

基于多任务 Faster R-CNN 车辆假牌套牌的检测方法*

陈朋¹, 汤一平^{1,2}, 何霞¹, 王辉², 袁公萍¹

(1. 浙江工业大学信息工程学院 杭州 310023; 2. 银江股份有限公司 杭州 310000)

摘要:针对现有车辆假牌套牌各种检测方法存在计算复杂度高、检测精度低、鲁棒性欠缺等问题,提出一种基于多任务的高速区域卷积神经网络(Faster R-CNN)车辆假牌套牌的检测方法。首先利用时空约束得到疑似套牌车辆,接着用Faster R-CNN定位分割出车辆前脸部分图像,然后对疑似套牌车辆的车脸公脸部分(车辆的基本特征)的特征进行比对;在此基础上再对高仿套牌车辆的车脸私脸部分(车检标)的细微特征进行检测比对。这种分层次的、从车辆宏观特征到微观特征的视觉检测方法,具有检测速度快、鲁棒性高、泛化能力强、实施部署方便、检测精度高等优点。实验研究表明,在Vehicle ID数据集和杭州卡口数据集中分别取得了99.39%、99.22%的检测精度。

关键词: 车辆假牌套牌检测;多任务高速区域卷积神经网络;车辆脸部特征;分层特征比对

中图分类号: TP391 TH-39 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Detection method for fake plate vehicles based on multitask Faster R-CNN

Chen Peng¹, Tang Yiping^{1,2}, He Xia¹, Wang Hui², Yuan Gongping¹

(1. School of Information Engineering, Zhejiang University of Technology, Hangzhou, 310023, China;

2. Enjoyor Co., Ltd, Hangzhou 310000, China)

Abstract: Current vehicle detection method for fake plate vehicles has a high computational complexity, low detection accuracy, lack of robustness. This paper presents of fake plate vehicles detection method based on multitask Faster R-CNN (region-based convolutional neural network). Firstly, spatio-temporal constraint is used to obtain the suspected fake plate vehicle. Then, front part of the vehicle is located in the image using Faster R-CNN. Next, the public face (basic characteristics of a vehicle) of suspicious fake plate vehicles is contrasted. In further, the subtle features of a private face (Annual inspection certificate for vehicles) is contrasted. This hierarchical visual inspection method, detected from macroscopic features of vehicles to microscopic features, has the advantages of fast detection speed, high robustness, strong generalization ability, convenient deployment and high detection precision. Experimental results show that detection accuracy are 99.39% and 99.22% on the Vehicle ID data set and the Hangzhou bayonet data set, respectively.

Keywords: fake plate vehicles detection; multitask faster region-based convolutional neural network (Faster R-CNN); vehicle facial features; hierarchical feature matching

0 引言

随着汽车的保有量的急剧增长,所带来一系列交通问题也日益突出,套牌、假牌等机动车违法乱纪现象更是屡禁不止,随之带来的交通事故、法律纠纷势、走私犯罪给社会造成极大的不稳定因素。虽然公安、交管等部门

花费大量人力财力进行整治,但是收效甚微。目前对车辆套牌车的检测研究中主要有如下两大类方法。

一类是先对查询车辆进行车牌号识别^[1-2],通过车牌号从数据库中读取车型信息,然后与检测识别到的车型做比对,从而判断是否为套牌车。文献[3]提出了一种多特征组合的车脸特征匹配方法,通过对车脸特征细微差别的强化实现了一种新的车辆识别方法,这种舍弃特

征单一冗余的车体区域,强调特征丰富的车脸区域,对外观相似度较高的套牌车检测具有一定的意义。文献[4]将统一放缩后的输入图像分割成16个模块组件,分别利用方向梯度直方图(histogram of oriented gradient, HOG)进行特征提取,最后利用支持向量机(support vector machine, SVM)进行分类。由于引入了机器学习模型在分类精度上有了较大的提高,但是由于HOG方法对噪声敏感,在环境复杂、车外观相似的条件很难获取到真正有区分性的特征。文献[5]利用区域卷积神经网络(regions with convolutional neural network, R-CNN)模型对车体不同的区域组件分别进行位置的检测,然后通过卷积神经网络(convolutional neural network, CNN)进行各组件的特征提取最后实现融合,由于涉及到多个模型、多个车体组件,因此该方法存在着较高的计算复杂度。此外,上述方法都是基于识别分类后进行套牌车的鉴定,对于高仿真车辆和从未训练过的车辆,这类方法将失效。

另一类套牌识别方法,主要是依托大数据技术,通过判断同一车牌号车辆在不同监控点视频中出现的时间关系,结合不同监控点的位置关系,根据其时间空间逻辑一致性原则来判断是否为套牌车。文献[6]提出了一种交互式套牌车可视识别分析系统,通过设定相应的规则和筛选条件对卡口数据进行分析得出套牌车检测结果。文献[7]从时间、空间和属性3个部分对路口交通数据进行了分析。文献[8]提出了一种基于融合策略的套牌车识别方法,通过时空特性的约束,充分融合路径、速度、交通状况等多种信息对疑似车辆进行可信度评估,从而确定其中是否存在套牌。但是,受地理位置、车速流量的影响,这类方法时空约束复杂、数据量大、车速估计不精准,要想实现对套牌车的精确锁定,仍存在较大的难度,实用性欠缺。

套牌车检测的难点在于同一车牌号同一款型的高仿车的检测。鉴于此,本文将车脸进一步分为公有脸和私有脸两个部分。

所谓公有脸是指车辆颜色、车辆类型、车辆款式、车辆大小等一系列车体宏观特征。这些属性特征相对容易模仿。所谓私有脸主要包括车检标、车挂件、摆件等一系列人为添加的车脸属性特征,车检标的粘贴国家有明文规定,《中华人民共和国道路交通安全法》第95条和《中华人民共和国道路交通安全法实施条例》第13条都明确规定,道路行驶的机动车必须放置检验合格标志、保险标志,并且应当粘贴在机动车前窗右上角^[9-10]。车检标完全是由人工粘贴,并且相对固定;所以每辆车的车检标在数量上的、顺序上的、间隔上都会存在着一些细微差异。然而根据这些细微差异来甄别高仿套牌假牌车辆是具有一定挑战性的研究工作。

