

DOI:10.19431/j.cnki.1673-0062.2022.01.012

基于 SqueezeNet 和 YOLOv2 的交通违法证据评价

刘洪龙¹, 李向阳^{1*}, 徐正华¹, 卢朝晖²

(1. 南华大学 资源环境与安全工程学院, 湖南 衡阳 421001; 2. 浙江力嘉电子科技有限公司, 浙江 绍兴 311800)

摘要:针对交通监控反向抓拍交通违法行为图像预判率高的问题,提出了一种基于迁移学习的多尺度交通违法行为证据评价方法。构建了以 SqueezeNet 为特征提取层、YOLOv2 为目标检测层融合高分辨率细粒度特征的检测网络。通过卷积神经网络算法训练该模型学习抓拍车辆图像特征,识别图像中唯一交通违法行为车辆,再次训练识别驾驶员所在中心区域。在保证特征识别提取能力不变的条件下,采用迁移学习的方式重新训练 SqueezeNet,对驾驶员中心区域图像进行好坏二分类,将成像清晰的交通违法行为图像提交给人工审核。实验结果表明此方法将违法车辆检测准确率提升到 99.3%,驾驶员所在关键区域检测准确率提升到 96.3%,图像质量评价准确率提升到 92.6%,大幅降低了人工审核工作量。

关键词:交通违法行为;图像识别;检测卷积网络;图像质量;迁移学习

中图分类号:TB24 **文献标志码:**A

文章编号:1673-0062(2022)01-0080-08

Traffic Violation Evidence Evaluation Based on YOLOv2 and SqueezeNet

LIU Honglong¹, LI Xiangyang^{1*}, XU Zhenghua¹, LU Zhaohui²

(1. School of Resource Environment and Safety Engineering, University of South China, Hengyang, Hunan 421001, China; 2. Zhejiang Lijia Electronic Technology Co., Ltd., Shaoxing, Zhejiang 311800, China)

Abstract: Aiming at the problem of the high predictive rate of traffic violation images captured by reverse traffic monitoring, a multi-scale traffic violation evidence evaluation method based on migration learning is proposed. A detection network is constructed that uses SqueezeNet as the feature extraction layer and YOLOv2 as the target detection layer to fuse high-resolution fine-grained features. Through the convolutional neural network algorithm, the model is trained to learn the characteristics of the captured vehicle image, identify the

收稿日期:2021-09-14

基金项目:湖南省社会发展领域重点研发项目(2019SK2011);湖南省教育厅重点项目(20A440);国防科工局“十三五”技术基础科研项目(403C001);国防科工局“十三五”技术基础科研项目(2018年)(403B01)

作者简介:刘洪龙(1993—),男,硕士研究生,主要从事道路交通安全与图像处理等方面的研究。E-mail:1962078028@qq.com。*通信作者:李向阳(1962—),男,教授,主要从事核通风与空气净化、铀矿冶辐射防护等方面的研究。E-mail:LCCZCLxy@126.com

only illegal traffic vehicle in the image, and retrain to identify the central area where the driver is located. Under the condition that the ability of feature recognition and extraction remains unchanged, SqueezeNet is retrained by transfer learning to classify the image of the driver's central area as good or bad, and the clear image of traffic violations is submitted to manual review. Experimental results show that this method improves the detection accuracy of illegal vehicles to 99.3%, the detection accuracy of key areas where the driver is located to 96.3%, and the image quality evaluation accuracy to 92.6%, which greatly reduces the workload of manual review.

key words: traffic violation; image recognition; detection of a convolutional network; image quality; transfer learning

0 引言

图像视频处理是当今计算机应用的重要领域,在交通安全监管领域有着重要应用。交通管理部门的非现场执法系统,针对不同的交通违法行为,识别原理各不相同,有些是基于视觉进行预识别,如不系安全带、驾车打电话等违法行为。

由于大量高清监控设备所监控的区域较大,一张针对特定车辆抓拍的照片中,往往存在多辆无关的车辆。在数据审核中,仅根据照片中特定目标区域图像质量,来决定该照片可否作为车辆的违法证据。如针对不系安全带、驾车打电话的违法行为,前排驾驶员所在区域的图像质量,决定了违法证据的可靠性。

图像质量评价一般分为两大类,主观评价方法和客观评价方法。客观评价方法一般分为,全参考、半参考和无参考三类。全参考和半参考需借助参考图像进行对比,无参考评价方法采用提取统计特征并利用分类器判断失真类型,然后利用对应回归模型估算图像质量^[1-5]。

