

基于卷积神经网络的污损遮挡号牌分类

李林¹ 赵凯月¹ 赵晓永² 魏帅琴³ 张兵⁴

1 燕山大学信息与科学工程学院 河北 秦皇岛 066004

2 北京信息科技大学信息管理学院 北京 100097

3 空军指挥学院教研保障中心 北京 100097

4 公安交通警察支队指挥中心 河北 沧州 061000

(lilin@ysu.edu.cn)

摘要 作为智能交通的重要组成部分之一,车牌识别在人们的日常生活中发挥着不可替代的作用,例如,生活中违规车辆常常因号牌污损或者故意遮挡等来躲避处罚,进一步加大了执法的难度,因此提高污损或遮挡车牌的识别效率依然是当今自动识别系统中的一项至关重要的任务。文中主要集中解决遮挡号牌的识别问题,其主要分为正常号牌、部分遮挡号牌、完全遮挡号牌以及未悬挂4种情况。传统的OCR算法在汉字、字符以及数字之间的识别上具有很高的准确性,当将其运用到识别车牌上后,虽然在正常号牌和部分遮挡号牌的检测上也体现出了很好的识别效果,但对全部遮挡和未悬挂车牌的识别效果依然很差,随着人工智能的发展,使得其在全遮挡和未悬挂号牌的识别上也能有更好的效果。因此,结合传统算法的优点,采用OCR技术并结合现今的深度学习算法,优化对污损车牌的识别效果。

关键词: 智能交通;目标检测;污损车牌;深度学习;OCR

中图分类号 TP311.5

Contaminated and Shielded Number Plate Recognition Based on Convolutional Neural Network

LI Lin¹, ZHAO Kai-yue¹, ZHAO Xiao-yong², WEI Shuai-qin³ and ZHANG Bing⁴

1 College of Information and Science Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China

2 School of Information Management, Beijing University of Information Technology, Beijing 100097, China

3 Teaching and Research Support Center of Air Force Command College, Beijing 100097, China

4 Command Center of Police Traffic Detachment Cangzhou, Cangzhou, Hebei 061000, China

Abstract As one of the important components of intelligent transportation, license plate recognition plays an irreplaceable role in people's daily life. For example, in daily life, illegal vehicles often avoid punishment because of the number plate contamination and occlusion, which further increases the difficulty of law enforcement. Therefore, improving the recognition efficiency of contaminated license plate is still a crucial issue in today's automatic recognition system. The paper mainly focuses on the recognition of shielded number plate. There are four main cases: normal number plate, partially shielded number plate, completely shielded number plate and not hanging plate. The traditional OCR algorithm has a high accuracy in the recognition of Chinese characters, characters and numbers. When it is applied to the recognition of license plates, although the detection of normal and partial shielded license plates shows a good recognition effect, the recognition effect of completely shielded number plates and not hanging license plates is still very poor. With the development of artificial intelligence, it is possible to get better recognition on completely shielded plates and not hanging plates. Therefore, combined with the advantages of traditional algorithms, this paper adopted OCR technology and the current deep learning algorithm to optimize the recognition effect of stained license plate.

Keywords Intelligent transportation, Target detection, Contaminated license plate, Deep learning, OCR

1 引言

自21世纪以来,随着国家综合国力的提高,交通道路建设的需求开始逐步增加,汽车逐渐成为人们不可或缺的生活

必需品,这导致汽车的数量急剧增加,进而引发了城市道路拥堵、交通事故等一系列问题,给城市的发展带来了巨大的压力。为了解决这些交通问题,各种智能交通系统^[1](Intelligent Transport Systems, ITS)应运而生,不仅能够实现对交

基金项目:北京市教育委员会科技计划一般项目(KM201711232018);国家自然科学基金项目(61502039);软件定义能源互联网的模型与优化控制研究项目;北京信息科技大学2019年度科技计划一般项目(KM201911232002)

This work was supported by the General Project of Science and Technology Plan of Beijing Municipal Education Commission (KM201711232018), National Natural Science Foundation Project (61502039), Research Project of Model and Optimization Control of Software Defined Energy Internet and General Project of Science and Technology Plan of Beijing University of Information Technology in 2019 (KM201911232002).

通信作者:赵凯月(zhaokaiyue9421@foxmail.com)

通状况的实时监控,也更加方便管理,提高工作效率。

车辆号牌智能违法识别系统是智能交通系统中的重要组成部分,也是实现违法号牌识别的最重要的一个环节。该系统利用卷积神经网络算法来实现目标检测、视频图像处理以及模式识别等不同技术的相互融合,进而完成对号牌的定位和识别,最终实现对车牌信息的提取和分类。目前车牌识别系统已经在许多不同的场景下得到了应用。例如,利用遮挡号牌识别算法能够有效地识别道路上污损遮挡号牌的车辆,公安交警通常会在交通卡口架设监控摄像头,利用车牌识别系统对过往车辆中违反交通规则以及涉及不法案件的肇事车辆号牌进行识别,以此达到识别相关车辆是否具有违法行为的目的。但遗憾的是,目前在检测车辆号牌领域,研究人员主要针对清晰车牌图像来设计识别系统,而很少对模糊或者污损的车牌图像分类的相关技术进行研究。

