$:

$$r_{t_{AGT}} = Rand(r_1, r_2)$$

$$H_{t_{AGT}} = \sqrt{S_{t_{AGT}} \times r_{t_{AGT}}}$$

$$W_{t_{AGT}} = \sqrt{S_{t_{AGT}} / r_{t_{AGT}}}$$

其中 r_1 , r_2 为灰度变换矩形高宽比的最大值和最小值;

在原始图像 x^v 中随机初始化一个点 $Q(x_{t_{AGT}}, y_{t_{AGT}})$,满足下列条件:

$$\begin{cases} x_{t_{AGT}} = Rand(0, W) \\ y_{t_{AGT}} = Rand(0, H) \\ x_{t_{AGT}} + W_{t_{AGT}} \leq W \\ y_{t_{AGT}} + H_{t_{AGT}} \leq H \end{cases}$$

H 为图像的高, W 为图像的宽;

选定的LAGT区域为 $RandPosition$:

$$RandPosition = (x_{t_{AGT}}, y_{t_{AGT}}, x_{t_{AGT}} + W_{t_{AGT}}, y_{t_{AGT}} + H_{t_{AGT}})$$

对于每个 x^v , 所选定的LAGT区域 $rect$ 为:

$$rect = RandPosition(x^v)$$

LAGT算法最终可表示为:

$$x^g = LAGT(x^v, x^g, rect)$$

其中, $LAGT(g)$ 是将 x^g 图像对应矩形中的像素给予到 x^v 图像, x^g 是LAGT变换后的样本;

将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内, 提取特征, 输出得到图像特征;

所述嵌入有级联自注意力模块包括SSAM和CSAM;

所述SSAM对预处理图像数据提取的中间特征图为 $A \in R^{C \times H \times W}$, 其中 C 为特征通道数, $H \times W$ 为中间特征图的大小, 对 A 进行 1×1 卷积操作, 得到 $Q \in R^{r \times H \times W}$, $K \in R^{r \times H \times W}$, $V \in R^{C \times H \times W}$, 其中 $r = 8$, 对 Q, K, V 重塑操作后, 求得空间自注意力亲和矩阵 $B = Q^T \times K$, $B \in R^{N \times N}$, 其中 $N = H \times W$, 该过程表示如下:

$$b_{i,j} = \frac{\exp(Q_i g K_j)}{\sum_{j=1}^N \exp(Q_i g K_j)}, i, j \in \{1, \dots, N\}$$

其中 $b_{i,j}$ 表示空间中第 i 个位置对第 j 个位置的注意权重, 将 V 与 B 相乘嵌入注意力权重后, 再与原始特征图像素叠加得到空间自注意力加权的特征图 X :

$$X = \lambda V g B + A$$

其中 λ 是调整SSAM影响的超参数, 将空间自注意力加权的特征图 X 通过CSAM进行处理;

所述CSAM对于输入带有空间自注意力加权的特征图 $X \in R^{C \times H \times W}$, 重塑操作后, 求得通道自注意力亲和矩阵 $H = X \times X^T$, $H \in R^{C \times C}$, 该过程表示如下:

$$h_{i,j} = \frac{\exp(X_i g X_j)}{\sum_{j=1}^C \exp(X_i g X_j)}, i, j \in \{1, \dots, C\}$$

其中 $h_{i,j}$ 表示通道 i 对通道 j 的注意权重, 对于 H , 初始化一个与 H 大小相同且值全为

H 最大值的矩阵 H_{\max} , 新的通道自注意力权重亲和矩阵为 $H' = H_{\max} - H$, $H' \in R^{C \times C}$, 将 H' 与 X 相乘嵌入注意力权重后再与 X 对应位置像素叠加得到通道自注意力加权的特征图 Y ;

$$Y = \delta H' gX + X$$

其中 δ 是调整CSAM影响的超参数;

将通道自注意力加权的特征图 Y 经过OSNet骨干网后, 得到图像特征;

将图像特征通过全连接层进行分类, 将图像特征映射到相应的类别标签上, 得到分类后的图像特征;

使用分类后的图像特征计算带有标签平滑的身份损失, 通过反向传播更新梯度, 对预先建立的轻量行人重识别网络模型进行优化训练, 得到优化后的轻量行人重识别网络模型;

获取行人重识别数据集的测试集, 将行人重识别数据集的测试集输入至优化后的轻量行人重识别网络模型内, 输出得到轻量行人重识别结果。

2. 根据权利要求1所述的基于随机颜色丢弃和注意力的轻量行人重识别方法, 其特征在于, 所述使用分类后的图像特征计算带有标签平滑的身份损失 L_{ID} 的计算过程:

$$L_{ID} = -\sum_{i=1}^N q_i \log(p_i) \begin{cases} q_i = 0, y \neq i \\ q_i = 1, y = i \end{cases}$$

其中 $i \in \{1, 2, \dots, N\}$ 表示行人类别, N 表示训练集行人图像数量, y 为真实标签, p_i 为网络的预测逻辑值。

3. 基于随机颜色丢弃和注意力的轻量行人重识别系统, 其特征在于, 包括:

图像处理模块, 用于接收图像数据, 对图像数据进行预处理, 得到预处理图像数据;

所述对图像数据进行预处理的过程包括:

数据增强: 采用随机翻转, 随机擦除以及基于LAGT的随机颜色丢弃策略进行执行, 最终得到预处理图像数据;

所述基于LAGT的随机颜色丢弃策略采用聚合灰度变换对图像灰度化, 计算过程如下:

$$t_{AGT}(R, G, B) = \frac{1}{\phi} [I_R(e, f) + I_G(e, f) + I_B(e, f)]$$

其中, R, G 和 B 表示红色、绿色和蓝色通道, $e \in \{1, \dots, H\}$, $f \in \{1, \dots, W\}$, $I_R(\cdot)$, $I_G(\cdot)$ 和 $I_B(\cdot)$ 分别表示红色通道、绿色通道和蓝色通道的特定位置处的像素值, 权重系数表示为 ϕ , H 为图像的高, W 为图像的宽;

所述LAGT的实现过程如下:

在数据加载过程中, 采用随机身份采样器, 随机选取 K 种身份, 每个身份选取 M 张行人, 图片训练批次大小为 $B = K \times M$, 集合表示为 $x_i^v = (x_i^v | i = 1, 2, \dots, K \times M)$, 其中 $x_i^v = (x_i^v | y_i)$ 表示训练批次中的第 i 个图像, y_i 表示第 i 个图像的样本标签, LAGT以概率 P 将原始图像转变为灰度图像, 从原始图像随机选择一个矩形区域, 将灰度图像对应位置区域的灰度值替换到原始图像之中, 给定一张原始行人图片 x^v , 以概率 P 进行聚合灰度变换, 对应的AGT图像被定义为:

$$\mathbf{x}^g = t_{AGT}(\mathbf{x}^v)$$

原始图像 \mathbf{x}^v 面积大小 S :

$$S = H \times W$$

其中, H 为图像的高, W 为图像的宽;

AGT矩形的面积 $S_{t_{AGT}}$:

$$S_{t_{AGT}} = Rand(s_l, s_h) \times S$$

其中, s_l 、 s_h 为AGT图像面积与原始图像面积比的最小值与最大值;

AGT矩形的高宽比 $r_{t_{AGT}}$ 及高 $H_{t_{AGT}}$ 和宽 $W_{t_{AGT}}$:

$$r_{t_{AGT}} = Rand(r_1, r_2)$$

$$H_{t_{AGT}} = \sqrt{S_{t_{AGT}} \times r_{t_{AGT}}}$$

$$W_{t_{AGT}} = \sqrt{S_{t_{AGT}} / r_{t_{AGT}}}$$

其中 r_1 、 r_2 为灰度变换矩形高宽比的最大值和最小值;

在原始图像 \mathbf{x}^v 中随机初始化一个点 $Q(x_{t_{AGT}}, y_{t_{AGT}})$, 满足下列条件:

$$\begin{cases} x_{t_{AGT}} = Rand(0, W) \\ y_{t_{AGT}} = Rand(0, H) \\ x_{t_{AGT}} + W_{t_{AGT}} \leq W \\ y_{t_{AGT}} + H_{t_{AGT}} \leq H \end{cases}$$

H 为图像的高, W 为图像的宽;

选定的LAGT区域为 $RandPosition$:

$$RandPosition = (x_{t_{AGT}}, y_{t_{AGT}}, x_{t_{AGT}} + W_{t_{AGT}}, y_{t_{AGT}} + H_{t_{AGT}})$$

