

图像增强对低光照场景语义分割影响研究

艾羽丰, 郭继昌, 安冠华, 张 怡

(天津大学电气自动化与信息工程学院, 天津 300072)

摘要: 在低光照环境下获取的图像通常会出现图像亮度低、颜色失真、细节信息丢失以及对比度低等问题。为了满足主观视觉体验的需求, 往往会对图像进行增强处理。然而, 图像增强对机器视觉应用性能的影响缺乏系统研究。本文以语义分割这一机器视觉应用为例, 首先对主流的语义分割方法和低光照图像增强方法进行归纳总结, 然后对经图像增强方法处理的低光照图像进行语义分割, 从而探究图像增强方法对低光照场景语义分割性能的影响。实验结果表明, 增强处理可以改善图像的人眼视觉效果, 但是可能会引入噪声等影响, 并且图像增强方法和语义分割方法关注的重点和特征不完全一致。图像增强对于低光照场景语义分割性能的促进作用并不明显, 甚至会带来负面影响。

关键词: 语义分割; 图像增强; 低光照场景; 图像处理; 机器视觉

中图分类号: TP391.4 **文献标志码:** A

Effect of Image Enhancement on Semantic Segmentation of Low-Light Scene

AI Yufeng, GUO Jichang, AN Guanhua, ZHANG Yi

(School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: Images acquired in low-light environments always suffer from low brightness, color distortion, loss of detail information, low contrast, and other problems. To meet the needs of subjective visual experience, researchers often enhance the images. However, the impact of image enhancement on the performance of machine vision applications is not systematically researched. In this paper, we first summarize typical low-light image enhancement methods and semantic segmentation methods. Next, we take a machine vision application (i.e., semantic segmentation) as an example and select the low-light scene to investigate the effect of image enhancement methods on the semantic segmentation performance of the low-light scene. The experimental results show that enhancement processing can improve the visual effect of images, but may introduce noise. In addition, image enhancement methods and semantic segmentation methods do not concentrate exactly on the same focus and features. Therefore, image enhancement does not contribute significantly to the performance of semantic segmentation in low-light scenes, and even brings negative effects.

Key words: semantic segmentation; image enhancement; low-light scene; image processing; machine vision

引言

语义分割是一种典型的机器视觉任务,它建立了输入图像和输出类别的像素级别的映射,在多个重要领域发挥着重要的作用,如医疗诊断^[1-2]、自动驾驶^[3-4]和机器人导航^[5-6]等。传统的图像分割方法^[7-16]根据图像的颜色、纹理、灰度和形状等简单特征划分图像区域,使区域间有明显的差异性,而区域内各像素呈相似性,从而得到最终的分割结果。随着深度学习技术飞速发展,其在语义分割领域取得了突破性进展,Long等^[17]提出了全卷积神经网络(Fully convolutional networks, FCN)首次将深度学习技术应用于语义分割任务中。此后,一些研究者提出的一系列基于神经网络的语义分割方法均实现了较高的分割性能^[18-23]。随着移动设备的发展,轻量级语义分割网络^[24-30]成为了研究的重点,可以较好地达到效率和精度的权衡,适用于内存受限的移动设备部署。当前的语义分割方法主要针对光照充足的高质量图像。然而在实际应用中,由于图像在各种光照条件下获取,并非所有图像都具有理想的质量。当在光线不足的情况下进行拍摄时,像素值处于低动态范围内,整个图像亮度下降,会导致纹理缺失、图像模糊、颜色失真、对比度低以及细节缺失等问题^[31-32]。对于这些图像,若直接采用当前的语义分割方法对其进行分割将难以获得理想的分割结果。一般认为,对低照度图像进行增强处理可以提高其肉眼视觉效果,但是对于增强处理是否可以提升语义分割性能却缺乏系统的研究和验证。本文针对图像增强对于低光照场景语义分割性能的影响问题展开研究。首先介绍了当前典型的语义分割方法和低光照图像增强方法。在此基础上,设计了以多种经典图像增强方法作为预处理步骤的低光照图像语义分割任务的实验方案,研究图像增强方法对于低光照场景语义分割方法性能的影响。最后通过对比语义分割指标和图像增强指标、观察增强和分割的主观图,总结了图像增强方法对低光照场景语义分割性能的影响规律。

