引文格式:卢朝晖,齐国栋,于慧敏,等.交通目标特定区域的图像质量评估[J].武汉大学学报(信息科学版),2025,50(10):2064-2071.DOI:10.13203/j.whugis20230192

Citation: LU Zhaohui, QI Guodong, YU Huimin, et al. Image Quality Assessment for Specific Areas of Traffic Targets[J]. Geomatics and Information Science of Wuhan University, 2025, 50(10):2064-2071.DOI:10.13203/j.whugis20230192

交通目标特定区域的图像质量评估

卢朝晖[1] 齐国栋[1] 于慧敏[1] 闫禹铭[1]

1 浙江大学信息与电子工程学院,浙江 杭州,310013

摘 要:针对交通监控场景下复杂道路环境干扰和真实参考图像缺失等因素导致的图像质量评估不佳问题,提出了一种两阶段图像质量评估方法,用于感兴趣交通目标中特定区域的图像质量评估。首先,为了减少背景等图像上下文的影响,提出使用目标检测算法提取图像中的交通目标,并根据轮廓、几何特征和车牌识别算法识别感兴趣交通目标中的特定区域;其次,为了弥补特定区域中缺失真实参考图像的问题,提出使用生成对抗网络生成伪参考图,设计一种新的损失函数,对生成对抗网络和图像质量评估器进行协同训练。在交通监控场景下的图像质量评估数据集上,对所提方法进行了评估实验。与其他图像质量评估方法相比较,所提方法模型小,并在斯皮尔曼秩相关系数和皮尔逊线性相关系数上均处于最优水平。因此,所提出的两阶段图像质量评估方法可以应用于复杂交通场景下的图像质量评估。

关键词:智能交通;特定区域;图像质量评估;目标识别;生成对抗网络

中图分类号:P237 **文献标识码**:A **收稿日期**:2024-05-18

DOI:10.13203/j.whugis20230192 **文章编号**:1671-8860(2025)10-2064-08

Image Quality Assessment for Specific Areas of Traffic Targets

LU Zhaohui[1] *QI Guodong*[1] *YU Huimin*[1] *YAN Yuming*[1]

1 College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310013, China

Abstract: **Objectives**: The study of image quality evaluation is aimed at the problems of poor image quality evaluation caused by the interference of complex road environment and the absence of real reference images in traffic monitoring scenarios. **Methods**: We propose a two-stage image quality assessment method for a specific area of the traffic target of interest. First, in order to reduce the influence of image context such as background, an object detection algorithm is proposed to extract the traffic target in the image, and the specific area of the interested traffic target is identified according to the contour, geometric features and license plate recognition algorithm. Second, in order to make up for the lack of real reference images in a specific region, a novel loss function is proposed to generate pseudo-reference images using generative adversarial networks, and to train the generative adversarial networks and image quality evaluators together. **Results**: The proposed algorithm is evaluated on the dataset of image quality evaluation in traffic monitoring scenario. Compared with other image quality assessment methods, the model of this method is small, and it is at the optimal level in Spearman rank correlation coefficient index and Pearson linear correlation coefficient index. **Conclusions**: The two-stage image quality assessment method can provide an effective method for the practical application of image quality assessment in complex traffic scenarios.

Key words: intelligent transportation; specific areas; image quality assessment; object recognition; generative adversarial networks

随着智慧城市的发展,人工智能在城市交通执法管理上的应用越来越广泛,其对视觉图像的质量要求也越来越高,因此图像质量评估任务在人工智能领域中起着关键的作用,例如遥感图像

基金项目:浙大-力嘉研发中心研究项目(2021-KYY-536010-0023)。

第一作者:卢朝晖,博士生,主要研究方向为计算机视觉。12131107@zju.edu.cn

通信作者:齐国栋,博士,腾讯集团高级研究员。guodong_qi@zju.edu.cn

应用[1]、监控场景[2-3]等。在道路监控场景中,图像质量的好坏能直接影响有关部门在执法任务中对图像证据的取舍。在交通视频监控场景下,智能监控系统往往仅抓拍到某一个目标的单张图像,而无法事先储存该目标的任何参考图像,这就要求实现无真实参考图像质量评估方法。

无真实参考图像质量评估是在没有任何参考图像信息的情况下预测图像质量[4-6],因其更具实用价值,吸引了大量的科学研究工作。文献[7-8]基于自然图像统计特征提出了无参考图像空间质量评估器(blind image spatial quality evaluator, BRISQUE)和无参考图像质量评估器(naturalness image quality evaluator, NIQE),然而交通监控图像和自然图像两者的统计特征存在着较大的差异性;文献[9]提出了深双线性卷积神经网络(deep bilinear convolutional neural network, DBCNN),然而其针对的特定失真场景和监控场景不同;文献[10]提出了基于Transformer的多尺度图像质量变换器(multi-scale image quality transformer, MUSIQ)模型,文献[11]提出了将图像分块的PAQ2PIQ(from patches to pictures)方法,文献[12]提出了学习图像失真因素之间内在关系的图像质量评估(learns distortion graph representations for blind image quality assessment, GraphIQA),文献[13]提出使用非局部交互方法以实现具有长依赖性的图像质量评估的非局部建模(non-local modeling for image quality assessment, NLNet-IQA),然而这些模型结构复杂,同时需要大规模的数据集和资源用于训练,不适用于实际监控场景的应用。此外,在交通监控场景下,图像分辨率大,干扰信息多,许多具体应用更关注感兴趣目标的某一区域的图像质量,而非整张图像的图像质量。

