



(12) 发明专利申请

(10) 申请公布号 CN 105184271 A

(43) 申请公布日 2015. 12. 23

(21) 申请号 201510599194. 6

(22) 申请日 2015. 09. 18

(71) 申请人 苏州派瑞雷尔智能科技有限公司  
地址 215000 江苏省苏州市苏州工业园区林泉街 399 号 2 号楼 410 室

申请人 青岛智能产业技术研究院  
江苏中科智能工程有限公司  
天问汇科(北京)技术有限公司

(72) 发明人 姚彦洁 田滨 王飞跃 杨柳青

(74) 专利代理机构 中科专利商标代理有限责任公司 11021

代理人 宋焰琴

(51) Int. Cl.

G06K 9/00(2006. 01)

G06K 9/62(2006. 01)

G06K 9/66(2006. 01)

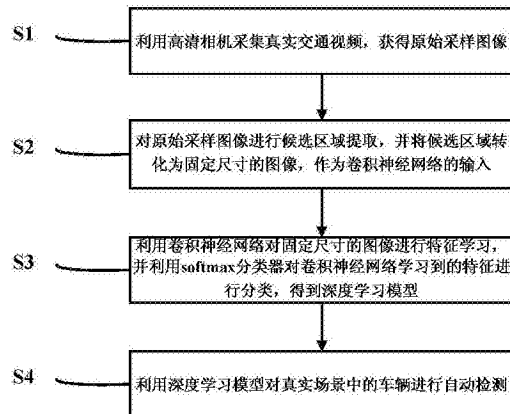
权利要求书2页 说明书5页 附图3页

(54) 发明名称

一种基于深度学习的车辆自动检测方法

(57) 摘要

本发明提供了一种高清图像中基于深度学习的车辆自动检测方法,利用卷积神经网络自动提取交通环境中的车辆特征,最终实现车辆目标的自动检测和识别。本发明可用于不同环境、光照强度及天气情况,对于复杂交通环境中的车辆,可以保证较高的检测准确性,可为后续视频监控中车辆目标的识别与跟踪提供一种新的实时可靠的预处理步骤。



1. 一种基于深度学习的车辆自动检测方法,其特征在于,包括:

S1,采集真实交通视频,获得原始采样图像;

S2,对所述原始采样图像进行候选区域提取,并将所述候选区域转化为固定尺寸的图像;

S3,利用卷积神经网络对所述固定尺寸的图像进行特征学习,并对学习到的特征进行分类,得到深度学习模型;

S4,利用所述深度学习模型对真实场景中的车辆进行自动检测。

2. 根据权利要求1所述车辆自动检测的方法,其特征在于,所述步骤S2包括:

对所述原始采样图像,利用训练后的二值正则梯度通用目标检测器,快速生成多尺度的、可能包含车辆目标的候选区域,并从所述候选区域中提取出多个目标子区域,根据所述目标子区域与车辆目标的重复部分的比重,决定其为车辆样本或背景样本,通过图像归一化将所述目标子区域转化为固定大小的图像。

3. 根据权利要求1所述车辆自动检测的方法,其特征在于,所述卷积神经网络包括1个输入层、4个卷积层、4个池化层、2个全连接层和1个输出层,其整体结构为:输入层—卷积层—池化层—卷积层—池化层—卷积层—卷积层—全连接层—全连接层—输出层,其中:

所述输入层用于将所述固定尺寸的图像输入至所述卷积神经网络;

所述卷积层用于采用多个卷积核得到图像特征;

所述池化层用于对所述图像特征进行均值池化操作,将所述图像特征的数据量降维原来的四分之一;

所述全连接层用于连接当层神经元与前一层神经元;

所述输出层用于计算分类结果,根据结果得分判定候选目标是小汽车、公交车或出租车。

4. 根据权利要求3所述车辆自动检测的方法,其特征在于,所述卷积层对输入图像进行二维卷积和偏置的操作后,再使用非线性激活函数,获得一个卷积结果,即图像特征 $O_j^n$ ,其表达式为:

$$O_j^n = f \left( \sum_{i=1}^U w_{i,j}^n O_j^{n-1} + b_j^n \right)$$

其中,n代表网络层数,U代表第n层的神经元个数, $w_{i,j}$ 表示第i个输入图像和第j个输出图像的卷积权重, $b_j$ 代表第j个输出图像的偏置, $f(\cdot)$ 为神经元激活函数, $f(x) = \max(0, x)$ 。