1 语义分割方法

语义分割任务旨在将图像分成若干个具有相似性质的区域。表1列举了一些经典的和最新的语义分割方法。

1.1 传统图像分割方法

传统的图像分割算法通过图像上一些简易的表层特征,如图像纹理和颜色等进行图像分割。典型的算法包括基于边缘检测的图像分割方法^[7-10]、基于阈值的图像分割方法^[11-12]、基于图论的图像分割方法^[13-14]以及基于区域的图像分割方法^[15-16]。这些传统方法所得到的分割结果相对粗糙,分割精度较低,且无相关的语义标签。

1.2 基于深度学习的语义分割方法

基于深度学习的语义分割方法因其强大的细节和语义信息提取能力可以得到更高的分割精度以及更好的分割效果,因此得到了普及和推广。

(1) ERFNet(Efficient residual factorized convNet)^[24]是一个对称的编解码器结构的网络。编码器主要由1D非瓶颈结构(Non-bottleneck-1D)和下采样模块构成,解码器采用步长为2的转置卷积操作将小尺寸的特征图恢复到原图大小,简化了运算和数据存储,同时保持了较高的精度。

(2) ICNet(Image cascade net work)^[25]将图像分为高中低3种不同分辨率,首先利用低分辨率的图像来完成语义分割,然后利用特征级联混合单元(Cascade feature fusion unit, CFF)和标签引导的级联策略将中分辨率和高分辨率的特征整合,逐步优化之前生成的粗糙分割结果。

(3) BiseNetV1(Bilateral segmentation network)^[29]由一条空间路径(Spatial path, SP)和一条语

表1 典型的语义分割方法
Table 1 Typical semantic segmentation methods

提出年份	方法名称	网络结构	优势	劣势
	基于边缘检测的图像分割方法		分割速度快,边缘分割较好。	图像噪声影响大,不适用多个目标边缘重叠的场景。
	基于阈值的图像分割方法		直接运用图像的灰度特性进行计算,计算量较小,简单易实现,运算速度快。	仅适用于分割目标简单且灰度差异较明显的图像。
	基于图论的图像分割方法		分割效果较好,速度快且稳定性高。	对场景复杂的图像分割效果较差。
	基于区域的图像分割方法		解决了空间不连续的缺点,有较好的区域特征,有效消除噪声的影响。	容易造成图像的物体轮廓丢失,速度较慢。
2017	ERFNet	编码-解码器结构	设计了非瓶颈-1D结构作为特征提取骨干,网络的内存占用和参数量较小。	分割精度较低,大目标物体分割不完整。
2018	ICNet	多分支结构	利用高中低3种分辨率的图进行分割,模型收敛速度较快,可以较好地分割高分辨率图。	模型参数量较大、分割速度较慢。
2018	BiseNetV1	双分支结构	采用双分支结构,为高分辨率的特征图补充了细节信息,模型分割精度较高。	模型计算占用大、参数量较大。
2019	LEDNet	编码-解码器结构	利用非对称的编码-解码器结构,在解码器中提出了注意力金字塔网络,模型计算复杂度和参数量较小。	分割精度较低,无法准确定位小目标物体。
2020	BiseNetV2	双分支结构	在训练时使用了一种增强训练策略来提高分割精度,不需要额外的推理成本,就可以得到较好的分割结果。	模型计算复杂度高且参数量较大。
2021	CGNet	轻量级编码器结构	采用轻量级编码器结构,结构简单,模型参数量极低。	分割精度较低,细节信息丢失,分割边缘较粗糙。

义路径(Context path, CP)组成。空间路径可以在生成高分辨率特征图的同时保留空间位置信息,语义路径拥有更大的感受野,从而获得更多的语义信息。同时,BiseNetV1还引入了一个特征融合模块(Feature fusion module, FFM)将两条路径生成的特征图进行融合,实现速度和精度的均衡。

(4) LEDNet(Lightweight encoder-decoder network)^[27]采用了不对称的编解码器结构,编码器采用ResNet作为骨干网络,并且使用通道分离(Channel split)和通道混洗(Channel shuffle)操作来降低计算量并提高分割精度;解码器使用了注意力金字塔网络(Attention Pyramid network, APN)来降低网络复杂度并获得与原图相同分辨率的输出。