目前,传统的应用方法在对模糊或者污损的车牌分类时仍然有很大的局限性,针对车辆图片中所包含的一些严重干扰因素,如由摄像头原因造成的低分辨率,雾霾天气、雨天和昼夜造成的图片曝光度过低或过高,国内的一些实验研究仍旧不能做到更加高效的处理。例如,对完全遮挡和未悬挂车牌这类复杂干扰的样本图片往往不能通过识别的手段来对结果进行分类,在传统算法上各种各样复杂和多样的样本图片往往不能满足对车牌高像素的基本要求。通过深度学习构建的深度网络在无监督特征提取方面表现出了优异性能^[2],为了弥补传统算法的缺点,本文提出利用卷积神经网络(CNN)^[3]来解决这一问题,卷积神经网络是根据生物神经网络而定义的一类包含卷积计算而抽象出的数学模型,其隐含层内的卷积核参数共享机制极大地减少了网络之间的参数数量,提高了计算效率,层间连接的稀疏性使得卷积神经网络能够以较小的计算量对格点化(grid-like topology)特征进行学习,越是大量的复杂多样的样本,能够学习到的特征越显著,模型对车牌点定位、分类的效果就越好,因此本文提出卷积神经网络和传统算法相结合的思路来提高分类效果。

针对国内日益不断增长的车辆,智能交通在车辆号牌信息提取方面的需求日益增长,识别违法污损遮挡号牌也日趋重要,传统分类方法在识别车辆号牌方面只能应对相对简单和规范的场景,而面对复杂环境下的号牌仍旧存在识别准确率低的特点。据此,本文提出一种结合传统 OCR 技术上的污损车牌识别的分类算法,污损遮挡违法号牌分为部分遮挡号牌、全部遮挡号牌和未悬挂车牌 3 类。

首先通过传统算法 OCR 对正常车牌和半遮挡车牌进行识别,利用对车辆号牌字符的识别,来实现对正常号牌和半遮挡号牌的分类。对未分类的车辆图片,进一步通过目标检测分类算法 YOLO 重新对图片上的车牌进行定位、分类,从而实现对全部遮挡与未悬挂号牌的分类,最终完成对污损遮挡号牌的分类。

2 本文工作

本文主要实现违法遮挡号牌的识别,并将模型打包移植到移动终端设备 Jetson Tx2 开发板上,最终实现终端对违法遮挡号牌的识别。主要过程分为样本标注、选择模型、训练数据、调整参数、算法封装及移植 6 部分。实验过程中算法部分分为两个步骤,首先利用卷积神经网络算法实现对全部遮挡和未悬挂号牌的分类,其次通过传统 OCR 技术进一步完成对

正常号牌和部分遮挡号牌的分类。硬件设计部分采用 Jetson Tx2 作为移植模型的终端设备。设计框图如图 1 所示。

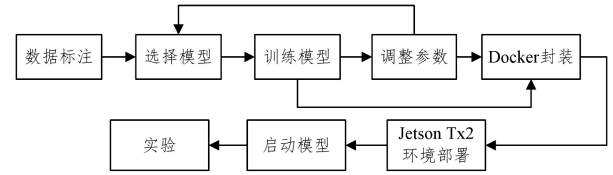


图 1 系统框图

Fig. 1 System block diagram

算法部分主要实现违法遮挡号牌的识别。违法遮挡号牌分为部分遮挡号牌、全部遮挡号牌和未悬挂号牌 3 种。实验采用传统 OCR 算法和卷积神经网络算法相结合的方法,实现对违法遮挡号牌的分类,具体实现流程如图 2 所示。

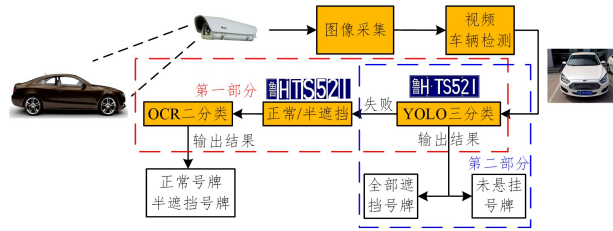


图 2 污损遮挡号牌识别流程图

Fig. 2 Flow chart for identification

整个系统的数据主要来自网络,实验中车辆号牌一共分为 4 类:正常号牌、部分遮挡号牌、未悬挂号牌和完全遮挡号牌。该系统分为如下两个部分:

(1)采用卷积神经网络分类。YOLO 算法在检测目标的速度和准确度上具有显著优势,本文拟采用 YOLO 算法将检测目标分为 3 类,首先完成对全部遮挡、未悬挂号牌和其他 3 类的划分,在其他类中进一步使用传统算法 OCR 区分正常号牌和半遮挡号牌两类。最终实现污损遮挡号牌的分类。

(2)采用 OCR 算法对图像上的车辆的号牌进行识别和分类。其主要实现过程如下:针对 YOLO 算法检测到的号牌位置进一步通过 OCR 算法来判断识别区域的字符,并按照设定的阈值来区分正常和部分遮挡号牌与未悬挂和完全遮挡号牌两大类,针对识别到字符的号牌通过字符二值化、字符分割等一系列操作给出每位不同的置信度,根据不同字符的最优阈值进一步将车牌判断为正常号牌和部分遮挡号牌。

