文章编号　1004-924X(2022)07-0854-15

基于改进Mask R-CNN＋LaneNet的车载图像车辆压线检测

孙建波[1,2]，张　叶[1*]，常旭岭[1]

(1. 中国科学院 长春光学精密机械与物理研究所 应用光学国家重点实验室，
吉林 长春 130033；
2. 中国科学院大学，北京 100049)

摘要：针对辅助驾驶或自动驾驶领域车载图像的车辆压线检测问题，以及检测过程中由于欠曝、阴影或实体遮挡等因素导致的漏检、误检问题，提出基于改进Mask R-CNN与LaneNet的车辆压线检测算法。在网络优化方面，在Mask R-CNN网络的基础上将RoI Align层的图像缩放算法(双线性插值)改进为双三次插值，将全连接层卷积化的VGG16网络取代LaneNet的E-Net共享解码器；在图像增强方面，改进Gamma校正算法以实现欠曝图像的自动校正；在训练数据方面，完成Tusimple数据集中车辆目标的标注并基于改进的随机擦除算法在网络训练过程中进行数据增强。实验结果表明：车辆检测速度保持不变的同时车道线检测速度提升了28%，车辆漏检率、误检率分别降低了38.93%，89.04%，车道线漏检率、误检率分别降低了67.21%，87.05%，算法的性能指标可满足车辆压线判断的应用需求。

关　键　词：图像校正；实例分割；车道线检测；数据增强；压线检测

中图分类号：TP394.1；TH691.9　　**文献标识码**：A　　**doi**：10.37188/OPE.20223007.0854

Vehicle pressure line detection based on improved Mask R-CNN + LaneNet

SUN Jianbo[1,2], ZHANG Ye[1*], CHANG Xuling[1]

(1. *State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun* 130033, *China*;
2. *University of Chinese Academy of Sciences, Beijing* 100049, *China*)
* *Corresponding author, E-mail: yolanda@spirits.ai*

Abstract: To address the problem of vehicle pressure line detection of on-board images in the field of assisted or automatic driving, as well as the problem of missed and false detection caused by underexposure, shadow, or solid occlusion in the detection process, a vehicle pressure line detection algorithm based on improved Mask R-CNN and LaneNet was proposed. In terms of network optimization, based on the Mask R-CNN network, the image scaling algorithm (bilinear interpolation) of the ROI alignment layer was improved to bicubic interpolation, and the convoluted VGG16 network of a full connection layer was replaced by LaneNet's E-Net shared decoder. For image enhancement, the Gamma correction algorithm

收稿日期：2021-09-22；**修订日期**：2021-10-19.

基金项目：吉林省与中国科学院科技合作高技术产业化专项基金资助项目(No. 2020SYHZ0031)

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

was improved to realize the automatic correction of underexposed images. In terms of training data, the vehicle target in the Tusimple data set was marked, and the data were enhanced in the network training process, based on the improved random erasing algorithm. The experimental results show that while the vehicle detection speed remains unchanged, the lane line detection speed is increased by 28%, and the vehicle missed and false detection rate are reduced by 38.93% and 89.04%, respectively. Further, the lane line missed and false detection rate are reduced by 67.21% and 87.05%, respectively. The achieved performance index can meet the requirements of vehicle line pressing judgement method.

Key words: image correction; instance segmentation; lane line detection; data enhancement ; line pressing detection

1　引　言

车辆压线检测是指对车辆在行驶过程中轧越实线（白色实线、单黄实线、双黄实线、黄色虚实线）通行的违章行为进行检测，因为车辆压线会占用同向车道或对向车道，造成拥堵、刮蹭甚至事故，危害公众的财产及生命安全[1]，因此，基于车载图像的车辆压线检测研究是辅助驾驶乃至自动驾驶技术中环境感知不可或缺的一部分，对于提高自动驾驶系统的可靠性和保障行车安全具有重要意义[2]。

在车载图像的车辆压线检测研究中，车道线的提取和车辆压线判断是关键。车道线检测是车辆压线判断的重要前提环节。其算法大体上分为：基于特征的算法、基于模型的算法和基于深度学习的算法。基于特征的车道线检测方法主要根据车道与周围环境的特征差异再结合其他方法提取车道线。常用的特征包括颜色[3]、灰度[4]、梯度[5]和边缘宽度[6]等，在根据特征识别出车道线的前提下结合 Hough 变换[7]、粒子滤波[8]或 Kalman 滤波[9]等进一步提取最终的车道线。这种传统的车道检测算法需要根据不同的街道场景人工手动调节滤波算子和参数，工作量大且鲁棒性较差，当行车环境出现明显变化时，易出现误检、漏检的问题，效果不佳，且对于车道特征的完整性依赖较大，适用性较差。

基于模型的车道线检测方法通过利用道路车道线的先验知识，建立合适的数学模型来描述真实的车道线。为提高对于弯曲车道线检测的适应性，文献[10]提出一种基于 Hough 变换的分段直线模型，以此检测拟合弯曲车道线，但此算法对于急弯道车道线的检测容易失效。为解决传统 Hough 变换方法在弯道工况下识别能力不足的问题，文献[11]提出了一种动态区域规划的双模型车道线检测算法，根据“微元”理论进行图像动态细化分区对各区域局部独立识别，利用直线和 B 样条曲线双模型完成车道线拟合。为解决实际交通场景中光照条件恶劣引起的车道线检测性能降低的问题，Yoo 等[12]提出了一种基于梯度增强转换的具有光照鲁棒性的车道线检测算法。然而，这种检测方法在车道线间断或存在遮挡的场景下不能很好地工作，针对具体场景需要建立对应的成像模型，具有较大的局限性。总的来说，这类算法对于特定道路的车道线检测准确性较高，比基于特征的车道线检测算法的鲁棒性强，但是有较强的局限性，计算成本过大，不适应复杂的环境。

与传统算法相比，基于深度学习的车道线检测算法表现出更准确、鲁棒性更强的特性。起初的深度学习算法大部分基于分类器进行车道检测，文献[13]从图像感兴趣区域提取车道线的 Haar 特征并引入级联的车道线分类器，对车道线进行粗检测，然后借助线段检测器对粗检测结果进行车道线拟合。基于卷积神经网络（Convolutional Neural Network，CNN）的深度学习方法通过强大的表现学习能力将场景理解推向了一个新的水平，但对于具有长结构区域并且可以被遮挡的车道线，仍然表现不佳。针对

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

这一问题，文献[14]提出了空间CNN（Spatial CNN，SCNN），SCNN将feature map的行或列视为层，并顺序地应用卷积、非线性激活和求和操作，形成了一个空间上的深度神经网络，因为空间信息可以通过层间传播来加强，在CNN不连续或杂乱的情况下，SCNN可以很好地保持车道标记的平滑性和连续性；但是其检测速度较慢（7.5 frame/s），实际适用性较差。文献[15]提出了一种自注意蒸馏模块，基于信息蒸馏解决大主干网络对速度的影响，可以聚合文本信息。该方法允许使用更轻量级的骨干，在保持实时效率的同时实现了高性能。类似YOLOv3或SSD的单阶段模型，文献[16]提出了基于锚点的车道检测注意机制，该机制表现SOTA达到了250 frame/s，但以上两种网络难以保证存在遮挡及光线条件恶劣场景下车道线的检测精度，应用场景受限。

现有的压线检测算法大部分基于卡口监控设备获取的固定区域图像进行检测判断。文献[17]根据压线重叠区域特征进行综合判断实现压线检测。文献[18]利用帧数统计判断当前图像的车道线边界状态，然后根据车道线的不同状态分别采用检测车前部的均匀灰度区域和检测车头复杂边界信息的方法，来实现车辆的越线检测。文献[19]针对道路中间的单双实黄线区域，采用相邻两幅图像逐行平均灰度值进行比较的方法实现压线判断。不同于固定卡口图像的压线检测，基于车载图像的车辆压线检测算法可以应用于更加复杂多变的交通场景，但相关研究较少。文献[20]综合利用车道线的几何特征与图像特征，基于小波变换提取车道线区域，提出了一种基于图像掩蔽局部匹配的越线行为判定方法，但该方法易受到非车辆遮挡车道线的干扰，从而造成误判。文献[21]在提取车道线的基础上提出根据车载图像中左右车道线夹角来判断车辆压线的方法，此方法简单易实现，但只限于直线、无大面积遮挡的车道，测试得出的压线车道线的夹角范围对于多车道（三车道以上）路况并不适用。文献[22]利用GoogleNet-FCN语义分割网络来识别车道线和wheel-line线，根据车道线和wheel-line线的交点位置设置阈值判断是否压线，但此方法中最佳阈值的设定取决于不同的压线数据，需要根据不同场景对模型进行针对性训练，泛化能力差。

