

融合 SIFT 和级联分类器的特种车辆 自动检测识别

唐海涛^{1,2}, 吴果林^{1,2}, 范广义³, 陈迪三¹

- (1. 桂林航天工业学院 理学院, 广西 桂林 541004;
2. 桂航大数据技术应用研究中心, 广西 桂林 541004;
3. 南京高精齿轮集团有限公司, 江苏 南京 210000)

摘要:针对特定场景中特种车辆因多环境影响因素下数据不均衡、检测精度和识别准确率低的问题,提出一种融合尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)和级联分类器的特种车辆自动检测及识别预测方法。首先,图像预处理后运用 SIFT 特征提取图像主体区域特征点及特征描述子;其次,结合 SIFT 特征点的密度调整优化算法实现目标车辆检测;最后,运用 KMeans 聚类算法获得目标检测框中 SIFT 特征描述子的中心聚类点,生成表征目标主体图像的 128 维特征描述子,并最终输入 RF-RBF(Random Forest-Radial Basis Function)级联分类器进行学习并识别预测。该文均采用 K 折交叉验证方法保证模型的稳定性和可靠性。实验结果表明,在特定场景下特种车辆目标检测获得了 75.47% 平均交并比,级联分类器在特种车辆识别的综合准确率、精确率、召回率、F1-Score 值以及 FPS 值分别为 87.35%、88.17%、97.27%、92.38% 以及 21。进一步验证融合 SIFT 和级联分类器模型具有较好的自动化检测准确性和识别分类能力。

关键词:尺度不变特征变换;KMeans;RF-RBF 级联分类器;K 折交叉验证;特种车辆

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2023)09-0182-08

doi:10.3969/j.issn.1673-629X.2023.09.027

Automatic Detection and Recognition of Special Vehicles Incorporating SIFT and Cascade Classifier

TANG Hai-tao^{1,2}, WU Guo-lin^{1,2}, FAN Guang-yi³, CHEN Di-san¹

- (1. School of Science, Guilin University of Aerospace Technology, Guilin 541004, China;
2. Research Center for Big Data Technology Application in Guat, Guilin 541004, China;
3. Nanjing High-Speed & Accurate Gear Group Co., Ltd., Nanjing 210000, China)

Abstract: In order to overcome the unbalanced data and low detection and recognition accuracy of special vehicles in specific scenes caused by a number of environmental influencing factors, a Scale-Invariant Feature Transform (SIFT) and cascade classifier method for special vehicle identification and recognition prediction is proposed. Firstly, the SIFT features are used to extract the feature points and feature descriptors of the image subject area after image pre-processing. Secondly, the density adjustment optimization algorithm of SIFT feature points is combined to realize the target vehicle detection. Finally, the KMeans clustering algorithm is used to obtain the central clustering points of the SIFT feature descriptors in the target detection frame to generate 128-dimensional feature descriptors characterizing the target subject image, and finally input RF-RBF (Random Forest-Radial Basis Function) cascade classifier for learning, identifying and prediction. All K-fold cross-validation methods are used to ensure the stability and reliability of the model. The experimental results show that 75.47% average cross-comparison ratio is obtained for special vehicle target detection in specific scenarios, and the combined accuracy, precision, recall, F1-Score and FPS values of the cascade classifier in special vehicle recognition are 87.35%, 88.17%, 97.27%, 92.38% and 21, respectively. Such model has better automatic detection accuracy and recognition classification ability.

Key words: scale invariant feature transformation; KMeans; RF-RBF cascade classifier; K-fold cross-validation; special vehicle

收稿日期:2022-10-23

修回日期:2023-02-23

基金项目:国家自然科学基金项目(62001134);广西高校中青年教师科研基础能力提升项目(2019KY0992);桂林航天工业学院校级基金项目(XJ21KT28)

作者简介:唐海涛(1989-),男,通讯作者,工程师,硕士,CCF会员(L7908M),研究方向为机器学习、自然语言处理与图像识别;吴果林(1977-),男,教授,博士,研究方向为模式识别、大数据与复杂网络。

0 引言

据公安部统计,截至 2022 年 3 月底,全国机动车保有量已突破 4 亿辆^[1],全国各地车辆的快速增长和普及使得当地交通监管部门面临更加严峻的管控和安全维护等问题;同时,全国各地区的车辆终端监测设备实时不间断的数据监测采集使得后端积累了海量、不确定、多样化的交通数据。传统人工或简单交通信息化系统在现有的数据规模面前已面临巨大压力。因此,智能交通系统应时而生且至关重要,其依托人工智能技术的发展,使得车辆自动化检测及识别成为智能交通系统中的重要一环,因而在公共突发事件预警、灾害救援以及军事等领域具有较好的应用前景;而其中特别是针对特种车辆^[2](即装置有特殊设备、具备特殊功能,用于承担特殊运输任务或执行特种作业任务的车辆)的自动化检测识别与管理,相较于其他类型车辆检测识别的研究更具有迫切性和现实性。