移动终端采用 NVIDIA Jetson Tx2 作为模型算法的主要载体。模型在终端设备上的移植操作分为算法封装和模型移植两部分。

首先,采用 Docker 技术将模型算法打包成镜像。在宿主机上安装和部署 Docker 以及调用 CUDA 和显卡的驱动程序,从镜像仓库中拉取所要部署环境的基础运行环境系统镜像 Ubuntu16.04,在镜像层上创建可读写文件层即容器,利用容器的映射功能和网络共享功能,可在容器中快速部署算法、安装环境依赖等。其次,在 Jetson Tx2 上安装和部署 Docker,将宿主机上打包的镜像移植到 Tx2 上,最终完成模型在 Tx2 的部署和安装。

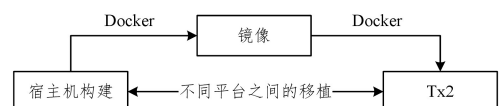
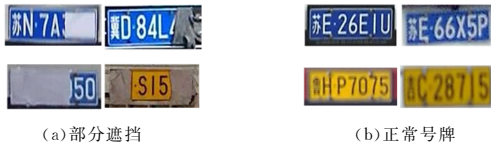


图 3 模型移植

Fig. 3 Model transplantation

3 基于 OCR 的正常与部分遮挡号牌识别

在人工智能领域,图像识别作为其重要的一个分支,具有更为重要的作用,尤其是 OCR 技术具有更加广泛的应用。正常车牌和部分遮挡号牌的识别依旧可以通过对车牌信息的提取来对其进行分类。因此,在选择识别正常车牌和部分遮挡号牌的算法时依旧采用传统的 OCR 技术来实现对其两种类型的识别和分类,以下只对这两种类型的车牌进行分类处理^[4]。



(a)部分遮挡 (b)正常号牌

图4 部分遮挡与正常号牌的对比

Fig. 4 Compare of partial occlusion and normal plate

从整体来说,按照过程 OCR 一般分为两个部分:图像处理以及文字识别。在识别车牌号上的数字、汉字和字母之前,需要首先对数据进行预处理,以提高模型后续训练中对原始图片的特征信息提取。这个过程通常包含以下几个方面:车辆号牌的识别和定位、号牌灰度化与二值化、号牌降噪、字符分割、字符倾斜矫正等。实现图片的预处理,能更好地滤除噪点,提取图片中的车牌号信息。在此基础上,根据读取到的车牌信息进一步对车牌号进行判断分类,对正常号牌以及部分遮挡号牌做出区分,从而达到区分正常号牌与部分遮挡号牌的目的。

OCR 算法的主要思想是通过识别汉字、字母或数字来达到分类的目的。其分为两部分:首先对样本图片中的车牌进行精确定位^[5],其次对定位区域上每一位字符或汉字进行分割识别,通过分析判断号牌的种类。详细过程如下:首先进行特征变换、提取等操作以对图像进行预处理;其次经过 YOLO 算法模型得到图像车辆号牌位置,从而实现对车辆号牌的定位;最后对图像截取的车辆号牌的位置区域进行灰度化、二值化、图像分割等操作,对定位的车牌位置区域进行灰度化,根据灰度图判别位置区域的字符,按照字符尺度大小信息进行字符切割,通过对识别结果中的每一位字符进行缩放来实现特征提取,并对识别字符的特征向量与字符数据库中标准字符的表达形式进行匹配区别,从而识别字符。对每位字符进行比对之后得到相应的置信度,通过判断每个字符的阈值来确定识别的准确率,流程图如图 5 所示。

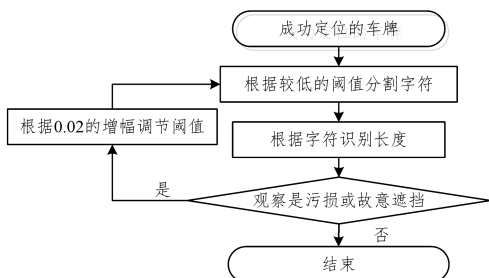


图5 阈值流程图

Fig. 5 Threshold flow chart

具体实现过程如下:车辆图像经过 OCR 处理之后,针对定位的车牌区域按照设定的阈值来识别字符,将低于阈值的舍弃,高于阈值的作为字符识别的最终结果。具体流程如图

5 所示,首先确定一个较低的识别阈值作为测试阈值,本文中该阈值取 0.5,再根据阈值对字符进行识别,以人眼识别的结果作为验证最优阈值的指标,不断按照 0.02 的阈值增幅进行调节。通过实验的不断迭代,得出最优阈值。最终实现正常号牌与污损遮挡号牌的分类。

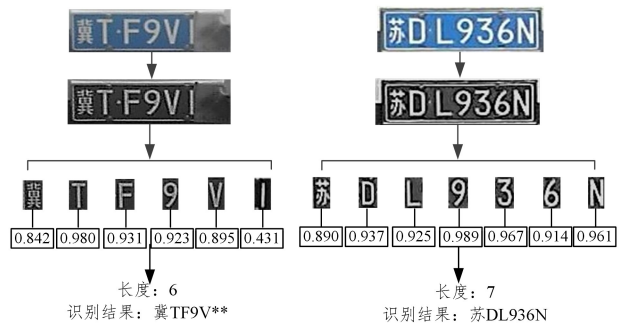


图6 阈值分析

Fig. 6 Threshold analysis

汉字和字母、数字阈值的设置都不尽相同,汉字较为复杂,实验过程中适当放低阈值能提高识别的准确率,字母和数字较为简单,适当提高阈值有助于降低识别错误率。因此,实验中的阈值设置如表 1 所列。