要比较车脸上车检标的细微差异,完全取决于车辆图像的分割定位以及车检标图像的分割定位。在深度卷

积神经网络技术出现以前,要从抓拍的图像中准确定位分割出车辆图像以及车检标图像是一项极其困难的事情^[11]。

近年来深度学习的研究取得了重大突破,从R-CNN^[12]到Fast R-CNN^[13]再到Faster R-CNN^[14],具体表现为进一步精度提升、加速、端到端及更实用化,几乎覆盖了从分类到检测、分割、定位各个领域。

一种优秀的套牌假牌车的检测方法,必须同时兼顾实时性、鲁棒性和准确性。为此,本文提出了一种基于多任务Faster R-CNN车辆假牌套牌的检测方法。让车检标分割定位、车辆公有脸的特征提取和车辆私有脸的特征提取都共享一个Faster R-CNN网络,这样能显著降低整个检测任务的复杂度,同时利用时空约束条件和逐层比对筛选的方式降低了计算量的同时,也有效提升了检测的实时性。

本文在现有Faster R-CNN的基础上,设计了一种能同时进行多任务、多层次特征提取功能的网络,利用共享卷积层,实现了网络多任务联合训练,联合特征提取的功能。提出了一种分层次的、从车辆的宏观特征到微观特征套牌假牌车视觉检测方法,兼顾了检测的实时性、鲁棒性和准确性,能有效检测出高仿套牌车。利用有限的数据集对网络进行训练,通过比对车脸的公有脸和私有脸的特征向量来甄别假牌套牌车辆,使得系统具备较好的泛化能力,实现对未训练车辆类型的特征提取与比对。

1 多任务Faster R-CNN车辆假牌检测整体架构

为了降低车辆假牌套牌的计算复杂度,提高检测精度,增强检测的鲁棒性,本文提出了一种分层次的、从车辆的宏观特征到微观特征假牌套牌车视觉检测方法。根据时空约束条件得到在某一个区域范围内(区域范围可以由用户设定)同一个的车牌号的疑似假牌套牌车辆图像。首先,从这些图像中分割定位出车辆对象部分图像,然后对这些车辆的宏观特征(车脸的公有部分)进行比对,稽查出低仿真的假牌套牌车辆;剩余的还需通过车辆的微观特征(车脸的私有部分)比对,进一步排查和比对高仿真的假牌套牌车辆,当然前提是能从车辆对象部分图像中精准的分割定位出车检标对象部分图像。

整体设计框架如图1所示,主要包括3个部分。第1部分实现了车辆的分割定位,使得能够从道路背景环境中分割出车辆图像;第2部分对应图1中的A区域,该部分主要实现车辆整体基本特征(车辆公有脸)的提取;第3部分对应图1中的B区域,该部分完成了车检标的检测定位和特征提取即车辆私有脸特征的提取。整个网络架构中,A和B两部分通过共用卷积层实现卷积权值共享,此外,两部分也有各自独立的全连接层和损失层。

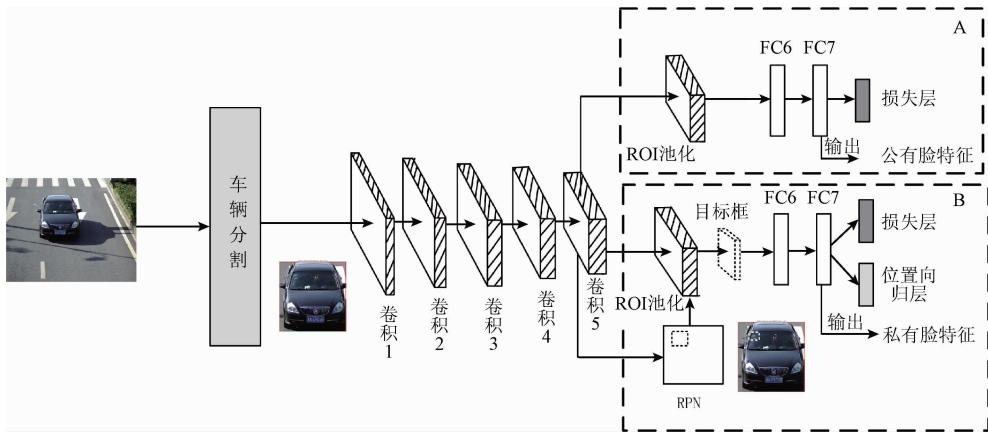


图 1 多任务卷积神经网络结构

Fig. 1 Structure of multitask convolutional neural networks

1.1 车辆对象的分割

通过时空约束条件提取到的同一车牌号的疑似套牌车辆包含了复杂的道路背景环境,如何从这些不确定的环境中精准地分割出车辆目标区域是极其重要的一步,传统的图像分割方法多数是基于区域、边缘的分割方法^[15-16]。Faster-RCNN 本身具备良好的目标定位分割检测功能,一般条件下可以直接从背景中检测出目标区域实现分割;但是直接从这种大范围背景下同时对车辆对象和车检标对象进行定位分割显然会影响到检测的效率与精度,因此本文通过网络级联方式来实现目标区域的分层次的定位分割。首先实现车辆的分割定位,在此基础上再进行车检标的分割定位,这种分层次的区域分割相对容易实现,且具有较高的鲁棒性。对于车辆对象的分割现有的基于深度神经网络的方法已实现了较高的精度,具备了较高的实时性,参见文献[17-18]。由于本文主要关注车辆的宏观特征车辆基本特征(公有脸)和微观特征车检标(私有脸)的比对,关于车辆分割定位的详细内容篇幅关系不再赘述。