客观评价方法在交通违法证据分类评价中面对整幅图像,关键区域为驾驶员所在小尺度区域,同时缺乏参考图像,实际测试不能得到很好的评价效果。主观评价方法的决策者是人,该种评价方法最准确可靠。近年来随着深度学习在图像识别领域技术的发展,卷积神经网络 AlexNet^[6]、GoogleNet^[7]、ResNet^[8]、SqueezeNet 等在图像分类任务中取得了较好的成果,目标检测算法以 R-CNN、Fast R-CNN^[9]、Faster-RCNN、YOLO 各类方法在车辆检测中得到广泛应用。YOLO 算法经过多次迭代从 YOLOv1^[10]、YOLO9000^[11]、YOLOv3、YOLOv4 到 YOLOv5, YOLO 算法在灵敏度和处理时间上得到了较好平衡。

机器学习应用于实际场景时面临无法获得该

领域大量训练数据的主要障碍,除此之外还面临场景数据量小、模型需要强鲁棒性、个性化定制、数据隐私安全等一系列问题^[1]。迁移学习是一种新的机器学习范式,即使在没有太多标签数据的场景下,也能够帮助提升模型性能^[1]。

本文数据来源于某市公安局交警警察大队,主要针对基于机器视觉的交通违法预判图像,根据不同的违法行为,针对性地提取交通违法预判图像中的关键区域,再对图像关键区域做出质量评价。本文设计了一种 SqueezeNet 与 YOLOv2 结合的目标检测网络,提出了一种基于 SqueezeNet 网络的图像质量主观评价方法。本方法对交警非现场执法系统所抓拍的图像进行预处理,在剔除其中不符合质量要求的违法图像后,再递交工作人员审核,提高违法数据的录用率。

1 YOLOv2 与 SqueezeNet 基本原理

1.1 YOLOv2 基本原理

YOLO 是一个端到端的目标检测网络^[12],能够达到实时的检测效果^[12]。YOLOv2 网络由输入层、卷积层、池化层、全连接层和输出层组成^[13]。YOLOv2 借鉴了 GoogLeNet 思想,采用固定框来预测边界框,引入批量标准化处理,保证了稳定训练与加速收敛^[14]。

YOLOv2 可检测出目标物体在图像中的矩形边界位置,识别物体类别、给出物体分类与物体位置的置信度^[15]。输入图像宽为 W , 高为 H , 将其划分为 $w \times h$ 个网格, 其中 $n = W/w = H/h$, n 为特征检测层分辨率缩放倍数。每个网格中设置 B 个锚框(包围检测物体的边界框)来检测不同大小的物体,每个锚框预测 5 个参数, (t_x, t_y) 代表物体中心坐标, (t_w, t_h) 表示物体的宽和高, t_o 代表物体的置信度。锚框参数以及置信度计算方式为

$$b_x = \sigma(t_x) + c_x, \quad (1)$$

$$b_y = \sigma(t_y) + c_y, \quad (2)$$

$$b_w = p_w \times e^{t_w} \quad (3)$$

$$b_h = p_h \times e^{t_h} \quad (4)$$

$$p_R(o_{\text{bjeect}}) \times I_{OU}(b, o_{\text{bjeect}}) = \sigma(t_o), \quad (5)$$

上式中 $\sigma()$ 为 sigmoid 函数; c_x, c_y 为锚框所在网格左上角坐标; p_w, p_h 为框宽和框高; b_x, b_y, b_w, b_h 为锚框参数, 乘特征检测层缩放倍数 n 可得到检测物体在输入图片中的像素坐标(见图 1)。

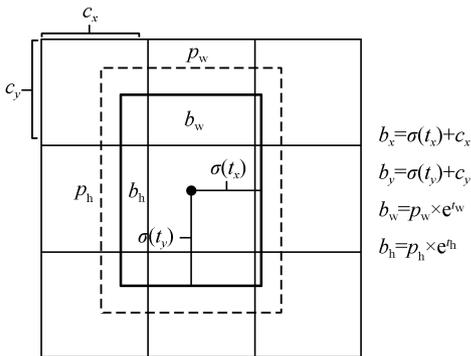


图 1 检测物体预测位置

Fig. 1 Predicted position of detected object

式(5)为检测物体置信度, $p_R(o_{\text{bjeect}})$ 为物体在锚框中的概率, $I_{OU}(b, o_{\text{bjeect}})$ 为锚框与物体的重合度, 即

$$I_{OU} = \frac{S_{\text{pre}} \cap S_{\text{gt}}}{S_{\text{pre}} \cup S_{\text{gt}}} \quad (6)$$

本文只检测机动车 1 类, 先利用公式(1) ~ (6) 求出每个检测物体包围边界框参数和置信度。通过 $p_R(o_{\text{bjeect}}) \times I_{OU}(b, o_{\text{bjeect}})$ 对所有结果排序, 并使用非极大值抑制筛选检测目标获得检测结果。

在训练期间, YOLOv2 对象检测网络优化了预测边界框和训练标签之间的 MSE (mean-square error) 损失^[10]。损失函数 L_{oss} 定义为

$$L_{\text{oss}} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] + \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 + \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2,$$