5. 根据权利要求4所述车辆自动检测的方法,其特征在于,输出层通过softmax函数计算分类结果 $O_k$ :

$$O_k = f \left( \sum_{i=1}^M w_{i,k}^N O_j^{N-1} + b_k \right)$$

其中,k代表输出层的类别数,M为最后一个全连接层的神经元数目,N代表网络的总层

数,  $f(\cdot)$  为 softmax 函数,  $f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}}$ 。

6. 根据权利要求 5 所述车辆自动检测的方法, 其特征在于, 在进行特征学习时, 采用梯度下降法调整所述卷积核的权重和偏置, 并对池化层的误差进行上采样操作。

## 一种基于深度学习的车辆自动检测方法

### 技术领域

[0001] 本发明属于模式识别技术领域和交通监控系统技术领域,特别涉及一种高清图像中基于深度学习的车辆自动检测的方法。

### 背景技术

[0002] 目前我国城市化水平已经超过 50%,汽车保有量达 1.4 亿辆,随之而来的交通拥堵,环境污染,交通安全等问题日益突出。为解决现存交通系统中人,车,路之间的矛盾,智能交通系统成为首选的理想方案。其中,智能视频监控系统是智能交通系统中必不可少的部分。基于智能视频监控系统的车辆检测,是利用模式识别技术对经过特定路段和路口的车辆进行自动检测,获取车辆基本信息,如车牌,速度,车型等,为之后的车辆行为分析,违章车辆图像抓拍,和违章取证等奠定基础。传统的智能监控系统通常将车辆检测视为一个二分类问题来处理,无法自动判定车辆的类型信息,而常见的车型识别方法也只是根据尺寸将车分为大型车、中型车,和小型车,无法满足现代监控的需求。另外,当前的智能视频监控系统主要通过人工提取车辆特征,如颜色特征、纹理特征,轮廓特征、几何特征等,来识别车辆。手工标记车辆特征的方法耗时费力,且不能的充分利用车辆信息,无法覆盖较多的行车环境,在遇到车辆遮挡,或车辆阴影时,无法达到较高检测的准确性。

[0003] 深度学习通过学习一种深层的非线性网络结构,实现复杂函数逼近和输入数据表征,展现出强大的数据集本质特征的学习能力。卷积神经网络,作为一种典型的深度学习方法,是为二维图像处理而特别设计的一个多层感知器。卷积神经网络不需要人为参与特征的选取过程,可以自动地学习大量数据集中的目标特征。其权值共享和局部连接机制使得它具备优于传统技术的优点:对几何变换、形变、光照具有一定程度的不变性,同时具有良好的容错能力、并行处理能力和自学习能力。这些优点使得卷积神经网络在处理环境信息复制,推理规则不明确情况下的问题时具有较大优势,并能容忍车辆的尺度变化,旋转变形等问题。因此,卷积神经网络可以用于解决智能视频监控系统中车辆的自动检测和车型识别,保证智能监控系统准确地、稳定地获取车辆信息。

### 发明内容

[0004] (一) 要解决的技术问题

[0005] 本发明的目的在于,提供一种基于深度学习的车辆自动检测方法,可以自动学习车辆特征,实现更高精度的车辆自动检测,和车辆类型识别。

[0006] (二) 技术方案

[0007] 本发明提供一种基于深度学习的车辆自动检测的方法,包括:

[0008] S1,采集真实交通视频,获得原始采样图像;

[0009] S2,对原始采样图像进行候选区域提取,并将候选区域转化为固定尺寸的图像;

[0010] S3,利用卷积神经网络对固定尺寸的图像进行特征学习,并对学习到的特征进行分类,得到深度学习模型;

[0011] S4,利用深度学习模型对真实场景中的车辆进行自动检测,并自动识别车辆类型(小汽车/出租车/公交车)。

[0012] (三)有益效果

[0013] 本发明具以下优点:

[0014] (1)可以自动逐层抽象出输入卷积神经网络的图像特征,减少人为标记特征的工作量,节约人工成本和时间成本;

[0015] (2)快速实现车辆目标定位,并能识别出车辆类型(小汽车/出租车/公交车);

[0016] (3)对不同环境,光照强度,和天气情况,均具有良好的适应性;

