

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 9 月 17 日 (17.09.2020)



(10) 国际公布号
WO 2020/181685 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2019/092749
- (22) 国际申请日: 2019 年 6 月 25 日 (25.06.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910185300.4 2019年3月12日 (12.03.2019) CN
- (71) 申请人: 南京邮电大学(NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) [CN/CN]; 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。
- (72) 发明人: 张登银(ZHANG, Dengyin); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。 金天宇(JIN, Tianyu); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。 丁飞(DING, Fei); 中国江苏省南京市鼓楼区新

模范马路 66 号, Jiangsu 210003 (CN)。 赵莎莎(ZHAO, Shasha); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。 刘锦(LIU, Jin); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。 薛睿(XUE, Rui); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。 聂涵(NIE, Han); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。 王雪纯(WANG, Xuechun); 中国江苏省南京市鼓楼区新模范马路 66 号, Jiangsu 210003 (CN)。

(74) 代理人: 南京纵横知识产权代理有限公司(NANJING ZONGHENG INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国江苏省南京市浦口高新区高新路 9 号自主创新广场 3 楼, Jiangsu 210032 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

(54) Title: VEHICLE-MOUNTED VIDEO TARGET DETECTION METHOD BASED ON DEEP LEARNING

(54) 发明名称: 一种基于深度学习的车载视频目标检测方法

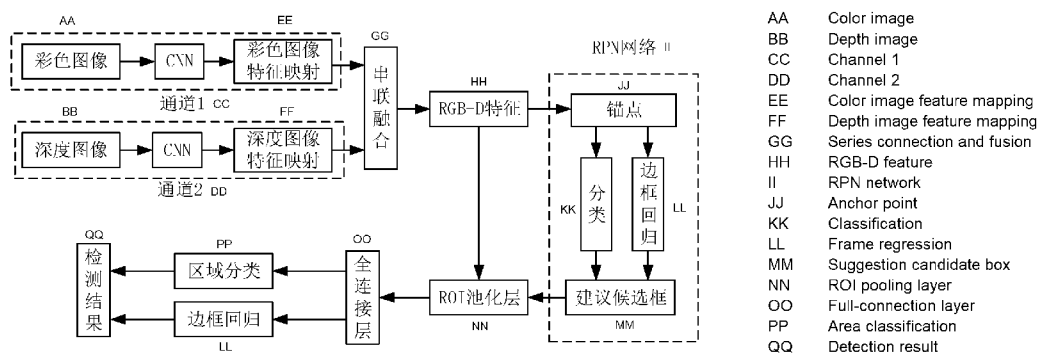


图 1

(57) Abstract: Disclosed is a vehicle-mounted video target detection method based on deep learning. The method employs an improved Faster R-CNN algorithm to implement target detection in a complex traffic environment, and to provide a driving safety auxiliary function. A serious problem of small target missed detection exists in an existing target tracking algorithm. According to the present invention, by adding a depth information channel, connecting the depth information channel to an original color image channel in parallel, performing fusion in a channel dimension, and performing candidate box extraction and target detection on a fused feature image, the detection rate of small targets is improved; in addition, by adding training of a difficult sample in training, the overall target recognition rate of the algorithm is improved. According to the present invention, the problem of small target missed detection existing in the Faster R-CNN algorithm can be fully considered, and the accuracy of vehicle recognition in a complex traffic scene is improved by means of depth image feature fusion and a difficult sample mining method.