其中 $\lambda_{\text{coord}}, \lambda_{\text{noobj}}$ 表示权重, x, y, w, h, c, p 为网络预值; $\hat{x}, \hat{y}, \hat{w}, \hat{h}, \hat{c}, \hat{p}$ 为标签标注; I_{ij}^{obj} 检测目标在第 i 格子内, I_{ij}^{obj} 检测目标在第 i 格子的第 j 个预测边界框内; I_{ij}^{noobj} 检测目标不在第 i 格子的第 j 个预测边界框内。

1.2 SqueezeNet 基本原理

SqueezeNet 由 F. N. Iandola 等人提出的一种轻量且高效的 CNN 模型, 可以在保证不损失精度的条件下, 将原始 AlexNet 压缩至原来的 1/510 (<0.5 MB)^[16]。SqueezeNet 具有模型参数少, 对于分布式平行数据训练更加有效^[17]等优点。

SqueezeNet 的核心是 Fire 模块, 包括 squeeze 和 expand 两部分。squeeze 采用 1×1 的卷积核对上一层 feature map 进行卷积, 降低 feature map 维数。expand 包括 1×1 和 3×3 的卷积核且 expand 层卷积核 (1×1 和 3×3 卷积核的总数) 数量大于 squeeze 层卷积核数量。SqueezeNet 主要包含 8 个 Fire 模块, 且以卷积层结束(见图 2)。

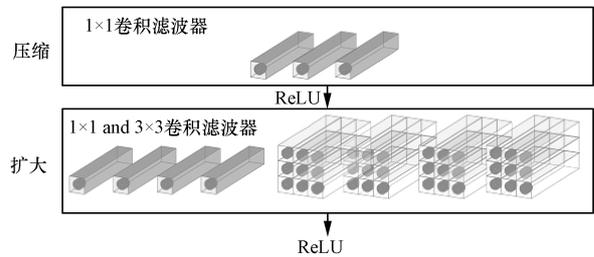


图 2 SqueezeNet Fire 模块

Fig. 2 SqueezeNet Fire module

2 交通违法证据评价模型设计

模型准确率的提高对训练样本标签数量有极大依赖性, 较少的训练样本会导致模型过拟合。ImageNet 数据集远大于交通违法图片样本集, 采用迁移学习的方法将 ImageNet 训练过能识别 1 000 类物体的 SqueezeNet 应用到交通违法车辆识别与图像评价中, 能够提高模型的泛化能力和防止模型过拟合。深度模型迁移学习有两种为基于特征和基于模型的方法, 本文采用一起使用的方式, 通过参数共享和微调, 再次训练模型以更好地适应目标域。

2.1 特征迁移强化目标识别和提取

本文训练标签数据集为 4 000 张交通违法预判图像, 由于训练集样本数量较少, 既要保证机动车检测准确率, 也要尽可能加快检测速度, 故采用

预训练后的轻量级网络模型 SqueezeNet 来增强特征提取能力。Squeezenet 初始网络深度为 18 层,通过 ImageNet 数据库的 100 多万张图像进行训练,可以将图像分类为 1 000 个对象类别,具有较强的特征识别能力。

改进后的 YOLOv2 目标检测网络采用在 ImageNet 数据库训练后的 SqueezeNet 网络替换 YOLOv2 原有骨干网络 Darknet-19。该网络在 SqueezeNet 第 2 个 Fire 模块后的最大池化层后增加一个步长为 2×2 的通道转换层,在 SqueezeNet 最后一个卷积层(conv10)后的 ReLU 层融合特征

后连接到检测子网。检测子网由两组卷积层、ReLU 层、BN 层,再加一个卷积层,YOLOv2 预测边框转换层和输出层构成。该网络将 28×28×128 特征转换为 14×14×512,和 conv10 后的 ReLU 层特征融合为 14×14×1 512。该网络融合了高分辨率特征,从而增强了对小尺度目标的检测准确率。此处借鉴 YOLOv2 passthrough layer^[13]的设计思路,将网络识别准确率提升 1%。本文使用改进后的网络检测交通违法行为,检测违法车辆车窗区域,改进后的 YOLOv2 网络结构如图 3 所示。