针对上述问题，本文提出基于改进的Mask R-CNN＋LaneNet的车载图像车辆压线检测方法。采用双立方插值取代Mask R-CNN网络中感兴趣区域（Region of Interest，RoI）Align层的双线性插值算法，将特征候选框内的图像缩放为固定大小的RoI feature；改进随机擦除（Random Erasing，RE）算法进行训练数据的增强，强化网络对车辆更多局部特征的认知，提高Mask R-CNN对于环境的鲁棒性和检测准确度；将VGG16全连接层卷积化取代LaneNet网络的E-Net解码器，以减少参数计算量和内存消耗，提升车道线的检测速度；采用Gamma校正算法根据不同图像的灰度分布自动调节校正算子，以适当扩大图像中高灰度区域的动态范围并提升图像亮度，提高车载图像中车辆车道线的辨识度从而降低漏检率、误检率；最后，结合车辆分割与车道线拟合结果，提出基于简化的RoI图像提取wheel_line和车道拟合直线的方法判断车辆压线。

2　算法原理

2.1　算法总体流程

基于车载图像的车辆压线检测算法分为两个阶段实现，如图1所示。在训练阶段，分别改进Mask R-CNN车辆检测网络和LaneNet车道线检测网络；为了应对实际交通场景下实体或阴影遮挡以及远景小目标带来的漏检问题，通过改进RE算法进行数据增强提升网络对于复杂环境的鲁棒性。在检测阶段，为解决欠曝光或光线不足等对车道线、车辆检测的干扰，改进Gamma算法校正图像，并基于改进的Mask R-CNN，LaneNet网络训练生成的权重模型进行车道线、车辆的提取，最后完成车辆压线的判断。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

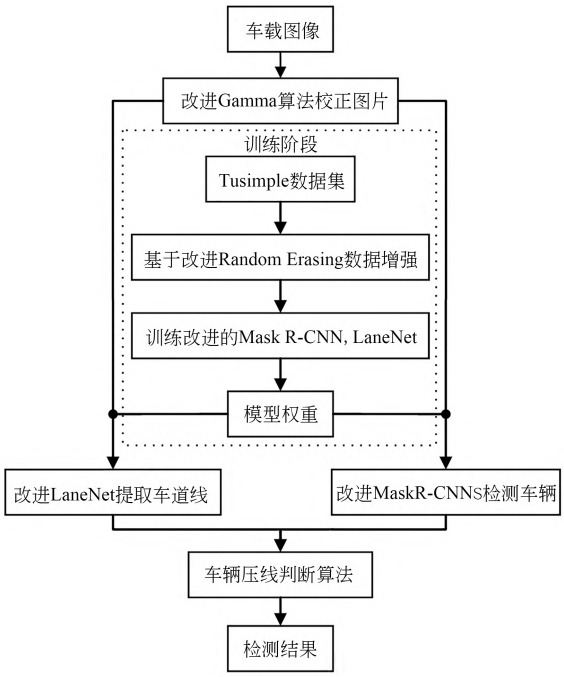


图 1　车辆压线检测算法流程

Fig. 1　Flow chart of vehicle pressing line detection algorithm

2.2　Random Erasing 数据增强

在实际交通场景中，车辆会存在阴影或实体遮挡的情况，而 Mask R-CNN 模型的实例分割建立在其目标检测分支获取预测框的基础上，所以本文参考 RE 数据增强算法[23]，在模型训练过程中，通过随机擦除目标车辆的局部特征模拟不同的遮挡效果，使模型在训练过程中仅通过局部特征识别车辆，强化模型对于车辆局部特征的辨识能力以提高车辆的查全率。

改进后的 RE 算法流程如图 2 所示。首先对数据集中的车辆目标进行标注得到标签文件，由标签内容获得车辆在图像中的位置信息，遍历图片中的目标车辆以概率 P 在其真实框内创建矩形区域，并用图像三通道的平均像素值擦除区域内的像素，最终得到擦除后的图像。其中，S_h，S_l 分别是需要随机擦除的矩形面积的上下阈值，r_1 是矩形长宽比 r 的阈值上限（$r_1 \in (0,1)$），(x_e, y_e) 是擦除区域左下角的坐标。经改进的 RE 算法处理后的增强图片如图 3 所示。

2.3　基于改进 Gamma 算法的图像校正

为了提高欠曝光或阴影场景下车道线和车辆检测的查全率和准确率，图像预处理阶段加入改进的自适应 Gamma 校正算法[24]进行图像正则

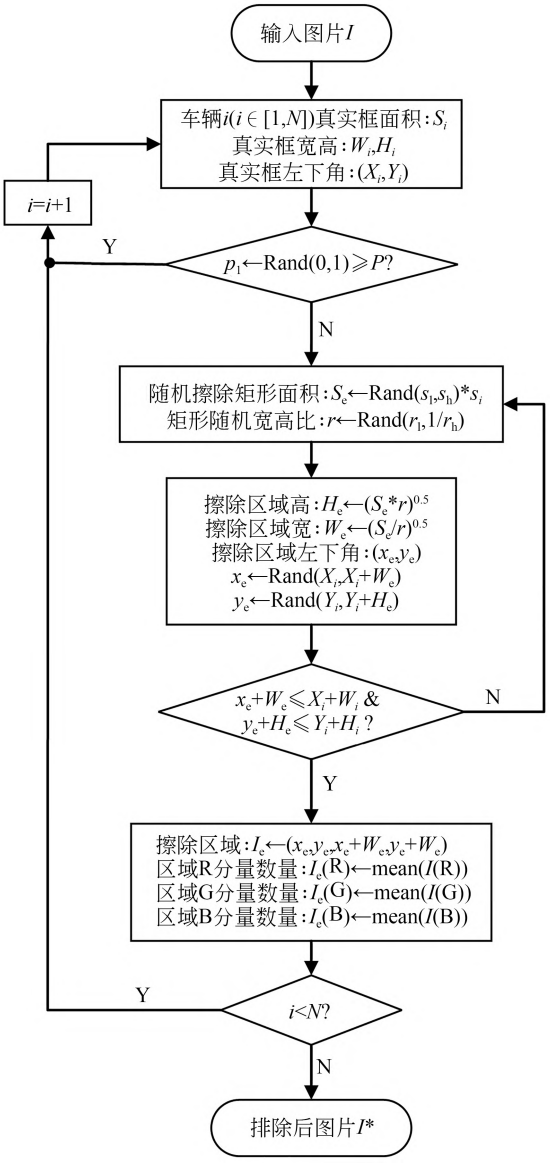


图 2　改进的随机擦除算法流程

Fig. 2　Flow chart of improved random erasing algorithm

化处理，步骤如下：

（1）将输入图像从 RGB 色彩空间转换到 HSV 空间记为图像 I，并提取 I 的亮度分量 I_V；

（2）利用双边滤波处理从图像亮度分量 I_V 中提取出光照分量 I_S，得到（包含空域核和值域核）：

$$I_S(i,j)=\frac{\sum_{mn} I_V(m,n)\omega(i,j,m,n)}{\sum_{mn}\omega(i,j,m,n)}. \tag{1}$$