对于每个 \mathbf{x}^v , 所选定的LAGT区域 $rect$ 为:

$$rect = RandPosition(\mathbf{x}^v)$$

LAGT算法最终可表示为:

$$\mathbf{x}^{lg} = LAGT(\mathbf{x}^v, \mathbf{x}^g, rect)$$

其中, $LAGT(\mathbf{g})$ 是将 \mathbf{x}^g 图像对应矩形中的像素给予到 \mathbf{x}^v 图像, \mathbf{x}^{lg} 是LAGT变换后的样本;

特征提取模块, 用于将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内, 提取特征, 输出得到图像特征;

所述嵌入有级联自注意力模块包括SSAM和CSAM;

所述SSAM对预处理图像数据提取的中间特征图为 $A \in R^{C \times H \times W}$, 其中 C 为特征通道数, $H \times W$ 为中间特征图的大小, 对 A 进行 1×1 卷积操作, 得到 $Q \in R^{r \times H \times W}$, $K \in R^{r \times H \times W}$,

$V \in R^{C \times H \times W}$, 其中 $r = 8$, 对 Q, K, V 重塑操作后, 求得空间自注意力亲和矩阵 $B = Q^T \times K$, $B \in R^{N \times N}$, 其中 $N = H \times W$, 该过程表示如下:

$$b_{i,j} = \frac{\exp(Q_i \mathbf{g} K_j)}{\sum_{j=1}^N \exp(Q_i \mathbf{g} K_j)}, i, j \in \{1, \dots, N\}$$

其中 $b_{i,j}$ 表示空间中第 i 个位置对第 j 个位置的注意权重, 将 V 与 B 相乘嵌入注意力权重后, 再与原始特征图像素叠加得到空间自注意力加权的特征图 X :

$$X = \lambda V \mathbf{g} B + A$$

其中 λ 是调整SSAM影响的超参数, 将空间自注意力加权的特征图 X 通过CSAM进行处理;

所述CSAM对于输入带有空间自注意力加权的特征图 $X \in R^{C \times H \times W}$, 重塑操作后, 求得通道自注意力亲和矩阵 $H = X \times X^T$, $H \in R^{C \times C}$, 该过程表示如下:

$$h_{i,j} = \frac{\exp(X_i \mathbf{g} X_j)}{\sum_{j=1}^C \exp(X_i \mathbf{g} X_j)}, i, j \in \{1, \dots, C\}$$

其中 $h_{i,j}$ 表示通道 i 对通道 j 的注意权重, 对于 H , 初始化一个与 H 大小相同且值全为 H 最大值的矩阵 H_{\max} , 新的通道自注意权重亲和矩阵为 $H' = H_{\max} - H$, $H' \in R^{C \times C}$, 将 H' 与 X 相乘嵌入注意力权重后再与 X 对应位置像素叠加得到通道自注意力加权的特征图 Y ;

$$Y = \delta H' \mathbf{g} X + X$$

其中 δ 是调整CSAM影响的超参数;

将通道自注意力加权的特征图 Y 经过OSNet骨干网后, 得到图像特征;

图像分类模块, 用于将图像特征通过全连接层进行分类, 将图像特征映射到相应的类别标签上, 得到分类后的图像特征;

模型训练模块, 用于使用分类后的图像特征计算带有标签平滑的身份损失, 通过反向传播更新梯度, 对预先建立的轻量行人重识别网络模型进行优化训练, 得到优化后的轻量行人重识别网络模型;

行人重识别模块, 用于获取行人重识别数据集的测试集, 将行人重识别数据集的测试集输入至优化后的轻量行人重识别网络模型内, 输出得到轻量行人重识别结果。

4. 一种终端设备, 包括存储器、处理器及存储在存储器中并能够在处理器上运行的计算机程序, 其特征在于, 所述存储器中存储有能够在处理器上运行的计算机程序, 所述处理器加载并执行计算机程序时, 采用了权利要求1至2中任一项所述的基于随机颜色丢弃和注意力的轻量行人重识别方法。

基于随机颜色丢弃和注意力的轻量行人重识别方法及系统

技术领域

[0001] 本发明涉及行人重识别技术领域,具体的是基于随机颜色丢弃和注意力的轻量行人重识别方法及系统。

背景技术

[0002] 随着深度学习技术的迅猛发展,行人重识别领域取得了显著的进展。残差网络(ResNet),应用于行人重识别领域并取得了显著的成果,ResNet网络通过引入残差连接,解决了深层网络训练中的梯度消失问题,使得网络更加容易优化。

[0003] 现有技术的识别方法的鲁棒性和泛化性不高,为提高识别系统的鲁棒性、泛化性,一些数据增强策略如颜色增强被广泛应用。颜色信息作为行人重要的区分特征,可以丰富行人特征的表达,尤其在面对光照变化等复杂场景时发挥重要作用,融合颜色信息的行人重识别系统相比传统方法在复杂环境下具有更明显的优势。但在某些情况下,颜色特征产生的颜色偏差在一定程度上限制了模型做出正确的预测,主要体现在两个方面:首先是同一行人图像之间的颜色差异,增加了误识别的可能性;其次,颜色偏差弱化了不同行人图像之间的特征差异,降低了系统的辨识度。目前,对减弱颜色偏差提高模型鲁棒性的研究较少。

发明内容

[0004] 为解决上述背景技术中提到的不足,本发明的目的在于提供基于随机颜色丢弃和注意力的轻量行人重识别方法及系统。

[0005] 第一方面,本发明的目的可以通过以下技术方案实现:基于随机颜色丢弃和注意力的轻量行人重识别方法,方法包括以下步骤:

[0006] 接收图像数据,对图像数据进行预处理,得到预处理图像数据;

[0007] 将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内,提取特征,输出得到图像特征;

[0008] 将图像特征通过全连接层进行分类,将图像特征映射到相应的类别标签上,得到分类后的图像特征;

[0009] 使用分类后的图像特征计算带有标签平滑的身份损失,通过反向传播更新梯度,对预先建立的轻量行人重识别网络模型进行优化训练,得到优化后的轻量行人重识别网络模型;

[0010] 获取行人重识别数据集的测试集,将行人重识别数据集的测试集输入至优化后的轻量行人重识别网络模型内,输出得到轻量行人重识别结果。

[0011] 结合第一方面,在第一方面的某些实现方式中,该方法还包括:所述对图像数据进行预处理的过程包括:

[0012] 数据增强:采用随机翻转,随机擦除以及基于LAGT的随机颜色丢弃策略进行执行,最终得到预处理图像数据。

[0013] 结合第一方面,在第一方面的某些实现方式中,该方法还包括:所述基于LAGT的随机颜色丢弃策略采用聚合灰度变换对图像灰度化,计算过程如下:

$$[0014] \quad t_{AGT}(R, G, B) = \frac{1}{\varphi} [I_R(e, f) + I_G(e, f) + I_B(e, f)]$$

[0015] 其中, R, G 和 B 表示红色、绿色和蓝色通道, $e \in \{1, \dots, H\}$, $f \in \{1, \dots, W\}$, $I_R(\cdot)$, $I_G(\cdot)$ 和 $I_B(\cdot)$ 分别表示红色通道、绿色通道和蓝色通道的特定位置处的像素值,权重系数表示为 φ , H 为图像的高, W 为图像的宽。

[0016] 结合第一方面,在第一方面的某些实现方式中,该方法还包括:所述LAGT的实现过程如下:

[0017] 在数据加载过程中,采用随机身份采样器,随机选取 K 种身份,每个身份选取 M 张行人,图片训练批次大小为 $B = K \times M$,集合表示为 $x_i^v = (x_i^v | i = 1, 2, \dots, K \times M)$,其中 $x_i^v = (x_i^v | y_i)$ 表示训练批次中的第 i 个图像, y_i 表示第 i 个图像的样本标签,LAGT以概率 P 将原始图像转变为灰度图像,从原始图像随机选择一个矩形区域,将灰度图像对应位置区域的灰度值替换到原始图像之中,给定一张原始行人图片 x^v ,以概率 P 进行聚合灰度变换,对应的AGT图像被定义为:

$$[0018] \quad x^g = t_{AGT}(x^v)$$

[0019] 原始图像 x^v 面积大小 S :

$$[0020] \quad S = H \times W$$

[0021] 其中, H 为图像的高, W 为图像的宽;

[0022] AGT矩形的面积 $S_{t_{AGT}}$:

$$[0023] \quad S_{t_{AGT}} = Rand(s_l, s_h) \times S$$

[0024] 其中, s_l 、 s_h 为AGT图像面积与原始图像面积比的最小值与最大值;

