



(12) 发明专利申请

(10) 申请公布号 CN 112784927 A

(43) 申请公布日 2021.05.11

(21) 申请号 202110177362.8

(22) 申请日 2021.02.09

(71) 申请人 中国人民解放军国防科技大学  
地址 410073 湖南省长沙市开福区德雅路  
109号

(72) 发明人 傅瑞罡 高颖慧 董小虎 李飏  
朱永锋

(74) 专利代理机构 国防科技大学专利服务中心  
43202

代理人 王文惠

(51) Int. Cl.

G06K 9/62 (2006.01)

G06K 9/46 (2006.01)

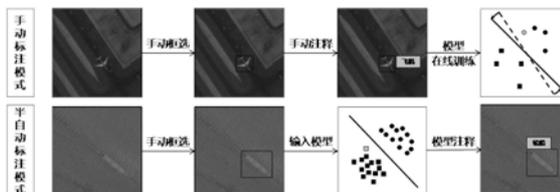
权利要求书3页 说明书6页 附图2页

(54) 发明名称

一种基于在线学习的半自动图像标注方法

(57) 摘要

本发明属于数据工程领域,具体公开了一种基于在线学习的半自动图像标注方法,用于为基于学习的图像目标检测方法制备训练数据。本方法针对目标检测领域中的手工制备训练数据耗时问题,通过边标注边学习的方式,提取并利用手动标注过程中存在的监督信息,提高了图像标注的自动化程度,提升了数据集制备的效率。



1. 一种基于在线学习的半自动图像标注方法,其特征在于,包括以下步骤:

(1) 确定待标注图像集中存在的目标类别数,初始化一个多分类的逻辑回归分类器;

所述步骤(1)通过以下步骤实现:

1) 利用人机交互界面,确定数据集中的目标类别数 $n$ ;

2) 初始化逻辑回归分类器:逻辑回归分类器的初始化即建立一个全零的 $m \times n$ 大小的参数矩阵 $w$ ,其中, $m$ 为预设参数,与目标HOG特征的维度相匹配, $n$ 为目标类别数;

(2) 输入待标注图像,执行手动标注模式:通过手动框选目标位置、手动注释目标类别,完成对图像中所有目标的标注,并通过这些已标注数据在线训练分类器;

所述步骤(2)通过以下步骤实现:

1) 利用人机交互界面,手动框选目标位置,手动注释目标类别;

2) 提取 $t$ 时刻目标样本(记 $t$ 时刻为标注第 $t$ 个目标样本的时刻)的HOG特征 $x$ , $x$ 为一个 $m$ 维的矢量,即 $x \in \mathbb{R}^{m \times 1}$ ;

3) 获取 $t$ 时刻目标样本的真实标签 $y$ ,并计算分类器的交叉熵损失:

$$L = -\log \left( \frac{e^{w_{t,y}^T x}}{\sum_{i=1}^n e^{w_{t,i}^T x}} \right) \quad (1)$$

其中, $w_{t,i}$ 、 $w_{t,y}$ 分别为 $t$ 时刻分类器参数矩阵 $w$ 中的第 $i$ 列和第 $y$ 列矢量;

4) 使用Ftrl方法在线训练分类器模型,以使得分类器的损失变小;记 $w_{t,i,j}$ 为 $t$ 时刻分类器参数矩阵 $w$ 第 $i$ 列第 $j$ 行元素的值, $g_{t,i,j}$ 为分类器交叉熵损失 $L$ 关于参数 $w_{t,i,j}$ 的梯度值,Ftrl方法的参数更新公式为:

$$w_{t+1,i,j} = \arg \min_w \left( \left( \sum_{s=1}^t g_{s,i,j} \right) w + \frac{1}{2} \sum_{s=1}^t \sigma_{s,i,j} \|w - w_{s,i,j}\|_2^2 + \lambda_1 \|w\|_1 + \frac{1}{2} \lambda_2 \|w\|_2^2 \right) \quad (2)$$