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本发明公开了一种基于深度学习的车载视频目标检测方法, 利用改进的Faster R-CNN算法实现复杂交通环境中的目标检测, 提供行车安全辅助功能。现有的目标跟踪算法存在严重的小目标漏检问题, 本发明通过增加一个深度信息通道, 将其与原有的彩色图像通道并联, 并在通道维度上进行融合, 在融合后的特征图像上进行候选框提取和目标检测, 提高小目标的检测率, 另外在训练中添加对难样本的训练, 提高算法整体的目标识别率。本发明能够充分考虑Faster R-CNN算法存在的小目标漏检问题, 通过深度图像特征融合和难样本挖掘方法, 提高复杂交通场景中车辆识别的准确率。

一种基于深度学习的车载视频目标检测方法

技术领域

本发明涉及一种基于深度学习的车载视频目标检测方法，属于视频图像处理技术领域。

背景技术

在行车过程中，对车辆前方的车辆、行人及其他障碍物进行目标检测与跟踪，并在此基础上进行前车的行为分析，是行车安全辅助系统的基础。传统目标检测方法的主要步骤一般为：提取目标特征，训练对应的分类器，滑动窗口搜索，重复和误报过滤。这种目标检测的滑窗选择策略没有针对性，时间复杂度高，窗口冗余，且手工设计的特征鲁棒性较差，分类器不可靠；同时现有的目标检测算法不能够灵活地训练数据以根据不同需求学习有效的特征完成具体的检测任务。

发明内容

本发明的目的在于解决以上现有技术的不足，提供一种基于深度学习的车载视频目标检测方法。

为了实现上述目标，本发明采用如下的技术方案：

一种基于深度学习的车载视频目标检测方法，包括如下步骤：

步骤 1) 将深度坐标下的像素对齐到彩色坐标下；再将深度图像和彩色图像各自通过 CNN 进行特征提取，并将各自卷积层输出的特征图在通道维度上进行串联融合得到最终的 RGB-D 特征作为卷积后的卷积特征映射；

步骤 2) 构建区域建议网络 RPN, 所述区域建议网络 RPN 包括一个 3×3 的卷积层和两个 1×1 的并行卷积层; 将融合后的卷积特征映射输入 3×3 的卷积层, 在输入的特征映射上以像素为单位滑动预设大小的网络, 则每个滑动位置产生特定尺度的锚点;

将产生的锚点输入两个 1×1 的并行卷积层进行位置回归和前后景判断, 分别输出锚点的前后景置信度和所有候选框位置并按照预设条件从所得的矩形后选框中筛选出前景置信度最高的前特定数量的区域, 得到最终的区域建议集合 C ;

构建 Fast R-CNN 模型, 所述 Fast R-CNN 模型由两个 ROI 池化层、一个全连接层和两个并联的全连接层组成, 分别输出该区域的置信度以及边框回归之后的候选框位置; 将融合后的卷积特征输入 Fast R-CNN 模型, 输出图像中目标的位置及其类别和置信度;

步骤 3): 构建训练 RPN 网络的代价函数和训练 Fast R-CNN 网络的代价函数;

步骤 4) 使用标准的 ZF 模型训练和微调网络的各项参数, 通过从设定的标准方差的零均值高斯分布中提取权重来随机初始化所有新层;

步骤 5) 利用反向传播算法和随机梯度下降算法, 采用对 RPN 和 Fast R-CNN 两个网络交替训练的方式对模型进行训练, 根据预先设置的参数依次调整每层神经网络的权值;

步骤 6) 使用预先获得的训练集测试初步训练好的 Faster R-CNN 模型, 根据难样本的判别公式筛选出难样本;

步骤7) 将步骤6) 中产生的难样本加入训练集中, 对网络再次进行训练, 重复步骤5) -步骤7), 最终得到最优的Faster R-CNN模型;

步骤8) 对实际中采集的车载视频图像进行处理, 输入训练好的Faster R-CNN模型中, 输出该图像中目标类别、置信度以及目标位置。

本发明所达到的有益效果:

第一, 本发明在基于建议的卷积神经网络模型基础上, 提出一种基于深度信息补全的目标检测模型, 改进的 Faster R-CNN 添加了一条深度信息通道, 将彩色图像和深度图像分别通过结构一样的 CNN 进行特征提取, 两个 CNN 采用并联的结构, 再将原有的彩色图像特征映射与深度图像特征映射进行串联融合, 得到最终的图像特征, 与原有算法相比, 本发明得到的图像特征更加丰富, 补充了车辆的细节信息, 并且不会增加时间开销, 满足提高复杂场景下目标检测的要求。