当前,国内外许多研究机构和学者对如何提高车辆自动化检测识别的准确性、实时性等方面已进行了大量的探索与研究。传统的车辆识别检测方法主要聚焦于图像浅层特征描述子,比如:方向梯度直方图(Histogram of Oriented Gradients, HOG)、尺度不变特征变换(SIFT)、梅尔倒谱系数(Mel-Frequency Cepstral Coefficient, MFCC)、Gabor 小波变换以及颜色边缘特征等^[3-5];在此基础上,进一步结合各分类器实现车辆的自动化识别,其常见的分类器包括贝叶斯(Bayes)分类器、支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)、高斯混合模型(Gaussian Mixture Model, GMM)以及自适应增强算法(Adaptive Boosting, AdaBoost)分类器等^[6-8]。上述多种研究方法的实验结果也都验证了其模型可降低车辆自动化检测识别的算力依赖,同时也能进一步提高检测识别的准确性以及实时性;但对于多领域特定场景(如高速出入口检查站、防疫检查点等)下部分数据不平衡、高维稀疏特征空间等现实问题,传统模型方法普遍存在泛化能力、抗噪能力以及目标识别准确率普遍弱化的趋势。如 Iqbal U 等人^[9]结合 SIFT 和 Sobel 特征实现 65% 的 rank-1 车辆类型识别准确率,但方法的特征维度高,运算量大;李成娟等^[10]提出一种加权的基频自适应梅尔倒谱系数特征提取算法,较传统 MFCC 算法提升了识别准确率,但该算法受噪声(包括旋转、光照亮度、拍摄视角、复杂背景等)影响较大,算法稳定性降低;黄应清等^[11]在装甲车辆识别场景中结合使用小波矩和支持向量机获得了较高的实验识别率。综合上述方法,主要是通过提取一般普通车辆(即非特种车辆)图像的低层次特征进行训练并学习,但这对于特种车辆图像的抽象表示

其可适应性和拓展性都不强;同时,近几年得益于分布式计算和存储技术的快速发展,计算机算力和协同性得以大幅提升,使得深度学习模型在计算机视觉领域的应用得到极大推动,尤其是面向深层图像特征的有效提取;其中,Chen Y 等^[12]提出一种基于深度迁移学习方法并将其应用于特种车辆的识别场景中,该方法借助迁移学习使得深度学习训练样本需求减少且获得较好的识别率,但该方法实际训练过程中存在一定程度的过拟合,弱化了算法鲁棒性和泛化能力;张陈晨等^[13]结合一种改进的 YOLOv3-Tiny 目标检测算法实现对车辆检测,通过将部分卷积层替换池化层思路提高了车辆识别率;但算法针对一般普通车辆数据集且同时需要较高的人工图像标注成本;王斌等^[14]在航空图像车辆检测中提出一种基于特征金字塔网络(Feature Pyramid Network, FPN)的改进 Faster RCNN(Faster Regions With CNN Features)算法,在 VEDAI 数据集测试中的平均检测精度(Mean Average Precision, mAP)为 77.2%,该方法具有较好的图像检测精度,但同样针对特种车辆的检测识别场景下的性能表现有待进一步实验数据的支撑和验证。

上述多种研究方法为车辆有效的检测识别打下了基础,但总体上都集中于对一般普通车辆的识别检测研究,对于特种车辆的检测识别研究较少;同时,当下较热门的深度学习模型在各类图像复杂检测任务中均获得了较好的检测识别效果,但由于深度学习模型往往包含一定规模量的参数,通常需要大量有监督的数据进行模型训练才能获得较好的效果,而真实应用场景下特种类型车辆出现的频率较低,市场保有量也远低于其他类型车辆且存在分布不均衡等现实特点,从而使得深度学习模型在极少量受监督的样本下其泛化性能难以获得较高的提升,这也进一步说明了特种车辆的检测识别难度也普遍高于一般普通车辆。但同时特种车辆的用途、价值以及所代表的形象等受损往往又会带来更大的社会生命财产安全警示和隐患,而历史的公共自然灾害警示人们,特种车辆在特定时间点以及多样化场景下的识别、管理以及深层次应用研究非常重要,它能为军事、特种专业作业等领域提供强大的工具力支撑。基于此,针对救护车、消防车、警车、校车四大类特种车辆在小样本数据集下的自动检测识别,该文提出一种融合 SIFT 和级联分类器的特种车辆自动检测及识别方法。该方法主要结合传统的 SIFT 图像特征,同时通过引入基于密度调整的优化算法实现目标车辆较优检测,最后通过引入 KMeans 聚类算法获得目标检测框中 SIFT 特征描述子的中心聚类点,并输入级联 RF-RBF 分类器实现特种车辆有效的自动化识别预测。

具体工作如下:

(1) 结合 SIFT 图像特征, 提出一种基于特征点密度优化的目标检测方法, 有效提高了无监督下的图像检测精度。

(2) 针对复杂背景下的特种车辆数据集, 引入 KMeans 聚类算法, 实现特征描述符中心点聚类及数据降维, 增强图像检测区域有效信息的表达及无用特征信息的摒弃。

(3) 提出一种针对特种车辆识别预测的 RF-RBF 级联分类器模型, 提高了模型识别准确率。

1 相关工作

1.1 SIFT 特征点检测及描述

高质量的特征点是特种车辆图像检测识别的重要基础依赖, 完成局部图像特征的提取与描述可以更好地辅助后续物体检测、识别以及比对等复合型应用。尺度不变特征变换 (Scale-Invariant Feature Transform, SIFT) 通过引入尺度参数模拟人在由近到远时目标在人眼视网膜上的成像过程, 最终在不同的尺度空间序列中查找并获得极值点 (即特征点) 以及相关信息, 并依此融合成特征向量描述符。因尺度空间需满足一定的不变性 (包括灰度不变性、对比度不变性等), 使得 SIFT 特征具备旋转、尺度缩放、亮度变化等不变性特点, 与此同时, 特征相对于视角变换、仿射变化、噪声也保持着良好的鲁棒性, 是一种优秀的局部特征描述算法。融合 SIFT 图像处理算法首先通过引入尺度参数, 构建连续变化的尺度参数以获取不同尺度下的空间序列表示, 图像的尺度空间 $L(x, y, \sigma)$ 可定义为输入图像 $I(x, y)$ 与变化的高斯函数 $G(x, y, \sigma)$ 进行卷积, 即:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