由于双边滤波算法处理图像时考虑了空域信息和灰度相似性，可在尽可能保留车道线、车辆的边缘信息的同时抑制图像噪声，滤波核 ω 包

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

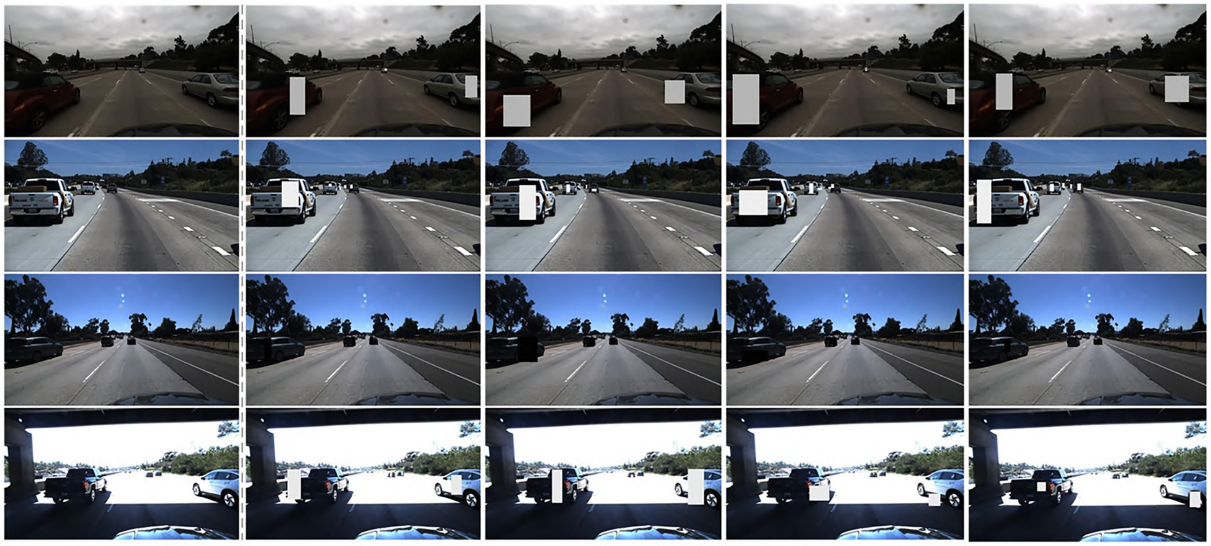


图 3　随机擦除后的训练图片

Fig. 3　Training images after random erasure

括空域核和值域核，其数学描述如下：

$$\omega(i,j,m,n)=\exp\left(-\frac{(i-m)^2+(j-n)^2}{2\sigma_d^2}-\frac{\|I(i,j)-I(m,n)\|^2}{2\sigma_r^2}\right),\tag{2}$$

式中：$I_S(i,j)$，$I_V(m,n)$分别为坐标(i,j)和(m,n)处的光照分量与亮度分量；σ_d，σ_r分别为空间邻近高斯函数标准差和像素值相似度高斯函数标准差。

(3)改进后的 Gamma 校正算法根据每一像素的光照分量$I_S(i,j)$自适应调整校正参数γ，如下：

$$\gamma=\frac{I_S(i,j)+\alpha}{1+\alpha},\tag{3}$$

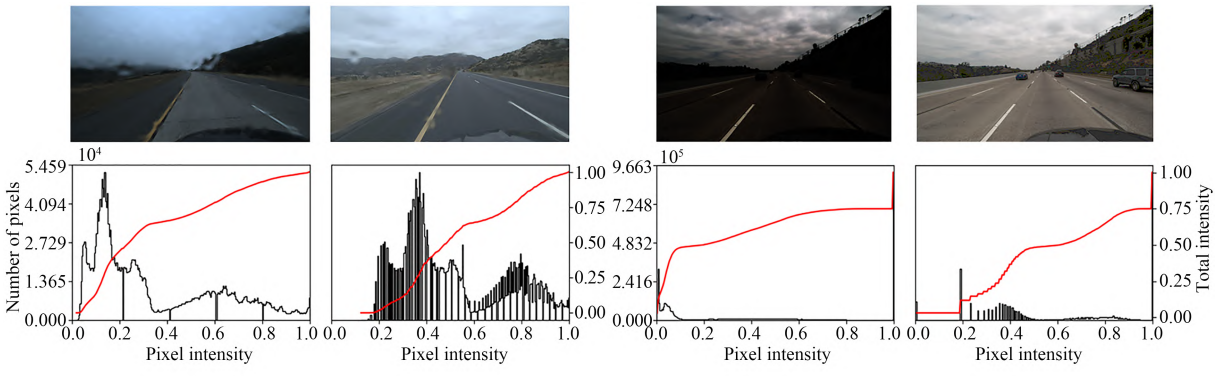
$$\alpha=1-\frac{2*\bar{I}_S}{I_{max}(i,j)+I_{min}(i,j)+1}.\tag{4}$$

校正后的像素光照分量为：

$$I_S'(i,j)=255\left[\frac{I_S(i,j)}{255}\right]^{\gamma},\tag{5}$$

式中：$\bar{I}_S$，$I_{max}(i,j)$，$I_{min}(i,j)$为校正前图像的平均光照分量、最小光照分量及最大光照分量。

图 4 为车载图片校正前后的效果、灰度分布直方图和累积曲线。显然通过改进 Gamma 算法可以增加中高灰度区域的动态范围，改善像素低灰度区域集中的现象，增强车辆、车道线与周围环境的对比，提升了图像的质量。



(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

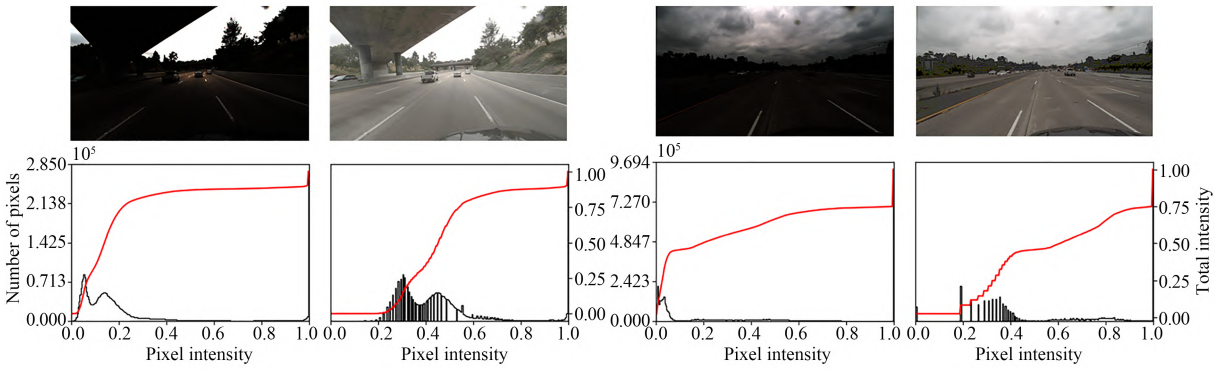


图 4　改进 Gamma 校正算法效果

Fig. 4　Effect of improved Gamma correction algorithm

2.4　基于改进 Mask R-CNN 的车辆检测

Mask R-CNN[25]的整体框架如图 5 所示。主干特征提取网络 Resnet101（结构见图 6）提取输入图片不同阶段输出的有效特征图；由于有效特征层的高层特征语义较强，但位置和分辨率较低，不利于小物体的检测，底层特征反之，特征金字塔网络（Feature Pyramid Network，FPN）[26]可以实现底层到高层有效特征的多尺度融合；融合后的多尺度特征经区域建议网络（RPN）[27]得到候选框（RoI）；RoIAlign 层通过双线性插值将生成的不同尺寸 RoI 进行区域匹配，即特征图和固定的 RoI 特征图的像素对应，原图和特征图的像