基于上述讨论,本文针对中国交通环境提出了感兴趣交通目标中特定区域的图像质量评估方法,该方法能够降低对自然图像统计特征的依赖,参数量相对较小;并提出了生成式伪参考图方法,利用了基于参考图的图像质量评估方法的优势。所提出的方法包含两个阶段,以交通目标特定区域为驾驶员所在位置的车辆前挡风玻璃区域为例:第一阶段为目标特定区域识别模块,使用目标检测的方法识别图像中车辆,并对指定车牌的前挡风玻璃区域进行检测与定位;第二阶段对特定区域进行基于伪参考图的图像质量评估,提出了一种基于质量评估协同学习的生成对抗模型,对特定区域进行重建生成一幅高质量的伪参考图像,通过特征融合和评估器进行图像质量评估。

1　方法原理

本文研究的目标特定区域为指定车辆中驾驶员所在位置的前挡风玻璃区域。

1.1　整体框架

本文方法的整体框架包含两个模块,如图1所示。

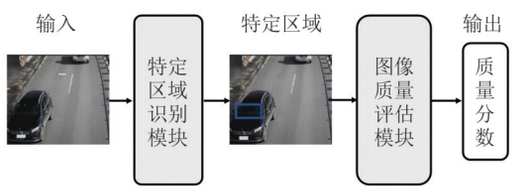


图1　交通目标特定区域的图像质量评估整体框架

Fig. 1　Overall Framework of Image Quality Assessment for Specific Areas of Traffic Targets

1.2　特定区域识别模块

特定区域识别模块是对指定目标中关键区域进行识别,包括交通目标检测、指定目标的匹配/识别及其关键区域的检测等初始任务,目标关键区域定义会因具体的应用而不同。现有的目标检测技术是检测所指定类别的所有目标,类别一般也是常用的类别,因此,无法将现有目标检测算法直接移植过来。同时,目标关键区域的识别也缺少相应的数据集或模型,特别是针对前挡风玻璃区域的检测,直接采用深度学习模型设计前挡风玻璃区域算法比较困难,需要一定规模的数据集。因此,在特定区域识别模块中,本文使用分而治之的策略,如图2所示,针对前挡风玻璃区域的检测,首先使用现有目标检测模型设计检测车辆算法,然后使用传统算法定位前挡风玻璃区域与车牌。

具体而言,图2(a)描述了车辆目标检测过程,为了能够实现算法的快速布置,本文采用现有通用的目标检测模型来设计车辆检测算法,用于提取车辆目标;图2(b)展示了特定区域检测过程,本文通过传统算法提取轮廓和长宽比等形状特征,用于筛选符合车牌及挡风玻璃区域,利用车牌识别算法进行车牌号码识别,筛选对应车牌号的挡风玻璃区域。如果车牌识别得到的结果与指定车牌号匹配,则判定当前车牌所对应的挡风玻璃区域为所需的特定区域。

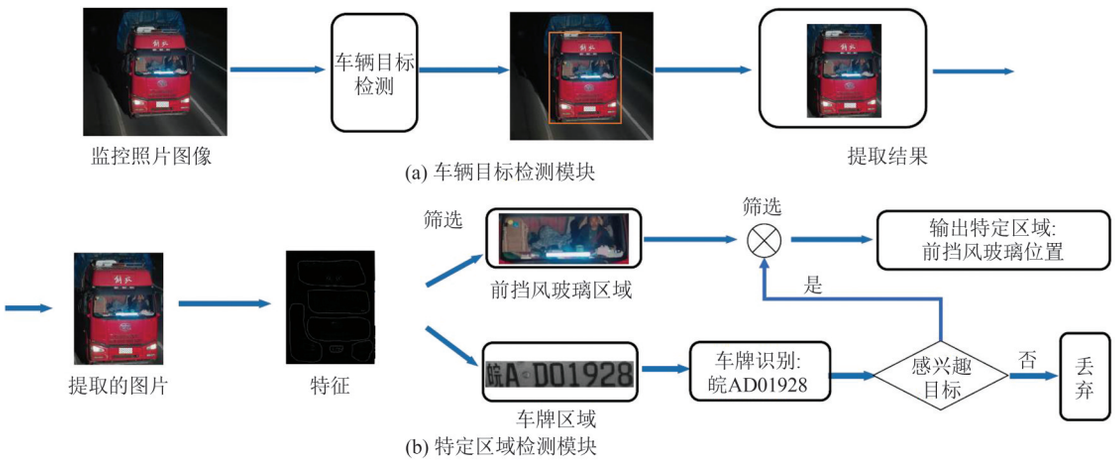


图2 特定区域识别模块示意图(指定车辆的挡风玻璃区域检测)