其中,

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

式中, $I(x, y)$ 表示图像灰度值, $G(x, y, \sigma)$ 为高斯核函数, (x, y) 为像素坐标, $*$ 表示卷积运算, σ 为尺度变化因子, 亦可称为模糊参数, 其值的大小决定图像的平滑程度, 值越小图像模糊程度越轻, 反之模糊程度越重。接下来, 为了更高效地在尺度空间检测稳定的关键点, 需要将每组内相邻尺度两组图像进行两两“相减”操作 (即下一层减去上一层), 从而生成高斯差分 (Difference Of Gaussian, DOG) 金字塔, 具体用 $D(x, y, \sigma)$ 表示, 其公式为:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (3)$$

式中, k 为两个相邻尺度空间的比例因子, 关键点的提

取则转化为在 DOG 尺度空间中进一步寻找局部极值点, 具体需要将每一个检测点与其同尺度的 8 个邻域像素点以及相邻上下两层邻域 9×2 个像素点共计 26 个点一一进行比较, 若像素点在 26 邻域中取得最大或最小值, 则生成了图像在该尺度下的一个极值点 (即特征点)。

在 DOG 尺度空间提取到的极值点是离散的, 并非真正的极值点, 需要进一步通过拟合三维二次函数来精确计算关键点的位置和尺度, 提高关键点的健壮性, 具体需要对 DOG 函数在尺度空间 $D(x, y, \sigma)$ 上进行 Taylor 式展开, 其矢量形式为:

$$\begin{aligned} D(x + \Delta x) &= D(\Delta x) + \frac{\partial D^T(x)}{\partial x} \Delta x + \\ &\quad \frac{1}{2} \Delta x^T \frac{\partial^2 D(x)}{\partial x^2} \Delta x \end{aligned} \quad (4)$$

上式对 Δx 求导数, 并令其等于 0, 则可以计算极值点的偏移量:

$$\Delta x = \frac{\partial^2 D(x)^{-1}}{\partial x^2} \frac{\partial D^T(x)}{\partial x} \quad (5)$$

当 Δx 偏移量大于阈值 (设定为 0.5) 时, 需要重新定位, 并在新的位置上进行差值计算, 反复操作直至收敛从而实现关键点的准确计算, 若迭代次数超过上限迭代值仍未收敛, 则删除该极值点。

其次, 为了使特征描述子具有旋转不变性, 还需要确定关键点的指定方向参数, 具体通过计算极值点邻近高斯金字塔图像的 3σ 邻域窗口内的梯度值, 并使用直方图统计邻域内像素的梯度和方向, 其中直方图的峰值所在的方向即表示为该关键点的主方向。

最后, 经上述计算, 获得各关键点的三个核心信息: 即尺度、位置以及方向。考虑到将上述三维度信息融合为表征关键点的统一描述符, 故用一组向量将关键点特征描述出来, 具体向量生成方法是通过将关键点周围图像进行区域 4×4 窗口分块, 并计算块内 8 个方向的梯度直方图, 最终生成一个 $4 \times 4 \times 8 = 128$ 维的表示向量, 上述向量即为关键点特征描述符。

1.2 基于特征点密度优化的目标检测

经 SIFT 检测获得图像所有特征点后, 主体目标区域的检测可以通过统计全部特征点坐标值进行标定。设二维随机遍历 (X, Y) 表示所有特征点坐标, 则目标检测框左上角坐标 $P_{lu}(X = x_{\min}, Y = y_{\max})$, 右下角坐标 $P_{rd}(X = x_{\max}, Y = y_{\min})$ 计算方式如下:

$$\forall x_i \in X, \exists x_{\min} \in X \& x_i \neq x_{\min}, \text{st. } x_{\min} < x_i$$

$$\forall x_i \in X, \exists x_{\max} \in X \& x_i \neq x_{\max}, \text{st. } x_{\max} > x_i$$

那么 x_{\min} 、 x_{\max} 分别是 X 随机变量严格的全局最小点、最大点; 同理, y_{\min} 、 y_{\max} 分别是 Y 随机变量严格的全局最小点、最大点。

$$\forall y_i \in Y, \exists y_{\min} \in Y \& y_i \neq y_{\min}, \text{st. } y_{\min} < y_i$$

$$\forall y_i \in Y, \exists y_{\max} \in Y \& y_i \neq y_{\max}, \text{st. } y_{\max} > y_i$$

但考虑到实际目标图片中背景特征的复杂性,导致一小部分特征点的坐标值偏离车辆主体覆盖区域,从而为后续实现准确、有效的特种车辆目标检测及识别带来极大干扰;即存在多个非主体特征点。基于此,采用基于特征点密度优化的目标检测算法,具体算法流程如下:

Step1:以 $x_{\text{split}} = x_{\min} + 2$ 为分割线起点,步长为 2 Pixel 向右进行移动切割;

Step2:每次切割,分别统计分割线左边区域 LArea 以及右边区域 RArea 的所有特征点个数 $\text{nums}(\text{LArea})$ 和 $\text{nums}(\text{RArea})$;其中 Larea 和 Rarea 具体表示如下:

$$\text{LArea} = \{(x_i, y_i) \mid x_{\min} \leq x_i \leq x_{\text{split}}, y_{\min} \leq y_i \leq y_{\max}\}$$

$$\text{RArea} = \{(x_i, y_i) \mid x_{\text{split}} \leq x_i \leq x_{\max}, y_{\min} \leq y_i \leq y_{\max}\}$$