第二, 本发明通过在训练阶段增加难样本挖掘策略, 使得模型在原有基础上更加关注难样本, 将车辆及疑似车辆的背景更好的区分, 达到提高准确性的目的;

第三, 本发明利用共享卷积网络来提取建议候选框的 Faster R-CNN 算法在实时性上得到了明显的提升, 该算法抛弃了传统的区域建议算法, 使用深度网络中的卷积层来提取候选框, 节约了大量时间开销。

附图说明

图 1 为本发明具体实施例方法的流程示意图。

图 2 为本发明具体实施例中改进的 Faster R-CN 算法的训练流程

图。

具体实施方式

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。

本发明目的在于提出了一种基于深度学习的车载视频目标检测方法，在 Faster R-CNN 的基础上，添加深度图像的特征映射来补充车辆细节信息，选用与提取彩色图像特征相同的卷积神经网络，彩色图像通道与深度图像通道采用并联的结构，提取出的特征经过串联融合得到最终的 RGB-D 特征，并在训练中添加难样本挖掘策略，提高算法在复杂交通场景中对小目标以及难目标的检测准确性。

如图 1 所示为本发明具体实施例的方法流程图。

在实施本发明方法时可以基于预先获取的训练集样本集和测试集样本，也可以根据需求制作训练集和测试集。在本实施例中，使用 KITTI 数据集构建训练样本集和测试样本集，包括以下步骤 1：采用 PASCAL VOC 数据集的格式及评价算法工具。首先，转换 KITTI 的类别：PASCAL VOC 总共有 20 个类别，在城市交通场景中，重点检测对象为车辆、行人、交通标志，因此将数据集分为上述 3 种类别；其次，转换标注信息：将标注文件从 txt 转化为 xml，去掉标注中的其他信息，只留下 3 类；最后，生成训练验证集和测试集。

如图 1 所述，本发明方法包括如下步骤：

步骤 2) 构建一种改进的 Faster R-CNN 模型，该模型综合了区域建议网络 (Regional Proposal Network, RPN) 和 Fast R-CNN 网

络；

2.1) 卷积神经网络 (Convolutional Neural Networks, CNN)

首先，将深度坐标下的像素对齐到彩色坐标下；CNN 选取 ZF 模型中特征提取的网络，将两个结构相同的 CNN 进行并联（原有的彩色图像通道为通道 1，并联的深度图像通道为通道 2）；两类图像通过 CNN 特征提取之后，特征图大小均为 hwc （这里 h 、 w 分别表示特征图的高、宽， c 为 RGB 三个通道），将彩色图像特征与深度图像特征作为两个通道进行特征融合，融合后的特征图大小为 $2hwc$ ；

2.2) 区域建议网络 RPN 包括一个 3×3 的卷积层和两个 1×1 的并行卷积层；

将融合后的卷积特征映射输入 3×3 的卷积层，在输入的特征映射上以像素为单位滑动一个小网络，本实施例中分别采用 3 个尺度和 3 个长宽比，则每个滑动位置产生 $k=3 \times 3=9$ 个不同尺度的锚点，则总共产生 hwk 个锚点，得到 hwk 个矩形候选框；

两个 1×1 的并行卷积层对上一层的锚点进行位置回归和前后景判断，分别输出锚点的前后景置信度和候选框位置。候选框位置包含候选框中心点坐标 x 和 y ，以及宽 w' 和高 h' 共四个参数；

2.2) 将 2.1) 所得的矩形候选框按照预设条件筛选出满足预设条件的预定数量的区域。本实施例中对所得的矩形候选框按照 softmax 的得分进行降序排序，保留前 2000 个区域，再进一步用非极大值抑制算法 (Non-Maximum Suppression, NMS) 筛选出前景置信度最高的前 300 个区域，得到最终的区域建议集合 C ；