[0025] AGT矩形的高宽比 $r_{t_{AGT}}$ 及高 $H_{t_{AGT}}$ 和宽 $W_{t_{AGT}}$:

$$[0026] \quad r_{t_{AGT}} = Rand(r_1, r_2)$$

$$[0027] \quad H_{t_{AGT}} = \sqrt{S_{t_{AGT}} \times r_{t_{AGT}}}$$

$$[0028] \quad W_{t_{AGT}} = \sqrt{S_{t_{AGT}} / r_{t_{AGT}}}$$

[0029] 其中 r_1 , r_2 为灰度变换矩形高宽比的最大值和最小值;

[0030] 在原始图像 x^v 中随机初始化一个点 $Q(x_{t_{AGT}}, y_{t_{AGT}})$,满足下列条件:

$$[0031] \quad \begin{cases} x_{t_{AGT}} = Rand(0, W) \\ y_{t_{AGT}} = Rand(0, H) \\ x_{t_{AGT}} + W_{t_{AGT}} \leq W \\ y_{t_{AGT}} + H_{t_{AGT}} \leq H \end{cases}$$

[0032] H 为图像的高, W 为图像的宽;

[0033] 选定的LAGT区域为 $RandPosition$:

$$[0034] \quad RandPosition = (x_{t_{AGT}}, y_{t_{AGT}}, x_{t_{AGT}} + W_{t_{AGT}}, y_{t_{AGT}} + H_{t_{AGT}})$$

[0035] 对于每个 x^v , 所选定的LAGT区域 $rect$ 为:

$$[0036] \quad rect = RandPosition(x^v)$$

[0037] LAGT算法最终可表示为:

$$[0038] \quad x^{lg} = LAGT(x^v, x^g, rect)$$

[0039] 其中, $LAGT(\mathbf{g})$ 是将 x^g 图像对应矩形中的像素给予到 x^v 图像, x^{lg} 是LAGT变换后的样本。

[0040] 结合第一方面, 在第一方面的某些实现方式中, 该方法还包括: 所述嵌入有级联自注意力模块包括SSAM和CSAM。

[0041] 结合第一方面, 在第一方面的某些实现方式中, 该方法还包括: 所述SSAM对预处理图像数据提取的中间特征图为 $A \in R^{C \times H \times W}$, 其中 C 为特征通道数, $H \times W$ 为中间特征图的大小, 对 A 进行 1×1 卷积操作, 得到 $Q \in R^{r \times H \times W}$, $K \in R^{r \times H \times W}$, $V \in R^{C \times H \times W}$, 其中 $r = 8$, 对 Q, K, V 重塑操作后, 求得空间自注意力亲和矩阵 $B = Q^T \times K$, $B \in R^{N \times N}$, 其中 $N = H \times W$, 该过程表示如下:

$$[0042] \quad b_{i,j} = \frac{\exp(Q_i \mathbf{g} K_j)}{\sum_{j=1}^N \exp(Q_i \mathbf{g} K_j)}, \quad i, j \in \{1, \dots, N\}$$

[0043] 其中 $b_{i,j}$ 表示空间中第 i 个位置对第 j 个位置的注意权重, 将 V 与 B 相乘嵌入注意力权重后, 再与原始特征图像素叠加得到空间自注意力加权的特征图 X :

$$[0044] \quad X = \lambda V \mathbf{g} B + A$$

[0045] 其中 λ 是调整SSAM影响的超参数, 将空间自注意力加权的特征图 X 通过CSAM进行处理。

[0046] 结合第一方面, 在第一方面的某些实现方式中, 该方法还包括: 所述CSAM对于输入带有空间自注意力加权的特征图 $X \in R^{C \times H \times W}$, 重塑操作后, 求得通道自注意力亲和矩阵 $H = X \times X^T$, $H \in R^{C \times C}$, 该过程表示如下:

$$[0047] \quad h_{i,j} = \frac{\exp(X_i \mathbf{g} X_j)}{\sum_{j=1}^C \exp(X_i \mathbf{g} X_j)}, \quad i, j \in \{1, \dots, C\}$$

[0048] 其中 $h_{i,j}$ 表示通道 i 对通道 j 的注意权重, 对于 H , 初始化一个与 H 大小相同且值全为 H 最大值的矩阵 H_{\max} , 新的通道自注意权重亲和矩阵为 $H' = H_{\max} - H$, $H' \in R^{C \times C}$, 将 H' 与 X 相乘嵌入注意力权重后再与 X 对应位置像素叠加得到通道自注意力加权的特征图 Y ;

[0049] $Y = \delta H' gX + X$

[0050] 其中 δ 是调整CSAM影响的超参数;

[0051] 将通道自注意力加权的特征图 Y 经过OSNet骨干网后, 得到图像特征。

[0052] 结合第一方面, 在第一方面的某些实现方式中, 该方法还包括: 所述使用分类后的图像特征计算带有标签平滑的身份损失 L_{ID} 的计算过程:

$$[0053] \quad L_{ID} = -\sum_{i=1}^N q_i \log(p_i) \begin{cases} q_i = 0, y \neq i \\ q_i = 1, y = i \end{cases}$$

[0054] 其中 $i \in \{1, 2, \dots, N\}$ 表示行人类别, N 表示训练集行人图像数量, y 为真实标签, P_i 为网络的预测逻辑值。

[0055] 第二方面, 为了达到上述目的, 本发明公开了基于随机颜色丢弃和注意力的轻量行人重识别系统, 包括:

[0056] 图像处理模块, 用于接收图像数据, 对图像数据进行预处理, 得到预处理图像数据;

[0057] 特征提取模块, 用于将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内, 提取特征, 输出得到图像特征;

[0058] 图像分类模块, 用于将图像特征通过全连接层进行分类, 将图像特征映射到相应的类别标签上, 得到分类后的图像特征;

[0059] 模型训练模块, 用于使用分类后的图像特征计算带有标签平滑的身份损失, 通过反向传播更新梯度, 对预先建立的轻量行人重识别网络模型进行优化训练, 得到优化后的轻量行人重识别网络模型;

[0060] 行人重识别模块, 用于获取行人重识别数据集的测试集, 将行人重识别数据集的测试集输入至优化后的轻量行人重识别网络模型内, 输出得到轻量行人重识别结果。

[0061] 一种终端设备, 包括存储器、处理器及存储在存储器中并能够在处理器上运行的计算机程序, 所述存储器中存储有能够在处理器上运行的计算机程序, 所述处理器加载并执行计算机程序时, 采用了如上所述的基于随机颜色丢弃和注意力的轻量行人重识别方法。

[0062] 本发明的有益效果:

[0063] 本发明采用OSNet作为模型的骨干网络, 参数量大幅减少;

[0064] 对于具有相同颜色衣物的行人, 背景颜色与行人颜色差异大的情形, 本发明通过LAGT算法能有效抑制颜色偏差对识别效果的负面影响, 鼓励模型发现并注重与颜色不相关的特征信息, 平衡神经网络对颜色特征与非颜色特征的权重, 提高了Baseline识别效果;

[0065] 对于具有复杂背景以及遮挡的行人图片, 本发明通过级联自注意力模块能有效聚合行人本身, 抑制不相关的背景, 从而使提取的特征更为精细且具判别力, 提高了Baseline

识别效果。

附图说明

[0066] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图;

[0067] 图1是本发明方法流程示意图;

[0068] 图2为本发明的基于随机颜色丢弃和自注意力的轻量行人重识别网络的总体框架图;

[0069] 图3为原始样本和本发明的LAGT样本示例;

[0070] 图4为本发明采用的SSAM;

[0071] 图5为本发明采用的CSAM;

[0072] 图6为Market-1501数据集上使用RGB图像与原灰度变换图像训练模型的检索结果;

[0073] 图7为数据集上绿色背景下行人图像示例;

[0074] 图8为Market-1501数据集上绿色背景行人图像使用RGB图像与原灰度变换图像训练模型的检索结果;

[0075] 图9为Market-1501数据集上使用RGB图像与不同灰度变换的灰度图像训练模型的检索结果;

[0076] 图10为本发明在颜色偏差下的行人重识别可视化对比结果;

[0077] 图11为本发明在复杂背景及遮挡下的行人重识别可视化对比结果;

[0078] 图12为本发明系统结构示意图。

具体实施方式

[0079] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0080] 实施例一:

[0081] 下面,对本申请实施例涉及的相关术语进行介绍:

[0082] 灰度变换是指根据某种目标条件按一定变换关系逐点改变源图像中每一个像素灰度值的方法。目的是为了改善画质,使图像的显示效果更加清晰。图像的灰度变换处理是图像增强处理技术中的一种非常基础、直接的空间域图像处理方法,也是图像数字化软件和图像显示软件的一个重要组成部分。

[0083] 自注意力(Self-Attention)机制是一种特殊的注意力机制,它允许模型在处理一个序列时,考虑到序列中每个元素与其他所有元素的关系。这种机制可以帮助模型更好地理解序列中的上下文信息,从而更准确地处理序列数据。(序列数据是一种数据类型,其中的元素存在特定的顺序。每个元素都有特定的位置,且位置之间的顺序关系对于数据的含义和处理方式有重要影响)

[0084] 如图1所示,基于随机颜色丢弃和注意力的轻量行人重识别方法,其特征在于,方法包括以下步骤:

[0085] 接收图像数据,对图像数据进行预处理,得到预处理图像数据;

[0086] 对图像数据进行预处理的过程包括:

[0087] 数据增强:采用随机翻转,随机擦除以及基于LAGT的随机颜色丢弃策略进行执行,最终得到预处理图像数据;

[0088] 基于LAGT的随机颜色丢弃策略采用聚合灰度变换对图像灰度化,计算过程如下:

$$[0089] \quad t_{AGT}(R, G, B) = \frac{1}{\varphi} [I_R(e, f) + I_G(e, f) + I_B(e, f)]$$

[0090] 其中, R, G 和 B 表示红色、绿色和蓝色通道, $e \in \{1, \dots, H\}$, $f \in \{1, \dots, W\}$, $I_R(\cdot)$, $I_G(\cdot)$ 和 $I_B(\cdot)$ 分别表示红色通道、绿色通道和蓝色通道的特定位置处的像素值,权重系数表示为 φ 。

[0091] LAGT算法实现如下:在数据加载过程中,采用随机身份采样器,随机选取 K 种身份,每个身份选取 M 张行人,图片训练批次大小为 $B = K \times M$ 。该集合表示为 $x_i^v = (x_i^v | i = 1, 2, \dots, K \times M)$, 其中 $x_i^v = (x_i^v | y_i)$ 表示训练批次中的第 i 个图像, y_i 表示第 i 个图像的样本标签。LAGT以概率 P 将原始图像转变为灰度图像,从原始图像随机选择一个矩形区域,将灰度图像对应位置区域的灰度值替换到原始图像之中。给定一张原始行人图片 x^v , 以概率 P 进行聚合灰度变换,其对应的AGT图像被定义为:

$$[0092] \quad x^g = t_{AGT}(x^v)$$

[0093] 原始图像 x^v 面积大小 S :

$$[0094] \quad S = H \times W$$

[0095] 其中, H 为图像的高, W 为图像的宽。

[0096] AGT矩形的面积 $S_{t_{AGT}}$:

$$[0097] \quad S_{t_{AGT}} = Rand(s_l, s_h) \times S$$

[0098] 其中, s_l 、 s_h 为AGT图像面积与原始图像面积比的最小值与最大值。

[0099] AGT矩形的高宽比 $r_{t_{AGT}}$ 及高 $H_{t_{AGT}}$ 和宽 $W_{t_{AGT}}$:

$$[0100] \quad r_{t_{AGT}} = Rand(r_1, r_2)$$

$$[0101] \quad H_{t_{AGT}} = \sqrt{S_{t_{AGT}} \times r_{t_{AGT}}}$$

$$[0102] \quad W_{t_{AGT}} = \sqrt{S_{t_{AGT}} / r_{t_{AGT}}}$$

[0103] 其中 r_1 , r_2 为灰度变换矩形高宽比的最大值和最小值。

[0104] 在原始图像 x^v 中随机初始化一个点 $Q(x_{t_{AGT}}, y_{t_{AGT}})$, 满足下列条件:

$$[0105] \quad \begin{cases} x_{t_{AGT}} = Rand(0, W) \\ y_{t_{AGT}} = Rand(0, H) \\ x_{t_{AGT}} + W_{t_{AGT}} \leq W \\ y_{t_{AGT}} + H_{t_{AGT}} \leq H \end{cases}$$

[0106] H 为图像的高, W 为图像的宽;

[0107] 选定的LAGT区域为 $RandPosition$:

$$[0108] \quad RandPosition = (x_{t_{AGT}}, y_{t_{AGT}}, x_{t_{AGT}} + W_{t_{AGT}}, y_{t_{AGT}} + H_{t_{AGT}})$$

[0109] 对于每个 x^v , 所选定的LAGT区域为:

$$[0110] \quad rect = RandPosition(x^v)$$

[0111] LAGT算法最终可表示为:

$$[0112] \quad x^{lg} = LAGT(x^v, x^g, rect)$$

[0113] 其中, $LAGT(\mathbf{g})$ 是将 x^g 图像对应矩形中的像素给予到 x^v 图像, x^{lg} 是LAGT变换后的样本。原始样本与LAGT样本如图3所示。

[0114] 将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内, 提取特征, 输出得到图像特征;

[0115] 嵌入有级联自注意力模块包括SSAM和CSAM, 分别如图4、5所示;

[0116] SSAM可以帮助网络聚合空间中语义相关的特征, 并突出个体特征, 使网络更加关注行人的细节。

[0117] 神经网络提取的中间特征图为 $A \in R^{C \times H \times W}$, 其中 C 为特征通道数, $H \times W$ 为中间特征图的大小。对 A 进行 1×1 卷积操作, 得到 $Q \in R^{r \times H \times W}$, $K \in R^{r \times H \times W}$, $V \in R^{C \times H \times W}$, 在推断注意力矩阵的过程中, 对通道降维可以提取更加精炼和抽象的特征表示, 减少计算的复杂性的同时不会影响网络的性能, 其中 $r = 8$ 。对 Q, K, V 重塑 (reshape) 操作后, 求得空间自注意力亲和矩阵 $B = Q^T \times K$, $B \in R^{N \times N}$, 其中 $N = H \times W$, 该过程表示入下:

$$[0118] \quad b_{i,j} = \frac{\exp(Q_i \mathbf{g} K_j)}{\sum_{j=1}^N \exp(Q_i \mathbf{g} K_j)}, \quad i, j \in \{1, \dots, N\}$$

[0119] 其中 $b_{i,j}$ 表示空间中第 i 个位置对第 j 个位置的注意权重。将 V 与 B 相乘嵌入注意力权重后, 再与原始特征图像素叠加得到空间自注意力加权的特征图 X :

$$[0120] \quad X = \lambda V \mathbf{g} B + A$$

[0121] 其中 λ 是调整SSAM影响的超参数。

[0122] CSAM则能够帮助网络学习通道间的关联信息, 进一步提取图像中不同通道之间的有效特征表示。

[0123] 对于输入带有空间自注意力加权的特征图 $X \in R^{C \times H \times W}$, 重塑 (reshape) 操作后, 求得通道自注意力亲和矩阵 $H = X \times X^T$, $H \in R^{C \times C}$, 该过程表示如下:

$$[0124] \quad h_{i,j} = \frac{\exp(X_i \mathbf{g} X_j)}{\sum_{j=1}^C \exp(X_i \mathbf{g} X_j)}, i, j \in \{1, \dots, C\}$$

[0125] 其中 $h_{i,j}$ 表示通道 i 对通道 j 的注意权重。对于 H , 我们使用归一化 (Normalize) 方法, 初始化一个与 H 大小相同且值全为 H 最大值的矩阵 H_{\max} , 新的通道自注意权重亲和矩阵为 $H' = H_{\max} - H$, $H' \in R^{C \times C}$, 此方法有效避免最大激活值受到噪声或异常值的影响, 通过对每个通道减去最大激活值, 鼓励其他通道提供互补的信息, 增加模型对不同特征通道的感知多样性, 从而提高模型的鲁棒性。将 H' 与 X 相乘嵌入注意力权重后再与 X 对应位置像素叠加得到通道自注意力加权的特征图 Y :

$$[0126] \quad Y = \delta H' \mathbf{g} X + X$$

[0127] 其中 δ 是调整CSAM影响的超参数。

[0128] 经过带有级联自注意力模块的OSNet骨干网后, 输出大小为 $b \times 512 \times 16 \times 8$ 的特征图, 其中 b 为批次大小, 512 为通道数, 16 为特征图高度, 8 为特征图宽度。

[0129] 将图像特征通过全连接层进行分类, 将图像特征映射到相应的类别标签上, 得到分类后的图像特征;

[0130] 大小为 $b \times N$ 的特征分数, N 为类别数目;