Step3:当 $\text{nums}(\text{LArea})$ 计算出来时,则 $\text{nums}(\text{RArea}) = \text{Total} - \text{nums}(\text{LArea})$, Total 表示 SIFT 检测的所有特征点总数;并将其 $\text{nums}(\text{LArea})$ 值追加写入 NH 数值列表中;

Step4:向右不断移动分割线,直到 $x_{\text{split}} \geq x_{\max}$ 则停止横向移动,并生成最终的 NH;

Step5:在等间距情况下计算 NH 数值列表的一阶向前差分,获得特征点数量的增量值 ΔNH_i ,其中 $i \in (0, \text{size}(\text{NH}))$,步长为 1;

Step6:从索引值 j (其中 $0 \leq j < \lceil \frac{\text{size}(\text{NH})}{2} \rceil$),步长为 1 开始检索增量值 ΔNH 列表,当第一次出现 $\Delta \text{NH}_i \geq 2$ 时,停止检索,并记录此时索引值 j ;

Step7:根据索引值 j ,反推计算新的 $x_{\min} = x_{\min} + (j + 1) * 2$;

Step8:将图像以 $\frac{x_{\min} + x_{\max}}{2}$ 中轴进行翻转,重新执行步骤 1 ~ 步骤 7,反推计算新的 $x_{\max} = x_{\max} - (j + 1) * 2$;

Step9:综上,完成了新的 x_{\min} 、 x_{\max} 计算,同理,以 $y_{\text{split}} = y_{\min} + 2$ 为分割线起点,步长为 2 Pixel 向上进行移动切割;反推计算完成了新的 y_{\min} 、 y_{\max} 计算;

Step10:至此,完成了基于特征点密度优化的目标检测全部算法流程。

1.3 基于 RF-RBF 级联分类器

考虑到特种车辆数据集占总体数据规模仅约 11% (特种车辆 430 张图像,非特种车辆 3 420 张),为了弱化正负样本数据集不平衡带来的后续识别影响,采用基于 RF-RBF 级联分类器的设计结构,可使得降低算力依赖的同时可极大程度地避免过拟合以及提高分类识别的准确率和效率。因此,在获得图像 SIFT 特征描述符后,继续将其输入两层级联分类器 RF-RBF 模型。其中第一层 RF 分类器是一个二分类器,主要将图像 Label 识别为特种车辆和非特种车辆;第二层 RBF 分类器是一个多分类器,在第一层获得特种车辆的标签后继续输入第二层 RBF 分类器^[15],实现消防车 (XiaoFangChe, XFC)、警车 (JingChe, JC)、邮政车

(YouZhengChe, YZC) 以及校车 (XiaoChe, XC) 四大特种车辆类型识别。同时第二层 RBF 分类器识别到的非特种车辆类型可逆向校正第一层 RF 分类器的识别准确率,进而相互之间形成反馈机制以提高整体模型的综合预测准确率。

2 融合 SIFT 和级联分类器的特种车辆检测识别建模

该文融合 SIFT 和级联分类器的特种车辆自动识别检测模型主要包括 SIFT 特征点提取、基于特征点密度优化的目标检测、SIFT 特征描述向量生成、RF-RBF 级联分类模型,其中 SIFT 特征点提取实现图像主体区域分布大致标定;接着,结合特征点密度优化算法实现主体区域 (即车辆区域) 的准确检测;其次,生成主体区域特征描述向量,完成主体区域数字化特征表示;最后,输入级联分类器 RF-RBF 学习训练后实现特种车辆的识别,且 RBF 可逆向校正 RF 模型的识别准确率。综合上述检测识别预测,最终获得不同类型的特种车辆划分。具体融合 SIFT 和级联分类器的特种车辆自动识别检测算法流程如下:

Step1:输入原始图像数据;

Step2:图像预处理,具体包括去噪、尺寸统一;

Step3:将数据集分为训练集和测试集,其二者占比为 8 : 2;

Step4:图像的 SIFT 特征提取,包括极值点 (即特征点) 和关键点特征向量;

Step5:将 Step4 中提取的所有特征点坐标,结合基于特征点密度优化的目标检测算法,实现目标车辆检测;

Step6:在目标检测内,提取训练集 SIFT 特征描述符,并采用 KMeans 聚类算法,对一张图像中 N 个关键点的特征描述符进行中心点聚类,并最终获得一张表征图像的 128 维特征描述符向量;

Step7:经 Step6 获得图像的数字表征向量后,并将数据集标签划分为特种车辆和非特种车辆,继续将其输入 RF 分类器进行二分类识别;

Step8:将 RF 分类器输出结果为特种车辆数据集接着输入 RBF 分类器,而 RF 判定为非特种车辆数据集则会直接判定为非特种车辆标签;

Step9:完成 RBF 模型对特种车辆类别的判定,若判定为某一类型特种车辆,则获得车辆类别标签;若不属于特种车辆类型,则逆向校正 RF 模型的识别准确率,进而通过反馈机制提高 RF 分类器识别预测准确率。

具体检测识别流程参见图 1。

3 实验及结果分析

3.1 图像数据集及预处理

实验数据均来源于网络采集,共四类特种车辆以及非特种车辆类型数据 (特种车辆图片数据已脱敏),特种车辆分别是消防车 (XiaoFangChe, XFC)、警车

(JingChe, JC)、邮政车(YouZhengChe, YZC)以及校车(XiaoChe, XC),共计 430 张图像;其中非特种车辆类型数据(Other, OH)则包括诸多类型车辆,如家用小轿