其中,待优化函数中的第一项 $\left( \sum_{s=1}^t g_{s,i,j} \right) w$ 与分类器损失关于参数的梯度有关;第二项

$\frac{1}{2} \sum_{s=1}^t \sigma_{s,i,j} \|w - w_{s,i,j}\|_2^2$ 为参数的约束项,用于避免 $t$ 时刻的迭代结果与从前结果产生过于严

重的偏差,稳定损失的收敛;第三项 $\lambda_1 \|w\|_1$ 、第四项 $\frac{1}{2} \lambda_2 \|w\|_2^2$ 分别为参数的 $L_1$ 、 $L_2$ 正则项,

用于保证参数的稀疏性; $\sigma_{s,i,j}$ 、 $\lambda_1$ 、 $\lambda_2$ 为损失权重;

若已知 $w_{t,i,j}$ ,则 $t+1$ 时刻的 $w_{t+1,i,j}$ 求解过程如下:

1. 计算损失 $L$ 关于参数 $w_{t,i,j}$ 的梯度 $g_{t,i,j}$ ;

$$g_{t,i,j} = \frac{\partial L}{\partial w_{t,i,j}} \quad (3)$$

2. 计算中间变量 $\sigma_{t,i,j}$ :

$$\sigma_{t,i,j} = \frac{1}{\alpha} \left( \sqrt{\sum_{s=1}^t g_{s,i,j}^2} - \sqrt{\sum_{s=1}^{t-1} g_{s,i,j}^2} \right) \quad (4)$$

其中, $\alpha$ 为人为设定的超参数;

3. 计算中间变量 $z_{t,i,j}$ :

$$z_{t,i,j} = \sum_{s=1}^t g_{s,i,j} - \sum_{s=1}^t \sigma_{s,i,j} w_{s,i,j} \quad (5)$$

4. 计算中间变量 $\eta_{t,i,j}$ :

$$\eta_{t,i,j} = \frac{\beta + \sqrt{\sum_{s=1}^t g_{s,i,j}^2}}{\alpha} \quad (6)$$

其中, $\beta$ 为人为设定的超参数;

5. 更新 $t+1$ 时刻的参数值 $w_{t+1,i,j}$ :

$$w_{t+1,i,j} = \begin{cases} 0 & \text{if } |z_{t,i,j}| \leq \lambda_1 \\ -\frac{1}{\eta_{t,i,j} + \lambda_2} (z_{t,i,j} - \text{sgn}(z_{t,i,j}) \lambda_1) & \text{otherwise} \end{cases} \quad (7)$$

其中, $\text{sgn}()$ 表示符号函数, $\lambda_1, \lambda_2$ 为人为设定的超参数;

(3) 测试分类器性能,并基于分类器性能决定是否由手动标注模式切换至半自动标注模式;

所述步骤(3)通过以下步骤实现:

从所有已标注的目标样本中随机采样 $d$ 个样本组成验证集,记为 $\{x_1, x_2, \dots, x_d\}$ ,对应真实标签为 $\{y_1, y_2, \dots, y_d\}$ ;使用分类器对验证集中的所有目标做类别预测,对于任意样本 $x_k, k \in \{1, 2, \dots, d\}$ ,其类别预测结果为:

$$\tilde{y}_k = \arg \max_i (\mathbf{w}_{t,i}^T \mathbf{x}_k) \quad (8)$$

其中, $w_{t,i}$ 为 $t$ 时刻 $w$ 中的第 $i$ 列矢量;

对比预测类别以及真实标签,分类器在验证集上的分类正确个数为:

$$\tilde{d} = \sum_{k=1}^d c_k, \quad c_k = \begin{cases} 1 & \tilde{y}_k = y_k \\ 0 & \tilde{y}_k \neq y_k \end{cases} \quad (9)$$