Fig. 2 Illustration of Specific Areas Recognition Module (Windshield Detection of Designated Vehicles)

1.3 无参考图像质量评估模块

为了结合基于参考图和无真实参考图的图像质量评估方法的优点,受生成对抗网络[14-15](generative adversarial network, GAN)启发,本文提出了一种协同质量评估的对抗学习网络和新的对抗训练方式,通过生成重建图像为伪参考图像,从而将较困难的无参考图像质量评估问题转变成相对简单的有参考图像质量评估问题。

本质上,本文方法未借助真实参考图像,仍然属于无参考图像质量评估的范畴。同时,尽管本文所生成的伪参考图可能图像质量较好,但是它不能代替原始图像作为后续任务(例如有关部门执法时对证据的选取)的输入。一方面因为这些后续任务需要原始图像保证证据的真实性,另一方面因为本文中的模型比较关注伪参考图对图像质量评估任务的贡献,保留对图像质量有益的特征,而忽略伪参考图本身的某些其他图像特征,这可能导致伪参考图的部分信息丢失。

1.3.1 结合质量评估和协同学习的伪参考图的生成

本文设计的对抗网络联合了图像质量评估网络,通过两者的协同学习,以对抗训练的方式生成有助于提升图像质量评估的伪参考图。因此,要求该对抗网络具有以下特点:(1)生成图像和理想图像的像素距离较小;(2)生成图像和理想图像的特征距离较小;(3)生成图像能够提升真实图像的质量评估的性能。

基于这3点,本文提出了像素损失、特征损失和对抗损失。像素损失是理想图像和重建图像的像素点之间的损失,惩罚重建图像和理想图像之间每个像素点的不同;特征损失的目的是降低重建图像的特征和理想图像的特征的差距,一方面,为了能够提供结构化和纹理化的特征信息,要求特征提取器具有良好的特征提取能力,另一方面,为了保留重建图像的噪声,需要选取具有弱的抗噪能力的特征提取器;对抗损失是为了学会鉴别输入图像是否是真样本或假样本,真假样本标签的判断与图像质量评估任务协同,即如果生成图像能帮助提升对实际图像的质量评估性能,那么它被赋予真样本标签。

具体而言,分别记理想图像为$\boldsymbol{I}_r$、退化图像为$\boldsymbol{I}_d$、重建图像为$\boldsymbol{I}_c$,为了产生成对的退化图像与重建图像,如图3(a)所示,在生成器模块中,先使用理想图像(图像质量高的图像,从数据集中选取)通过随机退化的方式转变成退化图像;然后通过生成器将退化图像恢复为重建图像。

为了训练生成器,采用像素损失和特征损失,如图3(b)所示,像素损失是理想图像和重建图像的像素点的L_1损失:

$$L_1=\sum\|g(\boldsymbol{I}_d)-\boldsymbol{I}_r\|_2^2 \tag{1}$$

式中,$g()$为生成器。特征损失为重建图像和理想图像特征之间的L_2损失:

$$L_2=\sum\|\phi(g(\boldsymbol{I}_d))-\phi(\boldsymbol{I}_r)\|_2^2 \tag{2}$$

式中,ϕ为特征提取器,可以使用预训练的浅层网络,比如VGG-19[16],预训练的网络保证其具有良好的特征提取能力,浅层的网络保证其具有弱的去噪能力,确保能捕捉失真信息。

此外,本文还提出了与鉴别器模块相对应的对抗损失,如图3(c)所示。目的是学会鉴别输入图像是否是真样本或假样本。为了得到输入图像的真值标签,本文结合了图像质量评估任务,并设置两组基线。即使输入图像不是特别清晰的重建图像,与基线相比较,如果它能帮助提升对实际预测图像的质量评估器的性能,那么它可被赋予真样本标签。