车、皮卡、面包车等,共计 3 420 张,整体数据分布均与真实场景基本一致。

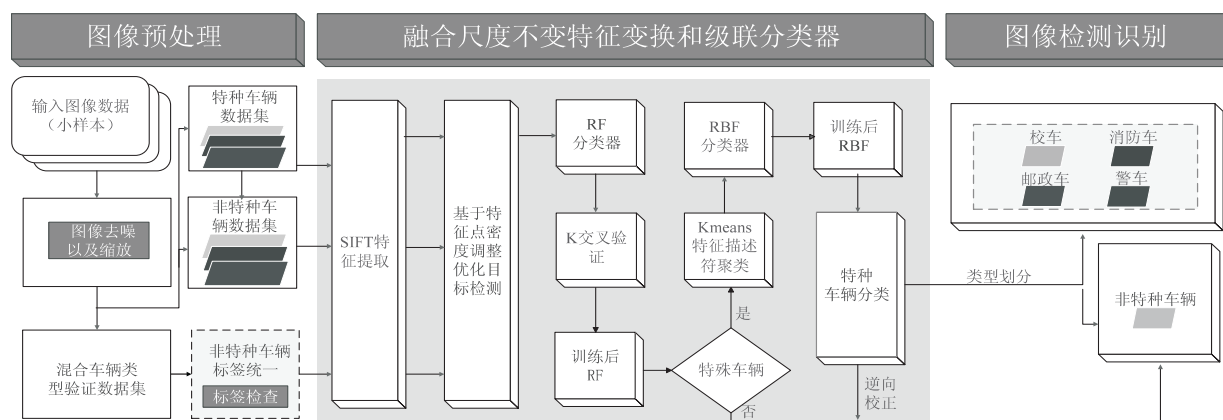


图 1 融合 SIFT 和级联分类器的特种车辆预测模型流程

在图像数据预处理阶段,考虑网络图像采集多源化以及环境复杂化,导致原始图片数据集存在尺寸大小不一、清晰度低、成像模糊等问题。因此,为提高图像真实信息表达,增强检测和识别的准确性和可靠性,该文首先对图像尺寸做归一化处理,将所有图片尺寸大小调整为 300X250 dpi(参考了一般图像的长宽比例,即 4:3);其次考虑到检测识别物体时,关键真实信息在于图像边缘(即梯度信息),所以进一步将 RGB 彩色图像做灰度化处理,可实现算力成本低依赖且能最大程度降低图像失真所带来的负影响,同时综合考虑到人眼对绿色信息的敏感性最高,且对蓝色敏感则最低的这一特性,进一步采用 RGB 三分量加权平均获得灰度图像;具体计算公式如下:

$$\text{Gray}(i,j) = 0.299 * R(i,j) + 0.578 * G(i,j) + 0.114 * B(i,j) \quad (6)$$

最后,通过引入典型的低通滤波器——中值滤波^[16](3×3),统计图像中以某点(i,j)为中心的小窗口内(3×3)的所有像素的灰度值,按从大到小顺序排列,并将中间值作为该点的新灰度值。从而实现图像边缘的保护且能去除突变的噪声,以此提升图片的清晰度。

在图像数据检测识别阶段,为更客观评价目标检测模型,使用了 Label Studio 工具对车辆图像进行人工标注,针对每个特种车辆类别(共计 4 类)标注 50 张,同时其他类型车辆也标注 50 张,共计 250 张人工预标定图像。其中 Label Studio 人工标注后返回的标记区域由四个参数 x、y、width 以及 height 确定,它们依次表示:起始点 x 轴坐标值、起始点 y 轴坐标值、标记区域的宽度和标记区域的高度(其中实际图像的宽度和高度分别用 W 和 H 表示)。但考虑到后端返回的标签值需与前端设置的标签值保持一致,还需要进一步通

过换算公式生成实际图像标记区域,具体左上角坐标(x_{\min} , y_{\min})以及右下角坐标(x_{\max} , y_{\max})换算公式如下(round 表示对浮点数进行四舍五入取值):

$$\begin{aligned} x_{\min} &= \text{round}(x * W/100) \\ y_{\min} &= \text{round}(y * H/100) \\ x_{\max} &= \text{round}((x + \text{width}) * W/100) \\ y_{\max} &= \text{round}((y + \text{height}) * H/100) \end{aligned}$$

3.2 实验参数与评估指标

文中模型与算法的实现主要以 Python 编程语言为主(版本 3.8.8),并结合开源的跨平台计算机视觉库 OpenCV 为框架(基于 Python),编程的集成开发环境为 Pycharm2021.2 Community Edition。实验采用了图像处理领域常用的局部特征检测算法 SIFT 作为后续特种车辆目标检测识别的基础依赖,为了使得算法能够更好地实现特种车辆局部特征的提取,采用了 cv2.SIFT_create().detect 做特征点检测,其中特征点个数设置为 100,同时设置特征点角点的程度(即特征点的强度)阈值大于等于 0.03;同时,在多分类器学习时,采用 K 交叉验证方法(选取 $k=5$)对训练集和测试集按 4:1 比例分拆并轮询互换,且分类器模型设置每次迭代次数上限为 10 000,正则化系数为 $1e-3$ 。

为了验证检测和识别模型算法的有效性,采用了图像检测识别常用的衡量指标:交并比 IOU (Intersection Over Union)、平均交并比 MIOU (Mean Intersection Over Union)、准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall)、F1-score 值以及每秒传输帧数 FPS (Frames Per Second, FPS)。其中,用以衡量两区域之间的重叠程度 IOU 和 MIOU,其具体计算公式如下:

$$\text{IOU} = \frac{S(A \cap B)}{S(A \cup B)} \quad (7)$$

$$\text{MIOU} = \frac{1}{k+1} \sum_{i=0}^k \frac{S_k(A \cap B)}{S_k(A \cup B)} \quad (8)$$