其中, $y_k$ 是样本 $x_k$ 对应的真实标签;分类器的分类准确率最终确定为 $\tilde{d}/d$ ,准确率越高,说明分类器的性能越好;当分类器在验证集上达到设定的分类准确率后,执行半自动标注模式以提升标注的速度;

(4) 输入待标注图像,执行半自动标注模式:手动框选目标位置,通过分类器自动完成对目标类别的注释;

所述步骤(4)通过以下步骤实现:

1) 利用人机交互界面,手动框选目标位置;

2) 提取目标的HOG特征 $x$ ;

3) 使用分类器对目标的类别做预测:将目标特征 $x$ 输入分类器,目标的类别置信度分布如下:

$$\mathbf{p} = \frac{e^{\mathbf{w}^T \mathbf{x}}}{\sum_{i=1}^n e^{\mathbf{w}_i^T \mathbf{x}}} \quad (10)$$

其中, $n$ 为目标类别数, $\mathbf{p}$ 是 $n$ 维矢量,它的第 $i$ 个元素 $p_i$ 对应目标属于第 $i$ 个类别的置信

度,满足 $0 < p_i < 1$ ,且 $\sum_{i=1}^n p_i = 1$ ;特征 $x$ 的类别最终判决为:

$$\tilde{y} = \arg \max_i (p_i) \quad (11)$$

$\tilde{y}$ 即为目标的注释结果,这是由分类器自动生成的。

## 一种基于在线学习的半自动图像标注方法

### 技术领域

[0001] 本发明涉及一种基于在线学习的半自动图像标注方法,属于数据工程领域,用于为基于学习的图像目标检测方法制备训练数据。

### 背景技术

[0002] 近年来,以卷积神经网络为代表的深度学习技术凭借其强大的特征学习能力,在图像目标检测领域的应用越来越广泛。训练数据的制备是特征学习的必要条件。检测领域的标注数据包括框选和注释两个步骤。框选指选定目标,并标出目标的外界矩形框;注释指提供目标的类别信息。

[0003] 目前,训练数据的制备通常依靠人力标注,配合一些交互式的辅助标注工具可以在一定程度上减轻了标注者的负担。这些标注工具,如labelme、labelimg等,提供了在图片上画多边形边界、注释物体类别等基本功能。但尽管如此,整体的标注效率仍然非常低下,每个目标实例的标注需要花费数十秒。

[0004] 为了节省标注的人力成本,一些研究探索了更简单的标注方法,例如基于眼动的标注方法、基于点的标注方法以及基于框验证的标注方法。基于眼动的标注方法需要配备眼动仪,价格昂贵,适普性不强。基于点的标注方法和基于框验证的标注方法在鲁棒性上稍差,即当目标所处环境较复杂时,方法性能不可靠。因此,研究一种鲁棒性强、自动化程度高的标注方法很有必要。

### 发明内容

[0005] 本发明要解决的技术问题是:针对目标检测领域中的手工制备训练数据耗时问题,通过边标注边学习的方式,提取并利用手动标注过程中存在的监督信息,提高图像标注的自动化程度,提升数据集制备的效率。

[0006] 为解决上述技术问题,本发明提出一种基于在线学习的半自动图像标注方法,所述方法包括四个步骤:

[0007] (1) 确定待标注图像集中存在的目标类别数,初始化一个多分类的逻辑回归分类器;

[0008] (2) 输入待标注图像,执行手动标注模式:通过手动框选目标位置、手动注释目标类别,完成对图像中所有目标的标注,并通过这些已标注数据在线训练分类器;

[0009] (3) 测试分类器性能,并基于分类器性能决定是否由手动标注模式切换至半自动标注模式;

[0010] (4) 输入待标注图像,执行半自动标注模式:手动框选目标位置,通过分类器自动完成对目标类别的注释。

[0011] 所述步骤(1)通过以下步骤实现:

[0012] 1) 利用人机交互界面,确定数据集中的目标类别数 $n$ ;