[0017] (4)对大流量车辆的拥挤环境,可以在一定程度上处理遮挡问题,保证较高的检测准确率。

### 附图说明

[0018] 图1是本发明实施例提供的基于深度学习的车辆自动检测方法的流程图。

[0019] 图2是本发明实施例中训练样本的典型图像。

[0020] 图3是本实施例中最终检测结果的示意图。

### 具体实施方式

[0021] 本发明提供一种高清图像中基于深度学习的车辆自动检测方法,利用卷积神经网络自动提取交通环境中的车辆特征,最终实现车辆目标的自动检测。本发明可用于不同环境、光照强度及天气情况,对于复杂交通环境中的车辆,可以保证较高的检测准确性,可为后续视频监控中车辆目标的识别与跟踪提供一种新的实时可靠的预处理步骤。

[0022] 根据本发明的一种实施方式,车辆自动检测方法包括:

[0023] S1,采集真实交通视频,获得原始采样图像;

[0024] S2,对原始采样图像进行候选区域提取,并将候选区域转化为固定尺寸的图像;

[0025] S3,利用卷积神经网络对固定尺寸的图像进行特征学习,并对学习到的特征进行分类,得到深度学习模型;

[0026] S4,利用深度学习模型对真实场景中的车辆进行自动检测。

[0027] 根据本发明的一种实施方式,步骤S2包括:

[0028] 对原始采样图像,利用训练后的二值正则梯度(BING: Binarized Normed Gradients)通用目标检测器,快速生成多尺度的、可能包含车辆目标的候选区域,并从候选区域中提取出多个目标子区域,根据目标子区域与车辆目标的重复部分的比重,决定其为车辆样本或背景样本,通过图像归一化将目标子区域转化为固定大小的图像,优选地,图像大小可选为 $32 \times 32$ 像素。

[0029] 根据本发明的一种实施方式,卷积神经网络包括1个输入层、4个卷积层、4个池化层、2个全连接层和1个输出层,其整体结构为:输入层—卷积层—池化层—卷积层—池化层—卷积层—卷积层—全连接层—全连接层—输出层,其中:

[0030] 输入层用于将所述固定尺寸的图像输入至所述卷积神经网络;

[0031] 卷积层用于采用多个卷积核得到图像特征;

[0032] 池化层用于对所述图像特征进行均值池化操作,将所述图像特征的数据量降维原

来的四分之一；

[0033] 全连接层用于连接当层神经元与前一层神经元；

[0034] 输出层用于计算分类结果。

[0035] 根据本发明的一种实施方式，卷积层对输入图像进行二维卷积和偏置的操作后，再使用非线性激活函数，获得一个卷积结果，即图像特征 $O_j^n$ ，其表达式为：

$$[0036] \quad O_j^n = f \left( \sum_{i=1}^U w_{i,j}^n O_j^{n-1} + b_j^n \right)$$

[0037] 其中，n 代表网络层数，U 代表第 n 层的神经元个数， $w_{i,j}$  表示第 i 个输入图像和第 j 个输出图像的卷积权重， $b_j$  代表第 j 个输出图像的偏置， $f(\cdot)$  为神经元激活函数， $f(x) = \max(0, x)$ 。

[0038] 根据本发明的一种实施方式，输出层通过 softmax 函数计算分类结果：

$$[0039] \quad O_k = f \left( \sum_{i=1}^M w_{i,k}^N O_j^{N-1} + b_k \right)$$

[0040] 其中，k 代表输出层的类别数，M 为最后一个全连接层的神经元数目，N 代表网络的总层数， $f(\cdot)$  为 softmax 函数， $f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}}$ 。

[0041] 根据本发明的一种实施方式，在进行特征学习时，采用梯度下降法调整所述卷积核的权重和偏置，并对池化层的误差进行上采样操作。

[0042] 为使本发明的目的、技术方案和优点更加清楚明白，以下结合具体实施例，并参照附图，对本发明进一步详细说明。

[0043] 图 1 是本发明实施例提供的基于深度学习的车辆自动检测方法的流程图，如图 1 所示，方法包括：

[0044] S1，利用高清相机采集真实交通视频，获得原始采样图像；

[0045] 本发明实例中，相机以一定高度和角度对车辆流进行拍摄，尽量模拟真实的交通监控系统的视角，获取尽可能多的交通图像样本。

[0046] S2，对原始采样图像进行候选区域提取，并将候选区域转化为固定尺寸的图像，作为卷积神经网络的输入。

[0047] 首先，利用训练后的二值正则梯度检测器生成可能的目标区域，从原始图像中提取出 1500 个不同大小的目标子区域，给出可能的目标位置。根据子区域与车辆目标的重复部分的比重，决定其是车辆样本还是背景样本。本实施例中，比重阈值选取为 0.3，可以保证检测到部分遮挡的车辆。