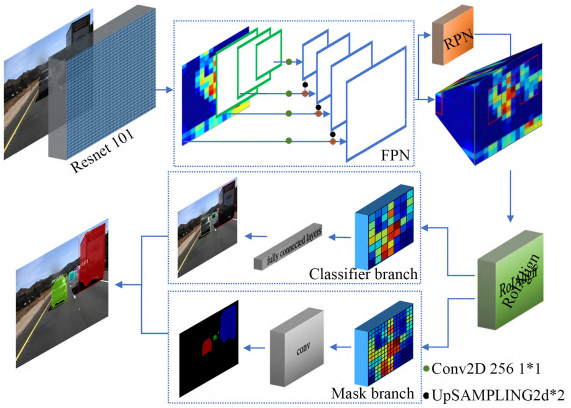


图 5　Mask R-CNN 整体框架

Fig. 5　Overall framework of Mask R-CNN

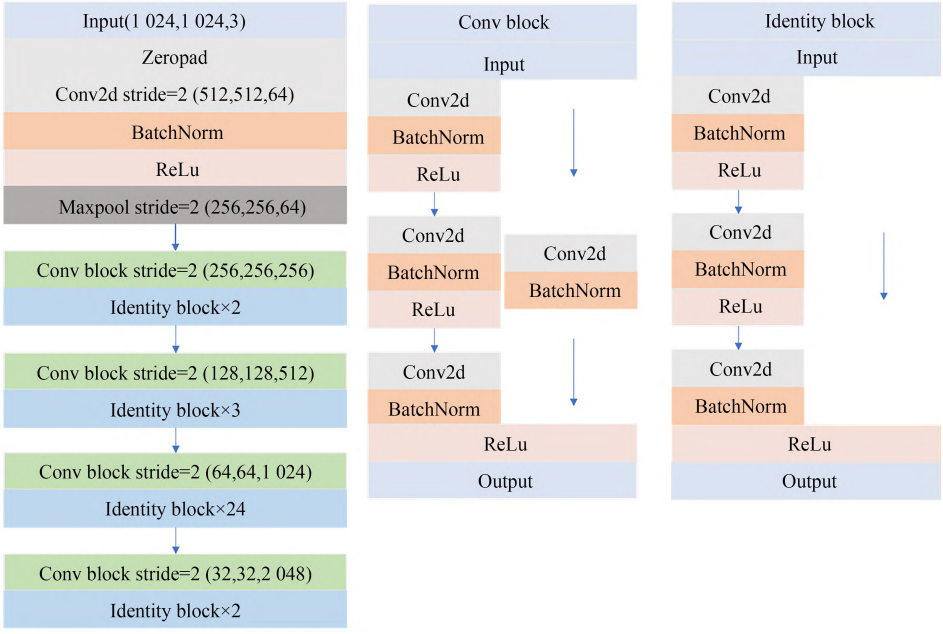


图 6　Mask R-CNN 主干特征提取网络结构

Fig. 6　Mask R-CNN backbone networks

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

素对应，从而将 RoI 映射为两种固定尺寸（7×7×256，14×14×256）的 RoI 特征。

如图 7 所示，目标检测分支利用一次通道数为 1 024，大小为 7×7 的卷积和一次通道数为 1 024，大小为 1×1 的卷积模拟两次 1 024 的全连接，对 RoIAlign 获得的 7×7×256 的 RoI 特征进行卷积，然后再分别全连接到类别总数（num_classes）和 4 倍的类别总数（num_classes×4）上（分别代表建议框内的物体、预测框的 4 个调整参数），实现目标的分类预测和回归预测；实例分割分支首先对 7×7×256 的 RoI 特征进行 4 次 3×3 的 256 通道的卷积，再进行一次反卷积提升分辨率、一次通道数为类别总数（num_classes）的卷积，最终输出 28×28×num_classes 的掩膜板，实现像素级实例分割。

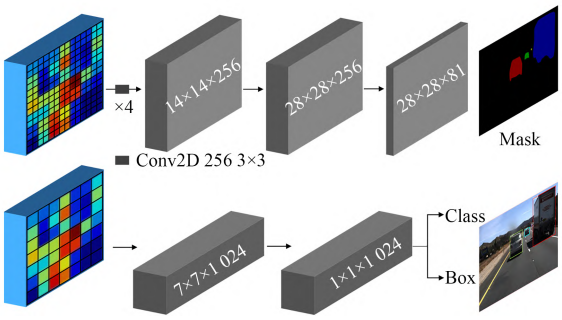


图 7　目标检测分支和实例分割分支

Fig. 7　Classifier branch and mask branch

2.4.1　RoI Align 层算法改进

为了从 RPN 网络确定的不同尺寸的候选框中导出固定大小的 RoI 特征图，Mask R-CNN 模型在 RoI Align 层采用双线性插值来解决 Faster R-CNN 模型采用的 RoI Pooling 操作中两次量化（特征图与原图对应像素的坐标取整，RoI 特征图与特征图对应像素的坐标取整）造成的原图与 RoI 特征图中的像素区域不匹配问题。但双线性插值算法未考虑各相邻像素点点间灰度值变化率的影响，而各邻点间灰度值的斜率不连续会造成缩放后图像的高频分量受到损失，细节信息退化，RoI Align 层输出的 RoI 特征图中的边缘信息较为模糊，进一步影响车辆分割精度和后续的车辆压线判断。

双立方（三次卷积）插值（Bicubic Interpolation，BI）[28] 不仅考虑到待插值点周围 4 个直接相邻像素点灰度值的影响，还兼顾了各相邻点间灰度值变化率的影响，所以利用插值点周围 4×4 邻域像素的灰度值进行插值，缩放后的图像边缘更为平滑，像质损失更少。故本文采用 BI 替代双线性插值，提高 RoI Align 层输出的 RoI 特征图的边缘质量，以提高车辆检测的精度和车辆的实例分割效果。

根据连续信号采样定理可知，对采样值用理论上最佳插值函数 $\sin \pi x/\pi x$ 进行插值，当采样频率不低于信号频谱最高频率的两倍时可以准确恢复原信号。BI 核函数 $\sin \pi x/\pi x$ 采用的三次近似多项式如下：

$$R(x)=\begin{cases}(\alpha+2)|x|^3-(\alpha+3)|x|^2+1, 0\leqslant|x|<1\\ \alpha|x|^3-5\alpha|x|^2+8\alpha|x|-4\alpha, 1\leqslant|x|<2\\ 0, 2\leqslant|x|\end{cases}. \tag{6}$$

参数 α 的取值为 -1，-0.5，0.25，0.75 等。卷积过程的矩阵乘积表示如下：

$$f(i+u,j+v)=\boldsymbol{ABC}, \tag{7}$$

其中：

$$\boldsymbol{A}=\left[h(1+u),h(u),h(1-u),h(2-u)\right], \tag{8}$$

$$\boldsymbol{B}=\begin{bmatrix}f(i-1,j-1) & f(i-1,j) & f(i-1,j+1) & f(i-1,j+2)\\ f(i,j-1) & f(i,j) & f(i,j+1) & f(i,j+2)\\ f(i+1,j-1) & f(i+1,j) & f(i+1,j+1) & f(i+1,j+2)\\ f(i+2,j-1) & f(i+2,j) & f(i+2,j+1) & f(i+2,j+2)\end{bmatrix}, \tag{9}$$

$$\boldsymbol{C}=\left[h(1+v),h(v),h(1-v),h(2-v)\right]^{\mathrm{T}}, \tag{10}$$

其中：(i,j) 表示待求的插值像素邻域中某个像素位置，$f(i,j)$ 为对应的灰度值，(u,v) 表示待插值点相对 (i,j) 的位移，即待插值点位置表示为 $(i+u,j+v)$，$f(i+u,j+v)$ 为待插点的灰度值，$h(x)$ 为插值核函数：

$$h(x)=\begin{cases}1-|x|, 0\leqslant|x|<1\\ 0, |x|>1\end{cases}. \tag{11}$$

以上插值过程可以用图 8 表示，实心圆点表

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

示已知原图像像素，虚线圆点是求出的中间插值点，网格圆点为待求的插值像素。

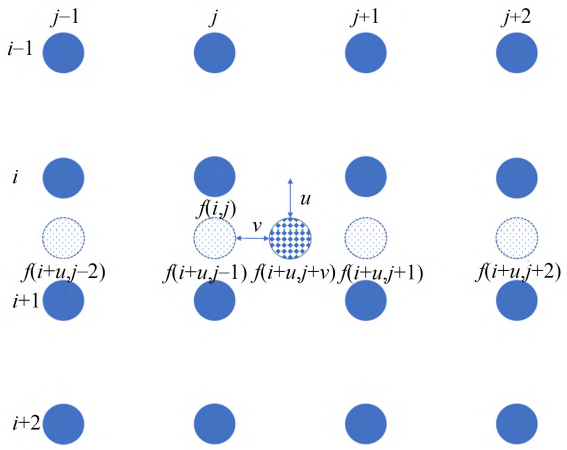


图 8　双立方卷积插值过程

Fig. 8　Bicube convolution interpolation process

2.5　基于改进 LaneNet 的车道线检测

本文通过改进 LaneNet[29] 网络的编码器解码器结构实现车道线检测，网络架构如图 9 所示。图像经过编码器提取到的共享特征分别流入到两个解码器来恢复细节和空间尺度，其中二值分割分支根据交叉熵损失函数生成车道线与背景的二值图像，像素嵌入分支基于聚类的判别损失函数得到区分不同车道线的特征图，对特征图做聚类得到车道线的数量和像素所属的车道线 id，结合两分支结果实现车道线的实例分割。神经网络 H-Net 根据原图片训练出预测变换矩阵 $\boldsymbol{H}$，并利用转换矩阵对同属一条车道线的所有像素点进行重新建模，完成车道线的拟合。