[0013] 2) 初始化逻辑回归分类器:逻辑回归分类器的初始化即建立一个全零的 $m \times n$ 大小

的参数矩阵 $w$ ,其中, $m$ 为预设参数,与目标HOG特征的维度相匹配, $n$ 为目标类别数。

[0014] 所述步骤(2)通过以下步骤实现:

[0015] 1) 利用人机交互界面,手动框选目标位置,手动注释目标类别;

[0016] 2) 提取 $t$ 时刻目标样本(记 $t$ 时刻为标注第 $t$ 个目标样本的时刻)的HOG特征 $x$ , $x$ 为一个 $m$ 维的矢量,即 $x \in \mathbb{R}^{m \times 1}$ ;

[0017] 3) 获取 $t$ 时刻目标样本的真实标签 $y$ ,并计算分类器的交叉熵损失:

$$[0018] \quad L = -\log \left( \frac{e^{w_{t,y}^T x}}{\sum_{i=1}^n e^{w_{t,i}^T x}} \right) \quad (1)$$

[0019] 其中, $w_{t,i}$ 、 $w_{t,y}$ 分别为 $t$ 时刻分类器参数矩阵 $w$ 中的第 $i$ 列和第 $y$ 列矢量;

[0020] 4) 使用Ftrl方法在线训练分类器模型,以使得分类器的损失变小;记 $w_{t,i,j}$ 为 $t$ 时刻分类器参数矩阵 $w$ 第 $i$ 列第 $j$ 行元素的值, $g_{t,i,j}$ 为分类器交叉熵损失 $L$ 关于参数 $w_{t,i,j}$ 的梯度值,Ftrl方法的参数更新公式为:

$$[0021] \quad w_{t+1,i,j} = \arg \min_w \left( \left( \sum_{s=1}^t g_{s,i,j} \right) w + \frac{1}{2} \sum_{s=1}^t \sigma_{s,i,j} \|w - w_{s,i,j}\|_2^2 + \lambda_1 \|w\|_1 + \frac{1}{2} \lambda_2 \|w\|_2^2 \right) \quad (2)$$

[0022] 其中,待优化函数中的第一项 $\left( \sum_{s=1}^t g_{s,i,j} \right) w$ 与分类器损失关于参数的梯度有关;第二

项 $\frac{1}{2} \sum_{s=1}^t \sigma_{s,i,j} \|w - w_{s,i,j}\|_2^2$ 为参数的约束项,用于避免 $t$ 时刻的迭代结果与从前结果产生过于

严重的偏差,稳定损失的收敛;第三项 $\lambda_1 \|w\|_1$ 、第四项 $\frac{1}{2} \lambda_2 \|w\|_2^2$ 分别为参数的 $L_1$ 、 $L_2$ 正则

项,用于保证参数的稀疏性; $\sigma_{s,i,j}$ 、 $\lambda_1$ 、 $\lambda_2$ 为损失权重;

[0023] 若已知 $w_{t,i,j}$ ,则 $t+1$ 时刻的 $w_{t+1,i,j}$ 求解过程如下:

[0024] 1. 计算损失 $L$ 关于参数 $w_{t,i,j}$ 的梯度 $g_{t,i,j}$ ;

$$[0025] \quad g_{t,i,j} = \frac{\partial L}{\partial w_{t,i,j}} \quad (3)$$

[0026] 2. 计算中间变量 $\sigma_{t,i,j}$ :

$$[0027] \quad \sigma_{t,i,j} = \frac{1}{\alpha} \left( \sqrt{\sum_{s=1}^t g_{s,i,j}^2} - \sqrt{\sum_{s=1}^{t-1} g_{s,i,j}^2} \right) \quad (4)$$

[0028] 其中, $\alpha$ 为人为设定的超参数;

[0029] 3. 计算中间变量 $z_{t,i,j}$ :

$$[0030] \quad z_{t,i,j} = \sum_{s=1}^t g_{s,i,j} - \sum_{s=1}^t \sigma_{s,i,j} w_{s,i,j} \quad (5)$$