表1 阈值设置表

Table 1 Threshold setting table

类型	阈值
汉字	1.870
字母或数字	0.920

根据识别的场景,部分遮挡号牌具体可以分为故意遮挡部分号牌和污损导致的遮挡号牌。半遮挡号牌的具体识别流程如下:

(1)针对故意部分遮挡号牌,通过设定较高的阈值筛选出具有较高置信度的字符,针对部分遮挡号牌,识别出的字符长度通过与正常号牌进行对比。根据长度与格式即可筛选出部分遮挡号牌。

(2)针对污损号牌,通过实验大量的图片,选择合适的阈值。针对污损车牌上的模糊字符会识别出较低的置信度,通过阈值可以筛选出具有较高识别准确率的字符,据此对车牌进行分类。

实际应用中,污损遮挡号牌的识别不仅和算法的识别率有关,更与所采集的车辆图像质量和实际车牌质量息息相关,车牌质量的好坏直接影响最终的识别性能,例如车牌会受到主观因素上的车辆套牌、车牌遮挡、多车牌等影响,也会受到客观因素上的生锈、字体脱落掉漆、号牌倾斜等影响。而车辆的图像质量在实际拍摄过程中会受到光照、拍摄方式、天气等各种因素的影响,这些因素的不同也不同程度的影响了最终的识别效果。

4 基于 YOLO 的未悬挂与完全遮挡号牌的识别

4.1 YOLO 的算法机制

YOLO 模型在思想上摒弃了一般目标检测中使用的滑动窗口方式来定位目标,而是使用了回归的思想,该网络摒弃了传统的滑动窗口定位目标的方法,直接读取整张图像,使用卷积算法提取整幅图像的特征,并对图像进行网格划分,从其划分区域中回归出待检测目标的类别和区域,相比之下虽然

牺牲了一些识别精度,但是大大加快了检测的速度,其设计过程如下:首先将输入图像的尺寸修改为模型需要的规定尺寸,车牌相对图片而言,目标较小,为了提高特征图的感受野,本文中采用的是416分辨率的图像,随即将处理之后的图像划分为 $S \times S$ 大小块的网格结构,目标预测的中心点落在网格中,检测其所落的网格,并负责预测其目标^[6],每个单元格预测对应目标的边界框和置信度,这个置信度表示两方面的内容,即区域内包含预测目标的置信度和预测区域的精确度,用 $pr(Object) \times IOU_{pred}^{truth}$ 来表示,其中包含两方面的内容:一方面是表示预测目标所在边界框的准确度,即预测框与实际目标边界框的交并比用 IOU_{pred}^{truth} 来表示;另一方面表示含有目标的可能性的大小 $pr(Object)$,当边界框中含有目标时, $pr(Ob-$

$ject)$ 为1,否则为0。每预测一个目标需要5个数值来表述,分别是归一化的预测区域的中心点相对左上角坐标的偏移量 (x, y) 、预测区域的宽高 (w, h) 以及置信度,除了对上述目标定位的问题,还有分类问题,与每个单元格预测的区域数目无关,每个单元格预测 C 个类别的后验概率 $pr(Class_i | Object)$,测试时,将后验概率和交并比相乘来计算各个边界框类别置信度的指标,即 $Pr(Class_i | Object) \times IOU_{pred}^{truth} = Pr(Class_i) \times IOU_{pred}^{truth}$,输出结果为 $S \times S \times (B \times (4 + 1) + C)$ 维的向量,这不仅直接完成了目标定位也完成了对号牌分类的任务。在网络结构方面,YOLO采用24个卷积层和2个全连接层,最后一层采用线性激活函数来实现分类。其网络结构如图7所示。