(5) BiseNetV2(Bilateral network)^[26]采用细节分支来捕获底层细节并生成高分辨率的特征图,并且采用语义分支来获取高层语义信息,随之采用了一个引导聚合层(Aggregation layer)来融合这两种类型的特征表示。此外,BiseNetV2还设计了一种增强训练策略,在不减慢推理速度的情况下提高分割精度。

(6) CGNet(Context guided network for semantic segmentation)^[28]主要由上下文网格化模块(Context grided block, CG block)构成,有效地学习了图像的局部特征,周边上下文信息和全局上下文信息,以实现在网络的所有阶段捕获上下文信息。CGNet的网络架构经过精心的设计,减少了参数量和内存占用。

2 低光照图像增强方法

为了改善低质图像的人眼视觉效果,图像增强对图像的局部或整体特征进行有目的地强调,如改善图像的颜色、亮度、对比度等。表2列举了本文采用的低光照图像增强方法,包括4种传统方法和4种基于深度学习的方法。

表2 典型的低光照图像增强方法
Table 2 Typical Low-light image enhancement methods

发表年份	方法名称	网络结构	优势	劣势
1971	Retinex		该算法较为合理地假设了低光照图像的成像原理,具有较好的增强效果。	图像亮度偏高,往往会引入不同程度的噪声。
1996	HE		对于原图的亮度与对比度提升较为明显,同时计算复杂度较小。	对于图像的局部与边缘细节信息处理效果欠佳。
1997	MSRCR		在保证图像高保真度的情况下对图像进行了动态范围的压缩,在色彩增强与颜色恒常等方面表现较好。	存在边缘锐化不足,纹理模糊等问题。
2002	ACE		对图像的色度和亮度进行了局部自适应滤波,在色彩和亮度的效果上有较好的表现。	算法的计算复杂度较高。
2018	MBLLEN	多分支结构	处理结果在对比度、亮度增强以及噪声去除等方面均有不错的效果,同时该网络可以进一步扩展至低光照视频增强领域,可实现性较强。	不能较好地处理真实低光照图像中存在的过曝问题。
2019	KinD	多分支结构	由于其采用了不同曝光程度的图片进行训练,使得该算法在达到较好增强效果的同时还具有较强的鲁棒性。	增强图像中往往会引入伪影与噪声。
2020	Zero-DCE	无监督网络	不需要成对合成数据,仅通过不同无参考的损失函数来实现低光照图像的增强,在保证增强效果的同时,也具有较快的推理速度。	对于所设计的损失函数有较高的要求,对于局部的细节信息处理效果一般,同时也存在一定程度的伪影问题。
2021	EnlightenGAN	对抗生成网络	避免了监督学习可能导致的过拟合问题,对于真实低光照图像的泛化能力较强,故训练时间较短。	存在引入不同程度的噪声以及颜色失真等问题。

2.1 传统低光照图像增强方法

典型的传统低光照图像增强方法包括 Retinex^[33],直方图均衡化(Histogram equalization, HE)^[34],多尺度 Retinex 图像增强(Multi-scale Retinex with color restore, MSRCR)^[35]和自动色彩均衡算法(Automatic color equalization, ACE)^[36]等,然而传统低光照图像增强方法鲁棒性不强,往往会存在图像失真的问题,增强效果较差。

2.2 基于深度学习的低光照图像增强方法

(1) MBLLEN(Low-light image/video enhancement using CNNs)^[37]是一种多分支的低光照增强网络。该网络主要由特征提取模块(Feature extraction module, FFM)、增强模块(Enhancement module, EM)和融合模块(Fusion module, FM)三个模块构成。该网络通过提取不同层次的特征图,并将其放在不同子网络中进行增强,最终将不同特征图的增强结果进行融合,使其相互补充,从而可以同时亮度、对比度以及伪影噪声等问题上取得较好的增强效果。