[0031] 4. 计算中间变量 $\eta_{t,i,j}$ :

$$[0032] \quad \eta_{t,i,j} = \frac{\beta + \sqrt{\sum_{s=1}^t g_{s,i,j}^2}}{\alpha} \quad (6)$$

[0033] 其中, $\beta$ 为人为设定的超参数;

[0034] 5.更新 $t+1$ 时刻的参数值 $w_{t+1,i,j}$ :

$$[0035] \quad w_{t+1,i,j} = \begin{cases} 0 & \text{if } |z_{t,i,j}| \leq \lambda_1 \\ -\frac{1}{\eta_{t,i,j} + \lambda_2} (z_{t,i,j} - \text{sgn}(z_{t,i,j}) \lambda_1) & \text{otherwise} \end{cases} \quad (7)$$

[0036] 其中, $\text{sgn}()$ 表示符号函数, $\lambda_1, \lambda_2$ 为人为设定的超参数。

[0037] 所述步骤(3)通过以下步骤实现:

[0038] 从所有已标注的目标样本中随机采样 $d$ 个样本组成验证集,记为 $\{x_1, x_2, \dots, x_d\}$ ,对应真实标签为 $\{y_1, y_2, \dots, y_d\}$ ;使用分类器对验证集中的所有目标做类别预测,对于任意样本 $x_k$ ,  $k \in \{1, 2, \dots, d\}$ ,其类别预测结果为:

$$[0039] \quad \tilde{y}_k = \arg \max_i (\mathbf{w}_{t,i}^T \mathbf{x}_k) \quad (8)$$

[0040] 其中, $w_{t,i}$ 为 $t$ 时刻 $w$ 中的第 $i$ 列矢量;

[0041] 对比预测类别以及真实标签,分类器在验证集上的分类正确个数为:

$$[0042] \quad \tilde{d} = \sum_{k=1}^d c_k, \quad c_k = \begin{cases} 1 & \tilde{y}_k = y_k \\ 0 & \tilde{y}_k \neq y_k \end{cases} \quad (9)$$

[0043] 其中, $y_k$ 是样本 $x_k$ 对应的真实标签;分类器的分类准确率最终确定为 $\tilde{d}/d$ ,准确率越高,说明分类器的性能越好;当分类器在验证集上达到设定的分类准确率后,执行半自动标注模式以提升标注的速度。

[0044] 所述步骤(4)通过以下步骤实现:

[0045] 1) 利用人机交互界面,手动框选目标位置;

[0046] 2) 提取目标的HOG特征 $x$ ;

[0047] 3) 使用分类器对目标的类别做预测:将目标特征 $x$ 输入分类器,目标的类别置信度分布如下:

$$[0048] \quad \mathbf{p} = \frac{e^{\mathbf{w}^T \mathbf{x}}}{\sum_{i=1}^n e^{\mathbf{w}_i^T \mathbf{x}}} \quad (10)$$

[0049] 其中, $n$ 为目标类别数, $\mathbf{p}$ 是 $n$ 维矢量,它的第 $i$ 个元素 $p_i$ 对应目标属于第 $i$ 个类别的置信度,满足 $0 < p_i < 1$ ,且 $\sum_{i=1}^n p_i = 1$ ;特征 $x$ 的类别最终判决为:

$$[0050] \quad \tilde{y} = \arg \max_i (p_i) \quad (11)$$

[0051]  $\tilde{y}$ 即为目标的注释结果,这是由分类器自动生成的。

[0052] 本发明的有益效果:

[0053] 与现有技术相比,本发明的优点在于通过将注释过程抽象为学习分类问题,实现了标注过程的部分自动化,提升了数据标注的效率。目标类别数越多,效率提升越明显。

## 附图说明

[0054] 图1为本发明两种标注模式的示意图;