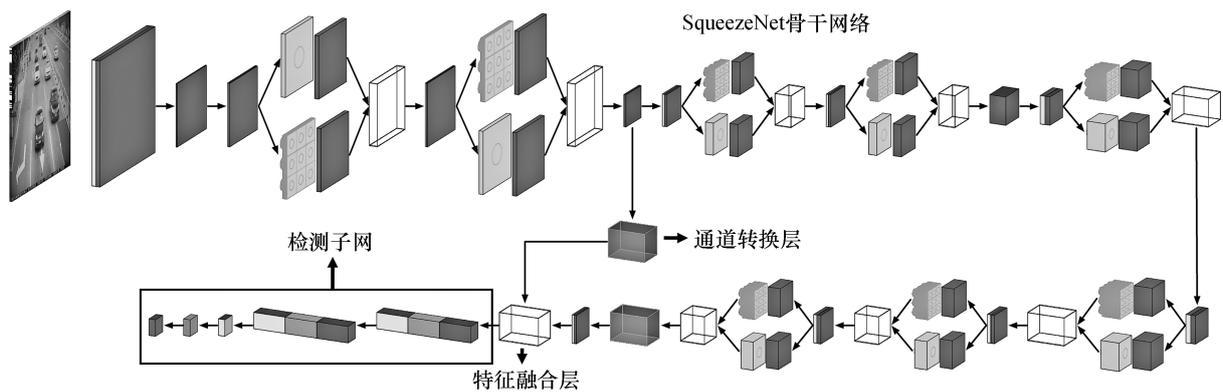


图 3 改进后的 YOLOv2 网络

Fig.3 Improved YOLOv2 network

2.2 模型迁移用于图像质量分类

模型迁移过程如图 4 所示,交通违法行为图像分类模型,选取 SqueezeNet 为骨干网络;1) 在 ImageNet 数据集上预训练 SqueezeNet,其最终分类器输出 1 000 类中的一类;2) 预训练之后,之前的最

终分类器被另一个随机初始化的 Softmax 分类器替换以预测 2 类图像质量,而分类器之前的层保持不变;3) 之前的卷积层都被冻结,继续微调模型最后几个全连接层。

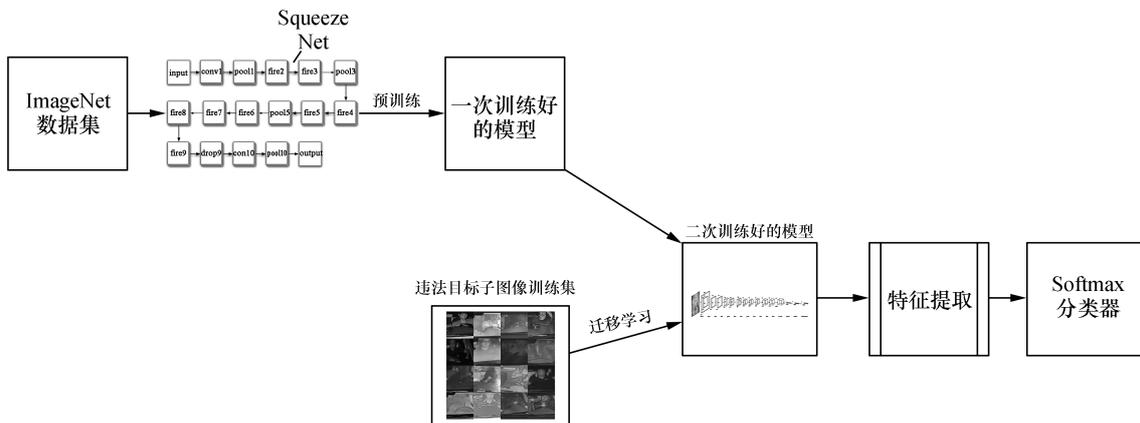


图 4 SqueezeNet 迁移学习流程

Fig.4 SqueezeNet migration learning process

2.3 交通违法证据评价过程

当驾驶员在驾驶过程中拨打接听电话或未按规定使用安全带时,驾驶员所在图像区域为交通违法证据审核的关键区域,本文采用 Softmax 分类器对交通违法预判图像关键区域进行图像质量分类。SqueezeNetYOLOv2_{v1} 为训练后的机动车检测网络,SqueezeNetYOLOv2_{v2} 为训练后的交通违法车辆车窗检测网络,SqueezeNet 网络为训练后的图像质量分类模型,交通违法证据评价过程包括下列步骤:

- 1) 输入原始图像,使用改进的 YOLOv2 检测算法(SqueezeNetYOLOv2_{v1} 网络)检测机动车;
- 2) 判断是否检测到机动车,若检测到则下一步,若未检测到则算法停止;
- 3) 对所有检测到的机动车锚框参数 t_y+t_h 进行排序,找出 t_y+t_h 的最大值所在锚框,并从输入图像中裁剪该锚框,即为唯一违法车辆;
- 4) 使用 SqueezeNetYOLOv2_{v2} 网络对 3) 中唯一违法车辆检测车窗,并裁剪出车窗;
- 5) 对 4) 中违法车辆车窗区域提取驾驶员所在关键区域;
- 6) 使用 SqueezeNet 网络对 5) 中驾驶员所在关键区域进行图像质量分类,给出图像质量分类结果和预测准确率。