(2) KinD(Kindling the darkness)^[38]是一种基于 Retinex 理论的低光照增强网络。该网络主要由层

分解网络(Layer decomposition net, LDN)及其两个分支子网络、反射率恢复网络(Reflectance restoration, RRN)、以及光照调节(Illumination adjustment net, IGN)网络所构成。由于该网络基于Retinex理论,故由层分解子网络包含的两个分支可将原始图像分解为反射图像与光强图像两个部分。其中IGN可以通过设置不同的参数灵活调整亮度,得到调整后的光强图像,而RRN可以将层分解网络得到的反射图像和调整后的光强图像结合,从而得到最终的增强图像。

(3) Zero-DCE(Zero-reference deep curve estimation)^[39]是一种无参考的基于曲线映射的轻量级低光照增强网络。曲线映射的增强方法可以根据输入图像和相应的曲线参数对原图像进行增强。DCE-Net可以根据输入的低光照图像生成像素级的曲线增强映射图,与传统曲线增强方法不同的是,该曲线映射不再是全局映射,而是像素级的映射,即在不同的局部区域进行映射,从而避免局部区域增强过度/不足的问题。在此基础上根据多个曲线增强映射图和对应的增强曲线对低光照图像进行迭代增强。

(4) EnlightenGAN^[40]:由于低光照图像很难收集到成对图像数据集,该方法首次应用了无监督的生成对抗网络,设计了无需成对训练数据集的低光照增强网络。其中生成器采用了注意力机制引导的U-Net,而判别器分别将全局以及局部图像与正常曝光图像进行双重判别,不仅可以从全局改善增强效果,同时也可以自适应地增强局部区域。

3 实验方案、数据集及评价指标

基于以上对于低光照图像增强算法和语义分割算法的理论研究,本节对基于增强处理的低光照图像场景语义分割研究设计实验方案,探究增强处理对低光照场景语义分割性能的影响,进而分析图像增强在语义分割中的作用。

3.1 实验方案

本文选取8种低光照图像增强方法,包括4种传统方法和4种基于深度学习的方法,对低光照图像进行预处理,获得对比度高、色彩丰富、边缘信息清晰以及视觉效果良好的增强图像。在此基础上,本文选取6种典型的基于深度学习的语义分割算法,以原始低照度图像以及增强后的图像为输入,得到图像的语义分割结果,该实验方案的流程如图1所示。为了避免网络层数过深导致的参数量大、训练速度过慢等问题,本文选取的语义分割算法均为轻量级网络,这些网络考虑了精度和效率的平衡,其分割结果规律和重量级语义分割网络一致。此外,本文对图像增强网络以及语义分割网络分别进行训练。

基于这个实验方案,本节将上述介绍的8种典型图像增强方法应用在低光照语义分割中进行对比实验分析,其目的在于:(1)验证有无增强两种条件下6种典型语义分割方法的效果;(2)验证不同图像增强方法对于语义分割性能的影响程度。

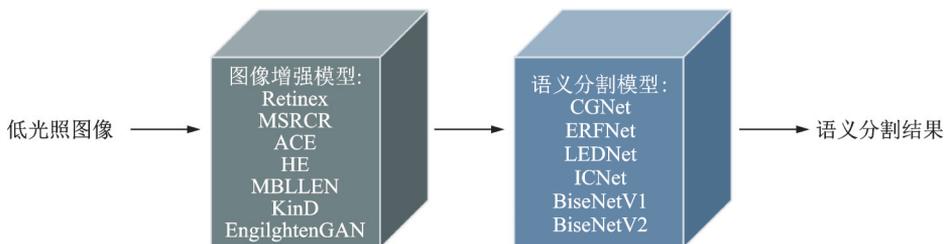


图1 基于图像增强的语义分割实验流程图

Fig.1 Experimental flowchart of semantic segmentation based on image enhancement

3.2 数据集

(1) ACDC 数据集^[41]是一个针对于不利的视觉条件下所提出的对驾驶场景进行语义理解的数据集。它拍摄于城市地区、高速公路和农村地区,包括了4 006张图像,均匀分布在雾天、夜间、雨天和雪天4种不利的视觉场景下,每张图像都带有19类高质量精细像素语义注释。本文选取其中的夜间图像,包括400张训练集和106张验证集,并采用随机翻转、随机缩放来进行数据增广。