[0055] 图2为基于本发明方法的图像标注示意图；

[0056] 图3为在线学习下的模型准确率收敛曲线。

### 具体实施方式

[0057] 为将本发明的目的、内容、优点阐述地更加清晰，以下将结合说明书附图对本发明作进一步详细说明。本发明涉及一种基于在线学习的半自动图像标注方法，包括如下步骤：

[0058] (1) 确定待标注图像集中存在的目标类别数，初始化一个多分类的逻辑回归分类器；

[0059] (2) 输入待标注图像，执行手动标注模式：通过手动框选目标位置、手动注释目标类别，完成对图像中所有目标的标注，并通过这些已标注数据在线训练分类器；

[0060] (3) 测试分类器性能，并基于分类器性能决定是否由手动标注模式切换至半自动标注模式；

[0061] (4) 输入待标注图像，执行半自动标注模式：手动框选目标位置，通过分类器自动完成对目标类别的注释。

[0062] 所述步骤(1)通过以下步骤实现：

[0063] 1) 利用人机交互界面，确定数据集中的目标类别数 $n$ ；

[0064] 2) 初始化逻辑回归分类器：逻辑回归分类器的初始化即建立一个全零的 $m \times n$ 大小的参数矩阵 $w$ ，其中， $m$ 为预设参数，与目标HOG特征的维度相匹配， $n$ 为目标类别数。

[0065] 所述步骤(2)执行手动标注模式，其流程如图1第一行所示。具体地，它通过以下步骤实现：

[0066] 1) 利用人机交互界面，手动框选目标位置，手动注释目标类别，如图2所示；

[0067] 2) 提取 $t$ 时刻目标样本（记 $t$ 时刻为标注第 $t$ 个目标样本的时刻）的HOG特征 $x$ ， $x$ 为一个 $m$ 维的矢量，即 $x \in \mathbb{R}^{m \times 1}$ ；

[0068] 3) 获取 $t$ 时刻目标样本的真实标签 $y$ ，并计算分类器的交叉熵损失：

$$[0069] \quad L = -\log \left( \frac{e^{w_{t,y}^T x}}{\sum_{i=1}^n e^{w_{t,i}^T x}} \right) \quad (12)$$

[0070] 其中， $w_{t,i}$ 、 $w_{t,y}$ 分别为 $t$ 时刻分类器参数矩阵 $w$ 中的第 $i$ 列和第 $y$ 列矢量。

[0071] 4) 使用Ftrl方法在线训练分类器模型，以使得分类器的损失变小；记 $w_{t,i,j}$ 为 $t$ 时刻分类器参数矩阵 $w$ 第 $i$ 列第 $j$ 行元素的值， $g_{t,i,j}$ 为分类器交叉熵损失 $L$ 关于参数 $w_{t,i,j}$ 的梯度值，Ftrl方法的参数更新公式为：

$$[0072] \quad w_{t+1,i,j} = \arg \min_w \left( \left( \sum_{s=1}^t g_{s,i,j} \right) w + \frac{1}{2} \sum_{s=1}^t \sigma_{s,i,j} \|w - w_{s,i,j}\|_2^2 + \lambda_1 \|w\|_1 + \frac{1}{2} \lambda_2 \|w\|_2^2 \right) \quad (13)$$

[0073] 其中，待优化函数中的第一项 $\left( \sum_{s=1}^t g_{s,i,j} \right) w$ 与分类器损失关于参数的梯度有关；第二

项 $\frac{1}{2} \sum_{s=1}^t \sigma_{s,i,j} \|w - w_{s,i,j}\|_2^2$ 为参数的约束项，用于避免 $t$ 时刻的迭代结果与从前结果产生过于

严重的偏差,稳定损失的收敛;第三项 $\lambda_1 \|w\|_1$ 、第四项 $\frac{1}{2}\lambda_2 \|w\|_2^2$ 分别为参数的 $L_1$ 、 $L_2$ 正则项,用于保证参数的稀疏性; $\sigma_{s,i,j}$ 、 $\lambda_1$ 、 $\lambda_2$ 为损失权重;