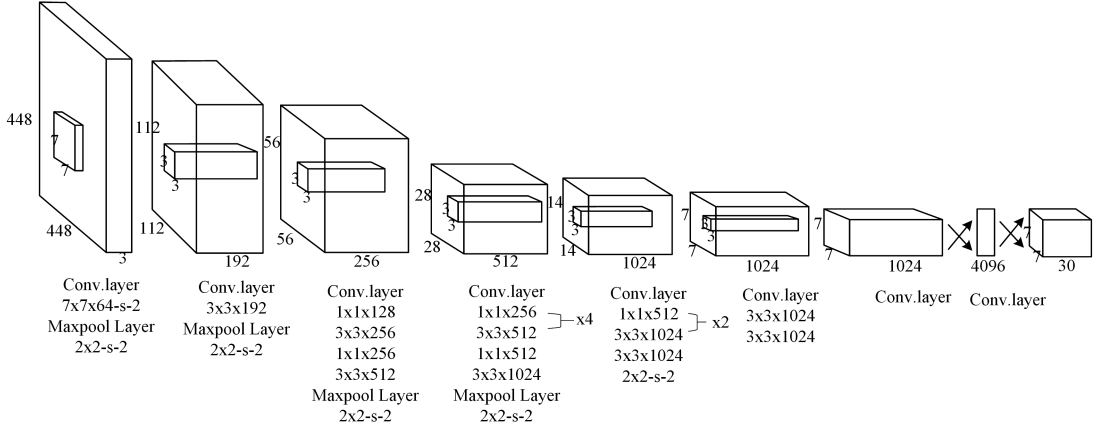


图7 YOLO网络结构

Fig. 7 Network structure of YOLO

4.2 模型训练

卷积神经网络中目标检测普遍存在运算速度慢的问题,其原因在于目标检测主要分为两步,首先对图像中的目标进行定位^[7],其次对定位后的目标进行分类。而在YOLO中则将two-stage的结构改为one-stage的结构,相对R-CNN则提高了算法的运算速度,而在YOLOv3中则将神经网络中的全连接层修改为batch normalization的网络结构,使目标分类的精确度即(mAP)有所提高。因此,基于本项目对速度和精度有一定的指标要求,选择YOLOv3作为此次项目的实验模型,YOLOv3相对于YOLOv1的神经网络设计主要有以下改变:

(1) 每个卷积层后添加了批标准化层,加速了网络的收敛。

(2) 在第16层开始分为两条路径,将低层的特征直接连接到高层,可提高模型性能。移除全连接层,最终的输出向量中保存了原来的位置信息。

(3) 输入尺寸变为 $416 \times 416 \times 3$,识别更高分辨率的图片。

该网络最终输入的图片尺寸为 $416 \times 416 \times 3$,输出向量尺寸为 $13 \times 13 \times 125$ 。

YOLOv3使用由坐标误差、IOU误差和分类误差3部分组成的均方和误差作为损失函数,不同误差对损失的影响不尽相同,简单相加时要计算不同损失的影响大小,例如,YOLO的坐标误差采用较大的权重值 $\lambda_{coord} = 5$,而在计算IOU误差时,需要区分包含检测目标的区域的置信度与不包含目标的置信度,两者之间的交并比(IOU)误差对网络的损失值即

Loss值的贡献值也不相同。在相同权值的情况下,不包含目标边界框的置信度对损失的误差影响更大,间接放大了其误差。

为了降低误差对网络参数梯度的影响,YOLO使用 $\lambda_{mobj_i} \cdot \lambda_{mobj_j} = 0.5 \cdot \lambda_{mobj_j} = 0.5$ 修正交并比误差。在卷积神经网络中除了不同的误差带给网络影响之外,对于相等的误差值,边界框的大小也会影响网络参数的梯度,较小目标区域对坐标误差的影响大于较大边界框对检测的影响。这是因为不同尺度的边界框在相同位置偏差的情况下,大目标位置区域的坐标误差占据的比例远小于小边界框。因此,为了解决这个问题,YOLO模型将预测网络的边界框宽和高的方法修改为目标的平方根的方法来降低误差,改善效果,即 $(x, y, \sqrt{w}, \sqrt{h})$ 。除此之外,这种改进方法也可以减小定位误差,但不能彻底消除对模型网络参数的影响。综上,YOLO在训练过程中的损失值计算如下。

其中,坐标误差为:

$$coordErr = \lambda_{coord} \cdot \sum_{i=0}^S \sum_{j=0}^B \prod_{ij}^{obj} [(x_i - \hat{x}_i)^2 - (y_i - \hat{y}_i)^2] + \lambda_{coord} \cdot \sum_{i=0}^S \sum_{j=0}^B \prod_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 - (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \quad (1)$$

分类误差为:

$$clsErr = \sum_{i=0}^S \prod_{i=0}^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2 \quad (2)$$

IOU误差为:

$$iouErr = \sum_{i=0}^S \sum_{j=0}^B \prod_{ij}^{obj} (C_i - \hat{C}_i)^2 + \lambda_{mobj} \cdot \sum_{i=0}^S \sum_{j=0}^B \prod_{ij}^{mobj} (C_i - \hat{C}_i) \quad (3)$$

训练的损失值分别为坐标、分类和 IOU 3 部分误差之和。

$$Loss = \sum_{i=0}^2 coordErr + iouErr + clsErr \quad (4)$$

其中,符号 $(\hat{x}, \hat{y}, \hat{w}, \hat{h}, \hat{C}, \hat{p})$ 为预测值,否则符号 (x, y, w, h, C, p) 则表示训练的标记值。为了能够使模型对样本图像中的车牌区域进行目标检测和区域定位,需要选取大量的样本对模型进行训练,而在这一过程中需要对图像中的车辆号牌区域进行标注,为了提高模型的泛化性,实验选取不同场景的图片并对其中的车辆号牌标注位置,标注信息包括类别和区域中心位置坐标以及边界框的宽和高。进入网络之前对数据进行预处理,即对区域位置进行归一化,训练目标检测模型时,需要首次预加载权重来初始化网络权重参数,根据实际网络收敛的效果来确定网络迭代参数,其中网络迭代参数、权重衰减率和动量采用默认值,选取收敛区间内的权重文件来评估模型实测效果,为了验证模型的泛化性和准确率,验证集为在原数据集中随机抽取的数据。