1.2 多任务检测网络设计

车辆的宏观特征车辆基本特征(公有脸)和微观特征车检标(私有脸)的比对主要依托于对这些特征的有效提取,这两个任务间组成了一个典型的多任务学习检测机制。

多任务学习(multitask learning, MTL)是一种归纳迁移机制,基本目标是提高泛化性能。在多个任务共享卷积层的基础上同时保持损失层独立。这种相互制约相互协同的学习方式,提高了系统的泛化能力,进而改善了因样本不足而造成的系统泛化能力弱的问题。在细粒度分类^[19]、面部关键点检测^[20]等领域都取得了不错的效果。此外,多个任务的同时进行在很大程度上提高了网络的处理效率,因此本文将利用该网络实现多层次特征的并行提取。

如图 1 所示,分割定位好车辆对象后,进入共享卷积层,由于这部分的设计和 Faster R-CNN 的卷积层一样,因此这里只做简单介绍。以 VGG16^[21]的卷积层为例,首先输入一张任意尺寸的图片,通过第一部分的卷积 1_1,卷积 1_2 两次卷积后输出 64 个特征图经过最大池化输出给第 2 部分的卷积层;同样第二部分卷积经过卷积 2_1,卷积 2_2 后得到 128 个特征图;以此类推,最后到第 5 部分的卷积 5_3 后输出 512 个特征图,到此为止网络完成了所有的卷积操作。通过如此的层层卷积池化使得输入图像一步步抽象,从而在最大化保留图片特征信息的同时减小了数据的计算量,提高了网络的泛化能力。

共享卷积层后,网络开始出现分支。图像经过 5 部分卷积后生成的特征图首先进入 A 区域,该区域主要实现车辆公有脸特征向量的提取。这一部分主要包括兴趣区域(region of interest, ROI)池化层、全连接层(fully connected layers, FC)以及损失层。由于输入的图像尺寸不固定而最后全连接层的输出的维数唯一,因此必须在 FC6 层前实现特征图大小的统一,所以在 FC6 层前增加了 ROI 池化层。所谓的 ROI 池化就是一种专门针对感兴趣区域的池化,由于 A 区域的主要功能是为了完成车辆整体基本特征的提取,因此 ROI 的坐标范围直接给定整幅车辆对象图片。要获取对应特征图上的坐标,首先要将 ROI 中的坐标映射到卷积后的特征图上,映射规则就是把各个坐标除以输入图片尺寸与特征图尺寸的比值,得到特征图上的包围框(bounding box, Bbox)坐标。然而由于输入图片大小并不是唯一的,所以这里需要对 Bbox 区域使用空间金字塔池化处理^[22]。该方法可以把任何尺度图像的卷积特征转化成相同的维度,从而满足输出维数的一致性。

网络的另一个分支是 B 区域,该部分主要实现了车

检标的位置检测和特征提取。和前面的介绍相类似,该部分也包含了一个 ROI 池化层,不同的是,该部分的 ROI 已经不再是车辆对象的整体区域,而是车检标对象的区域;因此需要借助区域建议网络(region proposal network, RPN)^[14]产生车检标的建议框(见图1RPN中虚线框),再通过后面的位置回归不断对这些建议框坐标做出新的调整,使得矩形框位置一步步接近真实区域位置(车检标对象位置)。RPN网络将一个任意尺度图像作为输入,输出矩形目标建议框的集合,每个框包含4个位置坐标变量和一个得分变量。为了生成区域建议框,在卷积5_3生成的特征图上进行多尺度的卷积操作,在每一个滑动窗口的位置使用3种尺度和3种长宽比的卷积核。卷积时以当前滑动窗口中心为中心,并对应一种尺度和长宽比的卷积核,在原图上映射得到9种不同尺度的候选区域,如对于长宽为 $w \times h$ 的特征图,则总共有 $w \times h \times 9$ 个候选区域。最后分类层输出 $w \times h \times 9 \times 2$ 个候选区域的得分,即对每个区域是目标/非目标两种情况的估计概率,回归层输出 $w \times h \times 9 \times 4$ 个参数,即候选区域的坐标参数。

1.3 多任务网络训练

相对单任务网络多任务网络之间的约束更加复杂,同时各个任务训练时收敛的速度也存在较大的差异,因此训练的难度会更大,对参数的合理性要求会更高。训练部分主要包含RPN网络的训练、特征提取网络的训练。其中特征提取网络又包含了车辆基本特征提取网络和车检标特征提取网络,在训练时这两个网络存在密切的联系。

训练RPN网络时,给每个候选区域分配一个二进制的标签,以此来标注该区域是不是目标。对于正标签候选区域规定为任意某个真正目标区域(ground truth, GT)包围框有大于0.7的交并比(intersection over union, IoU)。对于负标签候选区域规定为GT包围框的IoU比率低于0.3。对于IoU介于两者之间的统一舍弃。

训练车检标特征提取的网络包含了分类损失函数和位置回归损失函数,分别对应分类任务和检测任务。因此,总的损失函数用式(1)表示。

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (1)$$

分类损失函数由式(2)表示。

$$L_{\text{cls}}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (2)$$

区域回归损失函数由式(3)定义。

$$L_{\text{reg}} = (t_i, t_i^*) = R(t_i - t_i^*) \quad (3)$$

其中, R 是鲁棒的损失函数 smooth_{L1} ,用式(4)表示。

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (4)$$

式中: p_i 是候选区域是第 i 类的概率。若候选区域的标签为正样本(目标区域) $p_i^* = 1$,如果候选区域标签为负样本(背景区域) $p_i^* = 0$ 。 t_i 向量表示预测的包围框的4个参数化坐标, t_i^* 是对应的GT包围框的坐标向量。 N_{cls} 和 N_{reg} 分别为分类损失函数与位置回归损失函数的归一化系数, λ 为两者之间的权重参数。

由于车辆整体基本特征提取网络中目标区域坐标已知(整张车辆对象图片),因此在训练时不存在位置回归任务,只需要定义一个分类损失函数softmax,由式(5)给出。 P 由式(6)给出。

$$L(\theta) = -\frac{1}{N_b} \left[\sum_{i=1}^m \sum_j^k \mathbb{1}\{y^{(i)} = j\} \log p(y^{(i)} = j | \mathbf{x}^{(i)}; \theta) \right] \quad (5)$$

$$p(y^{(i)} = j | \mathbf{x}^{(i)}; \theta) = \frac{e^{\theta_j^T \mathbf{x}^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T \mathbf{x}^{(i)}}} \quad (6)$$