2.3) Fast R-CNN 由两个 ROI 池化层、一个全连接层和两个并联的全连接层组成, 分别输出该区域的置信度以及边框回归之后的候选框位置;

ROI pooling 层对区域建议集合 C 与融合后的卷积特征映射进行池化操作, ROI pooling 层根据输入的 image, 将 ROI 映射到特征映射的对应位置, 将映射后的区域划分为相同大小的 sections, 对每个 section 进行最大池化操作;

全连接层对 ROI pooling 层的输出结果进行合并, 最后输入并联的两个全连接层, 对候选框进行区域分类和边框回归, 输出图像中目标的位置及其类别、置信度;

步骤 3) 构建训练 RPN 网络的代价函数和训练 Fast R-CNN 网络的代价函数:

本实施例中训练 RPN 网络的代价函数为:

$$L_{RPN}(\{P_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(P_i, P_i^*) + \lambda \frac{1}{N_{reg}} \sum_i P_i^* L_{reg}(t_i, t_i^*),$$

其中, 锚点与 ground truth(即 标定过的真实数据)的交并比

(Intersection over Union, IoU) 最大或不低于 0.7 的标为正样本, P_i 为预测置信度; P_i^* 为标注值, 取 1 时表示正样本, 取 0 时表示负样本; i 表示锚点的索引; N_{cls} 为锚点总数量; N_{reg} 为正样本的数量; t_i 为预测的锚点边界框修正值; t_i^* 为实际的锚点边界框修正值; L_{cls} 为分类代价; L_{reg} 为边框回归代价; λ 为平衡权重;

本实施例中训练 Fast R-CNN 网络的代价函数为:

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda [u \geq 1] L_{reg}(t^u, v)$$

其中， u 为第 u 类； t^u 为第 u 类边框回归预测的修正值； v 为实际修正值； L_{cls} 为分类代价； L_{reg} 为边框回归代价； λ 为平衡权重；

步骤 4) 使用标准的 ZF 模型训练和微调网络的各项参数，通过从设定的标准方差的零均值高斯分布中提取权重来随机初始化所有新层；

步骤 5) 利用反向传播算法和随机梯度下降算法，采用对 RPN 和 Fast R-CNN 两个网络交替训练的方式对模型进行训练，依次调整每层神经网络的权值，网络初始学习率设为 0.01，最低学习率设为 0.0001，动量设置为 0.9，权重衰减系数为 0.0005，Dropout 值设置为 0.5，具体步骤如下：

(1) 采用反向传播算法和随机梯度算法训练 RPN 模型，该阶段迭代 80000 次；

(2) 使用 RPN 生成的候选框作为 Fast R-CNN 的输入，并进行独立训练，该阶段迭代 40000 次；

(3) 利用 (2) 中的结果对 RPN 网络结构进行初始化，固定共享卷积层（将共享卷积层的学习率设置为 0），更新 RPN 网络的参数，该阶段迭代 80000 次；

(4) 固定共享卷积层（将共享卷积层的学习率设置为 0），微调 Fast R-CNN 网络结构，更新其全连接层的参数，该阶段迭代 40000 次；

步骤 6) 使用训练集测试初步训练好的 Faster R-CNN 模型，根据

本发明的难样本判别公式筛选出难样本；

步骤7) 将步骤6) 中产生的难样本加入训练集中，对网络再次进行训练，重复步骤5)，从而加强网络对难样本的判别能力，最终得到最优的Faster R-CNN模型；训练过程可参见图2。