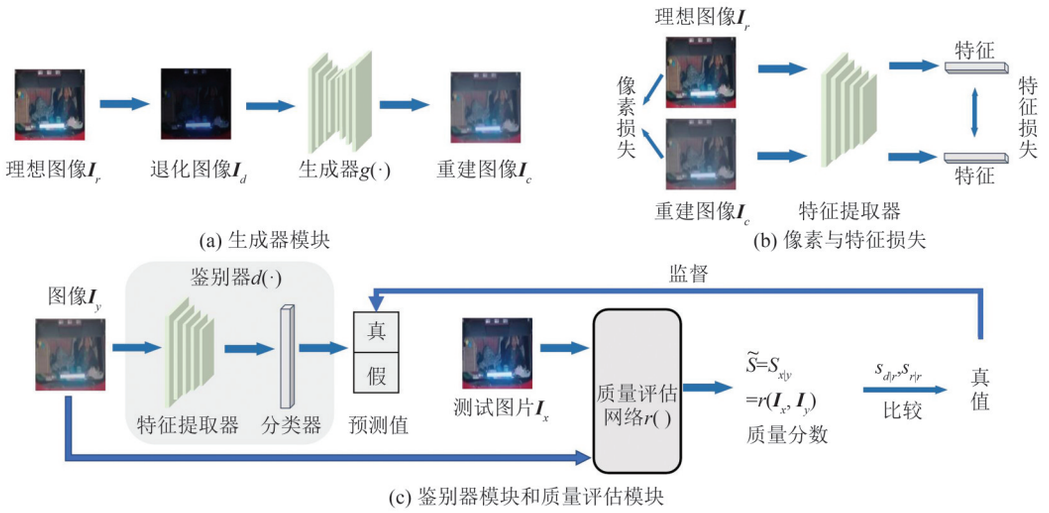


图 3　基于协同学习的生成式伪参考图的图像质量评估示意图

Fig. 3　Diagram of Image Quality Assessment of Generative Pseudo Reference Image Based on Collaborative Learning

具体而言,记$s_{x_1|x_2}$为测试图像I_{x_1}在以参考图像I_{x_2}为条件时所获得的质量评估分数,则这两组基线分别为:(1)在理想图像I_r作为参考图像时,退化图像I_d作为测试图像所获得的质量评估分数$s_{d|r}$;(2)在理想图像I_r作为参考图像时,理想图像I_r作为测试图像时所获得的质量评估分数$s_{r|r}$。同时,分别设置与两组基线相对照的两组实验,分别为:(3)在重建图像I_c作为参考图像时,退化图像作为测试图像所获得的质量评估分数$s_{d|c}$;(4)在重建图像I_c作为参考图像时,理想图像I_r作为测试图像所获得的质量评估分数$s_{r|c}$。

图像样本标签的判定与该图像作为输入是否能够提升质量评估器的性能有关,标签判定的规则为:对于理想图像I_r,始终将其定义为真样本标签;对于重建图像I_c,如果$s_{r|c}>s_{r|r}$或者$s_{d|c}<s_{d|r}$,说明其能帮助提升质量评估网络的性能,对其赋值为真样本标签。为了使鉴别器能够区分图像的真假性,同时能够结合质量评估任务学习,本文提出了一种对抗损失函数:

$$L_3=-\sum(\lg d(I_r)+\lg(1-a)) \tag{3}$$

$$a=\begin{cases} d(I_d), s_{r|c}>s_{r|r}或s_{d|c}<s_{d|r} \\ 1-d(I_d), 其他 \end{cases} \tag{4}$$

通过控制a,实现真假样本标签的赋值,协同学习 GAN 和质量评估器。同时,训练阶段的总损失函数为:

$$L=\alpha_1 L_1+\alpha_2 L_2+\alpha_3 L_3+L_4 \tag{5}$$

式中,α_1、α_2、α_3为平衡超参数;L为图像质量评估的损失函数。在训练时,特征提取网络ϕ保持不变,生成器$g(\ )$、鉴别器$d(\ )$及图像质量评估器$r(\ )$的参数被更新。

1.3.2　基于伪参考的图像质量评估模块

由于生成网络能够产生伪参考图,能够将图像质量评估器转变成基于伪参考的图像质量评估模型。为了与伪参考图的特征做对比分析,将两个图像的特征图看作双变化因子,通过双线性池化进行特征融合,对融合后的特征进行欧氏空间变换和回归分析,完成图像质量分数的评估。

具体而言,网络模型如图 4 所示,质量评估网络包含特征提取器$\varphi(\ )$、双线性池化以及回归器$h(\ )$。给定测试图像I_x和参考图像I_y,使用特征提取器φ,分别得到对应的图像特征$H_x=\varphi(I_x)\in \mathrm{R}^{(h_x\times w_x)\times d}$和$H_y=\varphi(I_y)\in \mathrm{R}^{(h_y\times w_y)\times d}$,其中$d$为特征图通道数,$h_x$、$w_x$、$h_y$、$w_y$为特征图的长宽维度,满足$h_x\times w_x=h_y\times w_y$。使用双线性池化对两个特征进行融合得到$H_{x|y}=H_y^{\mathrm{T}}H_x$,此时$H_{x|y}$维度为$d\times d$。对融合后的特征矩阵$H_{x|y}$做矩归一化和$L_2$归一化,使得在黎曼流形上的表征$H_{x|y}$变换到欧氏空间中:

$$\tilde{H}_{x|y}=\frac{\mathrm{sgn}(H_{x|y})\odot\sqrt{|H_{x|y}|}}{\|\ \mathrm{sgn}(H_{x|y})\odot\sqrt{|H_{x|y}|}\ \|_2} \tag{6}$$