式中: θ 是训练模型参数, N_b 表示每批次训练样本的个数, $y^{(i)}$ 表示类别标签, $\mathbf{x}^{(i)}$ 表示输入特征。其中,概率 p 表示将 x 分为类别 j 的概率。

结合多任务网络总损失函数可以表示为:

$$L(w^{(i)}) = L(\{p_i\}, \{t_i\}) + \lambda' L(\theta) \quad (7)$$

式中: λ' 为多任务的损失权重,主要为了解决不同网络任务之间收敛不一致的问题。

1.4 分层次特征提取与比对

车辆特征比对规则的设计直接影响到最后输出的精度,能否充分利用已提取得到的车辆宏观特征及微观特征进行分层次特征比对是提高甄别假牌套牌车辆精度的关键。

本文利用是车辆A和车辆B的(公有脸和私有脸)特征向量之间的夹角余弦实现逐层比对,几何中夹角余弦可用来衡量两个向量方向的差异,夹角余弦越大表示两个向量的夹角越小,夹角余弦越小表示两向量的夹角越大,该方法在图片的比对和人脸识别验证上都得到了广泛的应用^[23]。式(8)为比较两个特征向量的计算方法。

$$\cos(\theta) = \frac{\mathbf{AB}}{|\mathbf{A}| |\mathbf{B}|} \quad (8)$$

式中: \mathbf{A} 为车辆A的(公有脸和私有脸)特征向量, \mathbf{B} 为车辆B的(公有脸和私有脸)特征向量, $\cos(\theta)$ 为车辆A和车辆B的(公有脸和私有脸)特征向量之间的夹角余弦。在比对过程中,两个特征向量之间的夹角余弦值大于某个阈值就判定为车辆A和车辆B具有相似性。

在测试过程中,通过多任务网络分别获取疑似套牌

车辆的公有脸、私有脸检测网络的 FC7 层的特征向量。全连接层的一个作用是实现高维到低维度的变化,同时保留了有用的信息;另一个作用是隐含语义的表达,把原始特征映射到各个隐语义节点。对于最后一层全连接而言,就是分类的显示表达,由于本文只进行特征的比对,因此无需将 FC7 层映射到相应的类别上,而只进行特征的抽取。

提取特征向量后,首先用式(8)对公有脸进行余弦距离的计算得到比对疑似套牌车辆的相似度,如果相似度小于设定的第1个阈值,则认为基于公有脸的比对方法已经能够判定两辆车身份不一致,直接输出比对假牌套牌结果;否则,进一步排查高仿套牌车,用式(8)对车辆私有脸进行余弦距离的计算得到比对疑似套牌车辆的相似度。如果经过私有脸比对的相似度小于设定的第2个阈值则认为两辆车身份不一致,即为高仿套牌车,大于设定的第2个阈值则判定为两幅车辆图像为同一辆车。

在实际情况中多数套牌车公有脸都存在一定的差别,少量外观高度一致的高仿套牌车在车检标上也存在较为明显的差异,因此通过这种层次性的比对方式可以满足所有套牌车快速排查的要求。

最终,系统判定输出结果表达式如下,式(9)表示疑似套牌车辆的公有脸的比对输出结果,式(10)表示疑似套牌车辆的私有脸的比对输出结果。

$$O = \begin{cases} 0, & s_1 < t_1 \\ p, & s_1 \geq t_1 \end{cases} \quad (9)$$

$$p = \begin{cases} 0, & s_2 < t_2 \\ 1, & s_2 \geq t_2 \end{cases} \quad (10)$$

式中: O 表示系统的最终判定的输出结果,1 表示两幅图像中的车辆为同一辆车,0 表示两幅图像中的车辆为非同一辆车, s_1 、 s_2 分别为两幅车辆图像中的车辆公有脸和私有脸特征向量比对的相似值, t_1 为公有脸比对时所设定的相似度阈值, t_2 为私有脸比对时所设定的相似度阈值。

通过上述分层次的车辆特征比对能有效的检测出套牌假牌车辆,但是最终要确定那一辆是套牌假牌车辆还需要车主的配合。只要真车主将自己车辆上的年检标拍摄图像传输给管理部门,管理部门利用本文研制的私有脸比对方法进行检测,就能快速地锁定套牌或假牌车辆。

2 实验结果与分析

主要涉及硬件环境的介绍、训练和测试数据集的介绍以及评价指标的设定等内容,最终通过大量的测试数据对本文的提出的方法和常用方向梯度直方图(histogram of oriented gradient, HOG)、尺度不变特征变换

(scale-invariant feature transform, SIFT)、CNN 等车辆特征提取的方法进行了对比分析。

为了评估基于多任务 Faster R-CNN 车辆假牌套牌的检测方法的检测精度,本文分别在 Vehicle ID^[24] 数据集和杭州卡口数据集上进行了实验验证。

2.1 实验环境

在深度学习框架的选择与搭建上,目前可选择的环境有 Caffe、Torch、Theano 和 TensorFlow。考虑到到 Caffe 环境与 Python 语言的绑定,具有良好训练或调整前馈模型,Fast R-CNN 和 Faster R-CNN 内部有 Caffe,便于配置等优点,本文最终选择了 Caffe 环境。实验环境如表 1 所示。

表 1 实验环境

Table 1 Experimental environment

名称	参数
系统	LINUX64 Ubuntu14.04
框架	Caffe
语言	Python, C++
CPU	Intel® Core™ i7-2600 3.4 GHz × 8
GPU	TITAN X (Pascal)
RAM	8.00 GB

2.2 数据集

本文主要在 Vehicle ID 数据集和杭州卡口数据集进行了相关的训练和测试。Vehicle ID 数据集包含了 26 267 辆车,每辆车包含的图片张数不固定,总共包含 221 763 张图片。图片采集的角度、光照大小比较统一,如图 2(a) 所示。为了更好地考察模型的泛化能力,本文未进行大规模数据集的训练而是选取了其中部分图片张数较多的车辆类,通过剔除车辆非正面图片,最终获得 1 725 辆车 19 206 张图片的数据集。