步骤8) 对实际中采集的车载视频图像进行处理，输入训练好的Faster R-CNN模型中，输出该图像中目标类别、置信度以及目标位置。

本发明能够充分考虑 Faster R-CNN 算法存在的小目标漏检问题，通过深度图像特征融合和难样本挖掘方法，提高复杂交通场景中车辆识别的准确率。

本发明使用的基于卷积神经网络的目标检测算法可以在灵活的训练数据的情况下，根据不同需求学习有效的特征来完成具体的检测任务。R-CNN 算法是基于候选框建议与卷积神经网络相结合的一种目标检测算法，由于区域建议算法产生的数量众多的建议候选框以及较大的时间开销，该算法在实时性和准确性方面仍有较大的提升空间。利用共享卷积网络来提取建议候选框的 Faster R-CNN 算法在实时性上得到了明显的提升，该算法抛弃了传统的区域建议算法，使用深度网络中的卷积层来提取候选框，节约了大量时间开销，但在小目标较多以及较复杂的场景中，漏检的情况较为严重，仍有较大的改进空间。

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。

权利要求书

1. 一种基于深度学习的车载视频目标检测方法，其特征是，包括如下步骤：

步骤 1) 将深度坐标下的像素对齐到彩色坐标下；再将深度图像和彩色图像各自通过 CNN 进行特征提取，并将各自卷积层输出的特征图在通道维度上进行串联融合得到最终的 RGB-D 特征作为卷积后的卷积特征映射；

构建区域建议网络 RPN，所述区域建议网络 RPN 包括一个 3×3 的卷积层和两个 1×1 的并行卷积层；将融合后的卷积特征映射输入 3×3 的卷积层，在输入的特征映射上以像素为单位滑动预设大小的网络，则每个滑动位置产生特定尺度的锚点；

将产生的锚点输入两个 1×1 的并行卷积层进行位置回归和前后景判断，分别输出锚点的前后景置信度和所有候选框位置并按照预设条件从所得的矩形后选框中筛选满足特定条件的预设数量的区域，得到最终的区域建议集合 C；

步骤 2) 构建 Fast R-CNN 模型：

所述 Fast R-CNN 模型由两个 ROI 池化层、一个全连接层和两个并联的全连接层组成，分别输出该区域的置信度以及边框回归之后的候选框位置；将融合后的卷积特征输入 Fast R-CNN 模型，输出图像中目标的位置及其类别和置信度；

步骤 3)：构建训练 RPN 网络的代价函数和训练 Fast R-CNN 网络的代价函数；

步骤4) 使用标准的ZF模型训练和微调网络的各项参数, 通过从设定的标准方差的零均值高斯分布中提取权重来随机初始化所有新层;

步骤5) 利用反向传播算法和随机梯度下降算法, 采用对RPN和Fast R-CNN两个网络交替训练的方式对模型进行训练, 根据预先设置的参数依次调整每层神经网络的权值;

步骤6) 使用预先获得的训练集测试初步训练好的Faster R-CNN模型, 根据难样本的判别公式筛选出难样本;

步骤7) 将步骤6) 中产生的难样本加入训练集中, 对网络再次进行训练, 重复步骤5) -步骤7), 得到最优的Faster R-CNN模型;

步骤8) 对实际中采集的车载视频图像进行处理, 输入训练好的Faster R-CNN模型中, 输出该图像中目标类别、置信度以及目标位置。

2. 根据权利要求1所述的一种基于深度学习的车载视频目标检测方法, 其特征是, 所述步骤2) 中所述RGB-D特征作为RPN和Fast R-CNN共享的卷积特征映射, 其矩阵形式为:

$$Y_{merge}(i, j, k) = \begin{cases} Y_{RGB}(i, j, k), & 0 \leq k \leq c - 1 \\ Y_{depth}(i, j, k - c), & c \leq k \leq 2c - 1 \end{cases}$$

其中, i, j, k 为中间变量, $i \sim [0, h - 1]$, $j \sim [0, w - 1]$, $k \sim [0, 2c - 1]$, h 为特征图的高, w 为特征图的宽, c 为RGB三个通道; $Y_{RGB}(i, j, k)$ 是通道串联后的图像特征; $Y_{depth}(i, j, k - c)$ 是彩色图像特征; $Y_{merge}(i, j, k)$ 是深度图像特征。