(2) BDD100K 数据集^[42]是目前最大规模、内容最具多样性的公开驾驶数据集。它针对语义分割任务提供了7 000张训练集和1 000张验证集,涵盖不同时间、不同天气条件(包括晴天、阴天和雨天,以及白天和晚上的不同时间)的驾驶场景。本文选取出542张夜间图像,手动划分为430张训练集和112张验证集,利用随机翻转、随机缩放操作来增加数据鲁棒性。

3.3 评价指标

由于无法获取低光照图像的理想图像,所以本节选择均值、标准差、平均梯度、信息熵以及对比度等无参考评价指标作为低照度图像增强的评价指标,此类方法一般是基于图像统计特性。语义分割评价指标采用平均交并比(mean Intersection over union, mIoU)。

(1) 均值指图像像素的平均值,反映了图像的平均亮度。设图像均值为 μ ,待评价图像为 I ,大小为 $M \times N$,图像像素点为 i, j ,其计算公式为

$$\mu = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N I(i, j) \quad (1)$$

(2) 标准差指图像像素灰度值相对于均值的离散程度。标准差越大,图像的灰度级分布就越分散。设图像标准差为 σ ,其计算公式为

$$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I(i, j) - \mu)^2} \quad (2)$$

(3) 图像熵指图像的平均信息量,图像中的信息熵越大,说明图像包含的信息越多。设图像 $I(x, y)$ 具有 K 个等级,图像熵为 H ,其计算公式为

$$H = - \sum_{i=0}^{K-1} P_i \log_2 P_i \quad (3)$$

式中 P_i 为每一灰度出现的概率。

(4) 平均梯度反映图像的细节反差和纹理变换,平均梯度越大,图像层次越多。设平均梯度为 ∇G ,水平方向的梯度为 ∇_x ,垂直方向的梯度为 ∇_y ,其计算公式为

$$\nabla G = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \sqrt{\nabla_x F(i, j)^2 + \nabla_y F(i, j)^2} \quad (4)$$

(5) 对比度指图像黑与白的比值,比值越大,从黑到白的渐变层次就越多。设对比度为 C ,其计算公式为

$$C = \sum_{\delta} \delta(i, j)^2 P_{\delta}(i, j) \quad (5)$$

式中: $\delta(i, j) = \text{abs}(i - j)$ 表示相邻像素间的灰度差; $P_{\delta}(i, j)$ 为相邻像素间灰度差为 δ 的分布概率。

(6) 平均交并比(mIoU)为语义分割精确度评价指标,是计算真实值和预测值两个集合的交集和并集之比,其计算公式为

$$\text{mIoU} = \frac{1}{k} \sum_{i=1}^k \frac{P \cup G}{P \cap G} \quad (6)$$

式中: P 表示预测值; G 表示真实值; k 表示总类别个数。

(7)全变分(Total variation, TV)为梯度幅值的积分,用于衡量图像的平滑程度和噪声的大小,一般来说,其值越小,代表图像越平滑、噪声越小。设图像为 $U(x, y)$,其计算公式为

$$TV[U(x, y)] = \iint_{\Omega} |\nabla(U(x, y))| dx dy \quad (7)$$

式中 $\nabla(U(x, y))$ 代表图像的梯度。

4 实验结果及分析

为了公正客观地评估实验结果,实验中所有的方法都是在同一个硬件平台上进行训练和测试,并且所有算法均采用原文献所提供的参数设置。本文所有实验均在Ubuntu 18.04系统上完成,使用的显卡为单个1080Ti GPU,显存为11 GB。

4.1 ACDC数据集实验结果

4.1.1 图像增强结果

本文采用8种图像增强方法对低光照图像进行图像增强处理,并且计算增强后的低光照图像的相关图像评价指标,如表3所示。从表中可以看出,低光照图像经过ACE、EnlightenGAN、HE、MSRCR、Retinex和Zero-DCE方法增强处理后,其图像均值、标准差、平均梯度、信息熵和对比度都有了一定的提升,这说明增强后图像亮度提升、明暗差距变大、纹理特征和边缘信息更加充足;经KinD增强处理后的图像标准差略有下降,说明图像像素值的波动变小;经MBLLEN增强处理后的图像在平均梯度、对比度两个评价指标上相对于原始低光照图像较低,说明图像较为平滑,边缘特征相对不明显。此外,经过ACE、EnlightenGAN、HE、MSRCR、Retinex和Zero-DCE增强算法处理后的图像的全变分指标都增大了,说明增强后的图像引入了不同程度的噪声,而经MBLLEN增强算法处理后的图像全变分指标降低,说明增强后的图像变得相对平滑,图像噪声较少。