式中,$\odot$为元素乘积;sgn 为符号函数。将欧氏空间中的$\tilde{H}_{x|y}$输入到一个全连接神经网络层$h(\ )$,通过逻辑回归的方式得到图像质量分数$s_{x|y}=h(\tilde{H}_{x|y})$。

给定待测图像质量分数s,使用L_2损失函数训练图像质量评估网络:

$$L_4=\sum\|\ s_{x|y}-s\ \|_2^2 \tag{7}$$

在 GAN 训练中,当退化图像为测试图像时,其质量分数真值被设置为低,当理想图像为测试图像时,其质量分数真值被设置为高。

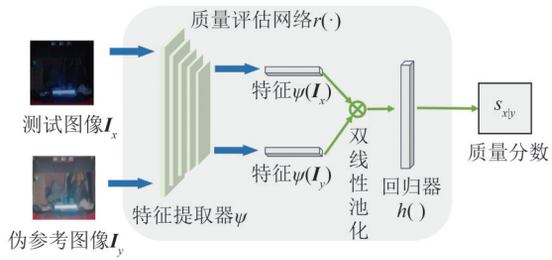


图4 基于伪参考的图像质量评估模块

Fig. 4 Image Quality Assessment Module Based on Pseudo Reference Image

GAN训练结束后,在自建数据集上调优时,使用GAN中的生成器网络$g(\cdot)$,对待测试的交通监控图像I_x生成伪参考图$g(I_x)$,评估I_x的图像质量为$s_{x|g(x)}=r(x,g(x))$,计算损失函数$L=\sum\| s_{x|g(x)}-s \|_2^2$。在调优过程中,固定生成器网络$g(\cdot)$参数,更新质量评估网络$r(\cdot,\cdot)$的参数。

2 图像质量评估实验与分析

2.1 实验设置

数据集主要涉及自建质量评估数据集和一些用于神经网络模型的预训练公开数据集。自建质量评估数据集由多个交通监控摄像头捕获的真实照片图像组成,累计4 241张图片,分辨率从1 920×1 080像素到4 096×2 208像素不等,每张图片包含多个目标和一个关键目标,并由4位经验丰富的专业标注人员对关键区域的图像质量进行标注。LIVE数据集[17]包含779张失真图片,这些失真图片全部由29张参考图像加入7~8种退化程度和5种失真类型生成,图像质量分数为百分制(0~100),值越高则说明质量越低;CID2013数据集[18]包含3 000张失真图像,由25张参考图像在5种退化程度和24种失真类型下生成。对于特定区域识别模块,车辆检测采用了基于YOLOv4[19]的目标检测算法,特定区域检测算法使用灰度调整、边缘轮廓检测和长宽比特征等方法,车牌识别使用RPnet算法[20],车牌号匹配算法使用字符串匹配方法[21]。

对于质量评估模块,使用两个常用的评估指标,斯皮尔曼秩相关系数(Spearman's rank correlation coefficient,SRCC)和皮尔逊线性相关系数(Pearson correlation coefficient,PLCC)[22],前者用于评估预测的单调性,而后者描述了算法的客观评价分与人眼主观打分之间的相关性,用于评估预测的准确度。另外,在计算PLCC之前,预测的质量评估分数需要代入$\widetilde{S}=\beta_1\left(\frac{1}{2}-\frac{1}{e^{\beta_2(\hat{s}-\beta_3)}}\right)+\beta_4\hat{S}+\beta_5$,其中$\beta_i(i=1,2,3,4,5)$为回归系数,$\hat{S}$为原始的图像质量评估分数,$\widetilde{S}$为回归操作的客观质量分数。

本文所有代码均在图形处理单元(graphics processing unit,GPU)1080TI服务器上进行训练和测试,训练或调优过程全部使用Adam优化器[23],学习率分别为1×10^{-3}和1×10^{-5},并在第0.5倍、0.25倍总迭代轮数时退变成原来的1/10;所有数据集被随机分成两部分,80%用于训练或调优,20%用于测试,保证这两部分图像内容独立。对于质量评估模块,输入图片大小统一为80×80像素,首先预训练GAN,然后在自建数据集上调优,调优过程中,GAN中的参数保持不变。超参数选择为$\alpha_1=0.5$、$\alpha_2=0.1$和$\alpha_3=0.5$。生成器的网络结构中编码器为5层下采样卷积神经网络,解码器为5层上采样卷积神经网络,鉴别器模块包含残差神经网络(residual neural network,ResNet)[24]特征提取器和一层全连接层的分类器,特征提取器ψ为浅层的ResNet。