其中, $S(A \cap B)$ 表示图像物体实际区域 (A) 与模型检测区域 (B) 的重合面积; $S(A \cup B)$ 表示两个区域整体所占有的面积; k 表示车辆的类别; S_k 表第 k 类别的面积。

$$\text{Accuracy (Acc)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (9)$$

$$\text{Precision (P)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (10)$$

$$\text{Recall (R)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (11)$$

$$\text{F1 - score} = \frac{2 \times P \times R}{P + R} \quad (12)$$

式中, TP、FP、TN 以及 FN 分别表示真阳性 (True Positive)、假阳性 (False Positive)、真阴性 (True Negative) 和假阴性 (False Negative)。

3.3 结果与分析

3.3.1 主流目标检测方法比较

为了进一步和目前主流的目标检测方法进行实验对比分析,该文引入了 YOLOv5s 算法和两类目标检测算法对不同车辆图像进行目标检测 (其中 LineStyle 为 · 表示人为标注, -- 表示覆盖全部特征点的目标检测,以及 - 表示基于特征点密度优化的目标检测) 以及进行 IOU、MIOU 指标统计,其中单个类别的 IOU 值以该类别多张图像的平均值作为最后评估的 IOU 值,而 MIOU 是多个类别的平均 IOU 指标统计值,具体统计结果如表 1 所示。

表 1 与主流目标检测方法在 5 类车辆检测图像的 IOU 统计结果

	IOU		
	基于全特征点覆盖的目标检测	YOLOv5s	基于特征点密度优化的目标检测
消防车	0.670 5	0.738 0	0.712 8
邮政车	0.585 0	0.781 0	0.788 1
警车	0.614 4	0.813 5	0.832 0
校车	0.598 9	0.762 4	0.671 6
其他	0.542 0	0.750 5	0.769 0
MIOU	0.602 2	0.769 1	0.754 7
训练时长/h	0	3.5	0
检测时长/ms	35	24	40

由表 1 可知, YOLOv5s (You Only Look Once: Unified, Real-Time Object Detection, YOLO) 在各类车辆的 IOU 指标统计中均取得了更优异的结果,整体的平均交并比 MIOU 达到 0.769 1,而基于特征点密度优化的目标检测算法的 MIOU 值为 0.754 7,但由于

YOLOv5s 模型是建立在大量有监督标注数据集的基础之上,而基于特征点密度优化的目标检测算法是一种无监督自动检测算法,二者在 MIOU 值相当的情况下,该文提出的图像检测算法在目标检测阶段的优越性主要表现在三方面:无需人为设置滑动窗口比例、无需大量人工标注、低算力成本依赖,且模型表现出较好的泛化能力。于传统下 $\text{IOU} \geq 0.5$ 即可认为目标检测模型获得了一个正确的预测结果,显然该文的于特定场景下基于特征点密度优化的车辆图像目标提取算法是适用的。

3.3.2 不同级联分类器比较

在目标检测基础上,选取目标检测框中的特征点生成对应的 128 维关键点特征描述符,而因优化后的目标检测主体区域包含的关键点数量 N ($0 < N \leq 100$) 随实际情况不同而不统一,故用 $N * 128$ 向量表征车辆数字特征则会带来高维度以及样本向量维度不一致的问题。为了降低数据维度以及统一训练分类样本,采用 KMeans 聚类算法,对一张图像中 N 个关键点的特征描述符进行中心点聚类,并最终获得一张表征图像的 128 维特征描述符向量,后续分类器将以该向量为输入数据进行分类训练和测试。

文中特种车辆识别采用级联分类器进行结果预测,分类器第一层设计首先将所有图像数据集分为两类:特种车辆和非特种车辆;然后分别使用常见的逻辑回归 (Logistic Regression, LR)、决策树 (Decision Tree, DT)、支持向量机 (Support Vector Machine, SVM)、随机森林 (Random Forest, RF) 以及径向基神经网络 (RBF) 进行二分类训练预测,其中 RBF 网络参数选择识别准确率较优的 453 个 (实验从 128 至 500 进行训练后挑选综合准确率较高的径向基神经元个数) 径向基神经元以及基函数的拓展常数 $\delta = 1.0$;最后结合准确率、精确率、召回率和 F1 值 4 方面综合比较各算法的性能,实验采用 K 折交叉验证 (选取 $k = 5$) 进行算法测试,实验结果如表 2 所示。

表 2 级联分类器 (一层) 各项性能指标对照

模型	准确率	精确率	召回率	F1 值
逻辑回归	0.880 5	0.887 3	0.990 2	0.935 9
决策树	0.863 5	0.892 0	0.961 5	0.925 4
随机森林	0.888 1	0.888 4	0.998 4	0.940 2
支持向量机	0.881 4	0.881 4	1.000 0	0.937 0
RBF	0.529 9	0.591 5	0.737 3	0.476 4

由表 2 可知,综合对比各类分类识别性能指标,其中随机森林分类模型的识别准确率、F1 值分别为 0.888 1 和 0.940 2,在综合对比各分类器中其性能表现最佳。

图像经过级联分类器第一层后,模型已对是否为特种车辆类型进行了识别,接下来将判别为特种车辆类型的图像继续输入级联分类器的第二层进行多特种类型识别细分预测分类。为了验证多分类模型在多特种类型识别预测中的性能表现,继续沿用表 2 中的前四类算法进行综合实验测试,具体特种车辆多分类识别预测结果参数见表 3。