表3 不同增强方法下ACDC数据集图像增强指标

Table 3 Image enhancement metrics of ACDC dataset by different enhancement methods

指标	均值	标准差	信息熵/(比特·像素 ⁻¹)	平均梯度	对比度	全变分
低光照图像	46.18	40.86	10.43	1.53	21.82	0.002 9
ACE	78.92	47.30	13.93	3.28	54.24	0.010 9
EnlightenGAN	141.93	55.25	14.08	3.41	41.09	0.007 8
HE	129.48	72.39	14.16	4.70	65.13	0.021 0
KinD	115.38	33.06	12.26	2.07	17.74	0.006 5
MBLLEN	120.30	42.27	10.99	0.91	7.46	0.001 4
MSRCR	126.83	57.32	15.42	8.22	240.57	0.049 6
Retinex	190.84	49.04	12.00	2.37	24.13	0.005 3
Zero-DCE	112.75	43.02	13.27	3.14	37.23	0.007 6

注:红色加粗表示增强后图像相比低光照图像的语义分割指标有所提升。

图2给出的主观效果图进一步说明了部分增强图像有更加丰富的色彩、明显的对比度、清晰的边缘以及更多的细节信息,但是同时也可以看出:经过Retinex算法增强的图像(H行)亮度过大,甚至出现曝光的情况;经ACE、EnlightenGAN、HE和MSRCR算法增强的图像(B、C、D、G行)放大了不必要的纹理轮廓并且存在明显的噪点;如D、E行所示,经HE和KinD增强处理后的图像出现了明显的颜色失真问题;如I行所示,经Zero-DCE增强的图像提升了图像亮度,可以较清晰地识别出杆子、道路边缘等细小