[0048] 然后，将获取的目标子区域转化为固定大小的图像，即图像的归一化，这里为  $32 \times 32$  像素大小，作为后续卷积神经网络的输入。

[0049] 图 2 为正样本的典型图像，包括公交车，小汽车，和出租车。其中，二值正则梯度的训练分两个步骤：1) 提取正负样本 64 维的二值正则梯度特征，训练线性 SVM 分类器；2) 对上步训练的分类器的结果，进行非极大值抑制，得到最终的输出。

[0050] S3,利用卷积神经网络对固定尺寸的图像进行特征学习,并利用 softmax 分类器对卷积神经网络学习到的特征进行分类,得到深度学习模型;

[0051] 本发明实例构建的卷积神经网络模型采用 10 层网络结构,包括输入层,4 个卷积层,4 个池化层,2 个全连接层,和输出层。整体结构为:输入层—卷积层—池化层—卷积层—池化层—卷积层—卷积层—全连接层—全连接层—输出层。

[0052] 输入层将  $32 \times 32$  像素大小的子图像输入卷积神经网络。

[0053] 卷积层的作用是提取图像特征。对输入图像进行二维卷积和偏置的操作后,再使用非线性激活函数,便可获得一个卷积结果,即获得一个图像特征  $O_j^n$ :

$$[0054] \quad O_j^n = f \left( \sum_{i=1}^U w_{i,j}^n O_j^{n-1} + b_j^n \right)$$

[0055] 其中,  $n$  代表网络层数,  $U$  代表第  $n$  层的神经元个数,  $w_{i,j}$  表示第  $i$  个输入图像和第  $j$  个输出图像的卷积权重,  $w_j$  代表第  $j$  个输出图像的偏置,  $f(\cdot)$  为神经元激活函数,  $f(x) = \max(0, x)$ 。并且,前 3 个卷积核的大小为  $5 \times 5$ ,最后 1 个卷积核的大小为  $3 \times 3$ ,各层卷积核的滑动步长为 1。

[0056] 池化层的作用是特征统计和数据降维,沿图像行和列的方向上,每两个像素进行统计,抽取四个像素的均值 (mean) 作为该区域的统计特征,同时将数据量降维原来的四分之一。

[0057] 全连接层是当层的神经元与前一次神经元全部连接,之后,输出层通过 softmax 函数计算分类结果  $O_k$ :

$$[0058] \quad O_k = f \left( \sum_{i=1}^M w_{i,k}^N O_j^{N-1} + b_k \right)$$

[0059] 其中,  $k$  代表输出层的类别数,  $M$  为最后一个全连接层的神经元数目,  $N$  代表网络的总层数,  $f(\cdot)$  为 softmax 函数,  $f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}}$ 。

[0060] 反向传播阶段,采用梯度下降法调整卷积神经网络各层的卷积核的权重和偏置。

输出层的误差函数为,  $E_i = -\log \left( \frac{e^{f y_i}}{\sum_j e^{f_j}} \right)$ 。

[0061] 卷积层的梯度:  $\delta_j^n = \beta_j^{l+1} (f'(z_j^l) \circ u(\delta_j^{l+1}))$ , 其中  $u(x) = x \otimes I_{n \times n}$ , 是上采样操作,这是由于池化层进行了下采样操作,为了补偿前一卷积层的数据损失,在对卷积层使用梯度下降算法时,需要对池化层的偏导误差进行上采样操作。

[0062] 对偏置  $b$ :  $\frac{\partial E}{\partial b_j} = \sum_{u,v} (\delta_j^l)_{uv}$ ;

[0063] 对卷积核权重  $k$ :  $\frac{\partial E}{\partial k_{ij}^l} = \sum_{u,v} (\delta_j^l)_{uv} (P_i^{l-1})_{uv}$ , 这里,  $P_i^{l-1}$ :  $x_i^{l-1}$  与  $k_{ij}^l$  逐元素相乘的 patch 结果。

[0064] 下采样层梯度： $\delta_j^l = f'(O_j^l) \circ \text{conv}(\delta_j^{l+1}, k_j^{l+1})$ ，其中， $\text{conv}(\cdot)$  是卷积操作。