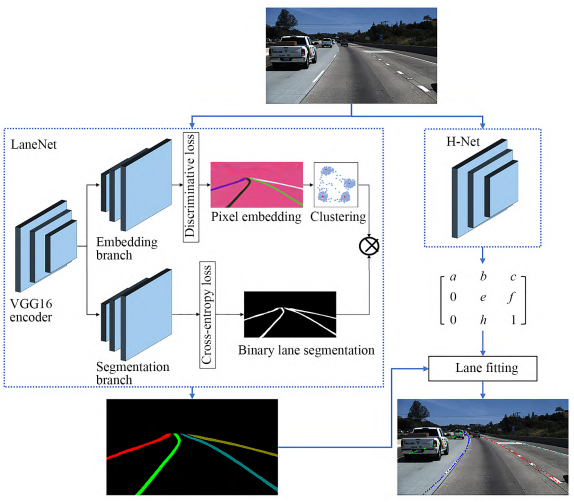


图 9　LaneNet 车道检测网络架构

Fig. 9　Framework of improved LaneNet for lane detection

为进一步提升车道检测速度，本文采用改进的 VGG16 代替原网络中的共享解码器 E-Net。VGG16 的网络复杂度远少于 E-Net，广泛应用于图像分割的特征提取上，将 VGG16 网络的最后 3 个全连接层改为卷积层，最后一卷积层中卷积核数量由 4 096 减少到 1 024，以降低参数的计算量和内存消耗；同时编码网络全部采用 3×3 小卷积核和 2×2 的最大池化层，相比于大卷积核（5×5 或 7×7），不仅获得相同的感受野且进一步减少了参数量，达到了更好的特征提取能力，改进的 VGG16 编码器结构如表 1 所示。

表 1　VGG16 共享编码器的结构参数

Tab. 1　Architecture parameters of VGG16 shared encoder

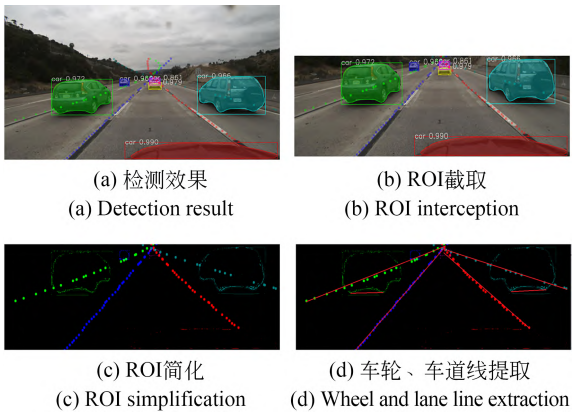
Block	Layers	Channels	Filters size/stride	Dilation rate
Input	—	3	—	—
Conv1	2	64	3×3	1
Maxpool1	1	—	2×2/2	—
Conv2	2	128	3×3	1
Maxpool2	1	—	2×2/2	—
Conv3	3	256	3×3	1
Maxpool3	1	—	2×2/2	—
Conv4	3	512	3×3	1
Maxpool4	1	—	2×2	—

2.6　车辆压线判断

如何利用二维空间的车载图像对压线车辆进行检测是本文的关键。本文在改进的 Mask R-CNN 和 LaneNet 提取出车载图像中车辆和车道线的基础上完成车辆压线的判断，具体步骤如下：

（1）考虑到车载摄像头与目标车辆距离较近，并且车道线的拟合准确度在远景处较差，存在相交现象会干扰压线的判断，故对检测完的图像（10(a)）截取 RoI 区域（图 10(b)），其宽度保持不变，高度取为图像底部与拟合的车道线交汇点的距离；

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net



(a) 检测效果
(a) Detection result

(b) ROI截取
(b) ROI interception

(c) ROI简化
(c) ROI simplification

(d) 车轮、车道线提取
(d) Wheel and lane line extraction

图 10　压线判断过程

Fig. 10　Line pressing judgement process

(2)采用零值像素遮滤掉所有未检测到的图像信息,只保留检测到的与车道线和车辆轮廓相关的特征,得到简化后的 RoI 图,如图 10(c)所示;

(3)根据实例分割结果提取车辆轮廓的点集合,筛选出高度方向上坐标数值最小的两点即为车辆后车轮与地面相切线段 wheel_line 的两端点,并采用改进的概率 Hough[30]变换对 RoI 内每条车道线上的点进行直线拟合,如图 10(d)所示;

(4)分析简化后 RoI 图中车辆的 wheel_line 与所有车道拟合线的位置关系,若相交则判断为压线违规。

3　实　验

3.1　数据集及实验环境配置

实验数据集采用自动驾驶研究中针对结构化道路的高速公路场景 Tusimple 数据集[31],在此基础上使用 Labelme 标注工具增加车辆目标的标注并基于原始图片做欠曝处理丰富数据集图片包含的场景。数据集样本共包含 3 626 张训练图像和 2 782 张测试图像,分辨率为 1 280×720,训练集的 80% 图像用作训练,20% 用作验证。实验的软件以及硬件配置如表 2 所示。

表 2　实验的软件及硬件配置

Tab. 2　Experimental software and hardware configuration

类　别	名　称	型号与参数
硬件	中央处理器	Intel(R)Core(TM)i7-10875H CPU@2.30GHz
	内存	RAM 32G
	图像处理器	NVIDIA GeForce RTX 2080Ti(11G)
软件	操作系统	Ubuntu 18.04
	深度学习框架	Pytorch-cuda 1.7.0
	开发语言	Python 3.6.6
	环境管理	Anaconda 3.6.5

3.2　实验结果

实验首先基于数据集 MS COCO[32]对改进后的 Mask R-CNN 网络进行预训练,预训练最大迭代次数设置为 3 000,初始学习率设置为 0.001。为了防止训练陷入局部最优,在迭代 2 000 次后将学习率调整为 0.000 1 完成预训练,并选取损失最小的最优权重对网络参数进行初始化;然后在 Tusimple 训练集上采用批处理的随机梯度下降法重新训练网络,学习率调整为 0.000 01,训练轮数设为 50,批处理数据量(batch size)设置为 4,最终损失函数值在 4.6 附近小幅波动不再下降,说明模型收敛,完成 Mask R-CNN 网络的训练。同时,利用 Tusimple 训练集对改进后的 LaneNet 网络进行训练,训练轮数设置为 12 000,batch size 为 4,训练权值衰减采用 L2 归一化,并采用 Adam 优化器优化训练效率,训练完成后保存损失最小的权重文件用于测试。

然后,分别利用原始网络和改进后的网络在相同的 Tusimple 测试集中进行检测,采用目标检测中常用的误检率(FP)、漏检率(FN)和每秒传输帧数(FPS)作为 Mask R-CNN 网络车辆检测

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

以及 LaneNet 网络车道线检测的性能评价指标。计算公式如下：

$$FP=\frac{F_{\text{pred}}}{N_{\text{pred}}}, \tag{12}$$

$$FN=\frac{M_{\text{pred}}}{N_{\text{gt}}}, \tag{13}$$

式中：F_{pred} 为误判为车道线或车辆的数目，N_{pred} 为检测到的车道线或车辆总数，M_{pred} 为未检测到的车道线或车辆数目，N_{gt} 为值标签中的车道线或车辆数量。改进前后两种网络的检测性能比较如表 3 所示，可见改进后的网络在误检率和漏检率上均大幅下降，车辆漏检率、误检率分别降低了 38.93%，89.04%，车道线漏检率、误检率分别降低了 67.21%，87.05%，车辆检测速度保持不变，车道线检测速度提升了 28%。相比于现有的实例分割主流算法（BlendMask[33] 和 YOLACT[34]），本文提出的算法在检测精度上性能更优；相比于现有的车道线检测算法（PINet[35] 和 Line CNN[36]），本文算法在检测精度和速度上均具有优势。