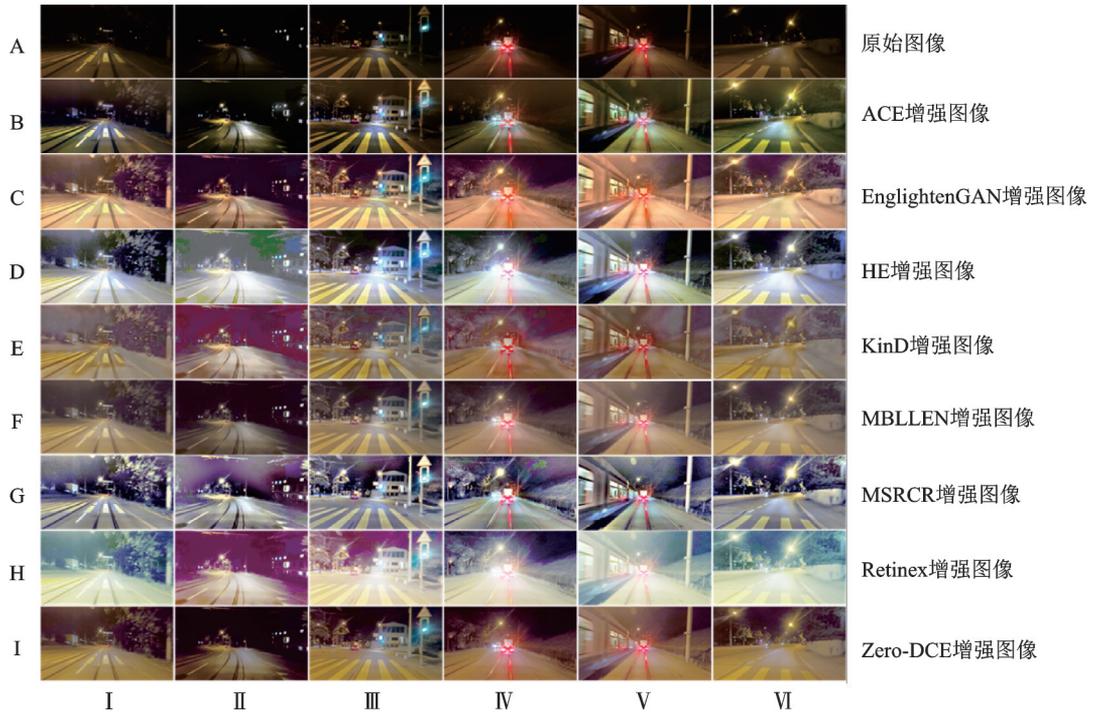


图2 不同图像增强算法用于ACDC数据集的主观结果

Fig.2 Results of different image enhancement algorithms on ACDC dataset

部分;从F行可以看出,经MBLLEN增强算法得到的图像对比度较低,梯度变化不明显,图像整体更加平滑。

4.1.2 语义分割结果

从表4可以看出,未经增强的低光照图像均取得了较高的mIoU,部分图像增强方法对部分语义分割标检测方法具有积极影响,如图像经MBLLEN增强处理后,利用ICNet、LEDNet、BiseNetV1和BiseNetV2方法的语义分割结果分别提升0.12%、0.84%、0.24%和0.79%;经过MSRCR增强处理后利用ICNet进行语义分割结果提升0.07%;经过EnlightenGAN增强处理后,利用ICNet进行语义分割结果提升0.48%;经过Zero-DCE增强处理后利用ICNet、BiseNetV2进行语义分割,结果分别提升0.16%、0.37%。但是大部分图像增强方法对于语义分割的性能有抑制作用。经过Retinex、HE、ACE和KinD进行图像增强处理之后采用不同语义分割算法模型进行语义分割任务,性能均有所下降,经过MSRCR、MBLLEN、EnlightGAN、Zero-DCE增强处理后部分语义分割算法的效果变差,如经MSRCR增强处理后CGNet、LEDNet、ERFNet、BiseNetV1和BiseNetV2的语义分割效果分别下降了0.78%、0.75%、1.02%、0.57%和0.55%。进一步地,表5给出了BiseNetV1在不同增强方法处理的ACDC数据集上的不同物体种类的mIoU。未经过增强的低光照图像在分割人行道、交通灯、单车时取得了最佳效果,而货车和巴士的分割效果最差,说明未经过增强的图像中大物体的特征易受到光线不足的影响无法得到充分的提取,导致分割不完整;经ACE处理后的图像在分割行人和摩托车时取得了最佳效果,而在分割墙和交通标志时效果最差,说明增强后的图像有较清晰的边缘轮廓,易分割边缘复杂的物体,但是图像引入的噪声导致无法将大物体分割完整;经过EnlightenGAN增强后的图像提升了对草地的分割性能,说明增强后的图像提升了亮度,可以更好地识别出黑暗场景的目标;经过HE处理的图像不能较好地分割植被和单车,其余种类的分割精度也较低,说明增强后的图像引入

表4 ACDC数据集中不同图像增强方法下低光照语义分割的mIoU性能分析

Table 4 mIoU for low-light semantic segmentation with different image enhancement methods on ACDC dataset %

分割方法	CGNet	ICNet	LEDNet	ERFNet	BiseNetV1	BiseNetV2
低光照图像	34.36	33.86	37.60	36.85	38.50	38.30
ACE	33.49	32.05	35.44	35.61	37.29	37.67
EnlightenGAN	33.78	34.34	36.46	36.07	37.58	37.24
HE	32.75	33.36	36.11	35.04	36.89	37.27
KinD	32.55	32.38	35.86	35.17	36.91	37.58
MBLLEN	33.94	33.98	38.44	35.71	38.74	39.09
MSRCR	33.58	33.93	36.85	35.83	37.93	37.75
Retinex	33.20	32.20	36.63	35.40	37.36	38.18
Zero-DCE	33.86	34.02	37.02	36.70	38.07	38.67

注:红色加粗表示增强后图像相比低光照图像的语义分割指标有所提升。

表5 ACDC数据集中BiseNetV1在经不同图像增强方法处理的不同种类物体mIoU性能分析

Table 5 mIoU of BiseNetV1 on ACDC dataset for different object types with different image enhancement methods %

