



(12)发明专利申请

(10)申请公布号 CN 109376572 A

(43)申请公布日 2019.02.22

(21)申请号 201810899841.9

(22)申请日 2018.08.09

(71)申请人 同济大学

地址 200092 上海市杨浦区四平路1239号

(72)发明人 徐子豪 黄伟泉 王胤

(74)专利代理机构 南京纵横知识产权代理有限公司 32224

代理人 董建林

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

权利要求书2页 说明书5页 附图4页

(54)发明名称

基于深度学习的交通视频中实时车辆检测与轨迹跟踪方法

(57)摘要

本发明公开了一种交通视频中实时车辆检测与轨迹跟踪方法,包括以下步骤,采用基于深度学习的目标检测算法,检测出交通视频中车辆的位置,提取到车辆的特征向量和类别;对车辆进行跟踪,获得出行驶轨迹。本发明整套方案鲁棒性强、漏检率低,而且容易扩展到多种车辆类别,满足监控视频中车辆检测和持续跟踪的要求。



1. 交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:包括以下步骤,
采用基于深度学习的目标检测算法,检测出交通视频中车辆的位置,提取车辆特征向量和类别;
对车辆进行跟踪,获得出行轨迹。
2. 根据权利要求1所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:检测交通视频中车辆位置,提取车辆特征向量和类别的过程为,
获取交通视频中单帧视频图像;
采用基于深度学习的目标检测算法,获取图像中各车辆的若干检测候选框坐标、框内车辆特征向量以及所属类别概率;
采用非极大值抑制算法,对车辆的若干检测候选框进行融合,得到各车辆的最优检测候选框;
各车辆的最优检测候选框即为各车辆最终的检测框,检测框坐标即为检测出的车辆位置,检测框内车辆特征向量即为检测出的车辆特征向量,概率最大的类别即为车辆的类别。
3. 根据权利要求2所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:深度学习的目标检测算法为改进的Faster R-CNN;
改进的Faster R-CNN为,在传统的Faster R-CNN结构上,用Resnet代替VGG作为骨干网络进行特征提取,在Resnet的中间卷积层的运算中使用带间隔的空洞卷积代替传统的相邻卷积。
4. 根据权利要求2所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:深度学习的目标检测算法为改进的SSD;
改进的SSD为,在传统的SSD结构上,用Inception代替VGG作为骨干网络进行特征提取,增加L层放大层,L等于依次经过卷积和池化处理得到的特征图层数,尺寸最小的特征图依次经过L层放大层,得到L层新的特征图,L层新的特征图与L层原来的特征图两两对应,对应的两个特征图大小相同,对对应的特征图进行像素求和,对像素求和后的L层特征图进行框坐标回归和分类。
5. 根据权利要求4所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:使用焦点损失训练改进的SSD。
6. 根据权利要求2所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:车辆跟踪的过程为,
获取当前帧的各车辆检测框,以及上一帧的各车辆检测框;
将上一帧一车辆检测框分别与当前帧所有车辆检测框进行相似度计算;若相似度最大值超过阈值,则相似度最大值对应的当前帧车辆检测框为当前跟踪车辆的框;若所有相似度值都不超过阈值,则判定当前帧没有检测到该车辆;
遍历所有上一帧车辆检测框,获得上一帧车所有车辆的跟踪情况;
遍历所有当前帧车辆检测框,未匹配成功的认为是新驶进监控范围的车辆。
7. 根据权利要求6所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:若某一车辆连续多帧没有检测到,则表面该车辆驶出,监控范围。
8. 根据权利要求6所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在于:两个检测框的相似度计算过程为,

计算两个检测框的IOU值；

计算两个框内车辆特征向量的相似度；

将IOU值和特征向量相似度的调和平均数作为两个检测框的相似度。

9. 根据权利要求1所述的交通视频中实时车辆检测与轨迹跟踪方法,其特征在於:所有检测结果和跟踪结构均存储在车辆信息表中。

基于深度学习的交通视频中实时车辆检测与轨迹跟踪方法

技术领域

[0001] 本发明涉及一种基于深度学习的交通视频中实时车辆检测与轨迹跟踪方法,属于交通视频监控领域。

背景技术

[0002] 车辆检测与轨迹跟踪任务是指在一段监控视频中,检测出通过该摄像头的 所有车,并绘制其运动轨迹。现有的车辆检测与轨迹跟踪方法是将传统的视频 中目标检测和跟踪结合而成,其中目标检测步骤为:训练目标分类器、提取单 帧图像、候选区域分类、目标区域融合,目标跟踪步骤为:提取单帧图像、给 定初始目标框、结合帧间图像变化预测目标位置。