3. 根据权利要求1所述的一种基于深度学习的车载视频目标检

测方法，其特征是，所述训练 RPN 网络的代价函数为：

$$L_{RPN}(\{P_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(P_i, P_i^*) + \lambda \frac{1}{N_{reg}} \sum_i P_i^* L_{reg}(t_i, t_i^*),$$

其中，将与标定过的真实数据的交并比最大或不低于 0.7 的锚点标为正样本， P_i 为预测置信度； P_i^* 为标注值，取 1 时表示正样本，取 0 时表示负样本； i 表示锚点的索引； N_{cls} 为锚点总数量； N_{reg} 为正样本的数量； t_i 为预测的锚点边界框修正值； t_i^* 为实际的锚点边界框修正值； L_{cls} 为分类代价； L_{reg} 为边框回归代价； λ 为平衡权重。

4. 根据权利要求 1 所述的一种基于深度学习的车载视频目标检测方法，其特征是，所述训练 Fast R-CNN 网络的代价函数为：

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda [u \geq 1] L_{reg}(t^u, v)$$

其中， u 为第 u 类； t^u 为第 u 类边框回归预测的修正值； v 为实际修正值； L_{reg} 为边框回归代价， p 是分类预测结果。

5. 根据权利要求 1 所述的一种基于深度学习的车载视频目标检测方法，其特征是，所述步骤 6) 中，所述难样本判别函数如下：

$$L(o, p) = L_{IoU}(o) + L_{score}(p),$$

$$L_{IoU}(o) = \frac{1}{1 + e^{k \times (o - 0.5)}},$$

$$L_{score}(p) = (1 - p),$$

其中， L_{IoU} 为边框误差； L_{score} 为分类误差； o 为样本与目标的相交率；

k 为对阈值的敏感系数； o 和 p 的取值范围均为 $0 \sim 1$ 。

6. 根据权利要求 1 所述的一种基于深度学习的车载视频目标检测方法，其特征是，步骤 5) 的具体步骤如下：

(1) 采用反向传播算法和随机梯度算法训练RPN模型，该阶段迭代80000次；

(2) 使用RPN生成的候选框作为Fast R-CNN的输入，并进行独立训练，该阶段迭代40000次；

(3) 利用(2)中的结果对RPN网络结构进行初始化，固定共享卷积层（将共享卷积层的学习率设置为0），更新RPN网络的参数，该阶段迭代80000次；

(4) 固定共享卷积层（将共享卷积层的学习率设置为0），微调Fast R-CNN网络结构，更新其全连接层的参数，该阶段迭代40000次。

7.根据权利要求1所述的一种基于深度学习的车载视频目标检测方法，其特征是，步骤5)中参数设定包括网络初始学习率设为0.01，最低学习率设为0.0001，动量设置为0.9，权重衰减系数为0.0005，Dropout值设置为0.5。