表 3 级联分类器(二层)各项性能指标对照

模型	准确率	精确率	召回率	F1 值
逻辑回归	0.652 1	0.603 7	0.607 7	0.597 7
决策树	0.547 2	0.513 7	0.503 7	0.503 6
随机森林	0.656 0	0.634 0	0.619 2	0.615 3
支持向量机	0.660 5	0.568 8	0.579 4	0.539 2

由表 3 可知,在多类型特种车辆识别预测中综合性能最优的仍旧是随机森林模型,但是其最高的准确率仅仅为 0.656 0,远远达不到实际场景对特种车辆类型识别所要求的一般准确率,故继续结合 RBF 模型对多特种类型识别进行预测。理论上,Poggio 和 Girosi 已证明 RBF 网络是连续函数的最佳逼近^[17],相较于传统分类识别算法,其对于多类型预测具有较大优势。因考虑到不同的 RBF 神经元个数会影响模型的预测效果,所以将神经元个数作为优化输入参数(其中基函数的拓展常数固定为 1.0),将数据集按 8:2 的比例分为训练集和测试集,并以测试集 F1 值达到最大化作为优化目标,并引入了神经元参数 NP (Neuron Parameters) 从 128(即数据的输入维度)到 500 的 RBF 模型对测试集 F1 值的收敛曲线,随着参数的增加,收敛曲线的波动在降低,实验结果最终确定 NP=282 后,总体趋于收敛于 1。

当确定 RBF 模型参数后,采用 K 折交叉验证重新对特种类型图像数据进行特种车辆多类别预测,详细

识别预测统计结果可参考表 4。

表 4 RBF 特种车辆多类别预测统计

值	准确率	精确率	召回率	F1 值
最小值	0.953 5	0.979 6	0.923 3	0.947 5
最大值	1.000 0	1.000 0	1.000 0	1.000 0
平均值	0.983 6	0.992 5	0.974 3	0.982 6

综合分析表 3 和表 4,K 折交叉验证下 RBF 模型在多类识别预测中的准确率、精确率、召回率以及 F1 值分别为 0.983 6、0.992 5、0.974 3 以及 0.982 6,相较于传统较优的分类模型(随机森林)算法在总体上准确率、精确率、召回率和 F1 值分别提高了 32.76 百分点、35.85 百分点、35.51 百分点和 36.73 百分点。故 RBF 在特种车辆多类型识别预测中具有较好的准确率和可靠性;同时,侧面也验证了 RBF 网络在连续多类型预测应用场景中具有较好的优势。

3.3.3 可视化展示

SIFT 特征点提取、基于特征点分布的目标检测以及基于特征点密度优化的目标检测识别算法预测部分可视化效果细节展示可参考图 2(a)~(d),其中+表示特征点,--表示基于特征点最大最小坐标点检测框,·表示基于密度优化的特征检测框,-表示人工标注检测框。从结果来看,调整角点阈值的 SIFT 局部特征提取算法可以很好地检测车辆主体区域,具有较强的特征点敏感性;同时基于覆盖全部特征点的目标检测策略在实际图像中能很好地保证覆盖车辆主体区域,但由于该方法过于放大检测区域,导致交集区域的增长速度低于并集的增长速度,反而会给优化目标 IOU 带来负影响,一定程度上影响了模型的目标检测效果;而基于特征点密度优化的目标检测算法相较于上述特征点分布的目标检测算法在检测精度方面带来更好的提升作用,图 3 为实际图像目标检测效果图。

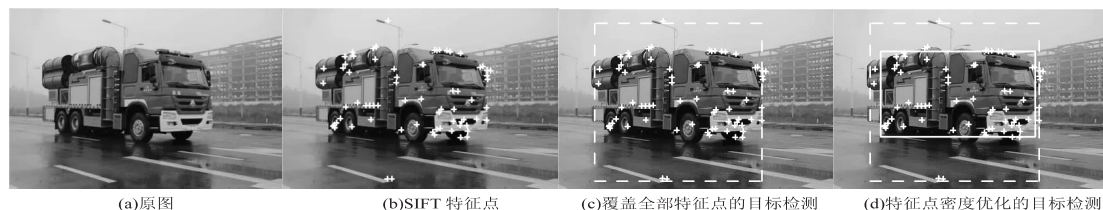


图 2 SIFT 特征点提取以及车辆目标检测效果图

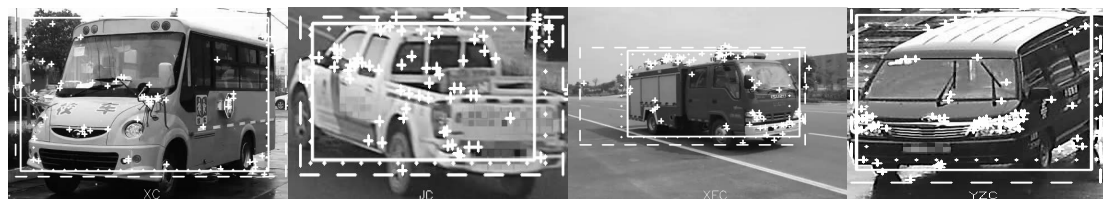


图 3 基于特征点密度优化的不同特种车辆类型目标检测效果图

为了进一步验证分析 RF-RBF 级联分类器的识别性能,特引入当前主流的图像检测识别算法

YOLOv5s 进行横向对比分析,结果如表 5 所示。

通过表 5 中的实验测试指标数据可知,RF-RBF

级联分类器在小样本数据集上仅实时识别速度性能低于 YOLOv5s,其余多个识别性能指标均高于 YOLOv5s 算法,且文中模型的训练时间远远低于对比的检测识别算法;进一步分析各指标值,级联分类器其召回率相较于识别准确率高出近 10 百分点,这也使得模型大大降低了类型误识别的风险,增强了模型的稳定性,同时由于 RBF 特种车辆识别准确率已达 0.983 6,测试集中仅有两个预测值为负数(即无法识别的特种车辆类型,认定为非特种车辆类型),所以逆向校正效果并不明显,倘若在 RBF 网络模型识别率不高的情况下,其逆向修正的效果才会更加突出。故在目标识别分类阶段,提出的识别模型相较于 YOLOv5s 的优势在于更少的训练时间成本以及较好的分类综合性能。