[0003] 现有的车辆检测与轨迹跟踪方法计算速度快,可以满足视频实时播放的要 求,但由于手工设计的传统特征本身灵活性较差,导致针对稍复杂场景时,鲁 棒性差,在车辆较多时漏检率高,跟踪不到位,而且其一般将车辆作为一大类 检测,无法对车辆进行细分类,如:小汽车、出租车、卡车等等。

发明内容

[0004] 本发明提供了一种基于深度学习的交通视频中实时车辆检测与轨迹跟踪方 法,解决了现有方法在稍复杂场景时,鲁棒性差、漏检率高,同时无法对车辆 进行分类的问题。

[0005] 为了解决上述技术问题,本发明所采用的技术方案是:

[0006] 交通视频中实时车辆检测与轨迹跟踪方法,包括以下步骤,

[0007] 采用基于深度学习的目标检测算法,检测出交通视频中车辆的位置,提取 车辆的特征向量和类别;

[0008] 对车辆进行跟踪,获得行驶轨迹。

[0009] 检测交通视频中车辆位置,提取车辆特征向量和类别的过程为,

[0010] 获取交通视频中单帧视频图像;

[0011] 采用基于深度学习的目标检测算法,获取图像中各车辆的若干检测候选框 坐标、框内车辆特征向量以及所属类别概率;

[0012] 采用非极大值抑制算法,对车辆的若干检测候选框进行融合,得到各车辆 的最优检测候选框;

[0013] 各车辆的最优检测候选框即为各车辆最终的检测框,检测框坐标即为检测 出的车辆位置,检测框内车辆特征向量即为检测出的车辆特征向量,概率最大 的类别即为车辆的类别。

[0014] 深度学习的目标检测算法为改进的Faster R-CNN;

[0015] 改进的Faster R-CNN为,在传统的Faster R-CNN结构上,用Resnet代替 VGG作为骨干网络进行特征提取,在Resnet的中间卷积层的运算中使用带间隔 的空洞卷积代替传统的相邻卷积。

- [0016] 深度学习的目标检测算法为改进的SSD;
- [0017] 改进的SSD为,在传统的SSD结构上,用Inception代替VGG作为骨干网络进行特征提取,增加L层放大层,L等于依次经过卷积和池化处理得到的特征图层数,尺寸最小的特征图依次经过L层放大层,得到L层新的特征图,L层新的特征图与L层原来的特征图两两对应,对应的两个特征图大小相同,对对应的特征图进行像素求和,对像素求和后的L层特征图进行框坐标回归和分类。
- [0018] 使用焦点损失训练改进的SSD。
- [0019] 车辆跟踪的过程为,
- [0020] 获取当前帧的各车辆检测框,以及上一帧的各车辆检测框;
- [0021] 将上一帧一车辆检测框分别与当前帧所有车辆检测框进行相似度计算;若相似度最大值超过阈值,则相似度最大值对应的当前帧车辆检测框为当前跟踪车辆的框;若所有相似度值都不超过阈值,则判定当前帧没有检测到该车辆;遍历所有上一帧车辆检测框,获得上一帧车所有车辆的跟踪情况;
- [0022] 遍历所有当前帧车辆检测框,未匹配成功的认为是新驶进监控范围的车辆。
- [0023] 若某一车辆连续多帧没有检测到,则表面该车辆驶出,监控范围。
- [0024] 两个检测框的相似度计算过程为,
- [0025] 计算两个检测框的IOU值;
- [0026] 计算两个框内车辆特征向量的相似度;
- [0027] 将IOU值和特征向量相似度的调和平均数作为两个检测框的相似度。
- [0028] 所有检测结果和跟踪结构均存储在车辆信息表中。
- [0029] 本发明所达到的有益效果:1、本发明利用深度学习的方法进行车辆检测,同时检测出车辆位置和类别,不需要手工设计特征,省去了特征选择的过程,提取特征质量更好,面对复杂场景更加鲁棒、漏检率低;2、本发明对传统的深度学习的目标检测算法进行了改进,在不改变检测速度的情况下,复杂场景下检测效果更好;3、本发明检测结果与IOU跟踪器相结合,凭借跟踪器速度快的优势实现对车辆的实时跟踪;4、本发明整套方案鲁棒性强、漏检率低,而且容易拓展到多种车辆类别,满足监控视频中车辆检测和持续跟踪的要求。