[0131] 使用分类后的图像特征计算带有标签平滑的身份损失, 通过反向传播更新梯度, 对预先建立的轻量行人重识别网络模型进行优化训练, 得到优化后的轻量行人重识别网络模型;

[0132] 使用分类后的图像特征计算带有标签平滑的身份损失 L_{ID} 的计算过程:

$$[0133] \quad L_{ID} = -\sum_{i=1}^N q_i \log(p_i) \begin{cases} q_i = 0, y \neq i \\ q_i = 1, y = i \end{cases}$$

[0134] 其中 $i \in \{1, 2, \dots, N\}$ 表示行人类别, N 表示训练集行人图像数量, y 为真实标签, p_i 为网络的预测逻辑值。

[0135] 标签平滑改变了 q_i 的结构:

$$[0136] \quad q_i = \begin{cases} q_i = 1 - \frac{N-1}{N} \varepsilon & \text{if } i = y \\ \varepsilon / N & \text{otherwise} \end{cases}$$

[0137] 其中 $\varepsilon = 0.1$, 鼓励模型不要太相信训练集, 增加泛化能力。

[0138] 获取行人重识别数据集, 将行人重识别数据集输入至优化后的轻量行人重识别网络模型内, 输出得到轻量行人重识别结果。

[0139] 在Market-1501和DukeMTMC-reID两大行人重识别数据集上进行测试。采用AMSGrad优化器, 初始学习率设为0.0015, 采用余弦退火策略更新学习率。批次大小和权重衰减分别为64和 $5e-4$ 。训练epoch设为250, 调整LAGT算法概率为0.6, SSAM和CSAM模块影响的超参数 λ 和 δ 为1, 标签平滑的ID Loss用于监督, 行人匹配使用余弦距离。

[0140] 具体的, 下面通过实施例对本发明方案作进一步阐述:

[0141] 实验环境参数如表1所示。

[0142] 表1

| Configuration | Parameter |
|-------------------|--------------------------------------|
| Operating System | Windows 11 |
| CPU | Inter i7-11700 |
| Memory | 32G |
| GPU | NVIDIA RTX3080 |
| Software Platform | Python3.8、 Pytorch1.12.0、cuda11.3 |

[0144] 在实验中,本发明选择ImageNet上预训练的OSNet作为Baseline,为了验证LAGT和SAM单独使用以及组合使用对网络模型的影响,在Market-1501和DukeMTMC-reID两个数据集上进行消融实验,结果如表2所示。由表2可知,LAGT和SAM对Baseline都有一定的提升。将两种模块同时加入Baseline,LAGT弱化颜色偏差对模型识别的负面影响,SAM使模型在空间和通道上提取更为细粒的特征,因此网络达到了最好的识别性能,在Market-1501数据集中,Rank-1和mAP分别达到了95.5%和87.7%,相较于Baseline提升了0.7%和1.0%,在DukeMTMC-reID数据集中,Rank-1和mAP分别达到了89.2%和77.2%,相较于Baseline提升了0.8%和0.5%。此外,我们的网络较为轻便,参数量仅为2.8M,为应用和部署提供了便利。

[0145] 表2

| 方法 | Market-1501 | | DukeMTMC-reID | | 参数量 |
|-------------------|-------------|------|---------------|------|------|
| | Rank-1 | mAP | Rank-1 | mAP | |
| Baseline | 94.8 | 86.7 | 88.4 | 76.7 | 2.2M |
| Baseline+LAGT | 95.1 | 87.5 | 88.8 | 77.0 | 2.2M |
| Baseline+SAM | 95.1 | 86.9 | 88.9 | 76.8 | 2.8M |
| Baseline+LAGT+SAM | 95.5 | 87.7 | 89.2 | 77.2 | 2.8M |

[0147] 为了验证本发明的有效性,在Market-1501,DukeMTMC-reID两个数据集上与近几年的先进行人重识别方法对比,其中方法有基于局部特征提取的PCB+RPP,基于注意力机制的HA-CNN,Manacs,AAANet,IANet。除此之外,我们还选取了一些其他性能较好的方法作比较,如BDB,BoT等,统计结果如表3所示。

[0148] 由表3可知,本发明的Rank-1指标在两个数据集上分别达到了95.5%、89.2%,mAP指标分别达到了87.7%、77.2%,本发明仅专注于全局特征的提取,相较于局部特征提取的PCB+RPP网络,在Market-1501上Rank-1和mAP分别提升了2.4%和6.7%,在DukeMTMC-reID上分别提升了6.3%和9.7%。相较于复杂的属性注意网络AAANet,在Market-1501上Rank-1和mAP分别提升了1.6%和4.3%,在DukeMTMC-reID上分别提升了1.5%和2.9%。相较于仅专注全局特征的强大基线BoT方法,本发明有参数量小的优势,在Market-1501上Rank-1和mAP分别提升了1.0%和1.8%,在DukeMTMC-reID上分别提升了2.8%和0.8%。

[0149] 表3

| 方法 | 骨干 | Market-1501 | | DukeMTMC-reID | |
|------------------|--------------|-------------|-------------|---------------|-------------|
| | | Rank-1 | mAP | Rank-1 | mAP |
| SVDNet | ResNet | 82.3 | 62.1 | 76.7 | 56.8 |
| HA-CNN | HA-CNN | 91.2 | 75.7 | 80.5 | 63.8 |
| Mancs | ResNet | 93.1 | 82.3 | 84.9 | 71.8 |
| PCB+RPP | ResNet | 93.1 | 81.0 | 82.9 | 68.5 |
| AAANet | ResNet | 93.9 | 83.4 | 87.7 | 74.3 |
| [0150] BDB | ResNet | 94.2 | 84.3 | 86.8 | 72.1 |
| IANet | ResNet | 94.4 | 83.1 | 87.1 | 73.4 |
| OSNet | OSNet | 94.8 | 84.9 | 88.6 | 73.5 |
| BoT | ResNet | 94.5 | 85.9 | 86.4 | 76.4 |
| DGNet | ResNet | 94.8 | 86.0 | 86.6 | 74.8 |
| Baseline | OSNet | 94.8 | 86.7 | 88.4 | 76.7 |
| Ours | OSNet | 95.5 | 87.7 | 89.2 | 77.2 |
| Ours (RK) | OSNet | 95.9 | 94.3 | 91.9 | 89.1 |

[0151] 表4展示了本发明网络与主流网络参数量对比结果,对照表格可以看出,本文网络在OSNet的基础上,加入LAGT和SAM模块仅在增加少量参数和计算时间的前提下,有效地提升了模型的识别性能。相比于ResNet网络,本发明模型参数精简,消耗的计算资源较少,缩短了模型的训练时间,使得模型能够更快地适应任务,提高了模型在实际应用中的效率。且与OSNet相比较,本发明在提高了行人的识别精度和泛化能力的同时,参数量仅增加了0.6M,说明设计的网络具有较好的轻量化。

[0152] 表4

| 网络 | 参数量 |
|-------------------|-------------|
| ResNet-50 | 23.5M |
| ResNet-101 | 42.5M |
| [0153] ResNet-152 | 58.1M |
| OSNet | 2.2M |
| Ours | 2.8M |

[0154] 数据集中包含了各种复杂且多变的颜色偏差,导致模型在应对颜色变化时的鲁棒性不足。为了解决这一问题,平衡颜色特征与其他关键判别特征之间的权重显得尤为重要。图6分别展示了在Market-1501数据集上用RGB图像与灰度图像训练模型的检索结果,从图中可以看出,当查询图片与图库图片中存在颜色偏差时会影响检索的结果,而忽略颜色信息后,样本的检索效果有了一定提升。

[0155] 基于此问题,通过随机颜色丢弃生成具有局部灰度图的训练图像,提高了模型对颜色偏差的鲁棒性。RCD策略中关键的一步是在图像中用其对应的局部灰度图像替换局部彩色图像,采用的灰度变换为:

[0156] $t(R, G, B) = 0.299R + 0.587G + 0.114B$

[0157] 其中, R 、 G 、 B 分别表示红色、绿色、蓝色通道。从式中可以看出, 为了体现人眼对不同颜色的敏感度, 红、绿、蓝三通道设置了不同的转换权重, 该转换策略更符合人眼视觉的直观映像, 但在基于深度学习网络的行人重识别任务中存在以下问题:

[0158] (1) 与人眼视觉对颜色的敏感度不同, 网络模型对三种颜色通道的信息具有相同的识别能力, 因此采用基于人眼视觉感知特征的颜色权重分配策略具有一定的局限性。