图 2 数据集示例

Fig. 2 Examples of dataset

杭州卡口数据集主要来源杭州各个卡口不同时间段高清摄像头抓拍的图片,经过分割归类,最终包含 621 辆车 8 215 张图片。由于来自实地的拍摄,相对标准数据集在光照、角度、清晰度都存在明显的差异,如图 2(b) 所示。

训练集测试集的划分本文遵循如下约定:对于 Vehicle ID 数据集训练集设定 1 215 辆车 15 935 张车辆图片,测试集设定 510 辆车 3 271 张车辆图片。对于卡口数据集训练集设定 470 辆车 6 312 张车辆图片,测试集设定 151 辆车 1 903 张图片。与一般数据集划分不同的是,由于对车辆进行成对比对分析,因此测试图片必须成对存在。具体地,将测试图片随机组成正负样本对(Vehicle ID 测试集 3 271 对,杭州卡口数据集 1 903 对),正样本对表示两张测试图片为同一辆车(不属于同一张图片),负样本对表示两张测试图片为不同辆车,测试集中正负样本对各一半。

为了进一步考查算法对高仿套牌车的判别能力,在负样本对的构成上,又规定了普通负样本和高仿负样本(外观一样的不同辆车)各为一半,最终测试集构成如表 2 所示。另外,为了衡量系统对未曾训练类别的比对能力,实验中所设定的测试集都是未参加训练的全新车类。

表 2 测试样本的构成
Table 2 Testing samples

测试集	正样本对	负样本对		合计
		普通	高仿	
Vehicle ID	1 636	818	817	3 271
卡口	951	476	476	1 903

2.3 实验设计与分析

传统基于车辆公有脸的提取主要依赖 HOG、SIFT 等

特征提取算法,在识别分类中这些方法结合 SVM 得到了广泛的应用^[25-26],并且取得了不错的效果,近些年来随着深度学习的兴起,基于 CNN、R-CNN 的方法也在车辆的识别检测上有了一定的突破^[27-28]。本文将结合这些特征提取算法与所提出的方法进行相关的对比分析。为了更好地比对实验结果,定义如下评价指标。

1) 比对精度,比对精度是指在最佳阈值条件下,系统判断正确的样本数占总样本数的比例,比对精度越高表明系统可靠性越大。

2) 距离差,距离差是指测试集中所有正样本相似度的平均值与负样本相似度的平均值的差,由式(11)定义。

$$D = \left| \sum_i s_i/m - \sum_j s_j/n \right| \quad (11)$$

式中: m 、 n 分别表示正负样本对数, s 表示样本比对的相似度, D 表示正负样本平均距离差。由式(11)可知,距离差值越大表明正负样本间的平均距离越大,系统的鲁棒性越强。

3) 系统最终精度,系统最终精度是利用本文分层算法后最终得出的比对精度,由于本文中的公有脸、私有脸比对检测是互相独立的,因此系统的最终精度可以近似表示为式(12)。

$$Acc_s = 1 - (1 - Acc_p)(1 - Acc_c) \quad (12)$$

式中: Acc_s 表示系统最终检测精度, Acc_p 和 Acc_c 分别表示私有脸和公有脸的检测精度。

实验中利用传统的 HOG、SIFT 特征提取方法以及基于深度神经网络的多组件 CNN^[5]、单任务 Faster R-CNN 方法分别对车辆的公有脸进行了比对分析,同时进一步实验了本文提出的基于多任务网络的公有脸、私有脸的方法,实验结果如表 3 所示。

表 3 对比实验

Table 3 Comparison experiments

方法	Vehicle ID 数据集			杭州卡口数据集		
	距离差	精度/%	阈值	距离差	精度/%	阈值
HOG	0.114	68.24	0.51	0.069	62.79	0.49
SIFT	0.202	77.25	0.15	0.099	68.63	0.13
多组件 CNN	0.385	80.34	0.71	0.251	74.72	0.54
Faster R-CNN	0.416	82.15	0.71	0.392	82.81	0.68
本文 + 公有脸	0.517	90.80	0.71	0.472	90.64	0.72
本文 + 私有脸	0.523	93.43	0.34	0.514	91.69	0.49

由表 3 可知,无论是在正负样本的距离差上还是在比对精度上本文公有脸算法明显优于 HOG、SIFT 算法,尤其是在不够标准的卡口数据集上,传统特征提取算法

表现的极其糟糕。相比较文献[5]中多组件 CNN 方法以及单任务 Faster R-CNN 方法本文算法也有较大提高。另外利用本文私有脸算法比本文公有脸的比对结果精度提

高了 3% 左右,距离差也略有提高,这种精度上的提高、鲁棒性的增强得益于私有特征的唯一性,面对测试集中存在的高仿的套牌车,基于公有脸的比对方法已经力所不及,但是私有脸比对没有因此受到束缚,依然能表现出良好的比对能力。

结合两个数据集作进一步分析发现,利用深度卷积神经网络的方法对数据集的适应能力更强,即使是复杂的卡口数据集在距离差和精度上也未出现太大的差异,这说明卷积神经网络的泛化性能极强。

通过表 3 中数据可知,本文公有脸的比对方法在两个数据集上的检测精度分别为 90.80% 和 90.64%,私有脸的检测精度分别为 93.43% 和 91.69%。利用式(12)计算得到系统的最终精度分别为 99.39% 和 99.22%。实验结果表明,引入分层比对后,使得“漏网之鱼”大大减少,精度上有了较大的提升。

为了直观地分析各种算法的优劣,根据 Vehicle ID 数据集上的实验结果绘制了正负样本相似度的分布直方图,如图 3 所示。图中浅色区域表示正样本对,深色区域表示负样本对,横坐标表示相似度(0.0~1.0),纵坐标表示某个相似度区间上(共 150 个区间)样本对存在的个数。由图 3 可知,两曲线交点的横坐标即对应系统的最佳阈值,两曲线的波峰距离近似上文中提出的样本距离差。理想情况下,深色浅色区域的重叠面积(错误率)越小波峰距离(鲁棒性)越大算法的性能越好。