附图说明

- [0030] 图1为本发明的流程图;
- [0031] 图2为本发明检测的流程图;
- [0032] 图3为改进的Faster R-CNN结构图;
- [0033] 图4为改进的SSD结构图;
- [0034] 图5为本发明跟踪的流程图。

具体实施方式

- [0035] 下面结合附图对本发明作进一步描述。以下实施例仅用于更加清楚地说明本发明的技术方案,而不能以此来限制本发明的保护范围。
- [0036] 如图1所示,交通视频中实时车辆检测与轨迹跟踪方法,包括以下步骤:

[0037] 步骤1,采用基于深度学习的目标检测算法,检测出交通视频中车辆的位置、车辆的特征向量和类别。

[0038] 如图2所示,具体检测过程如下:

[0039] 11) 获取交通视频中单帧视频图像,并对图像进行大小变换,以适应目标检测算法。

[0040] 12) 采用基于深度学习的目标检测算法,获取图像中各车辆的若干检测候选框坐标、框内车辆特征向量以及所属类别概率。

[0041] 深度学习的目标检测算法采用改进的Faster R-CNN或者改进的SSD (Single Shot MultiBox Detector)。

[0042] 其中,

[0043] 改进的Faster R-CNN是在传统的Faster R-CNN结构上,用Resnet代替VGG作为骨干网络进行特征提取,由于如今的视频图像已基本达到超清画质,超清图像尺寸大、冗余像素较多,所以在Resnet的中间卷积层的运算中使用带间隔的空洞卷积代替传统的相邻卷积。

[0044] 如图3所示,为改进的Faster R-CNN结构,具体的过程如下:

[0045] A1) 单帧视频图像先经过Resnet (残差神经网络) 得到特征图,Resnet的中间卷积层的运算中使用带间隔的空洞卷积。

[0046] 原卷积核在进行一般卷积运算时,是与特征图的相邻像素点做乘法运算,而带间隔的空洞卷积运算允许卷积核与固定间隔 l 的像素点做乘法运算,这样在不增加额外运算量的同时,增大感受野。

[0047] 带间隔的空洞卷积运算可用以下公式表示:

$$[0048] \quad (F *_l k)(p) = \sum_{s+t=p} F(s)k(t)$$

[0049] 其中, F 是 $\mathbb{Z}^2 \rightarrow \mathbb{R}$ 的离散函数, \mathbb{Z}, \mathbb{R} 分别是整数域和实数域, $*_l$ 为空洞卷积运算, k 为 $\Omega_r \rightarrow \mathbb{R}$ 的大小为 $(2r+1)^2$ 的离散滤波器, $\Omega_r = [-r, r]^2 \cap \mathbb{Z}^2$, r 为离散滤波器半径, l 为空洞卷积间隔因子, p 为整个运算函数的变量, $F(s)$ 为以 s 为变量的离散函数 F , $k(t)$ 为以 t 为变量的离散滤波器 k 。