表 5 特种车辆多类别综合识别结果对比

模型	准确率	精确率	召回率	F1 值	FPS
YOLOv5s	0.840 3	0.851 8	0.903 6	0.870 8	38
RF-RBF 分类	0.873 5	0.881 7	0.972 7	0.923 8	21

4 结束语

提出了一种融合尺度不变特征变换和级联分类器的特种车辆自动检测及识别预测方法,一方面,通过优化图像检测算法提高检测精度;另一方面,通过改进级联分类器设计以提高识别准确率和稳定性。仿真实验表明,融合 SIFT 特征和级联分类器实现了特定场景下特种车辆高效准确的检测识别,特种车辆综合目标检测 MIOU 为 75.47%,识别准确率达到 87.35%,基本能满足特定场景下特种车辆的实时自动化检测识别需求。该体系框架后续可进一步拓展至针对特定需求和部门的车辆类型检测识别应用领域。但是,由于方法聚焦于单一特定的检测识别场景,使得复杂场景下同一时间刻度多目标检测识别存在误检、漏检的问题。下一步工作重点将在 SIFT 特征点的基础上,引入拓扑数据分析技术^[18](Topological Data Analysis, TDA),利用复形捕捉特种车辆主体区域数据空间的拓扑 s 特征,实现多领域复杂场景下多目标的检测识别。

参考文献:

- [1] 公安部:全国机动车保有量突破 4 亿辆[EB/OL]. [2022-04-07]. http://www.gov.cn/xinwen/2022-04/07/content_5683950.htm.
- [2] 徐 达,蔡崇贤. 专用汽车结构与设计[M]. 北京:北京理工大学出版社,1998.
- [3] LIU S, YAN X, LI P, et al. Radar emitter recognition based on SIFT position and scale features[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2018, 65(12): 2062-2066.
- [4] YANG Z, HUANG Y. Algorithm for speech emotion recognition classification based on mel-frequency cepstral coefficients and broad learning system[J]. Evolutionary Intelligence, 2022, 15(4): 2485-2494.
- [5] BOUGHIDA A, KOUAHLA M N, LAFIFI Y. A novel approach for facial expression recognition based on Gabor filters and genetic algorithm[J]. Evolving Systems, 2022, 13(2): 331-345.
- [6] SHVAI N, HASNAT A, MEICLER A, et al. Accurate classification for automatic vehicle-type recognition based on ensemble classifiers[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(3): 1288-1297.
- [7] ZHAO J, WANG X. Vehicle-logo recognition based on modified HU invariant moments and SVM[J]. Multimedia Tools and Applications, 2019, 78(1): 75-97.
- [8] XIE J, ZHU M. Maneuver-based driving behavior classification based on random forest[J]. IEEE Sensors Letters, 2019, 3(11): 1-4.
- [9] IQBAL U, ZAMIR S W, SHAHID M H, et al. Image based vehicle type identification[C]//2010 international conference on information and emerging technologies. Karachi: IEEE, 2010: 1-5.
- [10] 李成娟, 易 强, 李宝清, 等. 基于基频的梅尔倒谱系数在车辆识别中的应用[J]. 重庆大学学报, 2021, 44(11): 17-23.
- [11] 黄应清, 赵 锴, 蒋晓瑜, 等. 基于小波矩和支持向量机的装甲车辆识别研究[J]. 装甲兵工程学院学报, 2012, 26(3): 61-64.
- [12] CHEN Y, YANG C, YANG S Y. A method for special vehicle recognition based on deep-transfer model[C]//2016 sixth international conference on instrumentation & measurement, computer, communication and control (IMCCC). Harbin: IEEE, 2016: 167-170.
- [13] 张陈晨, 靳 鸿. 基于改进 YOLOv3-Tiny 的目标检测技术研究[J]. 兵器装备工程学报, 2021, 42(9): 215-218.
- [14] 王 斌, 顾银娟. 基于 FPN 的航空图像车辆检测技术研究[J]. 工业控制计算机, 2020, 33(7): 88-90.
- [15] 滕少华, 唐海涛, 张 巍, 等. 混合 PNN 和 RBF 多神经网络模型的局域降雨类型识别及雨量预测[J]. 小型微型计算机系统, 2016, 37(11): 2571-2576.
- [16] TANG H, NI R, ZHAO Y, et al. Median filtering detection of small-size image based on CNN[J]. Journal of Visual Communication and Image Representation, 2018, 51: 162-168.
- [17] WANG Fang. RBF networks and applications based on evolutionary niche[D]. Shanghai: Shanghai Jiaotong University, 2008.
- [18] JASSIM S, ASAAD A. Automatic detection of image morphing by topology-based analysis[C]//2018 26th European signal processing conference (EUSIPCO). Rome: IEEE, 2018: 1007-1011.