3 实验与分析

3.1 数据集

交通违法车辆数据采用交通监控反向抓拍得到,经过交通违法服务器进行一次判别,图像分辨率从 $4\ 096 \times 2\ 208$ 到 $8\ 192 \times 2\ 208$ 。本文取 4 000 张预判图像进行标签标注,标注内容为违法机动车、非违法机动车、违法机动车驾驶员所在区域和违法机动车牌照(见图 5)。

根据驾驶员所在区域图像筛选评价,将该区域图像成像质量进行质量好和差二分类标注。成像质量差的分类定义为:人眼不能清晰地分辨出驾驶员所系安全带特征,不能分辨驾驶员是否在打手机,此类图像不能作为交通违法证据。图像成像质量差的分类主要包括:图像亮度过暗,亮度过曝,驾驶员所在区域有炫彩遮挡和成像模糊等情况,反之则认为图像质量好,可作为交通违法证据。

3.2 数据预处理

图像数据预处理首先提取标签标注区域图像,进一步提取标注区域内标识特征,再将驾驶员所在区域图像归一化和标准化以及多角度旋转。

尺度归一化将提取的图像归一化为 $227 \times 227 \times 3$ 。图像标准化先去均值和调整方差,再利用图像的不变矩特性,寻找一组参数来消除变换函数对图像的影响,消除仿射变换的干扰,提高 SqueezeNet 网络识别率。图像中心区域提取,主要是截取车窗区域右边的 $1/2$ 为主驾驶员区域。设车窗区域图像 I 的宽和高为 W 和 H ,中心处裁剪区域计算方法为 $I_{co}=I(x,y),x \in \left[\frac{9W}{16}, \frac{15W}{16}\right],y \in \left[\frac{H}{8}, \frac{7H}{8}\right]$ 。

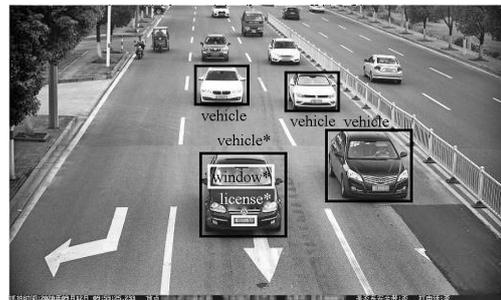


图5 数据集样本标注

Fig. 5 Data set sample annotation

图像进行随机旋转,旋转角度从 $-\frac{\pi}{6}$ 到 $\frac{\pi}{6}$ 并进行翻转。多角度旋转处理能增大样本数据量,提高模型的泛化处理能力,有助于增强网络对图像的识别能力和防止网络过度拟合。

3.3 模型训练

模型训练在 i5-10400@2.9GHz,16 G 内存,GTX1660Ti 6 GB 平台上对网络模型进行训练和测试。操作系统为 win10,测试环境为 CUDA development 11.0。机动车检测网络 SqueezeNetYOLOv2_{v1} 和车窗检测网络 SqueezeNetYOLOv2_{v2} 网络训练参数,输入图像分辨率为 $227 \times 227 \times 3$,初始学习率为 0.001,随机小批量 Mini Batch 为 16,最大迭代数 Max Epochs 为 30。驾驶员所在区域图像筛选评价 SqueezeNet 网络训练参数,输入图像分辨率为 $227 \times 227 \times 3$,初始学习率为 0.000 1,随机小批量 Mini Batch 为 30,最大迭代数 Max Epochs 为 70。

4 个模型使用训练迭代损失函数来提高模型的训练进度和训练效果,4 个模型的迭代训练损失函数整体呈下降趋势(图 6 为 4 个模型经过平滑处理的迭代训练函数曲线),前半段下降速度较快,后半段趋于平缓,在迭代 800 次后趋于稳定,趋向于极小值 0。训练结果表明随着训练迭代次数的增加,识别效果越来越准确并趋于最优。