[0050] A2) RPN (区域提取网络) 在特征图上进行滑窗计算,通过预先设置不同面积及尺寸目标框的方式实现检测候选框位置的预估。

[0051] A3) 对候选检测框进行分类,分成前景候选检测框和背景候选检测框,同时对前景候选检测框进行框坐标回归,框坐标回归即对位置进行修正。

[0052] A4) 不同大小的前景候选检测框经过ROI池化层调整成相同长度的特征向量。

[0053] A5) 经过全连接层连接,根据前景内容,对前景候选检测框进行分类,同时对内容为车辆的前景候选检测框进行框坐标回归。

[0054] 改进的SSD为,在传统的SSD结构上,用Inception代替VGG作为骨干网络进行特征提取,增加 L 层放大层, L 等于依次经过卷积和池化处理得到的特征图层的数,尺寸最小的特征图依次经过 L 层放大层,得到 L 层新的特征图, L 层新的特征图与 L 层原来的特征图两两对应,对应的两个特征图大小相同,对对应的特征图进行像素求和,对像素求和后的 L 层特征

图进行框坐标回归和分类。同时改进的SSD使用焦点损失替交叉熵损失对模型进行训练，解决各类别 车辆数量不平衡、样本难易区分程度不同的问题。

[0055] 假设n个训练样本，分类目标有C类，交叉熵损失CE定义如下：

$$[0056] \quad CE = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^C -y_j^{(i)} \log(\hat{f}(x^{(i)})_j)$$

[0057] 其中， $y_j^{(i)}$ 为第i个训练样本第j类的one-hot向量， $\hat{f}(x^{(i)})_j$ 为第i个训练 样本第j类预测概率。

[0058] 交叉熵本身将所有类别的物体均等对待，在遇到类别不平衡现象时容易造 成预测偏移，而且无法对难分样本进行加强训练。

[0059] 针对类别不平衡现象，可以针对不同类别引入一个权重因子 α 削弱大数量 类别对损失值的影响：

$$[0060] \quad CE = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^C -\alpha_j y_j^{(i)} \log(\hat{f}(x^{(i)})_j)$$

[0061] 其中， α_j 为j类的权重因子。

[0062] 针对难分样本问题，一个样本的预测概率越高，模型对该样本的识别力越 强，该样本成为易分样本，反之则为难分样本，可以以预测概率为基础，引入 一个权重因子 β 削弱易分样本对损失值的影响， β 定义如下：

$$[0063] \quad \beta_j^{(i)} = (1 - y_j^{(i)})^\gamma$$

[0064] 其中， $\beta_j^{(i)}$ 为第i个样本第j类的权重因子， γ 为一个可调节的超参数。

$$[0065] \quad \text{焦点损失FL定义为: } FL = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^C -\alpha_j \beta_j^{(i)} y_j^{(i)} \log(\hat{f}(x^{(i)})_j)。$$

[0066] 如图4所示，为改进的SSD结构，具体的过程如下：

[0067] 对于单帧视频图像，首先经过Inception计算，之后顺序经过四组卷积池化，每组卷积池化包括卷积层和池化层，分别得到四层特征图，编号分别为Mixed7c、Mixed6e、Mixed5d、Mixed4c，从Mixed4c到Mixed7c每一层特征图的大小是前 一层的一半（长宽都是前一层的1/2），四层放大层的放大倍数分别为1、2、4、8，将Mixed7c依次经过四层放大层，分别得到编号为UpMixed7c、UpMixed6e、UpMixed5d、UpMixed4c的新的特征图，其中，Mixed4c与UpMixed4c对应，Mixed5d与UpMixed5d对应，Mixed6e与UpMixed6e对应，Mixed7c与UpMixed7c 对应，每组对应的特征图大小相同，对每组对应的特征图进行像素求和，即特征融合，对像素求和后的四层特征图进行框坐标回归和分类，在训练时使用焦 点损失代替交叉熵损失进行训练。

[0068] 上述两种改进的方法经过离线训练可以准确地检测出图像中的车辆的位 置、车辆的特征向量和类别。前者运算速度较后者慢，但准确度较后者高，在 实际应用时，改进的SSD可以对每帧图像进行检测，而对于改进的Faster R-CNN，为了保证实时性，需要每隔2-3 帧检测图像。

[0069] 13) 经过检测后得到上百个不同大小的检测候选框,采用非极大值抑制算法,对车辆的若干检测候选框进行融合,得到各车辆的最优检测候选框。

[0070] 14) 各车辆的最优检测候选框即为各车辆最终的检测框,检测框坐标即为检测出的车辆位置,检测框内车辆特征向量即为检测出的车辆特征向量,概率最大的类别即为车辆的类别。