8. 根据权利要求1所述的一种基于深度学习的车载视频目标检测方法，其特征是，预先获得训练集和测试集的方法包括如下步骤：

使用KITTI数据集构建训练样本集和测试样本集；

按照PASCAL VOC格式转换KITTI的类别，将KITTI数据集分为车辆、行人和交通3种类别；

转换标注信息：将标注文件从txt转化为xml，去掉标注中的其他信息，只留下3类；最后，生成训练验证集和测试集。

9. 根据权利要求1所述的一种基于深度学习的车载视频目标检

测方法，其特征是，筛选出满足特定条件的预设数量的区域的方法如下：

将所得的矩形候选框按照 softmax 的得分进行降序排序，保留前 2000 个区域，再进一步用非极大值抑制算法筛选出前景置信度最高的前特定数量的区域。

10.根据权利要求 1 所述的一种基于深度学习的车载视频目标检测方法，其特征是，步骤 4) 所述设定的标准方差为 0.01。

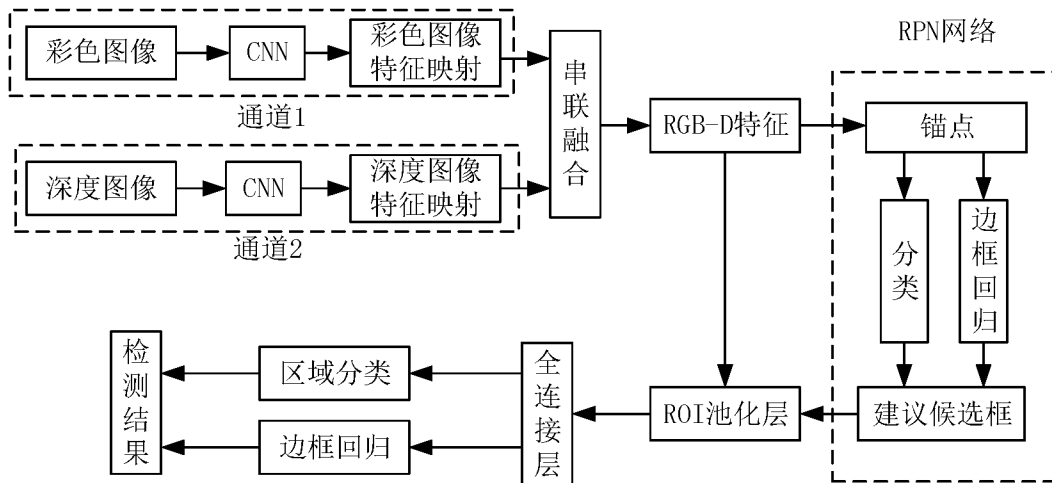


图 1

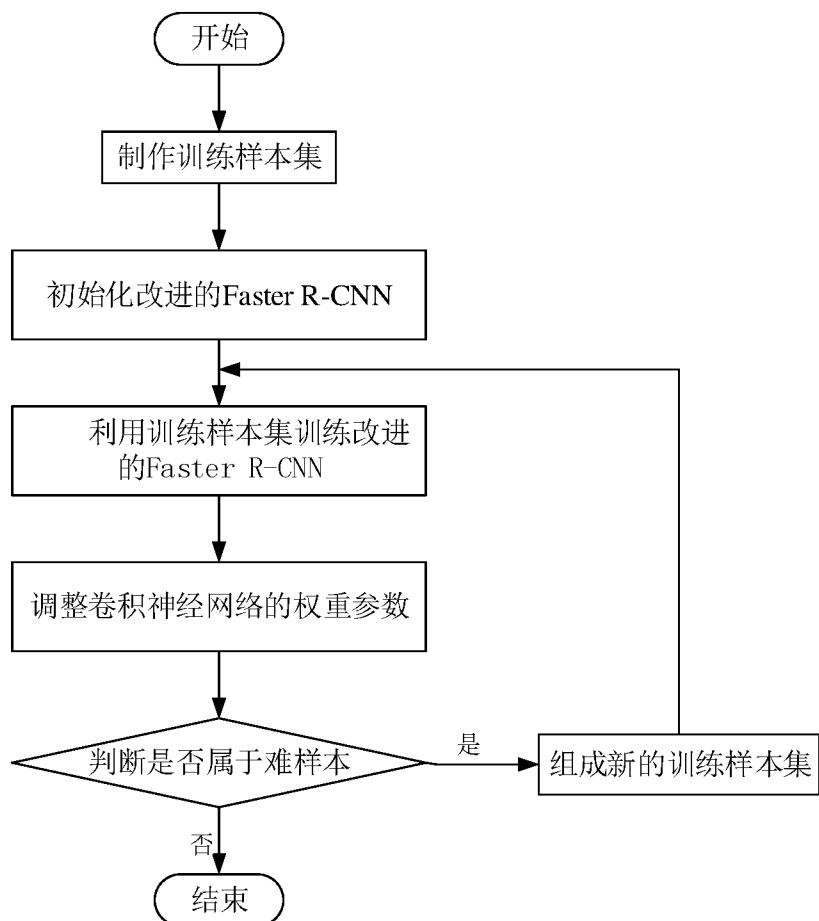


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/092749

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT; CNABS; VEN; CNKI: 深度学习, 卷积, 快速, 视频, 检测, 彩色, 车, 特征, 锚, 连接层, RPN, ROI, deep w learning, CNN, fast, video, detect, color, car, character, anchor, connection w layer