[0065] 对偏置  $b$ ： $\frac{\partial E}{\partial b_j} = \sum_{u,v} (\delta_j^l)_{uv}$ ；

[0066] 当误差  $E < \varepsilon$  时，其中， $\varepsilon$  为预先设置的最小误差阈值，训练结束，并保存各层的权重和阈值。

[0067] 最后，对训练好的网络进行微调，就可完成所述卷积神经网络模型的训练。注意，上述模型把检测问题转化成二分类问题，即车辆和背景。

[0068] S4，利用深度学习模型对真实场景中的车辆进行自动检测和识别。图 3 是本方法最终检测结果的示意图，其中，矩形框分别标出了出租车、小汽车和公交车。

[0069] 训练好模型后，就可将模型用于真实场景中的车辆进行自动检测，由于训练输入时包含了部分旋转和部分车辆的图像，所以此模型对旋转，部分遮挡有一定的容忍性。

[0070] 以上所述的具体实施例，对本发明的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本发明的具体实施例而已，并不用于限制本发明，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。



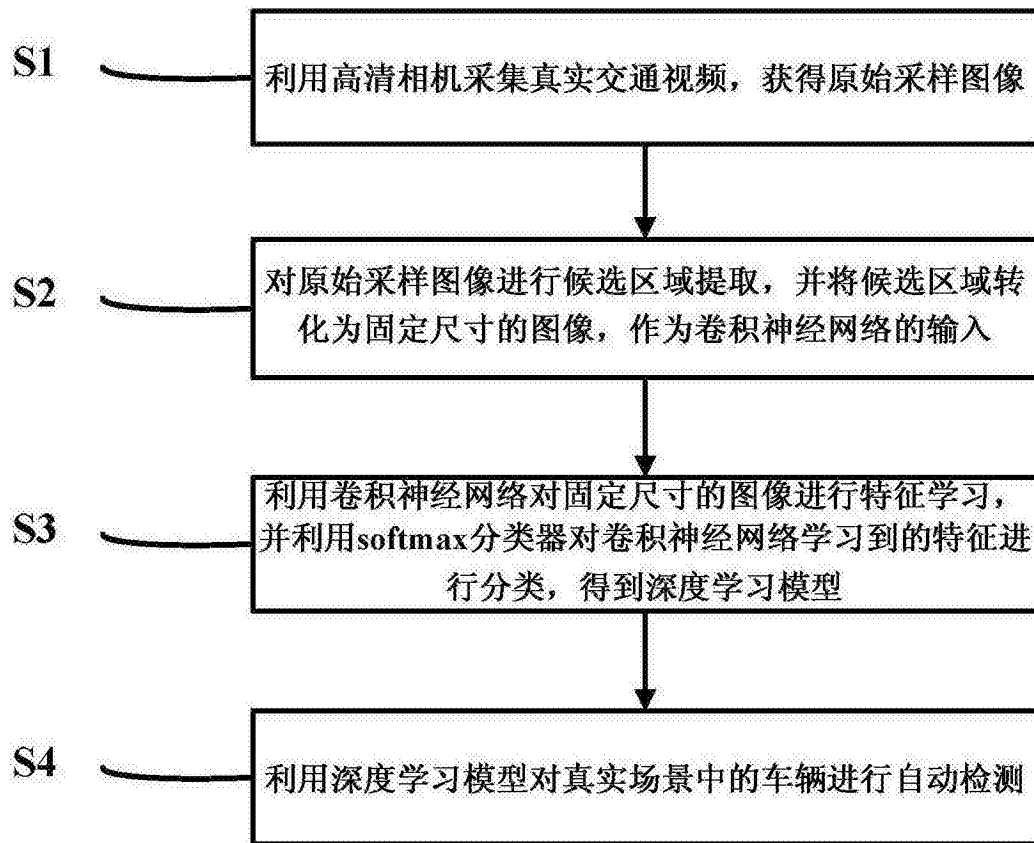


图 1

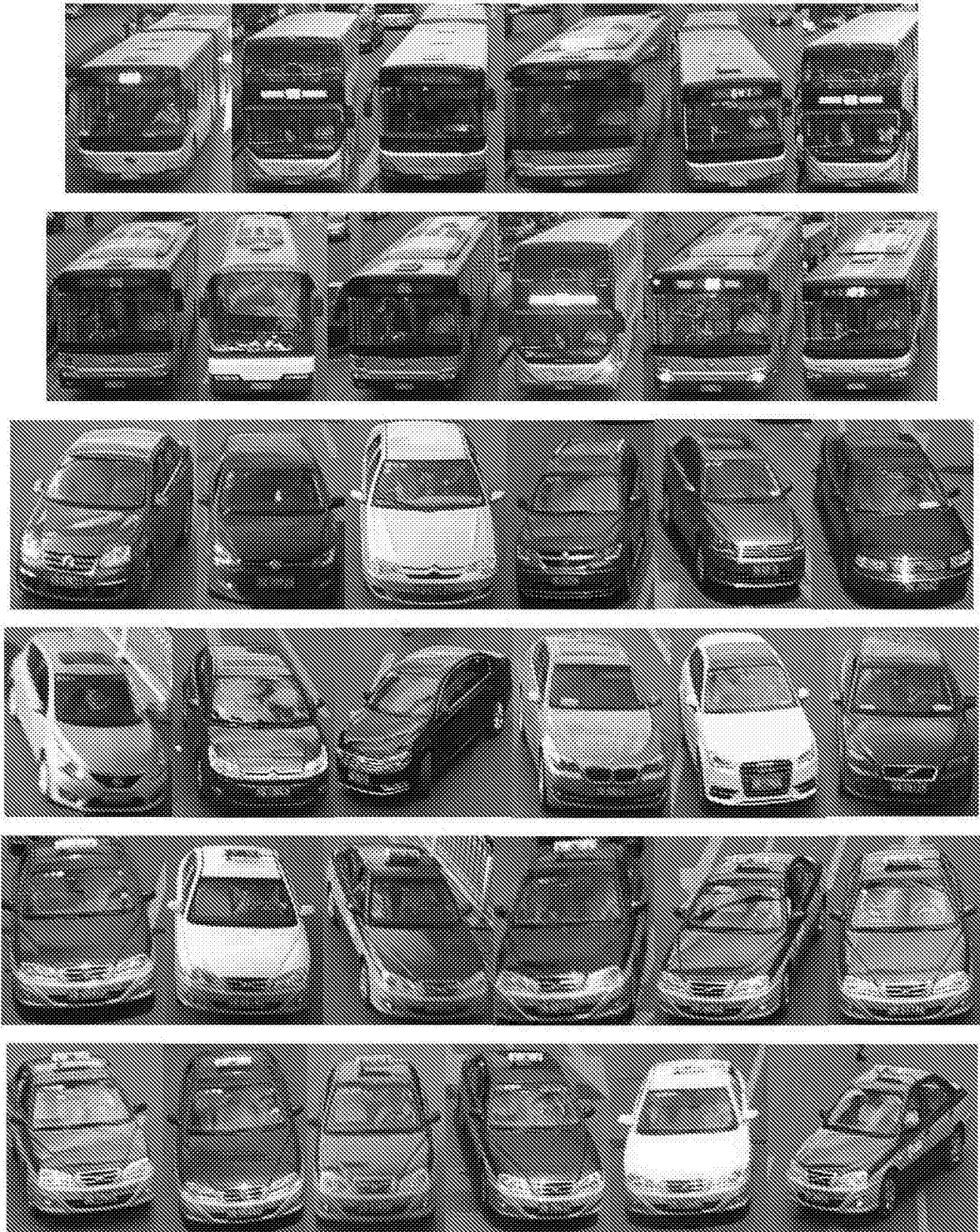


图 2

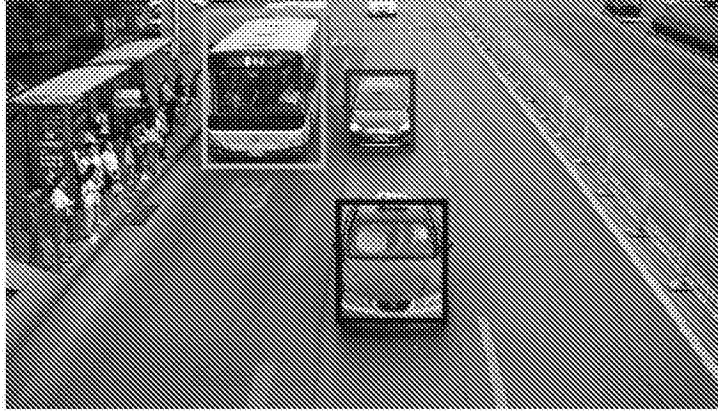


图 3