[0071] 步骤2,对车辆进行跟踪,获得行驶轨迹。

[0072] 如图5所示,具体车辆跟踪过程如下:

[0073] 21) 获取当前帧的各车辆检测框,以及上一帧的各车辆检测框。

[0074] 22) 将上一帧一车辆检测框分别与当前帧所有车辆检测框进行相似度计算;若相似度最大值超过阈值,则相似度最大值对应的当前帧车辆检测框为当前跟踪车辆的框;若所有相似度值都不超过阈值,则判定当前帧没有检测到该车辆。

[0075] 两个检测框的相似度计算过程为:计算两个检测框的IOU值,计算两个框内车辆特征向量的相似度;将IOU值和特征向量相似度的调和平均数作为两个检测框的相似度。

[0076] 23) 遍历所有上一帧车辆检测框,获得上一帧车所有车辆的跟踪情况,若某一车辆连续多帧没有检测到,则表面该车辆驶出,监控范围。

[0077] 24) 遍历所有当前帧车辆检测框,未匹配成功的认为是新驶进监控范围的车辆。

[0078] 上述放置设置有车辆信息表,该车辆信息表用以存储检测结果和跟踪结构,即每帧的检测结果均存储在车辆信息表,在进行车辆跟踪时,上一帧的各车辆检测框可直接从车辆信息表中获取,当有新驶进车辆时,在车辆信息表中添加对应车辆的信息,当车辆驶出监控范围,则将车辆信息表中对应车辆的信息全部移除(移除的信息可存入存储器),从车辆进入到驶出过程中,车辆信息表中该车辆的位置轨迹即为车辆的行驶轨迹。

[0079] 上述方法利用深度学习的方法进行车辆检测,同时检测出车辆位置和类别,不需要手工设计特征,省去了特征选择的过程,提取特征质量更好,面对复杂场景更加鲁棒、漏检率低。上述方法对传统的深度学习的目标检测算法进行了改进,在不改变检测速度的情况下,复杂场景下检测效果更好。上述方法将检测结果与IOU跟踪器相结合,凭借跟踪器速度快的优势实现对车辆的实时跟踪。

[0080] 上述整套方案鲁棒性强、漏检率低,而且容易拓展到多种车辆类别,满足监控视频中车辆检测和持续跟踪的要求。

[0081] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明技术原理的前提下,还可以做出若干改进和变形,这些改进和变形也应视为本发明的保护范围。