[0074] 若已知 $w_{t,i,j}$ ,则 $t+1$ 时刻的 $w_{t+1,i,j}$ 求解过程如下:

[0075] 1. 计算损失 $L$ 关于参数 $w_{t,i,j}$ 的梯度 $g_{t,i,j}$ ;

$$[0076] \quad g_{t,i,j} = \frac{\partial L}{\partial w_{t,i,j}} \quad (14)$$

[0077] 2. 计算中间变量 $\sigma_{t,i,j}$ :

$$[0078] \quad \sigma_{t,i,j} = \frac{1}{\alpha} \left( \sqrt{\sum_{s=1}^t g_{s,i,j}^2} - \sqrt{\sum_{s=1}^{t-1} g_{s,i,j}^2} \right) \quad (15)$$

[0079] 其中, $\alpha$ 为人为设定的超参数;

[0080] 3. 计算中间变量 $z_{t,i,j}$ :

$$[0081] \quad z_{t,i,j} = \sum_{s=1}^t g_{s,i,j} - \sum_{s=1}^t \sigma_{s,i,j} w_{s,i,j} \quad (16)$$

[0082] 4. 计算中间变量 $\eta_{t,i,j}$ :

$$[0083] \quad \eta_{t,i,j} = \frac{\beta + \sqrt{\sum_{s=1}^t g_{s,i,j}^2}}{\alpha} \quad (17)$$

[0084] 其中, $\beta$ 为人为设定的超参数;

[0085] 5. 更新 $t+1$ 时刻的参数值 $w_{t+1,i,j}$ :

$$[0086] \quad w_{t+1,i,j} = \begin{cases} 0 & \text{if } |z_{t,i,j}| \leq \lambda_1 \\ -\frac{1}{\eta_{t,i,j} + \lambda_2} (z_{t,i,j} - \text{sgn}(z_{t,i,j}) \lambda_1) & \text{otherwise} \end{cases} \quad (18)$$

[0087] 其中, $\text{sgn}()$ 表示符号函数, $\lambda_1$ 、 $\lambda_2$ 为人为设定的超参数。

[0088] 所述步骤(3)通过以下步骤实现:

[0089] 从所有已标注的目标样本中随机采样 $d$ 个样本组成验证集,记为 $\{x_1, x_2, \dots, x_d\}$ ,对应真实标签为 $\{y_1, y_2, \dots, y_d\}$ ;使用分类器对验证集中的所有目标做类别预测,对于任意样本 $x_k$ ,  $k \in \{1, 2, \dots, d\}$ ,其类别预测结果为:

$$[0090] \quad \tilde{y}_k = \arg \max_i (\mathbf{w}_{t,i}^T \mathbf{x}_k) \quad (19)$$

[0091] 其中, $w_{t,i}$ 为 $t$ 时刻 $w$ 中的第 $i$ 列矢量。

[0092] 对比预测类别以及真实标签,分类器在验证集上的分类正确个数为:

$$[0093] \quad \tilde{d} = \sum_{k=1}^d c_k, \quad c_k = \begin{cases} 1 & \tilde{y}_k = y_k \\ 0 & \tilde{y}_k \neq y_k \end{cases} \quad (20)$$

[0094] 其中, $y_k$ 是样本 $x_k$ 对应的真实标签。分类器的分类准确率最终可确定为 $\tilde{d}/d$ 。准确率越高,说明分类器的性能越好。当分类器在验证集上达到设定的分类准确率后,执行半自动标注模式以提升标注的速度。

[0095] 所述步骤(4)执行半自动标注模式,其流程如图1第二行所示。具体地,它通过以下步骤实现:

[0096] 1) 利用人机交互界面,手动框选目标位置;

[0097] 2) 提取目标的HOG特征 $x$ ;