[0159] (2) 在行人数据集中, 根据人们的着装习惯统计, 穿着颜色为绿色的行人较少, 而绿色特征往往出现在图片的背景信息上, 如图7所示, 树木、草坪作为行人图像的绿色背景信息, 当绿色通道赋值最大的转变权重时, 会增加模型对背景信息的泛化能力而忽略对行人本身的泛化能力, 产生识别偏差。

[0160] 针对此问题, 进行了相应的实验验证, 如图8所示。从图中可以看出, 当图片背景颜色特征与行人颜色特征偏差大, 特别是绿色背景图, 采用原灰度变换策略时, 绿色通道较大的转变权重会提高模型对背景的泛化能力, 忽略了对行人本身的泛化能力, 进而引入偏差降低了识别效果。

[0161] 根据以上分析, 本发明将在RCD基础上设计更适合学习模型的颜色丢弃策略, 改变原有的灰度变换, 新的变换方法应该满足以下三点要求:

[0162] (1) 减少网络模型对三种颜色通道的差异性, 弱化颜色偏差对识别带来的负面影响, 同时避免产生额外偏差;

[0163] (2) 颜色转换后, 要保留原始图像中更稳定的结构与纹理特征, 减少失真。

[0164] (3) 改进后的灰度变换方法不改变学习策略, 能避免训练时的过拟合现象, 且节省计算资源, 达到轻量和有效的目的。

[0165] 基于上述中阐述的灰度变换原理的分析, 采用加权平均法, 通过对分配给红色、绿色和蓝色通道的变换权重进行细致的平衡, 对灰度变换策略进行优化。对于 $H \times W$ 行人图像大小, 聚合灰度变换 (Aggregated Grayscale Transformation, AGT) 公式如下所示:

[0166]
$$t_{AGT}(R, G, B) = \frac{1}{\varphi} [I_R(e, f) + I_G(e, f) + I_B(e, f)]$$

[0167] 其中, R 、 G 和 B 表示红色、绿色和蓝色通道, $e \in \{1, \dots, H\}$, $f \in \{1, \dots, W\}$ 。 $I_R(\cdot)$ 、 $I_G(\cdot)$ 和 $I_B(\cdot)$ 分别表示红色通道、绿色通道和蓝色通道的特定位置处的像素值。权重系数表示为 φ , 是一个恒定的平衡因子, 选择 $\varphi = 3$ 。在上述公式中, 在从RGB图像到灰度图像的转换过程中, 将均匀的权重分配给三个通道。

[0168] 为了初步验证改进的有效性, 展示了在Market-1501上使用RGB图像与不同灰度变换生成的灰度图像训练模型的检索结果, 如图9所示, 其中 RGB 、 $t(R, G, B)$ 、 $t_{AGT}(R, G, B)$ 分别为输入为RGB图像时、采用原灰度变换后、采用本文的平均灰度变换后的检索结果。结果表明, AGT在消除颜色偏差的同时没有引入额外偏差, 且保留了RGB图像稳定的结构纹理特征, 识别效果优于原灰度变换。

[0169] 实施例二: 第二方面, 如图12所示, 为了达到上述目的, 本发明公开了基于随机颜色丢弃和注意力的轻量行人重识别系统, 包括:

[0170] 图像处理模块11,用于接收图像数据,对图像数据进行预处理,得到预处理图像数据;

[0171] 特征提取模块12,用于将预处理图像数据输入至预先建立的嵌入有级联自注意力模块的OSNet内,提取特征,输出得到图像特征;

[0172] 图像分类模块13,用于将图像特征通过全连接层进行分类,将图像特征映射到相应的类别标签上,得到分类后的图像特征;

[0173] 模型训练模块14,用于使用分类后的图像特征计算带有标签平滑的身份损失,通过反向传播更新梯度,对预先建立的轻量行人重识别网络模型进行优化训练,得到优化后的轻量行人重识别网络模型;

[0174] 行人重识别模块15,用于获取行人重识别数据集的测试集,将行人重识别数据集的测试集输入至优化后的轻量行人重识别网络模型内,输出得到轻量行人重识别结果。

[0175] 基于同一种发明构思,本发明还提供一种计算机设备,该计算机设备包括:一个或多个处理器,以及存储器,用于存储一个或多个计算机程序;程序包括程序指令,处理器用于执行存储器存储的程序指令。处理器可能是中央处理单元(Central Processing Unit, CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor、DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列(Field-Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等,其是终端的计算核心以及控制核心,其用于实现一条或一条以上指令,具体用于加载并执行计算机存储介质内一条或一条以上指令从而实现上述方法。

[0176] 需要进一步进行说明的是,基于同一种发明构思,本发明还提供一种计算机存储介质,该存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行上述方法。该存储介质可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是但不限于电、磁、光、电、磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本发明中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0177] 在本说明书的描述中,参考术语“一个实施例”、“示例”、“具体示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本公开的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0178] 以上显示和描述了本公开的基本原理、主要特征和本公开的优点。本行业的技术人员应该了解,本公开不受上述实施例的限制,上述实施例和说明书中描述的只是说明本公开的原理,在不脱离本公开精神和范围的前提下,本公开还会有各种变化和改进,这些变化和改进都落入要求保护的本公开范围内容。