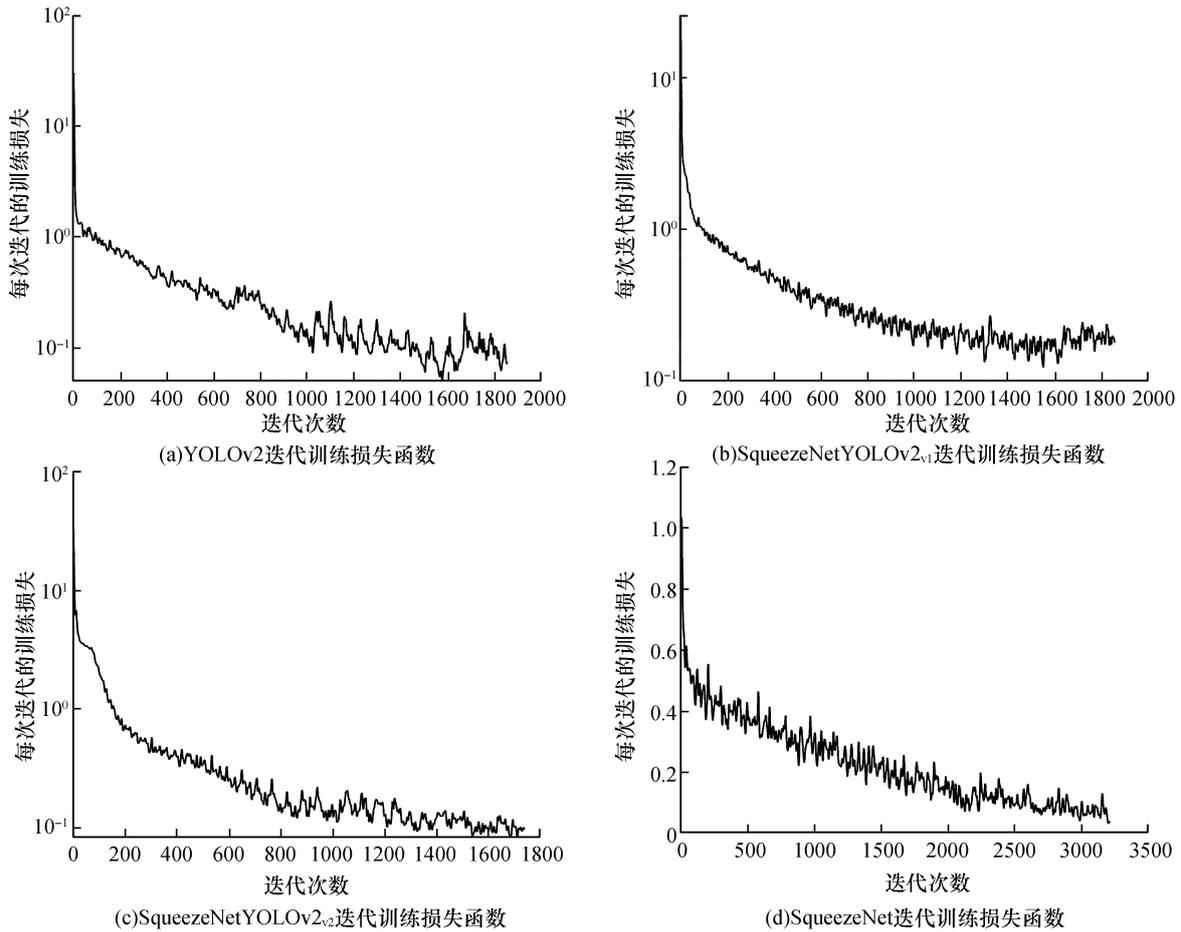


图 6 4 个模型迭代训练损失函数

Fig. 6 4 model iterative training loss function

3.4 训练结果

本文选取了 4 000 张方向抓拍交通违法行为图像作为 YOLOv2、SqueezeNetYOLOv2_{v1}、SqueezeNetYOLOv2_{v2}、SqueezeNet 网络的训练集与测试集。本文模型运行环境为 i5-10400@2.9GHz, 16 G 内存, GTX1660Ti 6 GB 的台式机。

使用 4 组不同的标注样本总计 4 000 个, 分别对 4 个模型进行检验, 设置 I_{OU} 置信阈值为 0.8, 绘制精度-召回曲线(图 7)与平均漏检-误报率曲线(图 8)并计算平均精度。如图 7、图 8 所示 4 个模型都经过 ImageNet 预训练后参数共享与微调, 传统的 YOLOv2 模型平均精度值为 0.85, 替换原有 Darknet-19 采用 SqueezeNet 骨干网络的模型平均精度分别为 0.966 与 1, 平均精度提升明显。当训练样本较少时, 采用 ImageNet 数据集微调方法比从零学习模型或纯监督学习方法能够达到更高的检测精度。

测试结果(见表 1)表明, 传统的 YOLOv2 网络经过训练对交通违法行为图像中机动车的识别准确率为

84.4%, 识别速度为 97.2 帧/s。使用 SqueezeNetYOLOv2_{v1} 结合的目标检测网络, 对交通违法行为图像中机动车的识别准确率为 99.3%, 识别速度为 49.8 帧/s。SqueezeNetYOLOv2_{v2} 网络对违法车辆驾驶员区域识别的准确率为 96.3%, 识别速度为 57.2 帧/s。SqueezeNet 驾驶员中心区域图像质量分类评价模型的分类准确率为 92.6%, 分类速度为 28.7 帧/s。