表 3　基于 TuSimple 测试集的算法性能对比

Tab. 3　Comparisions of algorithm performance on TuSimple testing set

Algorithm	Accuracy/%	FP	FN	FPS
Mask R-CNN[25]	94.85	0.042 3	0.084 0	5.46
BlendMask[33]	94.61	0.037 5	0.072 0	25
YOLACT[34]	95.36	0.032 7	0.068 9	33
Ours(for vehicle)	96.57	0.009 0	0.051 3	5.46
LaneNet[29]	96.38	0.078 0	0.024 4	50
Line CNN[36]	96.87	0.044 2	0.019 7	30
PINet[35]	96.70	0.029 4	0.026 3	30
Ours(for lane)	97.64	0.010 1	0.008 0	64

类激活图（Class Activation Map，CAM）是指网络对输入图像生成的类激活的热力图，是与特定输出类别相关的二维特征分数网络。网格每个位置的颜色信息表示对该类别判断依据的重要程度[37]（由蓝到红重要程度递减），可以显示模型更侧重于根据哪一部分特征进行类别和目标判别。为验证改进后的 RE 算法对于 Mask R-CNN 网络识别车辆特征能力的提升，如图 11 所示，获取有无增强数据训练的网络在不同场景下车辆检测输出的 3 种网格尺度（13×13，26×26，52×52）的类激活图，为避免干扰，不对输入图片进行 Gamma 校正预处理。网格尺度小的类激活图反映了网络对于远处目标车辆特征的识别情况，通过对比图 11(a2)与图 11(a5)及图 11(a3)与图 11(a6)，可以看出图 11(a5)、图 11(a6)中远处更多的车辆特征被捕捉到，说明采用增强数据训练后的模型对于小目标车辆的检测准确度有所提升。原始网络对于车体存在大面积遮挡的车辆存在无法识别的问题（如图 11(b2)与图 11(b3)左车道存在的两处被遮挡车辆），改进后的网络增强了车辆局部特征的识别能力，体现在图 11(b5)、图 11(b6)中，被遮挡车辆的局部特征被激活为其他颜色。阴影遮挡限制了网络对车辆特征的识别（图 11(c1)～11(c3)阴影下的车辆特征未被激活），通过数据增强网络，阴影环境的鲁棒性明显提升，如图 11(c4)～11(c6)阴影下的车辆特征均被捕捉到。

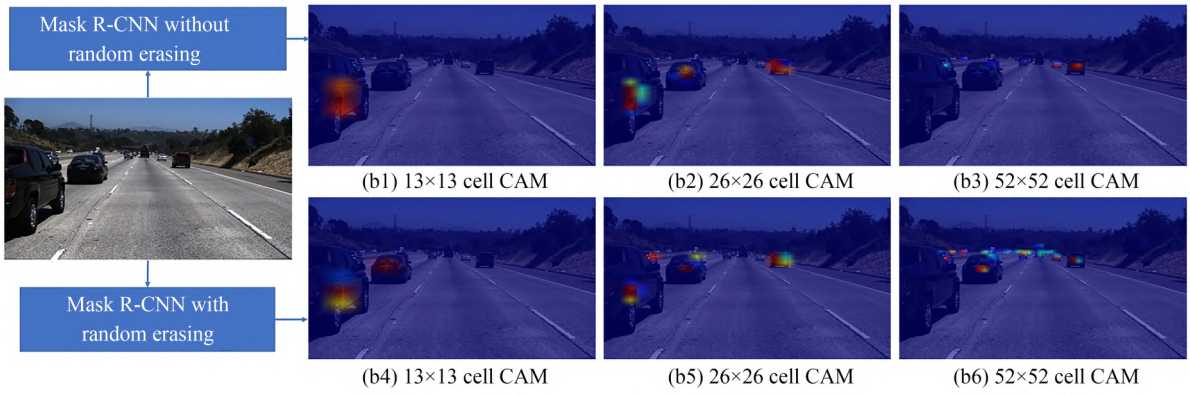
如图 12 所示，在道路场景 1，2 中原始网络（Baseline）均存在车辆误检的问题，场景 2，3，4 中均存在车辆漏检的现象。改进后的网络不仅改善了车辆的误检问题，而且对于欠曝图片中辨识度不高或远景的车辆的检测精度有所提升，场景 2，3，4 中原始网络漏检的车辆均被改进后的网络召回并检测到。在对 4 个场景中远景小车辆目标和处于阴影中车辆的检测上，BlendMask 与 YOLACT 算法都存在不同程度的漏检，本文算法表现出更高的检测精度，而且在场景 2，3 中，改进后的网络对于车辆边缘的分割效果要优于前两种网络。此外，通过在场景 2，4 中对比网络改进前后车道线的检测效果，改进后的网络对于车体或阴影遮挡、欠曝、弯曲的车道检测方面鲁棒性有所提升。图 13 进一步对比了本算法与其他主流算法的车道线检测效果。

如图 13 所示，5 个场景中的车道线存在欠曝、阴影遮挡和不同程度的缺损，导致原网络存在不同程度的漏检甚至完全检测不到（如场景 3），Line CNN 存在漏检（如场景 1，2，3）和单车道

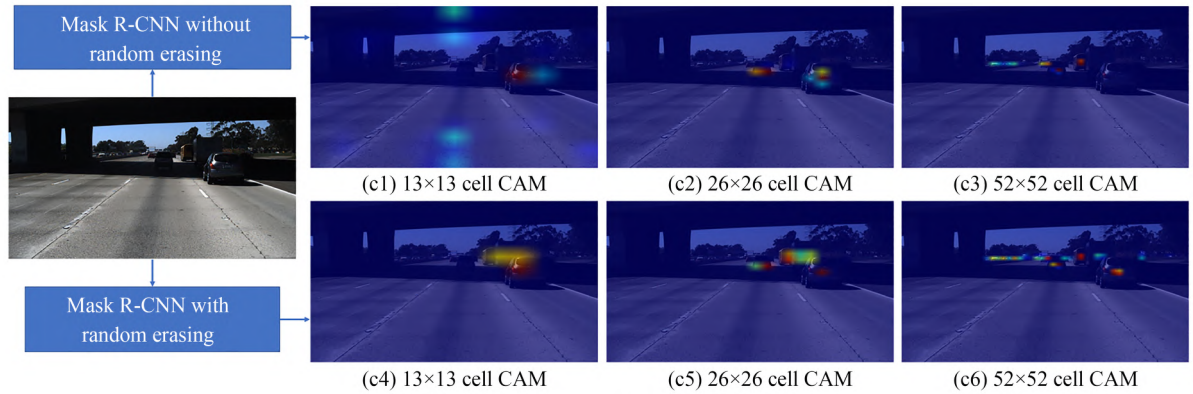
(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

不完全拟合(如场景4,5),PINet存在漏检(如场景1,2,3)、单车道拟合不完全(如场景2,3,5)和误检(如场景4)。本文基于LaneNet改进后的车道线检测算法一方面降低了漏检率和误检率,另一方面,完整的拟合效果为车辆压线判断提供了更加准确可行的前提条件。