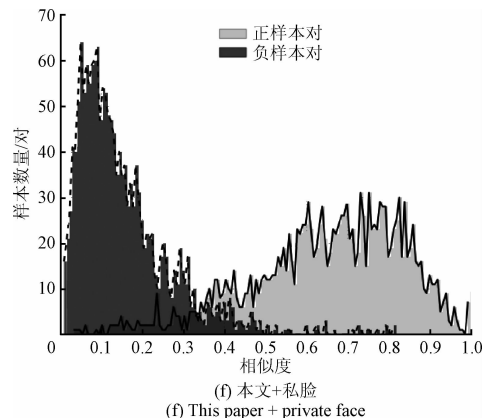
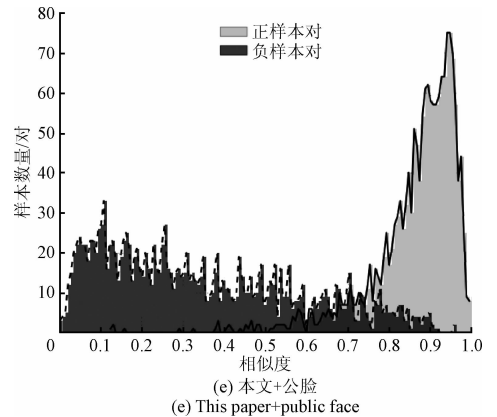
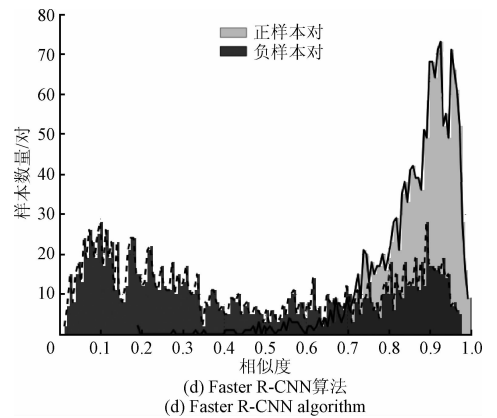
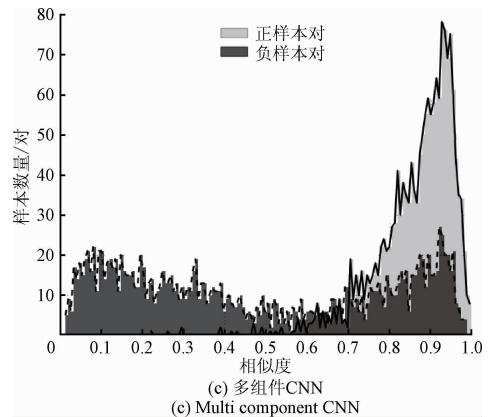
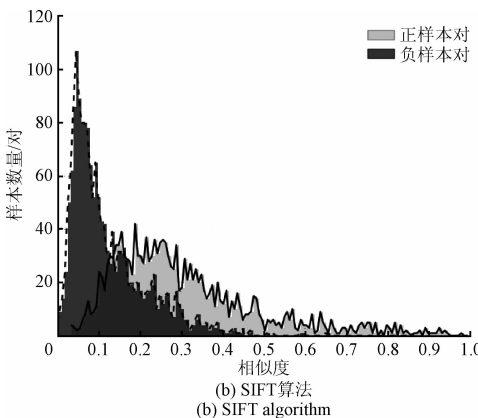
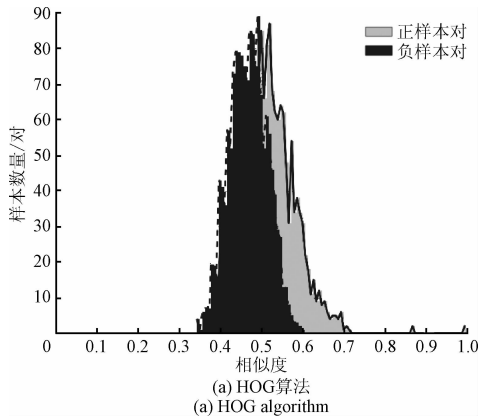


图 3 相似度分布

Fig. 3 Similarity distribution

图3与表3中的实验结果相吻合,传统特征的提取方法的缺点显而易见,尤其是HOG特征提取方法深色区域基本上和浅色区域已经重合,导致误识别率极高。由图3(e)可以看出本文+公有脸的方法对正样本对(浅色区域)的判别能力非常强,阈值选取0.71时可以达到较高的准确率,但是对负样本(深色区域)的判别能力明显的减弱,相似度可以达到0.8以上,这也是公有脸比对精度低的根本原因。

图3(f)对应本文提出的私脸比对方法,从图中可以明显看出私有脸比对方法的错误率要小于公有脸比对的方法,虽然在正样本的比对上效果略差,但是对于负样本的比对效果有了非常大的改善。

文献[5]中多组件CNN网络在虽然分类上有着较好的表现,但是由图3(c)可以看出在本文测试集上却不够理想,它和图3(d)中单任务Faster R-CNN缺点类似,对高仿车辆判别能力较弱,往往会得出较高的比对相似度,表现在图中右半轴深色区域数量有明显的上升。本文的多任务算法利用私有脸、公有脸两个任务之间的约束,使得整个网络一直朝着私有特征方向上优化,虽然训练难度增加,但是效果突出,如图3(e)所示。同样是基于公有脸的检测方法,本文方法对应的图中右半轴深色区域数量并没有明显的上升,表现出了对高仿车辆良好的分辨能力。

本文又提出了分层比对的方法,通过公有脸和私有脸的多级筛选,逐层比对,使得检测精度得到明显提高。

比对效率,在实际应用中有着及其重要的意义。本文利用同一对图片对几种算法进行了效率测试,如表4所示。相比较于传统的识别方法,本文算法有了较大的提高;相比较单任务Faster R-CNN,本文基于多任务的特征提取方法在效率上并没有出现太大差别,如果不采用多任务网络而是利用单任务网络依次进行公有脸、私有脸两次单独的特征提取,这样时间几乎要增加一倍;多组件CNN虽然单个网络简单规模小,但是由于采取了多网络模型分别对每个车体组件进行特征提取,在效率上的优势被大大弱化。