由此可以看出, SqueezeNetYOLOv2_{v2} 结合的目标检测网络相较于传统 YOLOv2 网络, 随着模型深度和参数的增加, 识别准确率提升了 14.9%。虽然驾驶员所在区域检测模型 SqueezeNetYOLOv2_{v2} 与违法车辆 SqueezeNetYOLOv2_{v1} 检测网络结构相同, 但模型的输入图像分辨率不同。SqueezeNetYOLOv2_{v1} 的输入为整幅图像, SqueezeNetYOLOv2_{v2} 的输入为 SqueezeNetYOLOv2_{v1} 检测到的唯一违法车辆, 输入图像分辨率尺度变小从而导致 SqueezeNetYOLOv2_{v2} 比 SqueezeNetYOLOv2_{v1} 检测速度有所提升。

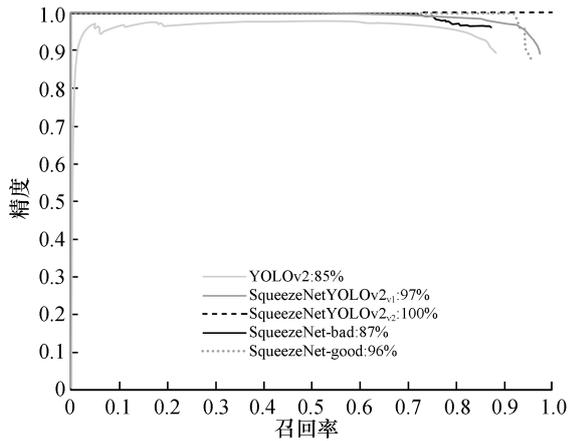


图7 精度-召回曲线

Fig.7 Data set sample annotation

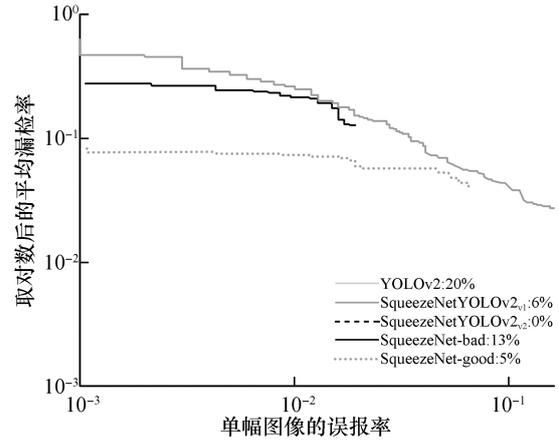


图8 平均漏检-误报曲线

Fig.8 Average missed-false positive curve

本文为查看 SqueezeNet 网络对驾驶员所在中心区域的图像质量分类效果引入类激活映射(class activation maps, CAM)^[18],生成驾驶员所在区域热力图,突出了图像类的特定区域。本文选取4张交通违法图像进行展示(如图9),通过热力图所在范围发现,SqueezeNet 网络类激活映射区域都集中在驾驶员区域,从侧面验证了模型的准确性。图9也展示了4张交通违法图像机动车检测、违法车辆检测和图像质量评价的结果和置信度。

表1 4 000 个样本集/1 000 个测试集结果

Table 1 4 000 sample set/1 000 test set results

网络模型	准确率/%	平均精度/%	平均漏检率/%	检测速度/(帧·s ⁻¹)
YOLOv2	84.4	85	20.3	97.2
SqueezeNetYOLOv2v1	99.3	96.6	5.7	49.8
SqueezeNetYOLOv2v2	96.3	100	0	57.2
SqueezeNet-bad	92.6	86.7	13.4	28.7
SqueezeNet-good	92.6	95.6	4.9	28.7