(a) 小目标场景
(a) Small object scene



(b) 车体遮挡场景
(b) Vehicle body occlusion scene



(c) 阴影遮挡场景
(c) Shadow occlusion scene

图 11　不同场景下的类激活图
Fig. 11　Class activation maps in different scenarios

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

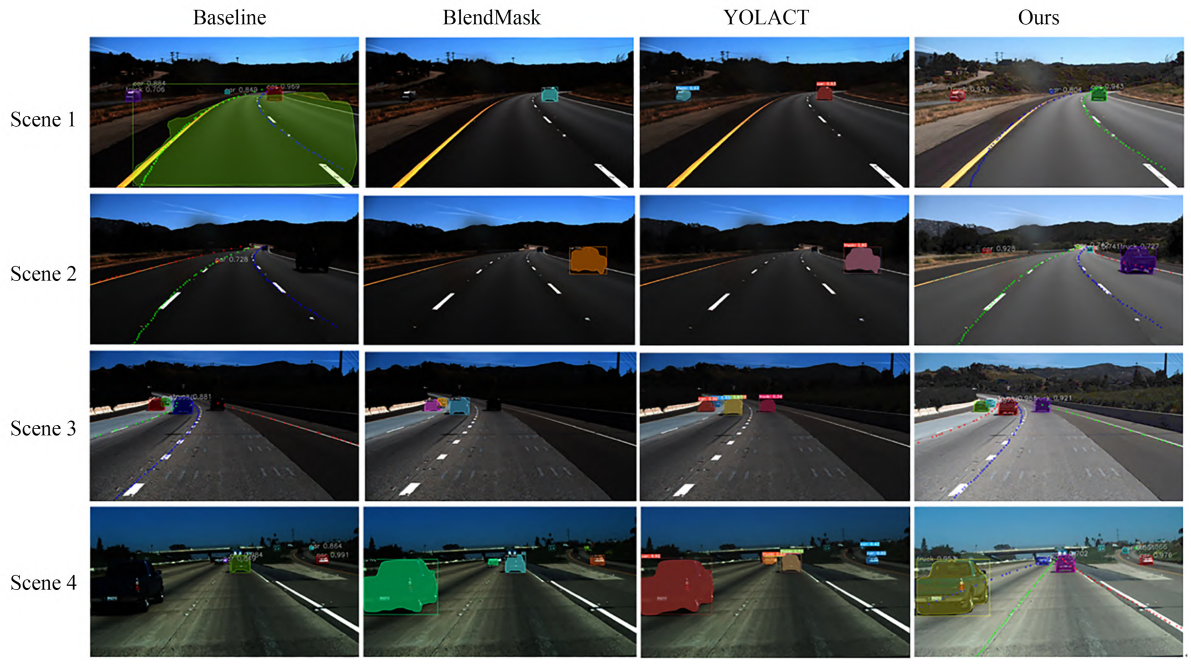


图 12　车辆检测效果对比

Fig. 12　Comparison of vehicle detection results

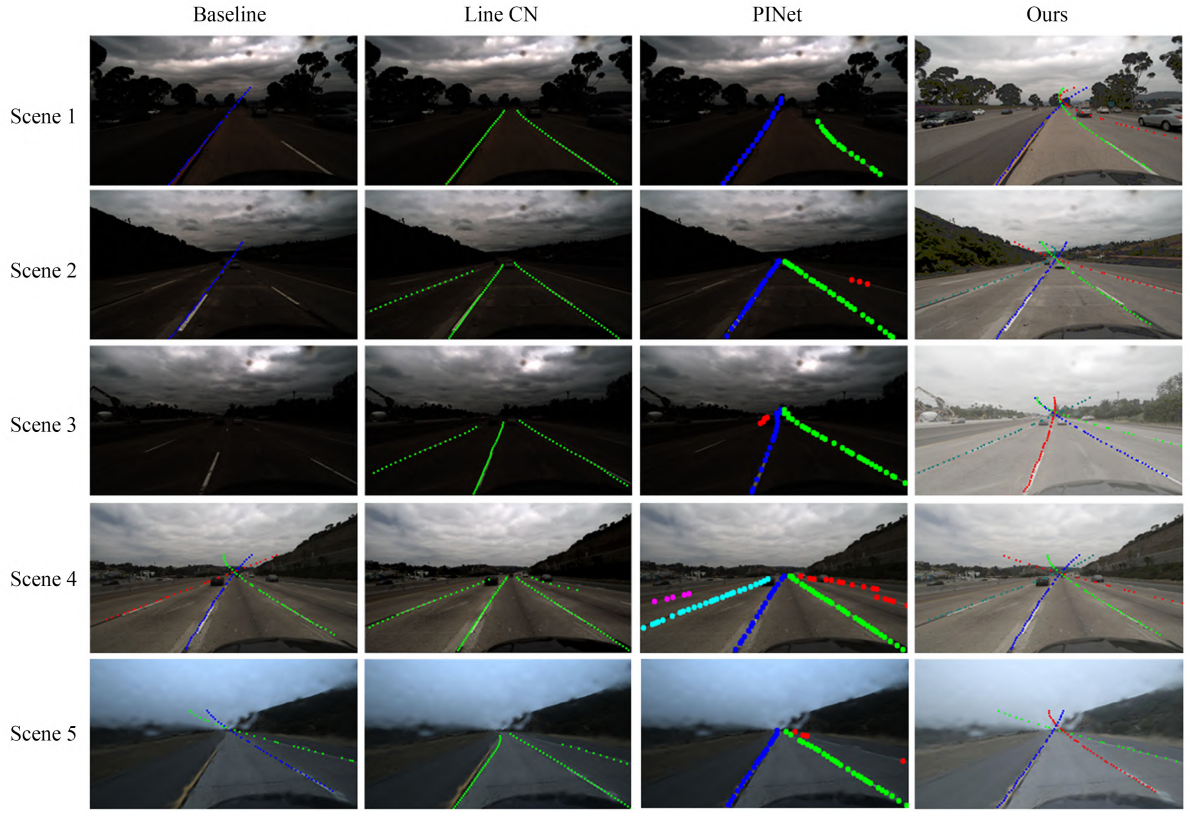


图 13　车道线检测效果对比

Fig. 13　Comparison of lane line detection results

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

4 结　论

本文以Mask R-CNN实例分割算法和LaneNet车道线检测算法为基础，实现了车载图像的车压线检测。通过改进RE算法扩展增强Tusimple数据集并训练Mask R-CNN网络，提升了网络对于阴影或实体遮挡的鲁棒性，改善了原模型在小目标车辆检测上存在的漏检问题；Gamma校正算法根据车载图像的灰度分布自动调整算子，从而扩大中高灰度区域的动态范围、提升欠曝区域的亮度，改善了欠曝图像中车辆及车道线的辨识度，进一步解决了车道线与车辆的漏检误检问题，使得最终车辆的漏检率、误检率分别下降了38.93%，89.04%，车道线的漏检率和误检率分别下降了67.21%，87.05%；采用双立方插值算法改进Mask R-CNN网络中RoIAlign层的双线性插值算法对候选框内的图像信息进行缩放，提高了RoI Align层输出的RoI特征图的边缘质量；将LaneNet的解码器E-Net替换为全连接层卷积化的VGG16，减少了参数计算量和内存消耗的同时车道线检测速度提升了28%，可以满足实时检测的需求；在保证车辆与车道线的检测效率的同时，对车载图像RoI区域进行简化处理，在此基础上提取车辆后轮的wheel_line线段，通过其与车道拟合直线相交点的个数判断有无车辆压线违规，该方法具有一定的可行性。

参考文献：

[1] 周跃琪. 基于卷积神经网络YOLO的车辆压实线检测方法研究[D]. 杭州：浙江科技学院，2020.
ZHOU Y Q. *Research on Vehicle Compaction Line Detection Method Based on Convolutional Neural Network YOLO*[D]. Hangzhou: Zhejiang University of Science & Technology, 2020. (in Chinese)

[2] 陈正斌，叶东毅. 带语义分割的轻量化车道线检测算法[J]. 小型微型计算机系统，2021，42(9)：1877-1883.
CHEN ZH B, YE D Y. Lightweight lane detection algorithm with semantic segmentation[J]. *Journal of Chinese Computer Systems*, 2021, 42(9): 1877-1883. (in Chinese)

[3] CHIU K Y, LIN S F. Lane detection using color-based segmentation[C]. *IEEE Proceedings. Intelligent Vehicles Symposium*, 2005. *June* 6-8, 2005, *Las Vegas*, *NV*, *USA*. *IEEE*, 2005: 706-711.

[4] 吴骅跃，赵祥模. 基于IPM和边缘图像过滤的多干扰车道线检测[J]. 中国公路学报，2020，33(5)：153-164.
WU H Y, ZHAO X M. Multi-interference lane recognition based on IPM and edge image filtering[J]. *China Journal of Highway and Transport*, 2020, 33(5): 153-164. (in Chinese)

[5] LOOSE H, FRANKE U, STILLER C. Kalman Particle Filter for lane recognition on rural roads[C]. 2009 *IEEE Intelligent Vehicles Symposium*. *June* 3-5, 2009, *Xi'an*, *China*. *IEEE*, 2009: 60-65.