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 108563977 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) 21 September 2018 (2018-09-21) description, paragraphs [0002]-[0179]	1-10
A	CN 107507172 A (STATE GRID SHANGHAI MUNICIPAL ELECTRIC POWER COMPANY) 22 December 2017 (2017-12-22) entire document	1-10
A	CN 107680106 A (NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS) 09 February 2018 (2018-02-09) entire document	1-10
A	谷頔 (GU, Qi). "基于深度学习的全天车辆检测 (All Day Vehicle Detection Based on Deep Learning)" 中国优秀硕士学位论文全文数据库 (Chinese Master's Theses Full-text Database), No. 6, 15 June 2018 (2018-06-15), entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

05 December 2019

Date of mailing of the international search report

13 December 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/092749

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	108563977	A	21 September 2018	None	
CN	107507172	A	22 December 2017	None	
CN	107680106	A	09 February 2018	None	

国际检索报告

国际申请号

PCT/CN2019/092749

A. 主题的分类 G06K 9/00 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																	
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNTXT; CNABS; VEN; CNKI: 深度学习, 卷积, 快速, 视频, 检测, 彩色, 车, 特征, 锚, 连接层, RPN, ROI, deep w learning, CNN, fast, video, detect, color, car, character, anchor, connection w layer																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 108563977 A (华南理工大学) 2018年 9月 21日 (2018 - 09 - 21) 说明书第2-179段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 107507172 A (国网上海市电力公司) 2017年 12月 22日 (2017 - 12 - 22) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 107680106 A (南京航空航天大学) 2018年 2月 9日 (2018 - 02 - 09) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>谷颢, “基于深度学习的全天车辆检测” 中国优秀硕士学位论文全文数据库, 第6期, 2018年 6月 15日 (2018 - 06 - 15), 全文</td> <td>1-10</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 108563977 A (华南理工大学) 2018年 9月 21日 (2018 - 09 - 21) 说明书第2-179段	1-10	A	CN 107507172 A (国网上海市电力公司) 2017年 12月 22日 (2017 - 12 - 22) 全文	1-10	A	CN 107680106 A (南京航空航天大学) 2018年 2月 9日 (2018 - 02 - 09) 全文	1-10	A	谷颢, “基于深度学习的全天车辆检测” 中国优秀硕士学位论文全文数据库, 第6期, 2018年 6月 15日 (2018 - 06 - 15), 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 108563977 A (华南理工大学) 2018年 9月 21日 (2018 - 09 - 21) 说明书第2-179段	1-10															
A	CN 107507172 A (国网上海市电力公司) 2017年 12月 22日 (2017 - 12 - 22) 全文	1-10															
A	CN 107680106 A (南京航空航天大学) 2018年 2月 9日 (2018 - 02 - 09) 全文	1-10															
A	谷颢, “基于深度学习的全天车辆检测” 中国优秀硕士学位论文全文数据库, 第6期, 2018年 6月 15日 (2018 - 06 - 15), 全文	1-10															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “D” 申请人在国际申请中引证的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																	
国际检索实际完成的日期 2019年 12月 5日		国际检索报告邮寄日期 2019年 12月 13日															
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 李宁 电话号码 62412342															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/092749

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	108563977	A	2018年 9月 21日	无	
CN	107507172	A	2017年 12月 22日	无	
CN	107680106	A	2018年 2月 9日	无	