YOLO 参数设置为 batch,即步长值为 64,每次读取图片批次(subdivisions)为 8 张,本文中 YOLO 算法用于对污损遮挡号牌中的全部遮挡、部分遮挡和正常号牌、未悬挂车牌进行 3 分类,因此设置分类参数 classes 为 3,根据 yolov3 的特性,卷积核 filters = (classes + 5) × 3 = 24,采用 24 层的卷积深度。学习率默认为 0.0001,提高精度和准确度的 random 参数设置为 1,多尺度训练模型,每迭代 100 次会存储一个新的权重文件,avg loss 是监督模型训练中错误率的指标,该指标越低越好,刚开始会很大,伴随着训练过程的不断迭代,错误率不断下降,最终降到 1 以下,此次模型中的每个类迭代 2000 次。

车牌识别算法 YOLO 在小批量(mini-batch)数据中使用随机梯度下降算法进行训练,使数据能快速达到收敛状态,学习率决定了数据权重在小批量数据中的收敛速度,即移动步长。如果学习率过低,训练达到收敛状态的耗时较长,但是训练结实可靠。相反,学习率过高会导致模型波动变大,不易达到最优解,甚至造成训练结果发散。因此,在保证不降低模型精度的情况下,选择合适的学习率能缩短训练时长,提高训练模型的效率。

深度学习的神经网络在学习过程中会逐渐提取图片信息的特征,随着训练过程的逐步深化,Loss 值不断下降,学习的特征会越来越多,这就可能引发信息提取的过拟合的现状。Loss 图如图 8 所示。

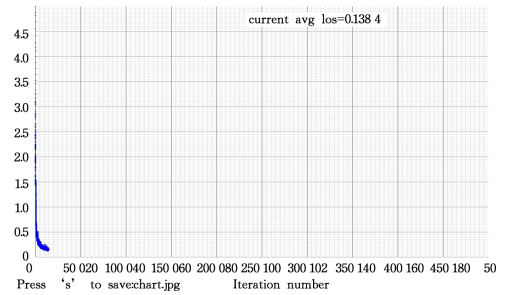


图 8 Loss 下降图

Fig. 8 Loss descent graph

如图 8 所示,Loss 值在开始一段时间内急速下降,随着迭代次数的不断增加,模型开始逐步收敛,并不断趋向最优解,梯度衰减也逐渐降低,最终得到一个最优模型,图 8 中的结果只能代表平均的损失率,并不能代表最优的效果,要使得各类的准确率达到最优,应根据每个类别的 AP 均衡才能获取最优的效果。

本文测试采用的权重文件为迭代 6000 次后产生的权重文件,评估模型指标采用各类样本数据的准确率(AP)、平均准确率 mean Average(mAP)、交并比(IoU)等,其中 AP 定义为平均精度,代表准确率和召回率(Precision-recall)曲线下面的面积,通常来说一个越好的分类器,Ap 值就越高。mAP 表示多个类别的 AP 的平均值,用来表达更好的模型的衡量指标,其越大表示整体的模型的识别精度越好,交并比表示系统预测出来的框与原来标记的框的重合程度。

本文标注了 100 张含有各种场景的车牌的车辆图片,其中包括各种可能情况对车牌造成影响场景,例如阴天、雾天等对像素造成的影响,泥泞对车牌的遮挡等。本文在 YOLOv3 的模型上采用迭代了 6000 次的权重进行对照实验,实验结果如图 9 所示。

```
detections_count = 3404, unique_truth_count = 1687
class_id = 0, name = normalPlate, ap = 90.69 %
class_id = 1, name = partialCoverPlate, ap = 90.51 %
class_id = 2, name = fullCoverPlate, ap = 88.80 %
class_id = 3, name = blankPlate, ap = 72.30 %
for thresh = 0.25, precision = 0.95, recall = 0.98, F1-score = 0.96
for thresh = 0.25, TP = 1646, FP = 80, FN = 41, average IoU = 80.41 %

mean average precision (mAP) = 0.855768, or 85.58 %
Total Detection Time: 41.000000 Seconds
```

图 9 6000 步迭代权重

Fig. 9 6000-step iteration weight

5 基于 Jetson Tx2 开发板的模型移植

5.1 硬件总体设计

如图 10 所示,Jetson Tx2 同时支持多种类型的接口,例如 USB2.0、UART 等,除此之外,也支持一口多用,其中 USB 3.0、PCIe 和 SATA 接口共享通道,Jetson Tx2 是一台基于 NVIDIA Pascal 架构的单模超级计算机,其由于外形小巧,节

能高效,在智能终端设备上具有很强大的功能,同时能够在嵌入式设备上构建更加大型、复杂的深度神经网络。

Jetson Tx2 提供了完整的适合嵌入式平台的开发软件工具,并且提供了具有能够构建神经网络的应用软件包,包括 TensorRT 加速包、CUDA 软件工具包、VisionWorks、Steamer 和 OpenCV,通过这些工具能够在嵌入式平台上快速地构建 AI 应用,这也是快速构建应用的一种更加全面的解决方案。