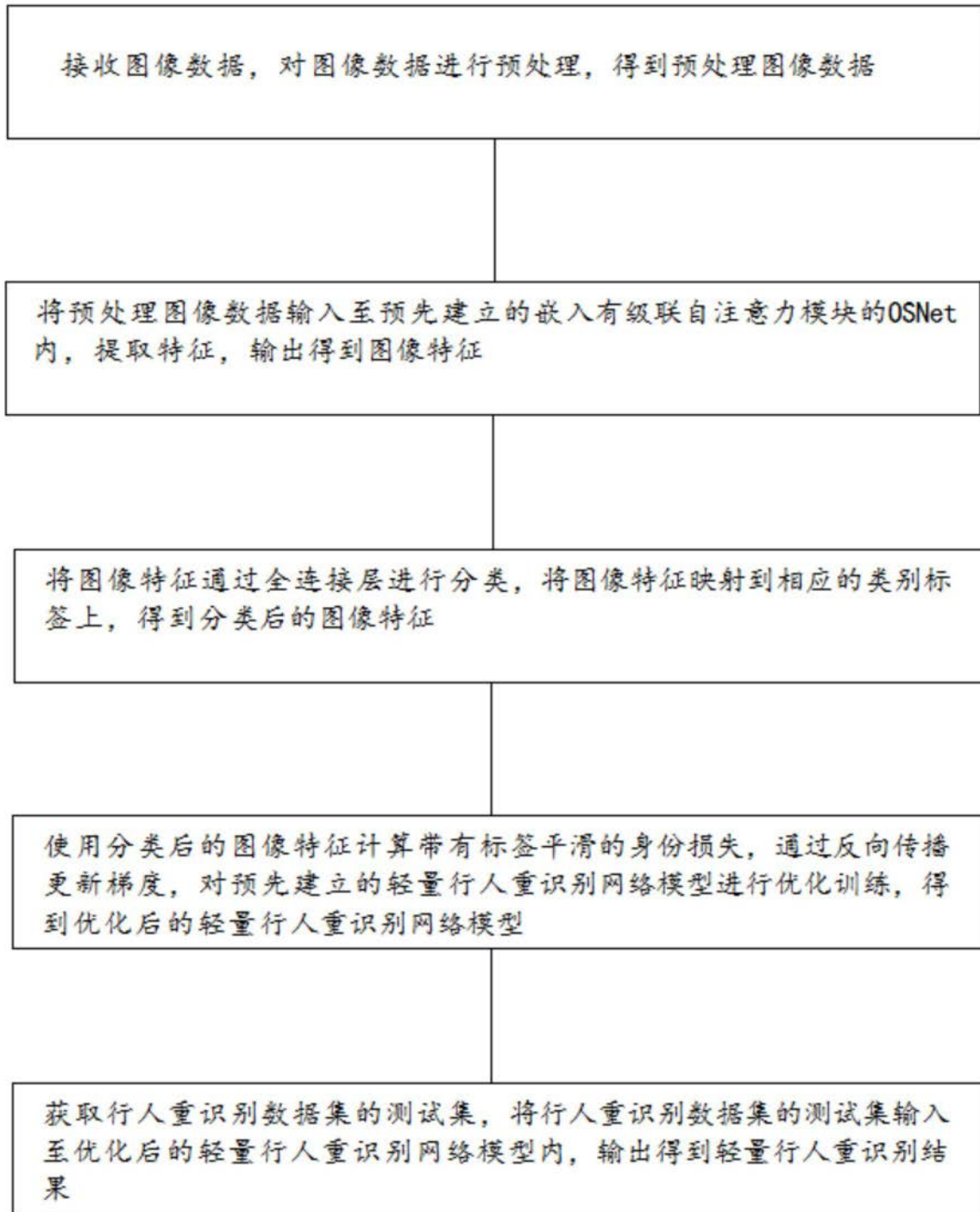


图1

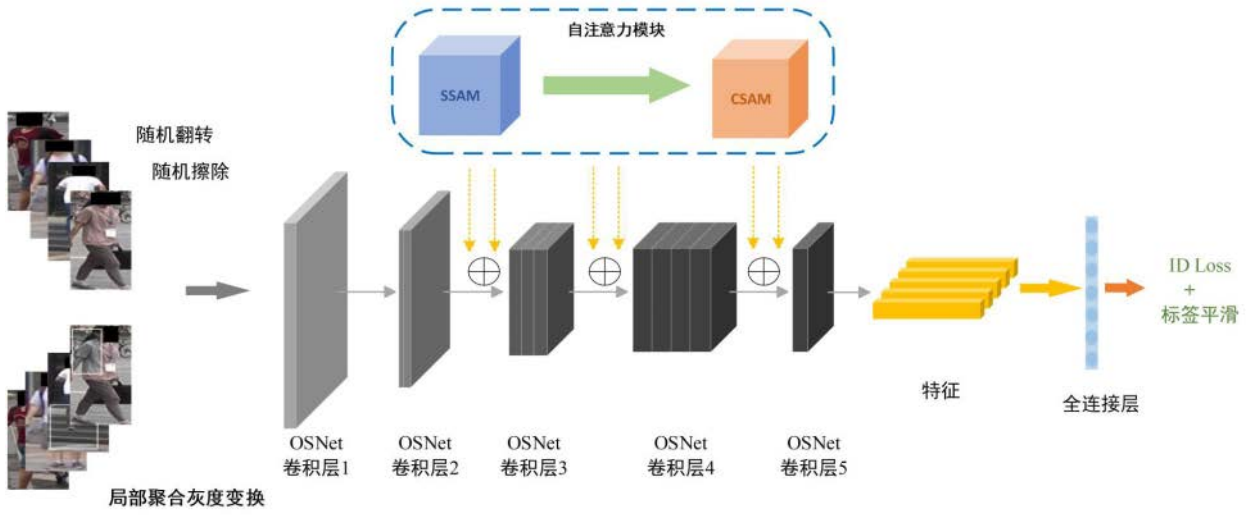


图2



图3

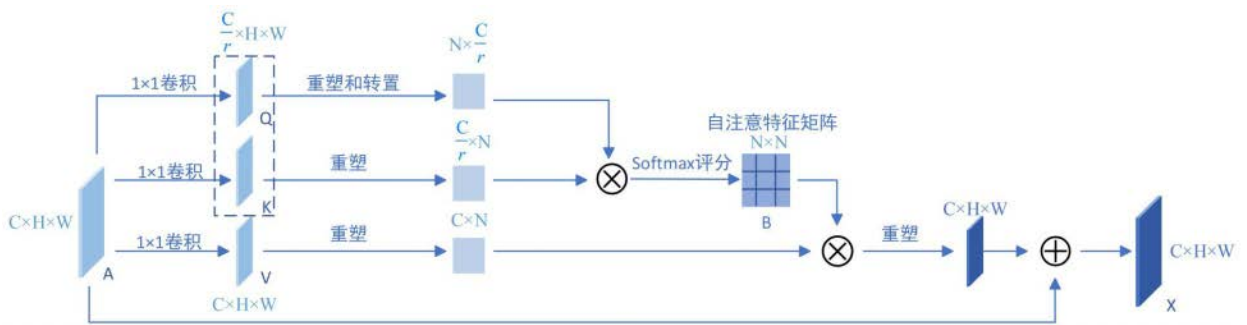


图4

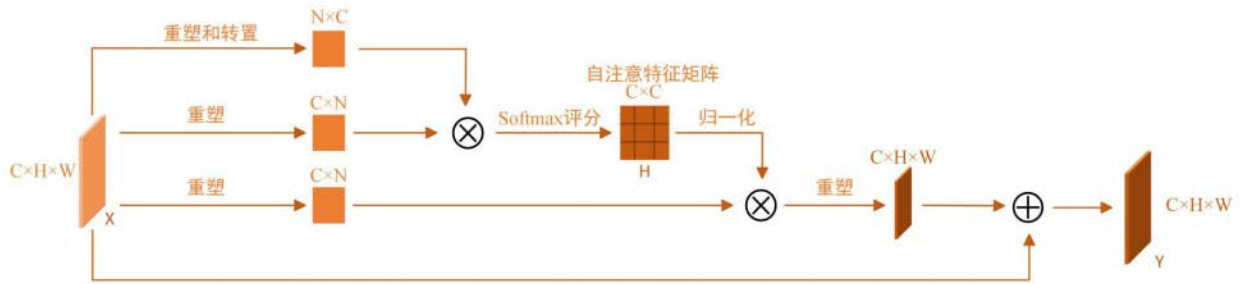


图5

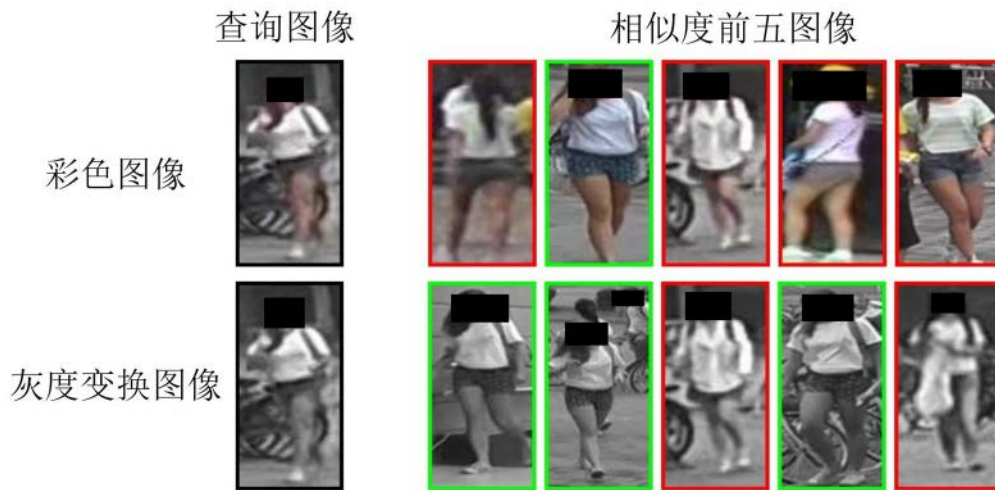