[6] TENG Z, KIM J H, KANG D J. Real-time lane detection by using multiple cues[C]. *ICCAS* 2010. *October* 27-30, 2010, *Gyeonggi-do*, *Korea* (*South*). *IEEE*, 2010: 2334-2337.

[7] WANG J Y, DUAN J M. Lane detection algorithm using vanishing point[C]. 2013 *International Conference on Machine Learning and Cybernetics*. *July* 14-17, 2013, *Tianjin*. *IEEE*, 2013: 735-740.

[8] DANESCU R, NEDEVSCHI S. Probabilistic lane tracking in difficult road scenarios using stereovision[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2009, 10(2): 272-282.

[9] KIM Z. Robust lane detection and tracking in challenging scenarios[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2008, 9(1): 16-26.

[10] 初雪梅，王珂娜，张维刚. 基于分段直线模型的弯道识别算法的研究[J]. 汽车工程，2012，34(12)：1141-1144，1149.
CHU X M, WANG K N, ZHANG W G. A study on a recognition algorithm of curved lane based on piecewise straightline model[J]. *Automotive Engineering*, 2012, 34(12): 1141-1144, 1149. (in Chinese)

[11] 王宝锋，齐志权，马国成，等. 基于动态区域规划的双模型车道线识别方法[J]. 北京理工大学学报，2014，34(5)：485-489.
WANG B F, QI ZH Q, MA G CH, *et al*. Lane detection system based on region dynamically planning and dual model fitting[J]. *Transactions of Bei-*

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

jing Institute of Technology, 2014, 34(5): 485-489. (in Chinese)

[12] YOO H, YANG U, SOHN K. Gradient-enhancing conversion for illumination-robust lane detection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(3): 1083-1094.

[13] 周宏宇，宋旭，刘国英. Haar特征耦合级联分类器的车道线检测算法[J]. 计算机工程与设计，2020，41(6)：1719-1724.
ZHOU H Y, SONG X, LIU G Y. Lane detection algorithm based on Haar feature based coupled cascade classifer[J]. *Computer Engineering and Design*, 2020, 41(6): 1719-1724. (in Chinese)

[14] PAN X G, SHI J P, LUO P, *et al*. Spatial as deep: spatial CNN for traffic scene understanding [EB/OL]. https://arxiv.org/abs/1712.06080

[15] HOU Y N, MA Z, LIU C X, *et al*. Learning lightweight lane detection CNNs by self attention distillation [J]. 2019 *IEEE/CVF International Conference on Computer Vision* (*ICCV*), 2019: 1013-1021.

[16] TABELINI L, BERRIEL R, PAIXÃO T M, *et al*. Keep your eyes on the lane: real-time attention-guided lane detection[C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). *June* 20-25, 2021, *Nashville*, *TN*, *USA*. *IEEE*, 2021: 294-302.

[17] 胡鹏. 基于机器视觉的交通压线判别方法研究[D]. 西安：西安科技大学，2017.
HU P. *Research on Vehicle Line Pressing Judgment Method Based on Machine Vision* [D]. Xi′an: Xi′an University of Science and Technology, 2017. (in Chinese)

[18] 洪琦，沙新华. 基于视频的车辆违章越线行为检测方法研究[J]. 深圳信息职业技术学院学报，2005，3(3)：12-16.
HONG Q, SHA X H. A study of a video-based method for detecting vehicles′ transgressing-lane behavior[J]. *Journal of Shenzhen Institute of Information Technology*, 2005, 3(3): 12-16. (in Chinese)

[19] 陆蔚. 一种快速检测违规压双黄线的算法[J]. 福建电脑，2009，25(10)：87，107.
LU W. A fast algorithm for detecting illegal double yellow lines [J]. *Fujian Computer*, 2009, 25(10): 87, 107. (in Chinese)

[20] 唐静. 基于车载视频的前方车辆越线违章行为监测技术研究[D]. 大连：大连海事大学，2018.
TANG J. *Research on Monitoring Technology of Crossing Violation of Vehicle in Front of Vehicle Video*[D]. Dalian: Dalian Maritime University, 2018. (in Chinese)

[21] 胡海辉. 基于车载摄像机的车辆压线与变道检测技术研究与实现[D]. 南京：东南大学，2015.
HU H H. *The Research and Implementation of the Detection Technology of Vehicle Pressing Line and Changing Lane Based on Onboard Camera* [D]. Nanjing: Southeast University, 2015. (in Chinese)

[22] 董俊. 基于深度学习的车辆违章压线检测算法研究[D]. 广州：广东工业大学，2020.
DONG J. *Research on Vehicle Violation Pressure Line Detection Algorithm Based on Deep Learning* [D]. Guangzhou: Guangdong University of Technology, 2020. (in Chinese)

[23] ZHONG Z, ZHENG L, KANG G L, *et al*. Random erasing data augmentation[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 13001-13008.

[24] 许少秋，杨群，钟小芸，等. 背光图像的检测与校正[J]. 计算机工程与应用，2017，53(21)：174-178.
XU SH Q, YANG Q, ZHONG X Y, *et al*. Detection and correction of backlight images [J]. *Computer Engineering and Applications*, 2017, 53(21): 174-178. (in Chinese)

[25] HE K M, GKIOXARI G, DOLLÁR P, *et al*. Mask R-CNN[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 386-397.

[26] LIN T Y, DOLLÁR P, GIRSHICK R, *et al*. Feature pyramid networks for object detection[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition*. *July* 21-26, 2017, *Honolulu*, *HI*, *USA*. *IEEE*, 2017: 936-944.

[27] REN S Q, HE K M, GIRSHICK R, *et al*. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.

[28] KEYS R. Cubic convolution interpolation for digital image processing [J]. *IEEE Transactions on Acoustics*, *Speech*, *and Signal Processing*, 1981, 29(6): 1153-1160.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

[29] NEVEN D, DE BRABANDERE B, GEORGOULIS S, *et al*. Towards end-to-end lane detection: an instance segmentation approach[C]. 2018 *IEEE Intelligent Vehicles Symposium*. *June* 26-30, 2018, *Changshu*, *China*. *IEEE*, 2018: 286-291.

[30] DUDA R O, HART P E. Use of the Hough transformation to detect lines and curves in pictures[J]. *Communications of the ACM*, 1972, 15(1): 11-15.

[31] Tusimple. Tusimple benchmark[Z/OL]. https://github.com/TuSimple/tusimple-benchmark, 2017.

[32] LIN T Y, MAIRE M, BELONGIE S J, *et al*. Microsoft COCO: common objects in context[J]. *Proceedings of the IEEE Conference on European Conference on Computer Vision*, 2014: 740-755.

[33] CHEN H, SUN K Y, TIAN Z, *et al*. BlendMask: top-down meets bottom-up for instance segmentation[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). *June* 13-19, 2020, *Seattle*, *WA*, *USA*. *IEEE*, 2020: 8570-8578.

[34] BOLYA D, ZHOU C, XIAO F Y, *et al*. YOLACT: real-time instance segmentation[C]. 2019 *IEEE/CVF International Conference on Computer Vision* (*ICCV*). *October* 27 - *November* 2, 2019, *Seoul*, *Korea* (*South*). *IEEE*, 2019: 9156-9165.

[35] KO Y, LEE Y, AZAM S, *et al*. Key points estimation and point instance segmentation approach for lane detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 8488, PP(99): 1-10.

[36] LI X, LI J, HU X L, *et al*. Line-CNN: end-to-end traffic line detection with line proposal unit[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 21(1): 248-258.

[37] SELVARAJU R R, COGSWELL M, DAS A, *et al*. Grad-CAM: visual explanations from deep networks via gradient-based localization[J]. *International Journal of Computer Vision*, 2020, 128(2): 336-359.

作者简介：

孙建波(1995—),男,吉林舒兰人,硕士研究生,2018年于长春工业大学获得学士学位,主要从事计算机视觉、深度学习图像处理方面的研究。E-mail: sunni18804305818@163.com

通讯作者：

张　叶(1981—),女,吉林长春人,研究员,博士生导师,主要从事计算机视觉、模式识别算法方面的研究。E-mail:yolanda@spirits.ai

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net