2.2 图像质量评估实验结果

本文比较了一些先进的无参考图像的图像质量评估方法,例如BRISQUE、NIQE、DBCNN、MUSIQ、PAQ2PIQ、GraphIQA和NLNet-IQA。这些方法均使用官方提供的代码,首先在公开数据集上预训练,再在交通监控场景下的数据集上调优。在调优时,将图像中的汽车前挡风玻璃区域缩放为统一尺寸作为模型的输入,结果如表1所示。其中,MUSIQ-AVA、MUSIQ-KONIQ、MUSIQ-PAQ2PIQ和MUSIQ-SPAQ分别指MUSIQ在公开数据集AVA、KONIQ、PAQ2PIQ(文献[11]中使用的数据集)、SPAQ上预训练后,再迁移到监控场景下的质量评估结果,其他方法均在LIVE和CID2013数据集上预训练后再迁移。

由表1可知,本文方法的SRCC高于其他算法,PLCC也处于最优水平。这可能是因为:(1)公开数据集和交通监控图像数据集之间的差异较大,作为无参考图像质量评估方法,其他方法依赖从公开数据集学到的知识,而本文方法能生成伪参考图,通过与伪参考图的特征对比降低了对从公开数据集学到的知识的依赖,更好地适应交通监控场景。(2)其他方法中的网络结构复杂,在调优时容易过拟合或陷入局部最优解,而本文

表 1　图像质量评估的SRCC和PRCC指标

Table 1　SRCC and PRCC Metrics for Image Quality Assessment

算法	SRCC	PLCC
BRISQUE	0.313	0.351
NIQE	0.451	0.480
DBCNN	0.764	0.786
PAQ2PIQ	0.719	0.754
MUSIQ-AVA	0.595	0.600
MUSIQ- KONIQ	0.787	0.808
MUSIQ- PAQ2PIQ	0.767	0.801
MUSIQ- SPAQ	0.789	0.813
GraphIQA	0.785	0.809
NLNet-IQA	0.781	0.812
本文方法	0.792	0.829

方法仅包含浅层的特征提取网络和质量分数回归器,结构简单,容易收敛到局部最优。

2.3　消融实验

2.3.1　特定区域识别模块的影响

为了评估所设计的基于单阶段的车辆前挡风玻璃特定区域的识别算法在图像质量评估实验中的重要性,进行了消融实验,结果如表 2 所示,“×”表示不采用此步骤,而“√”表示采取此步骤。特定区域识别模块包含感兴趣的车辆识别(在表 2 中记作“车辆识别”)和前挡风玻璃区域识别(在表 2 中记作“区域识别”),由于前挡风玻璃区域的识别依赖于感兴趣的车辆识别,根据是否采用这两种识别过程,产生额外两种方案。方案 1 为算法直接对原始图像做图像质量评估;方案 2 表示只进行感兴趣的车辆识别,评估整个车辆所在区域,而不是前挡风玻璃的图像质量。

表 2　特定区域识别模块的影响

Table 2　Impacts of Specific Area Identification Module

方案	特定区域识别模块		评估指标	
	车辆识别	区域识别	SRCC	PLCC
方案 1	×	×	0.513	0.439
方案 2	√	×	0.602	0.586
本文方法	√	√	0.792	0.829

由表 2 可以看出,方案 1 在 SRCC 和 PLCC 上表现均最差,方案 2 虽然也有提升,但是仍然落后于本文方法。其中原因主要在于背景区域或汽车非挡风玻璃区域图像质量优异,但是挡风玻璃区域由于镜面光反射而图像质量低下,或者黑夜等其他环境中背景区域图像质量差,但是挡风玻璃区域由于交通监控探头补光而图像质量好。由于背景区域和汽车非挡风玻璃区域占整幅图像比例较大,所产生的负面影响也大,因此必须提取前挡风玻璃区域,单独评估其图像质量。另外,实验结果说明了本文所设计的特定区域识别模块在交通监控场景下能准确识别前挡风玻璃区域,为后续质量评估模块奠定了基础。

2.3.2　图像质量评估模块中组件的影响

为了验证图像质量评估模块中各个组件的影响,本节研究了各个损失函数部分的效果。由于损失函数包含 4 个部分,通过不同组合产生 4 种额外方案,实验结果如表 3 所示。其中方案 1 为使用像素损失和对抗损失,方案 2 使用特征损失和对抗损失,方案 3 使用像素损失和特征损失,方案 4 使用像素损失、特征损失和对抗损失。在这些方案中,对抗损失均为一般的对抗损失,真假样本标签的判断完全取决于是否为生成的样本,不与图像质量评估任务协同学习,在图像质量评估器的训练时,固定生成器,式(6)中的特征提取器均使用预训练模型。

表 3　各个损失函数组件的影响

Table 3　Impacts of Each Component of Loss Function

方案	L_1	L_2	L_3	L_4	SRCC	PLCC
方案 1	√	×	√	×	0.661	0.654
方案 2	×	√	√	×	0.658	0.699
方案 3	√	√	×	×	0.405	0.593
方案 4	√	√	√	×	0.732	0.771
本文方法	√	√	√	√	0.792	0.829