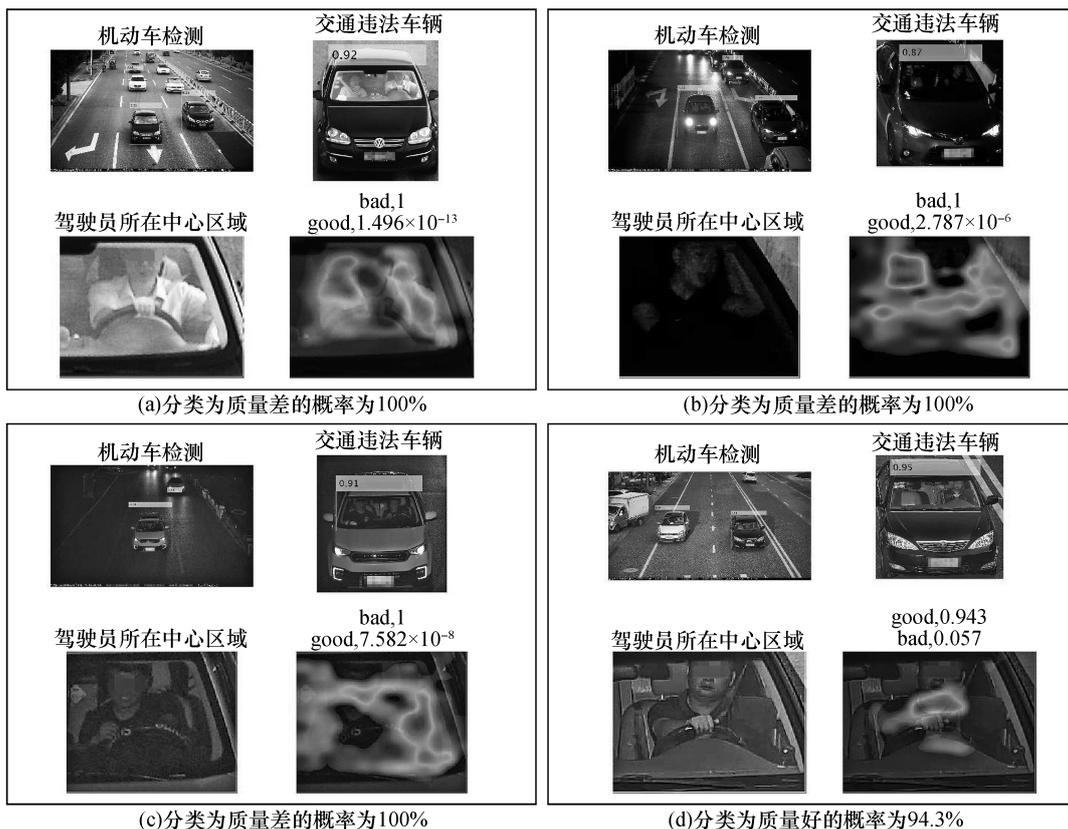


图9 多尺度交通违法证据筛选结果

Fig.9 Multi-scale traffic violation evidence screening results

4 结 论

本文为提高驾驶员开车打手机、开车不系安全带两种违法行为的预判准确率,减少交通违法证据审核无关区域,只对驾驶员所在关键区域进行判断的方法来提高模型运行速度和准确性。使用迁移学习的方式重新训练 SqueezeNet 与 YOLOv2 结合的模型对交通违法车辆和驾驶员所在区域进行目标检测,使用重新训练的 SqueezeNet 网络对驾驶员所在区域进行图像质量分类评价,提升了 YOLOv2 网络对机动车检测的准确率。使用 Squeezenet 为目标特征提取网络,减少了模型参数,既提高了模型的检测能力又保证了模型检测速度。通过改进后的 YOLOv2 模型,通过融合高分辨率细粒度特征提高了传统 YOLOv2 对小尺寸目标的检测敏感度。

参考文献:

- [1] 杨强,张宇,戴文渊,等. 迁移学习[M]. 北京:机械工业出版社,2020:194-204.
- [2] LAHSAINI I, DAHO M E H, CHIKH M A. Deep transfer learning based classification model for COVID-19 using chest CT-scans [J]. Pattern recognition letters, 2021, 152:122-128.
- [3] MA W, LIU Z, KUDYSHEV Z A, et al. Deep learning for the design of photonic structures[J]. Nature photonics, 2021, 15(2):77-90.
- [4] 钟睿,杨立,杜永成. 深度迁移学习预训练对红外尾流成像识别的影响[J]. 红外技术, 2021, 43(10):979-986.
- [5] 张雪芹,陈嘉豪,诸葛晶晶,等. 基于深度学习的快速植物图像识别[J]. 华东理工大学学报(自然科学版), 2018, 44(6):887-895.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84-90.
- [7] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Boston, USA:IEEE, 2015:1-9.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, USA:IEEE. 2016:770-778.
- [9] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision (CVPR). Boston, USA:IEEE. 2015:1440-1448.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, USA: IEEE. 2016:779-788.
- [11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Honolulu, USA: IEEE, 2017:7263-7271.
- [12] 王琳, 卫晨, 李伟山, 等. 结合金字塔池化模块的 YOLOv2 的井下行人检测[J]. 计算机工程与应用, 2019, 55(3):133-139.
- [13] 赖秋频, 杨军, 谭本东, 等. 基于 YOLOv2 网络的绝缘子自动识别与缺陷诊断模型[J]. 中国电力, 2019, 52(7):31-39.
- [14] 咎珊珊, 李波. 融合改进 YOLOv2 网络的视觉多目标跟踪方法[J]. 小型微型计算机系统, 2020, 41(12):2601-2606.
- [15] 龚静, 曹立, 亓琳, 等. 基于 YOLOv2 算法的运动车辆目标检测方法研究[J]. 电子科技, 2018, 31(6):5-8.
- [16] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [J]. Computer vision and pattern recognition, 2016, 20(3):1-22.
- [17] 鞠维欣, 赵希梅, 魏宾, 等. 基于深度学习的肝硬化识别[J]. 青岛大学学报(自然科学版), 2018, 31(4):54-61.
- [18] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, USA: IEEE. 2016:2921-2929.