图6



图7

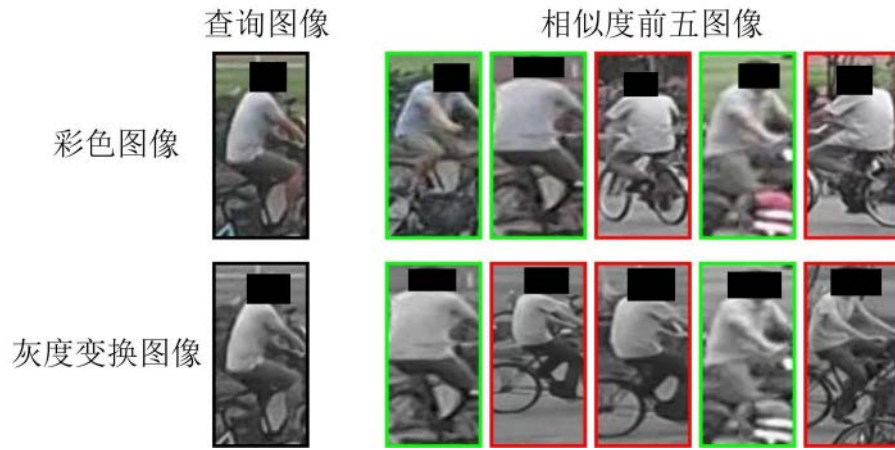


图8

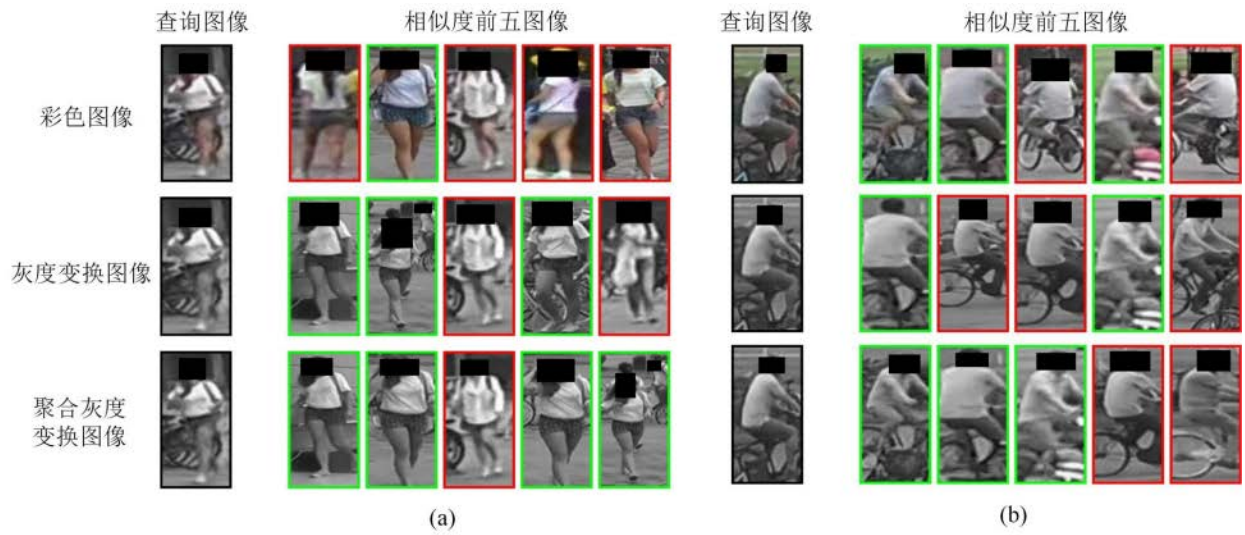


图9

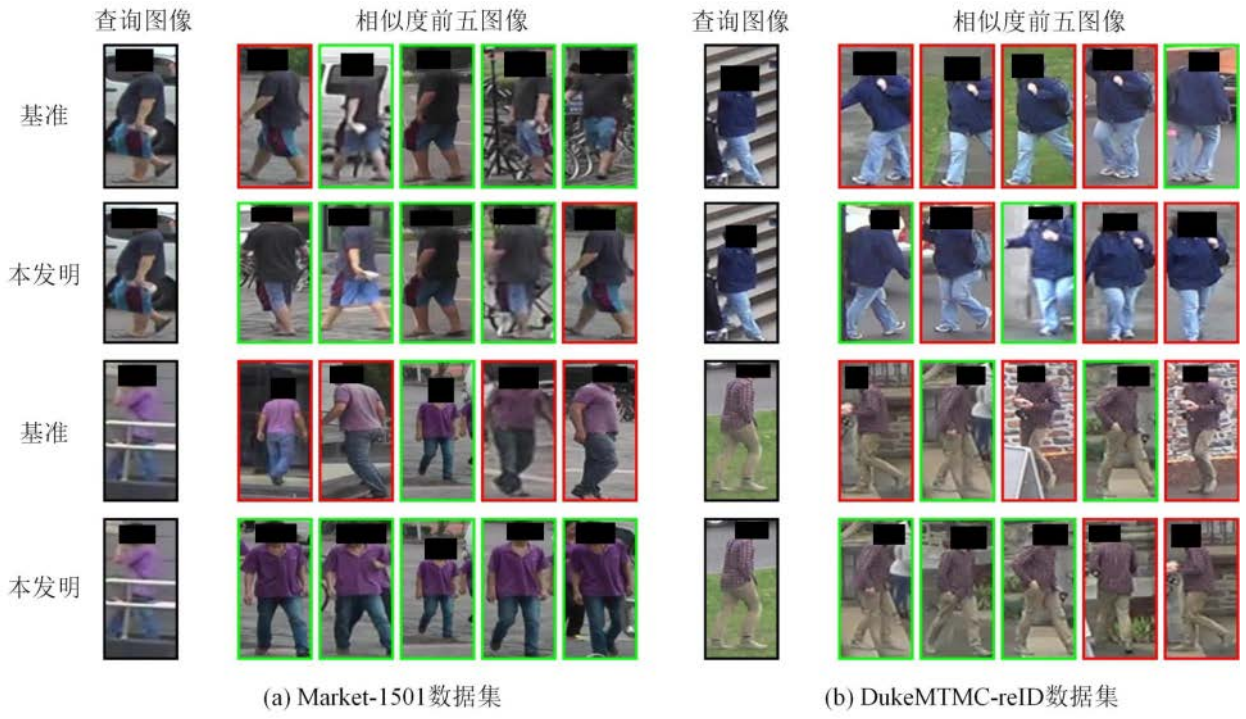


图10



图11

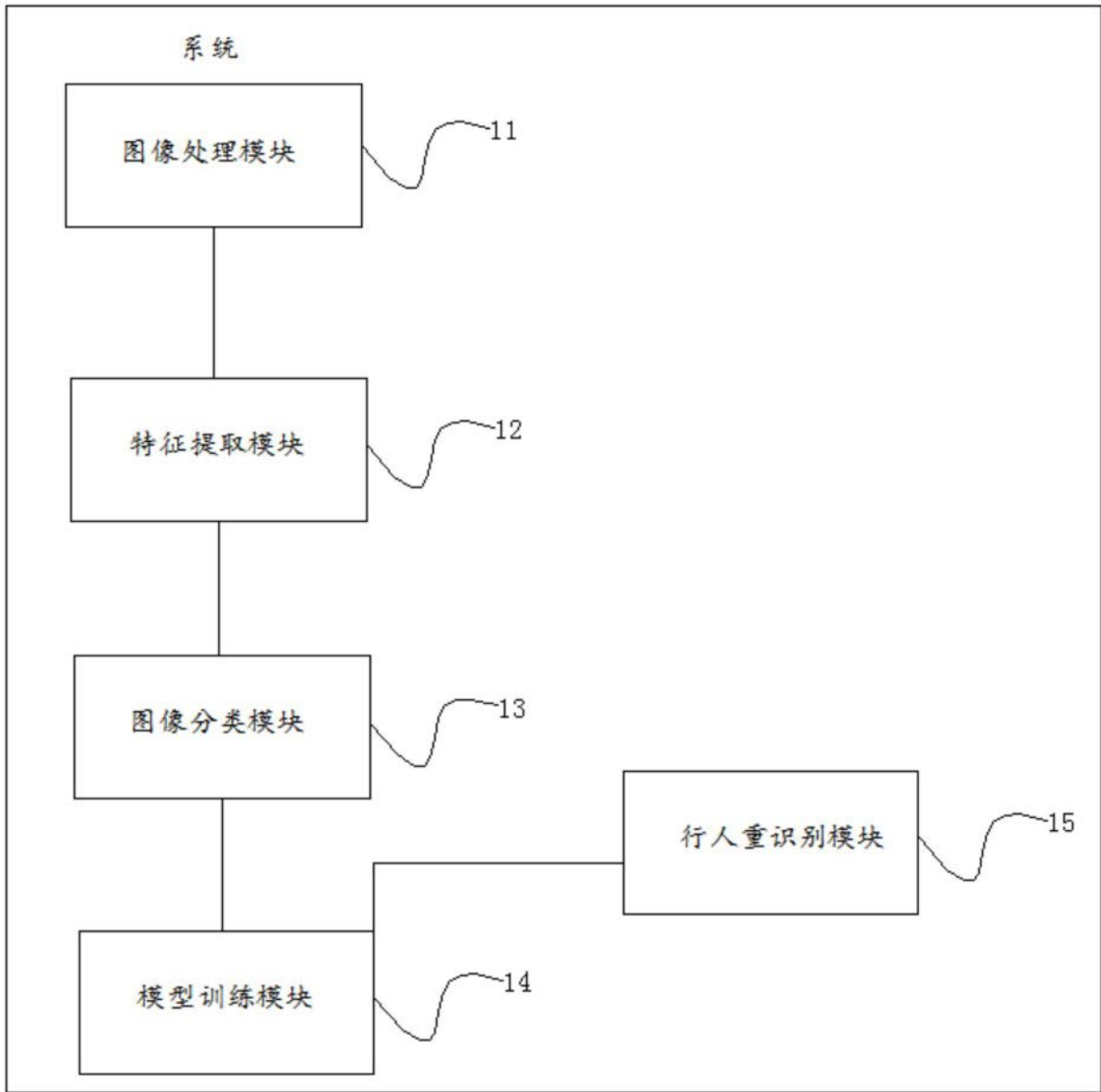


图12