表4 速度对比实验

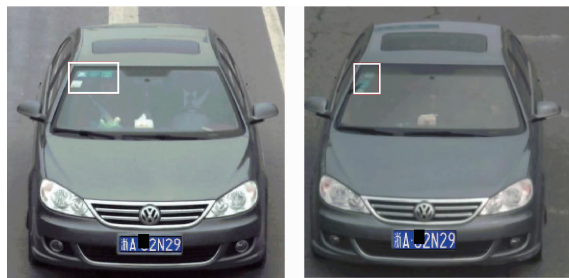
Table 4 Comparison of operating speed

方法	比对速度/s
HOG	2.193
SIFT	0.857
多组件 CNN	0.531
Faster R-CNN + 公脸	0.824
Faster R-CNN + 私脸	0.712
本文多任务网络	0.723

此外,由于本文采用了分层的比对方式,在私有脸比对时只是一些外观一致的车辆参与比对,在比对的图像数目上已经所剩无几,这样也提升了系统的比对效率,实

现了比对精度和效率的完美统一。

为了进一步考查本文提出的方法对高仿套牌车的实际比对的情况,收集了4对外观完全一样的高仿套牌车,运用Vehicle ID数据训练模型进行私有脸的比对分析;实验结果如图4所示,图中的方框为本文算法实际检测到的车检标。如果不根据车检标部分进行比对,人眼几乎不能判断真伪,但运用本文提出的检测方法能有效快速地检测出假牌套牌车辆。



(a) 相似度0.097
(a) Similarity degree is 0.097



(b) 相似度0.171
(b) Similarity degree is 0.171



(c) 相似度0.181
(c) Similarity degree is 0.181



(d) 相似度0.218
(d) Similarity degree is 0.218

图4 负样本测试结果

Fig. 4 Testing results of negative samples

图4(a)负样本对应的车辆外观一样,但是车检标有较大的不同,利用本文算法比对的相似度非常低,轻易得出正确的结果;图4(b)由于图片模糊反光等干扰,即使是具有个性化特征的车检标也显得十分相似,通过本文算法最后得出的相似度分为0.171远小于最佳阈值0.34。图4(c)车检标的个数非常少,包含的信息较少,即使在这种细微的差别上本文算法也有着不错的表现。图4(d)只是在车检标的排列顺序上有略微的差异,对应的比对的相似度也略有提高,即便如此相似度也远小于0.34。

在正样本图像的比对上,基于私有脸的比对方法也有着不错的表现,如图5所示。对同一辆车出现在不同卡口被抓拍的车辆图像进行了比对排查,图5(a)的私有脸比对相似度为0.692、图5(b)的私有脸比对相似度为0.820、图5(c)的私有脸比对相似度为0.731,每对正样本实验图像的比对相似度都远大于最佳阈值0.34。



图5 正样本测试结果

Fig. 5 Testing results of positive samples

综上所述,以上几种对比方法的不足主要体现在以下两点。

1) 常规的特征提取方法对噪声大、干扰强、外观区分度不高的车辆对象没有行之有效的特征提取能力,如HOG、SIFT方法。

2) 利用深度神经网络的方法对特征提取能力有了较大改进,但是只是基于公有脸的方法进行特征的提取对于本文关注的高仿套牌车依然存在较大不足,如多组件CNN、Faster R-CNN方法。

本文提出的多任务分层比对方法主要针对以上两点不足进行了改善,多任务卷积神经网络强化了对目标对象的特征提取能力,而私有脸车检标特征比对的引入进一步解决了高仿套牌车的问题。在私有脸比对上,几乎不受环境条件的约束,已具备较强的实战能力,可运用于跨域的套牌假牌车的排查中。

另外,由于车脸的私脸部分具有较高的唯一性,因此,本文提出的公有脸+私有脸的比对稽查方式对无牌车、不能识别车牌号的车辆同样具有强大的甄别能力。

3 结论与展望

本文首先提出了一种基于多任务 Faster R-CNN 车辆假牌套牌的检测方法,相对于传统基于图像逐个获取车辆信息的方法对其他约束条件的依赖性更低、鲁棒性更强、精确性更高、实时性更好。利用有限的数据集对检测网络进行训练,通过比对车辆公有脸和私有脸的特征向量来甄别套牌假牌车辆,使得系统具备较好的泛化能力。此外,采用了公有脸和私有脸的特征向量来描述车辆特征使得每一辆车都具有了唯一的视觉特征,这对于无牌车、无法识别车牌号的车辆也能进行跟踪与识别,可以有效协助公安机关、刑侦部门对违法犯罪案件中的车辆进行实时监控。双重车辆基本信息的比对方式有效提升了检测的鲁棒性和准确性。这种利用时空特性和车牌号码约束来缩小排查范围,用车检标精细比对方法精准的比较相同车牌相同车辆的异同点,在提升了排查效率的同时又提高了排查精度,实现了效率与精度的完美统一。

本文提出的方法虽然较现有方法有了较大的改善和提升,但是对于在车辆前窗右上角没有粘贴车检标的情况,私有脸比对方法就会失效。对于这个问题,准备在今后车检标识别时一并解决。

参考文献

- [1] 费继友, 谢金路, 李花, 等. 基于字符特征约束的自适应车牌校正提取[J]. 仪器仪表学报, 2016, 37(3): 632-639.
- FEI J Y, XIE J L, LI H, et al. Adaptive license plate correction and extraction based on character feature con-