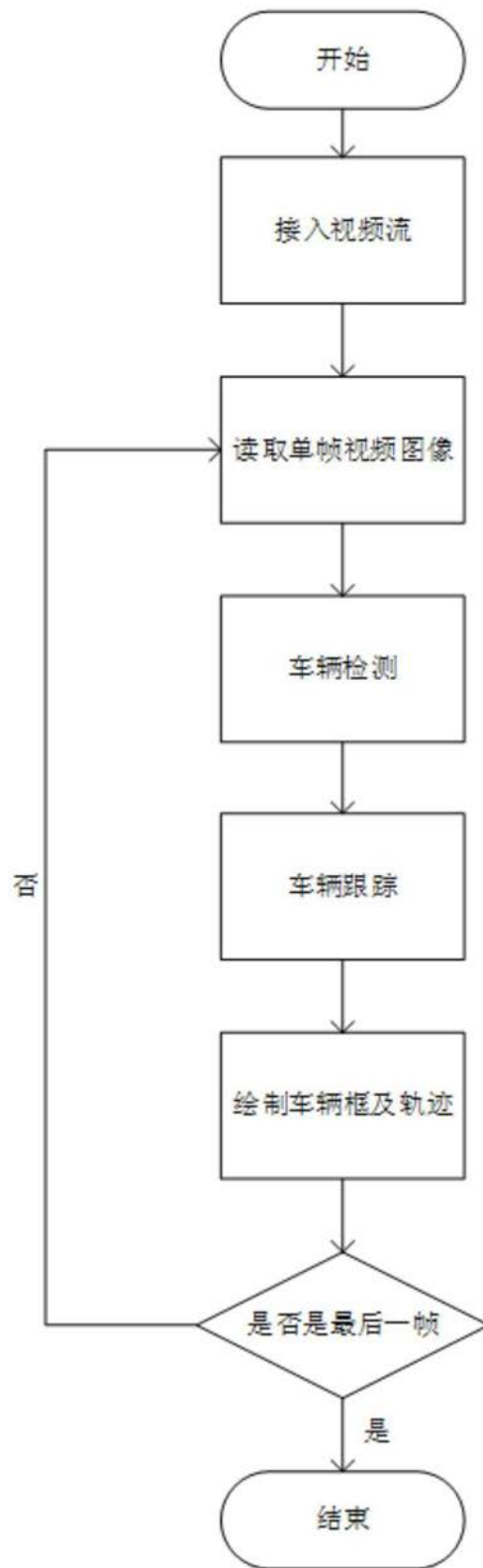


图1



图2

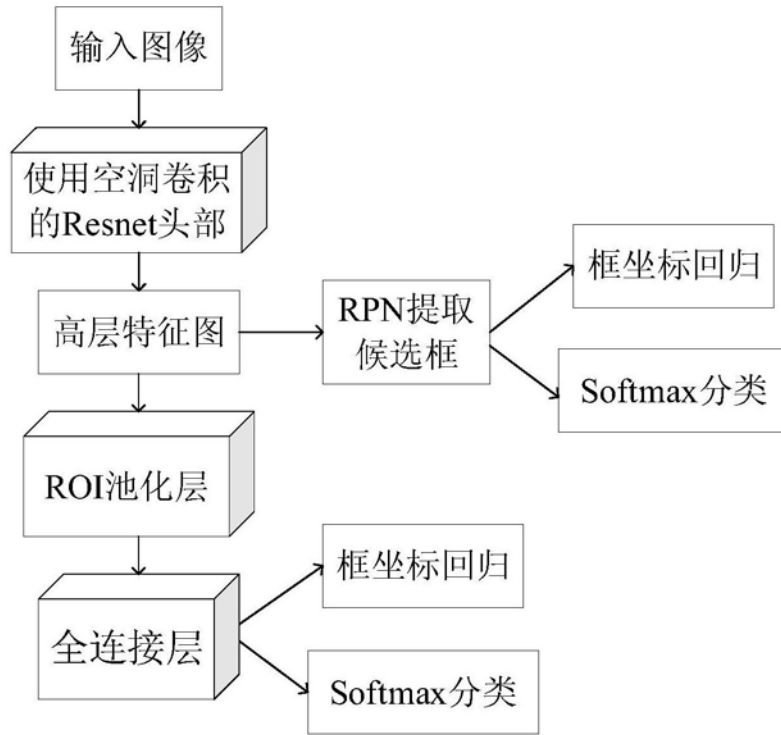