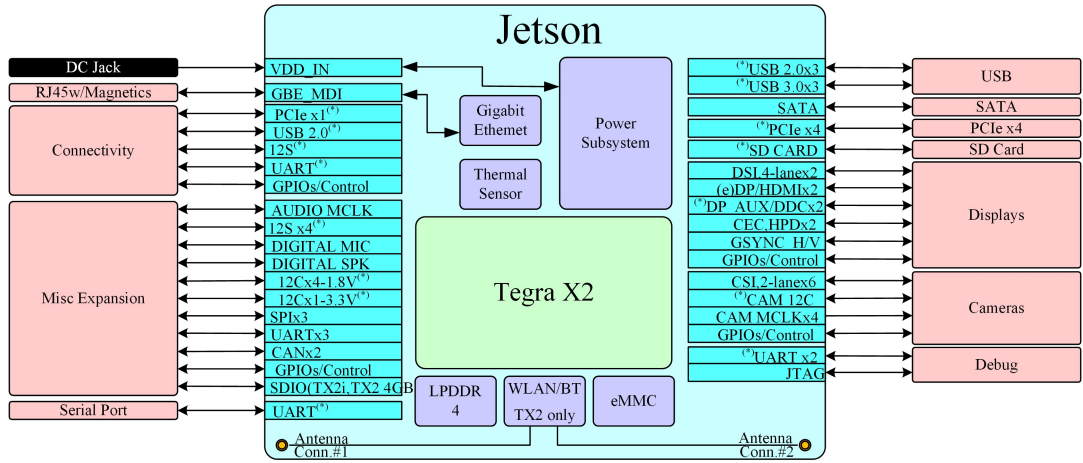


图 10 Jetson TX2 系列模块框图

Fig. 10 Block diagram of Jetson TX2 series modules

5.2 TensorRT 网络推理加速

TensorRT 是一款用于图像分类、分割和对象检测的神经网络。TensorRt 是一种高性能的深度学习推理库,主要是在运行过程中通过优化程序实现。它加速了深度学习推理,缩短了卷积和去卷积神经网络的运行时间。可为深度学习的应用提供低延迟、高吞吐量的推理。

(1)使用不同的精度进行快速推理和部署,满足在嵌入式平台上的不同要求,从而大幅缩短延迟。

(2)专注于开发新型的 AI 驱动的应用程序,而不是推理部署的性能调优。运行时能够确保最佳的推断性能,可以满足最苛刻的延迟和吞吐量要求。

(3)能够有效地减小内存的消耗。同时,对于同一输出的多个分支可以进行并行的运算。

5.3 Docker 模型打包

针对模型的运行和维护来说,使用更加便捷高效的管理工具能缩短维护时间,提升效率。针对以上问题,Docker 具有以下几个优势。

(1)提高部署效率。使用 Docker,开发人员可以使用镜像快速构建标准开发环境,可快速实现版本更迭管理,易于管理和部署。

(2)高效资源利用。与虚拟机不同,属于内核级别的操作,占用系统资源少,对计算机要求低。

(3)易于迁移扩展。Docker 具有很强的兼容性,几乎对平台不做限制,易于相互之间的迁移应用。

(4)便捷的更新管理。Docker 只需要使用 Dockerfile 对配置做部分修改,即可一键部署和更新,对容器的管理更加高效。

Docker 是在操作系统层面上直接复用主机的虚拟化,因此更加轻量级。图 11 是 Docker 在系统层面的架构图。

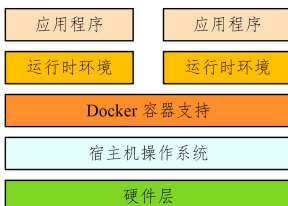


图 11 Docker 架构图

Fig. 11 Docker architecture diagram

Docker 分为镜像、容器、仓库 3 部分。其作用分别为:镜像是打包整个运行环境的基本单元,将算法部署在镜像中打包发布,便于在不同平台进行管理;容器是在镜像的基础上提供的额外可读写文件层,算法的部署在容器中进行;仓库是集中管理镜像的地方。

Docker 在运行前需要本地存在对应的镜像,其操作步骤如下。首先通过搜索命令从默认仓库中进行搜索和过滤,根据拉取镜像信息,例如版本号等,对所需镜像进行拉取。本文移植操作系统版本号为 Ubuntu server 16.04。其次,在镜像的基础上创建容器,在容器中构建所需环境,例如运行模型的框架、动态链接库和依赖等,最后将打包好的容器重新生成镜像并上传到仓库。并且在每次更迭版本之后都要对重新构建的镜像进行上传。仓库、镜像和容器之间的关系如图 12 所示。

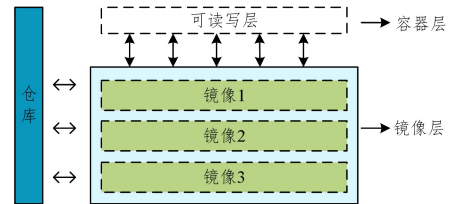


图 12 镜像、容器和仓库关系图

Fig. 12 Relation diagram of mirror image, container and warehouse

5.4 Jetson Tx2 模型移植

本文移植主要采用更加轻量级的框架 Darknet,其主要特点是易于安装,没有依赖项。不同平台之间的移植更加方便。

移植模型的主要过程包括镜像打包和 Tx2 模型移植两部分:

(1)在主机上采用 Docker 构建纯净的移植环境,从默认的镜像仓库中拉取与 Tx2 上搭载的相同的系统镜像: Ubuntu16.04。将算法模型打包在基础镜像上构造的容器中,导出新构建的镜像。