[0098] 3) 使用分类器对目标的类别做预测:将目标特征 $x$ 输入分类器,目标的类别置信度分布如下:

$$[0099] \quad \mathbf{p} = \frac{e^{w^T x}}{\sum_{i=1}^n e^{w_i^T x}} \quad (21)$$

[0100] 其中, $n$ 为目标类别数, $p$ 是 $n$ 维矢量,它的第 $i$ 个元素 $p_i$ 对应目标属于第 $i$ 个类别的置信度,满足 $0 < p_i < 1$ ,且 $\sum_{i=1}^n p_i = 1$ ;特征 $x$ 的类别最终判决为:

$$[0101] \quad \tilde{y} = \arg \max_i (p_i) \quad (22)$$

[0102]  $\tilde{y}$ 即为目标的注释结果,这是由分类器自动生成的。

[0103] 标注前期使用手动标注模式;随着分类器在线训练次数的增多,它的分类性能将越来越好;当分类器在验证集上达到设定的分类准确率后,执行半自动标注模式以提升标注的速度。Ftrl算法在飞机-轮船二分类问题上的收敛情况如图3所示。对于2000个测试样本,Ftrl算法在训练了200步左右就能得到80%的分类准确率,收敛速度较快。此外,随着数据量的增加,准确率越来越高,收敛也越来越稳定。

[0104] 以上所述仅是本发明的优选实施方式,本发明的保护范围并不局限于上述实施例,凡属于本发明思路下的技术方案均属于本发明的保护范围。应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理前提下的若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

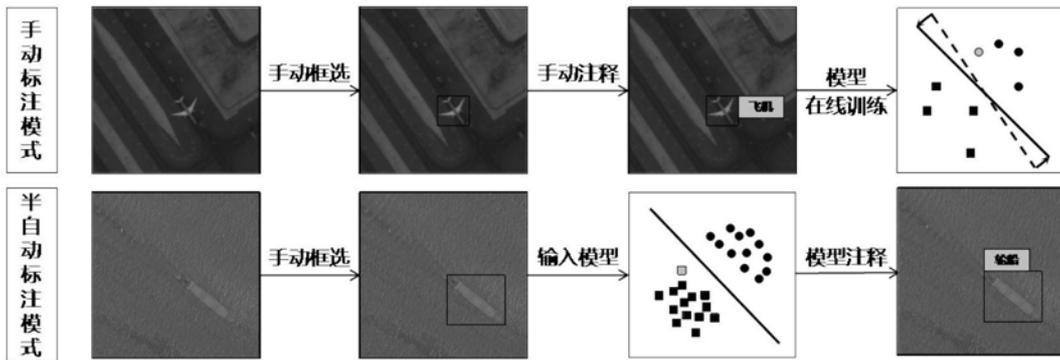


图1



图2

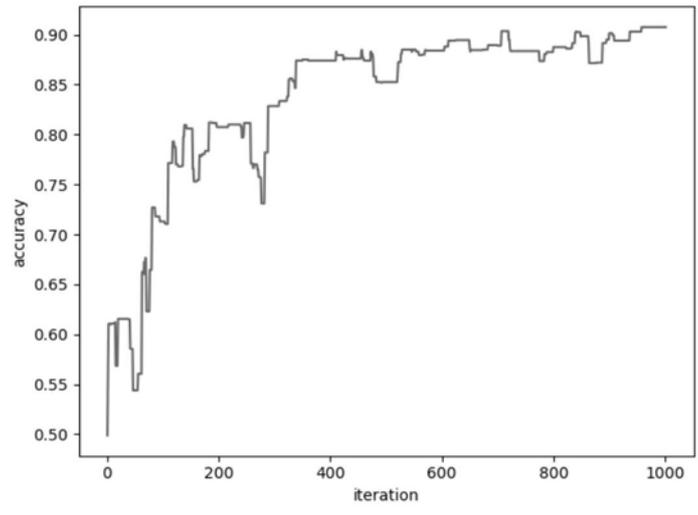


图3