在表 3 中,方案 3 结果最差,方案 1 和方案 2 与其相比,结果有较大提升,说明对抗训练对于交通监控图像的生成至关重要。对比方案 1 和方案 2,在 SRCC 度量下,方案 1 表现性能较好,而在 PLCC 度量下,方案 2 表现性能较好。这是因为一个关注像素级别的重建,而另一个关注特征级别的重建,两者在不同角度的度量下更能体现出优势。方案 4 同时考虑 3 种损失,在训练时能综合考虑特征级别和像素级别的重建,最终在 SRCC 和 PLCC 度量下均取得了最优结果。

本文方法与方案 4 对比,SRCC 提升了 0.060,PLCC 提升了 0.058。由此可以得出以下结论:通过协同学习,融合图像质量评估任务判断图像的真假样本标签,能够提升鉴别器的鉴别能力和生成器对图像质量评估任务的图像生成能力,进而产生更有帮助的伪参考图像,最终提升图像质量评估任务的性能。

2.4　模型复杂度分析

本节研究模型中各个组件的运行时间复杂度,计算在 4 000 张交通监控图像运行的平均时间,结果如表 4 所示。

表4 各个组件的运行时间/ms

Table 4 Time Cost for Each Component/ms

特定区域识别		质量评估		总耗时
车辆识别	区域识别	伪参考图生成	图像质量评估	
38.1	20.6	9.5	3.4	71.6

由表4可知,各个组件中,特定区域识别模块的运行占了81.9%的总时间,这是因为道路交通的平均输入图像的分辨率较大,例如本文中统一输入分辨率为640×640像素,一方面,这意味着卷积计算量较大,另一方面,模型将生成并筛选更多的候选区域,因此运行速度较慢。另外,由于前挡风玻璃区域识别使用的是传统方法,未使用GPU加速,速率较慢。对于图像质量评估模块,由于交通监控输入图像经过特定区域识别模块后输出的特定区域分辨率较低,例如本文中特定区域的分辨率为80×80像素,因此伪参考图的生成和图像质量评估算法计算量小,运行速度快。从总时间来看,模型的平均运行时间为71.6 ms,对道路监控图像能够较快地做出分析得到结果,快于人的主观判断速率(人的平均速率为2张/s)。

另外本节研究了本文方法和其他图像质量评估模型MUSIQ、GraphIQA的参数量大小。由于模型的输入均为车辆特定区域图像,不计算特定区域识别模块的参数量,结果如表5所示,本文模型参数量仅为19.8 MB,GPU占用资源较小,能提高多路交通监控视频输入数量。

表5 模型参数量的比较/MB

Table 5 Comparison of Model Parameters/MB

MUSIQ	GraphIQA	本文方法		
		生成网络	评估网络	总计
27	29.9	7.8	12.0	19.8

2.5 可视化实验结果

为了进一步验证本文针对监控场景下的质量评估所提出的两阶段框架的可靠性,对每个阶段的输出进行可视化验证,如图5所示。相较于其他方法直接将整体图片作为图像质量评估的输入,本文方法中特定区域识别模块能准确检测监控图像中的所有车辆,在给定感兴趣的车辆目标的车牌号后,其能输出所对应的感兴趣目标以及其中的挡风玻璃处的特定区域。前挡风玻璃区域图像输入到质量评估模块,回归得到图像质量分数,与主观图像质量评分接近。由此可见,本文的两阶段方法中,前一个阶段能为后一个阶段提供稳定可靠的输入。

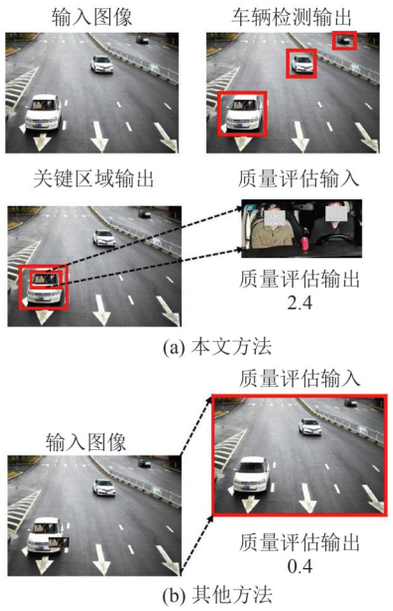


图5 可视化结果示例

Fig. 5 Example of Visualization Results

3 结 语

本文提出了一种两阶段的用于感兴趣车辆中特定区域的图像质量评估方法。为了消除背景等干扰因素的影响,使用目标检测、轮廓检测和车牌识别提取感兴趣车辆的特定区域;为了弥补参考图像的缺失,引入GAN用于生成伪参考图,并通过新的损失函数设计协同训练GAN和质量评估网络。为验证本文方法的有效性,在交通场景下的图像质量评估数据集上进行了算法测试。实验结果表明,本文方法在SRCC和PLCC上均取得最佳的结果;对算法进行了消融实验,结果证实了本文方法中各个组件的有效性;在图像质量评估模型的复杂度上,本文方法模型参数比其他算法少,GPU占用资源较小。实验结果验证了本文方法的有效性,可以更好地满足交通场景下的图像质量评估的实际应用要求。

参 考 文 献

[1] 郭从洲,李可,李贺,等. 遥感图像质量等级分类的深度卷积神经网络方法[J]. 武汉大学学报(信息科学版),2022,47(8):1279-1286.
GUO Congzhou, LI Ke, LI He, et al. Deep Convolution Neural Network Method for Remote Sensing Image Quality Level Classification[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(8): 1279-1286.