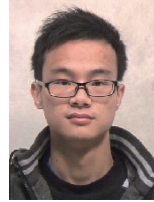
- straint [J]. Chinese Journal of Scientific Instrument, 2016, 37(3): 632-639.
- [2] 司朋举, 胡伟. 一种改进的神经网络车牌识别算法研究[J]. 电子测量技术, 2016, 39(10):100-103.
SI P J, HU W. Research on an improved neural network license plate recognition algorithm[J]. Electronic Measurement Technology, 2016, 39(10):100-103.
- [3] WANG Y, LI H, KIRUI C K, et al. Vehicle discrimination using a combined multiple features based on vehicle face[J]. Lecture Notes in Electrical Engineering, 2013: 503-511.
- [4] TEOH S S, BRÄUNL T. Performance evaluation of HOG and Gabor features for vision-based vehicle detection [C]. IEEE International Conference on Control System, Computing and Engineering, 2016:66-71.
- [5] HUANG K, ZHANG B. Fine-grained vehicle recognition by deep convolutional neural network[C]. Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2016: 465-470.
- [6] 任水林, 陶煜波, 林海. 交互式套牌车可视识别与分析[J]. 计算机辅助设计与图形学学报, 2016, 28(11): 1887-1898.
REN SH L, TAO Y B, LIN H. Interactive visual analysis of fake plate vehicles detection [J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(11):1887-1898.
- [7] GUO H, WANG Z, YU B, et al. TripVista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection[J]. IEEE Pacific Visualization Symposium, 2011,18(1):163-170.
- [8] 王峰, 靳小波, 于俊伟, 等. 基于融合策略的套牌车主动识别算法 [J]. 光电子·激光, 2015(11): 2209-2216.
WANG F, JIN X B, YU J W, et al. A positive recognition algorithm for fake plate vehicles based on fusion strategy [J]. Journal of Optoelectronics · Laser, 2015(11): 2209-2216.
- [9] 全国人大法工委行政法室. 中华人民共和国道路交通安全法: 案例应用版 [M]. 北京: 人民交通出版社, 2009.
Administrative law room of the National People's Congress Law Committee. People's republic of china road traffic safety law: Case study [M]. Beijing: China Communications Press, 2009.
- [10] 中华人民共和国道路交通安全法实施条例. 中华人民共和国道路交通安全法实施条例 [J]. 司法业务文选, 2004(6):20-28.
Regulations for the implementation of the road traffic safety law of the People's Republic of China. Regulations on the implementation of the road traffic safety law of the People's Republic of China [J]. Selected Works of the Administration of Justice, 2004(6): 20-28.
- [11] 宋晓琳, 邬紫阳, 张伟伟. 基于阴影和类 Haar 特征的动态车辆检测 [J]. 电子测量与仪器学报, 2015, 29(9):1340-1347.
SONG X L, WU Z Y, ZHANG W W. Dynamic vehicle detection based on shadow and Haar-like feature [J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(9):1340-1347.
- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [13] GIRSHICK R. Fast r-cnn [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [14] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. Advances in Neural Information Processing Systems, 2015: 91-99.
- [15] 李云红, 魏妮娜, 张晓丹. 基于多方向 Gabor 滤波器的图像分割 [J]. 国外电子测量技术, 2017, 36(3): 20-23.
LI Y H, WEI N N, ZHANG X D. Image segmentation based on multi-direction Gabor filters [J]. Foreign Electronic Measurement Technology, 2017, 36(3):20-23.
- [16] 赵泉华, 高郡, 李玉. 基于区域划分的多特征纹理图像分割 [J]. 仪器仪表学报, 2015, 36(11): 2519-2530.
ZHAO Q H, GAO J, LI Y. Multi-feature texture image segmentation based on tessellation technique [J]. Chinese Journal of Scientific Instrument, 2015, 36(11): 2519-2530.
- [17] TANG T, ZHOU S, DENG Z, et al. Vehicle detection in aerial images based on region convolutional neural networks and hard negative example mining [J]. Sensors, 2017, 17(2): 336.
- [18] DENG Z, SUN H, ZHOU S, et al. Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks [J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 10(8):3652-3664.
- [19] HE T, HUANG W, QIAO Y, et al. Text-attentional convolutional neural network for scene text detection [J]. IEEE transactions on image processing, 2016, 25(6):

- 2529-2541.
- [20] ZHANG Z, LUO P, LOY C C, et al. Facial landmark detection by deep multi-task learning[C]. European Conference on Computer Vision, 2014: 94-108.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014.
- [22] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [C]. European Conference on Computer Vision, 2014: 346-361.
- [23] SUN Y, CHEN Y, WANG X, et al. Deep learning face representation by joint identification-verification[C]. Advances in Neural Information Processing Systems, 2014: 1988-1996.
- [24] LIU H, TIAN Y, YANG Y, et al. Deep relative distance learning: Tell the difference between similar vehicles[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2167-2175.
- [25] LEE S H, BANG M S, JUNG K H, et al. An efficient selection of HOG feature for SVM classification of vehicle[C]. IEEE International Symposium on Consumer Electronics (ISCE), 2015: 1-2.
- [26] PEI W, JIA X, HUANG X, et al. A vehicle retrieval and recognition system based on smart phones[C]. 8th International Congress on Image and Signal Processing (CISP), 2015: 527-532.
- [27] HUANG K, ZHANG B. Fine-grained vehicle recognition

by deep Convolutional Neural Network[C]. International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2016: 465-470.

- [28] HUO Z, XIA Y, ZHANG B. Vehicle type classification and attribute prediction using multi-task RCNN[C]. International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2016: 564-569.

作者简介



陈朋, 2016 年于安徽工业大学获得学士学位, 现为浙江工业大学硕士研究生, 主要研究方向为计算机视觉、深度学习。

E-mail: chenpeng96357@foxmail.com

Chen Peng received his B. Sc. from Anhui University of Technology in 2016. Now he is a M. Sc. candidate in Zhejiang University of Technology. His main research interests include Computer Vision and Deep Learning.



汤一平(通讯作者), 1986 年于浙江大学获得硕士学位, 1995 年于日本国立埼玉大学获得博士学位, 现为浙江工业大学教授、银江股份有限公司资深顾问, 主要研究方向为全方位视觉传感器及应用、计算机视觉、深度学习。

E-mail: typ@zjut.edu.cn

Tang Yiping (Corresponding author) received his M. Sc. degree from Zhejiang University in 1986, Ph. D. degree from Saitama University (Japan) in 1995. Now he is a professor in Zhejiang University of Technology and a technical consultant in Enjoyer Co., Ltd. His main research interests include Omni-directional vision, Computer Vision and Deep Learning.