(2)在 Tx2 上安装 Docker, Docker 技术能将运行环境进行隔离,满足不同模型对环境的依赖^[8]。除了保证不同容器之间的模型互不干扰,也可通过高并发条件充分调用硬件资源。若将新构建的镜像移植到 Tx2 的 Docker 中,需要在硬件设备上部署完整的 NVIDIA-Docker2 来使用 CUDA 和显卡驱动。最后启动容器中的模型。

6 实验结果

本文使用 AP 和 mAP 作为评价模型精度的指标,其中 AP 定义为每个类别根据召回率和正确率绘制的 P-R 图下的面积,mAP 则为多个类别 AP 的均值,这两个指标均介于区间 $[0,1]$ 之间,AP 和 mAP 的值越大,分类效果越好,模型的识别精度就越高。但由于 AP 只能体现单类别的识别效果,mAP 只能观测模型的整体识别效果,因此使用两个指标来评

估模型。而在验证实验效率方面,使用召回率(Recall)来评估检测覆盖率,Recall 是指分类器中正确分类的正样本个数占该分类器中所有正样本个数的比例。

本次实验共选取 10000 张车辆测试图片,其中正常号牌图片 8500 张,未悬挂号牌图片 400 张,全部遮挡号牌 200 张,部分遮挡号牌图片 900 张。通过随机分配数量的样本进行测试,根据上述的评判标准,对样本的识别效果进行测试。测试结果如表 2 所列。

表 2 测试样本结果分布表

Table 2 Distribution table of test sample results

序号	类别	评测样本数	算法识别数	算法有效识别数	检出率/%	准确率/%	得分
1	部分遮挡	900	859	816	90.67	94.99	93.26
2	全部遮挡	200	200	136	68.00	68.00	68.00
3	未悬挂	400	658	347	86.75	52.74	66.34
4	正常号牌	8500	8283	8260	97.18	99.72	98.70

模型整体性能的总得分为: $81.39 = 93.26 \times 0.2 + 68.00 \times 0.2 + 66.34 \times 0.3 + 98.70 \times 0.3$ 。

本文中的总得分计算方式采用加权平均值的算法,其中未悬挂号牌在提取特征过程中相对于其他 3 种较为困难,正常号牌特征则更多且更为容易识别和区分,因此未悬挂和正常号牌采用权值 0.3,部分遮挡号牌和全部遮挡号牌则采用权值 0.2。

训练数据样本中全部遮挡号牌和未悬挂号牌的样本数量相对较少,很难包含针对这两种车牌的各种场景,造成因覆盖度不全而发生误检和漏检的情况。本文将 OCR 和深度学习两个模型相结合,一定程度上可以提高模型的检测准确率。

结束语 本文提出将深度学习中的目标检测算法与传统 OCR 技术相结合的方法,来解决污损遮挡号牌的识别问题,同时将模型移植到移动的终端设备上,这不仅提高了污损遮挡号牌中未悬挂和全部遮挡号牌的识别率,还将应用的场景扩展到移动的终端产品上。目前,卷积神经网络(CNN)由于其强大的特征提取能力和鲁棒性被逐渐应用于文字识别领域,使用这种方法来对污损遮挡号牌进行分类,在传统的 OCR 算法的基础上提高对未悬挂和全部遮挡号牌的识别率,甚至可以省略针对检测图片的一系列的预处理操作,例如增强图片对比度、二值化等。对污损遮挡号牌的分类可以简化为目标检测、图片分割两部分,甚至这一系列的操作在目标检测的特征图上只使用卷积神经网络就可以实现。除此之外,使用卷积神经网络算法可以提高系统的鲁棒性和识别精度。

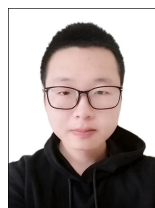
参 考 文 献

- [1] LI C Q. Research on Key Technologies of Fuzzy License Plate Image Recognition in Monitoring Video [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2017.
- [2] HU Z P, CHEN J L, WANG M, et al. New Progress of Convolutional Neural Network Classification Model in Pattern Recognition [J]. Journal of Yanshan University, 2015, 39(4): 283-291.

- [3] LE C Y, BOTTOU L, BENGIO Y, et al. Gradient-Based learning applied to document recognition[C]//Proc. of the IEEE. 1998: 2278-2324.
- [4] BU Y J. Research on Algorithms of License Plate Recognition System Based on Computer Vision [D]. University of Electronic Science and Technology, 2015.
- [5] WANG Z. Research and Improvement of License Plate Recognition System [D]. Shenyang: Shenyang University of Technology, 2016.
- [6] DING M L, DING M Y, NIU Y L, et al. A Method of Identifying Commodity Parameters in Pictures Based on Deep Learning [J]. Journal of Software, 2018, 29(4): 1039-1048.
- [7] ZHANG S L. Research on Classification of Vehicle Brands Based on Vision [D]. Dalian: Dalian University of Technology, 2015.
- [8] ZHANG S. Research on Docker Fault-tolerant Performance Monitoring and Adaptive Pre-replication Checkpoint Technology [D]. Nanning: Guangxi University, 2017.



LI Lin, born in 1977, Ph. D, associate professor. His main research interests include embedded system and signal processing.



ZHAO Kai-yue, born in 1993, postgraduate. His main research interests include machine learning and pattern recognition.