图3

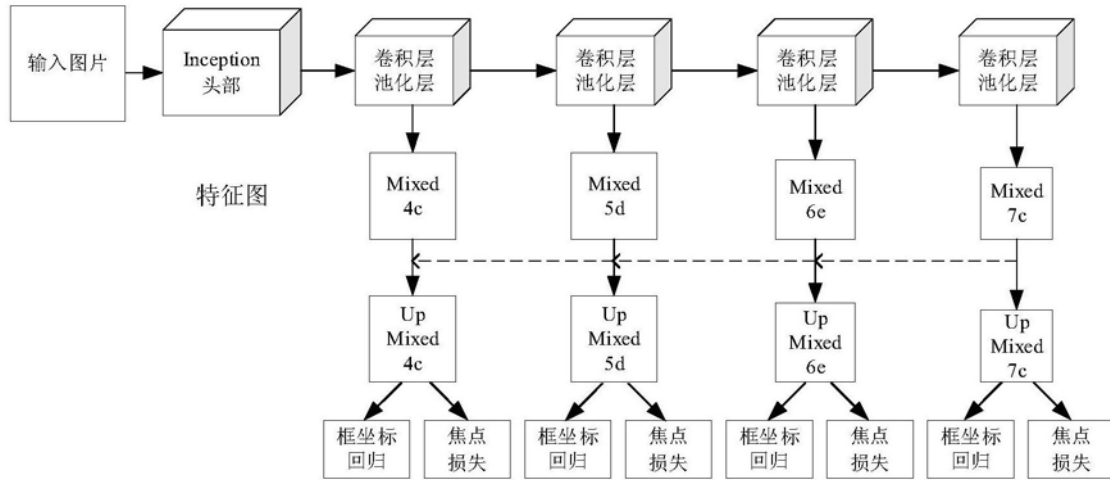


图4

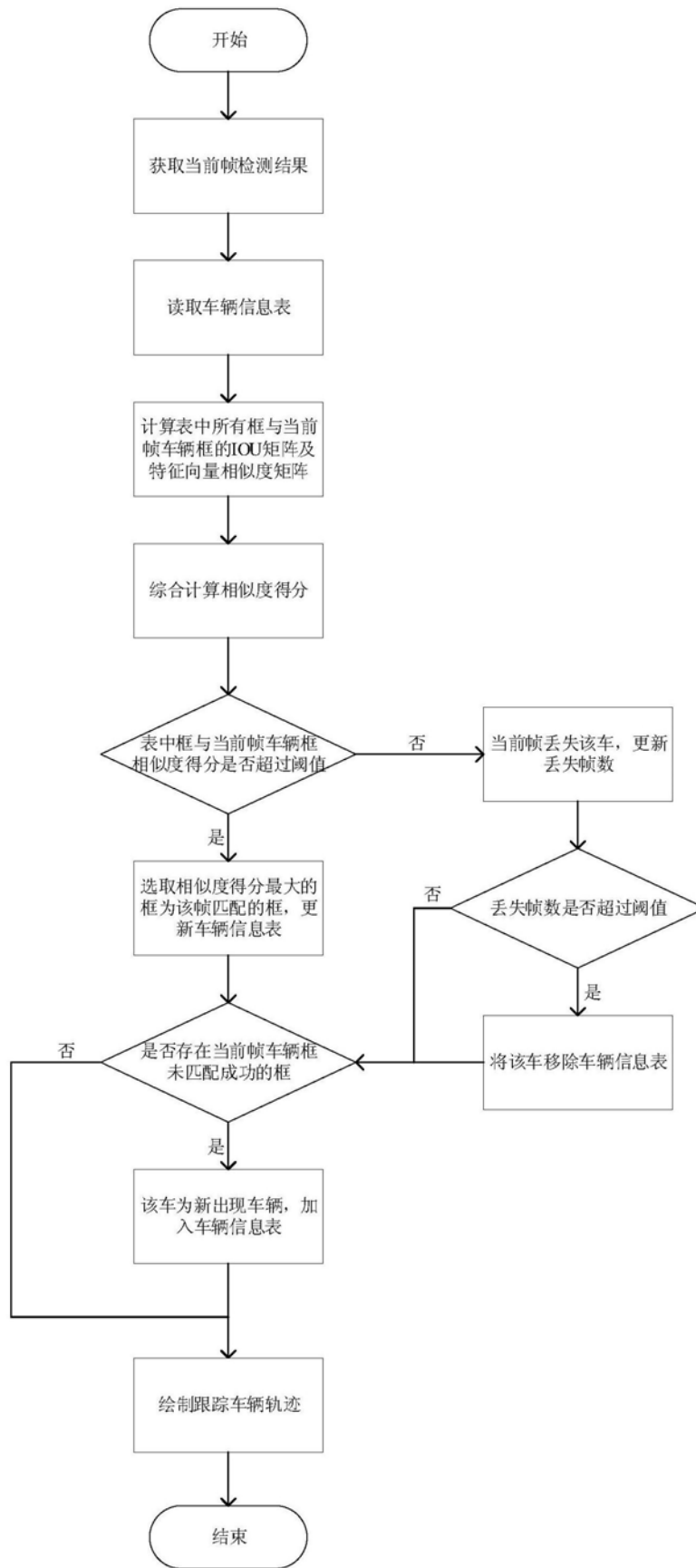


图5