[2] 许俊伟,刘永胜,吴达,等. 隧道与地下工程监控量测技术的发展与展望[J]. 测绘地理信息,2023,48(3):7-13.

XU Junwei, LIU Yongsheng, WU Da, et al. Development and Outlook of Monitoring and Measuring Technologies of Tunnel and Underground Engineering [J]. *Journal of Geomatics*, 2023, 48(3): 7-13.

[3] 张雪, 郭佳昕. 监控视频下的人脸图像质量分析技术[J]. 计算机系统应用, 2022, 31(9): 313-318. ZHANG Xue, GUO Jiaxin. Quality Analysis Technology for Face Images in Surveillance Videos[J]. *Computer Systems and Applications*, 2022, 31(9): 313-318.

[4] WANG Z, BOVIK A C. Reduced- and No-Reference Image Quality Assessment[J]. *IEEE Signal Processing Magazine*, 2011, 28(6): 29-40.

[5] TAO S, HE R Z, DAI L S, et al. MD-IQA: Learning Multi-scale Distributed Image Quality Assessment with Semi Supervised Learning for Low Dose CT[EB/OL]. (2023-11-14)[2024-01-05]. https://arxiv.org/abs/2311.08024.

[6] MOHAMMADI S, ASCENSO J. Predictive Sampling for Efficient Pairwise Subjective Image Quality Assessment[C]//The 31st ACM International Conference on Multimedia, Ottawa , Canada, 2023.

[7] MITTAL A, MOORTHY A K, BOVIK A C. No-Reference Image Quality Assessment in the Spatial Domain[J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708.

[8] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a "Completely Blind" Image Quality Analyzer[J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.

[9] ZHANG W X, MA K D, YAN J, et al. Blind Image Quality Assessment Using a Deep Bilinear Convolutional Neural Network[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(1): 36-47.

[10] KE J J, WANG Q F, WANG Y L, et al. MUSIQ: Multi-scale Image Quality Transformer[C]//IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 2021.

[11] YING Z Q, NIU H R, GUPTA P, et al. From Patches to Pictures (PaQ-2-PiQ): Mapping the Perceptual Space of Picture Quality[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020.

[12] SUN S M, YU T, XU J H, et al. GraphIQA: Learning Distortion Graph Representations for Blind Image Quality Assessment[J]. *IEEE Transactions on Multimedia*, 2022, 25: 2912-2925.

[13] JIA S, CHEN B L, LI D Q, et al. No-Reference Image Quality Assessment *via* Non-Local Dependency Modeling[C]//The 24th International Workshop on Multimedia Signal Processing, Shanghai, China, 2022.

[14] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[C]//The 28th International Conference on Neural Information Processing Systems, Montreal, Quebec, Canada, 2014.

[15] LI X L, LI G, ZHANG E Q, et al. Determinant Point Process Sampling Method for Text-to-Image Generation[J]. *Geomatics and Information Science of Wuhan University* , 2024, 49(2): 246-255.

[16] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[C]//The 3rd International Conference on Learning Representations, San Diego, CA, USA, 2015.

[17] GHADIYARAM D, BOVIK A C. Massive Online Crowdsourced Study of Subjective and Objective Picture Quality[J]. *IEEE Transactions on Image Processing*, 2016, 25(1): 372-387.

[18] VIRTANEN T, NUUTINEN M, VAAHTERANOKSA M, et al. CID2013: A Database for Evaluating No-Reference Image Quality Assessment Algorithms[J]. *IEEE Transactions on Image Processing*, 2015, 24(1): 390-402.

[19] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection [M]. (2020-04-23) [2024-01-05]. https://arxiv.org/abs/2004.10934.

[20] XU Z, YANG W, MENG A, et al. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline [C]//The European Conference on Computer Vision, Munich, Germany, 2018.

[21] NAVARRO G. A Guided Tour to Approximate String Matching [J]. *ACM Computing Surveys*, 2001, 33(1): 31-88.

[22] SPRENT P. Applied Nonparametric Statistics [M]. New York: Springer Science & Business Media, 2012.

[23] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization[C]//The 3rd International Conference for Learning Representations, San Diego, CA, USA, 2015.

[24] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016.