物体种类	低光照图像	ACE	EnlightenGAN	HE	KinD	MBLLEN	MSRCR	Retinex	Zero-DCE
道路	87.22	86.57	86.67	86.18	86.36	87.68	87.56	85.94	87.96
人行道	47.61	45.04	44.43	44.16	43.70	47.16	47.43	45.52	46.94
建筑	50.56	49.34	49.77	49.20	48.72	51.19	49.50	48.53	50.90
墙	31.48	24.69	29.59	28.84	24.86	28.99	31.94	29.94	27.92
围栏	12.55	11.86	13.52	14.64	10.89	12.80	11.20	18.05	14.81
杆子	37.61	35.63	36.70	36.83	35.34	35.96	35.80	33.98	36.93
交通灯	45.57	38.90	44.15	41.55	42.45	45.19	38.00	41.19	42.55
交通标志	35.49	29.00	32.11	32.58	33.51	35.57	32.17	32.06	34.75
植被	46.87	47.05	46.04	45.35	45.71	49.11	48.96	47.72	47.33
草地	11.85	7.55	15.59	7.94	10.54	6.94	6.87	13.61	8.39
天空	70.10	69.30	69.72	70.93	68.60	69.94	72.00	72.35	70.48
行人	30.17	33.96	28.40	25.55	33.71	29.54	29.74	21.31	22.28
骑行者	6.37	1.48	10.60	0.67	8.41	11.65	0.33	0.13	4.89
汽车	52.76	55.04	54.57	56.12	52.75	57.35	56.04	55.71	54.86
货车	17.23	21.01	18.82	22.28	20.43	20.87	22.42	22.34	20.76
巴士	27.12	32.23	30.82	31.56	31.12	30.42	35.17	30.12	32.98
火车	71.46	69.57	74.56	75.26	74.25	75.74	72.48	68.62	72.16
摩托车	4.77	17.86	2.71	7.04	3.97	5.65	0.14	0.68	3.08
单车	44.50	32.52	25.32	24.32	25.95	35.44	42.83	41.92	43.33
平均交并比	38.50	37.29	37.58	36.89	36.91	38.74	37.93	37.36	38.07

注:红色加粗表示最佳分割效果;蓝色加粗表示最差分割效果。

的大量噪声降低了图像的分割性能;经 KinD 处理的图像整体分割效果较差,对于人行道、围栏、天空、汽车的分割精度均为最低;经过 MBLLEN 增强处理的图像整体取得了最佳的分割效果,在分割建筑、交通标志、植被、骑行者、汽车和火车时效果最佳,说明增强后的图像较平滑,且提高了亮度,可以较好地定位和分割目标物体;经过 MSRCR 增强处理的图像对于大物体的分割效果较好,如墙、货车和巴士,然而对小物体如交通灯分割性能不佳,说明增强后的图像引入了噪声,干扰了分割模型对于小物体位置的判断;由于经 Retinex 处理的图像亮度较大,出现了图像曝光过度的情况,因此对于道路、建筑、杆子、行人和火车的分割效果都不尽人意,甚至骑行者的分割准确率仅有 0.13%,但是在分割较大物体如围栏和天空时,效果最佳;经 Zero-DCE 增强处理的图像对道路和杆子的分割效果较好,说明增强后的图像有较清晰的边缘。

图 3 给出的语义分割主观效果和图 4 给出的经 BiSeNetV1 模型提取得到的特征图进一步验证了以上分析。图 3 中,从第 I~II 列分别为采用 CGNet、ICNet、LEDNet、ERFNet、BiSeNetV1 和 BiSeNetV2 模型的语义分割效果。从第 A~K 行分别为低光照图像、真实标签、不同模型在低光照图像、经 ACE、EnglightenGAN、HE、KinD、MBLLEN、MSRCR、Retinex、Zero-DCE 增强后图像上的语义分割结果。图 4 中,从第 A~I 行分别为通过低光照图像以及经 ACE、EnglightenGAN、HE、KinD、MBLLEN、MSRCR、Retinex 和 Zero-DCE 增强后图像上提取的特征图。由于低光照图像的亮度较低、边缘模糊,无法很好地将不同物体分割开,而经过增强处理过的图像可以较好地区分不同物体的

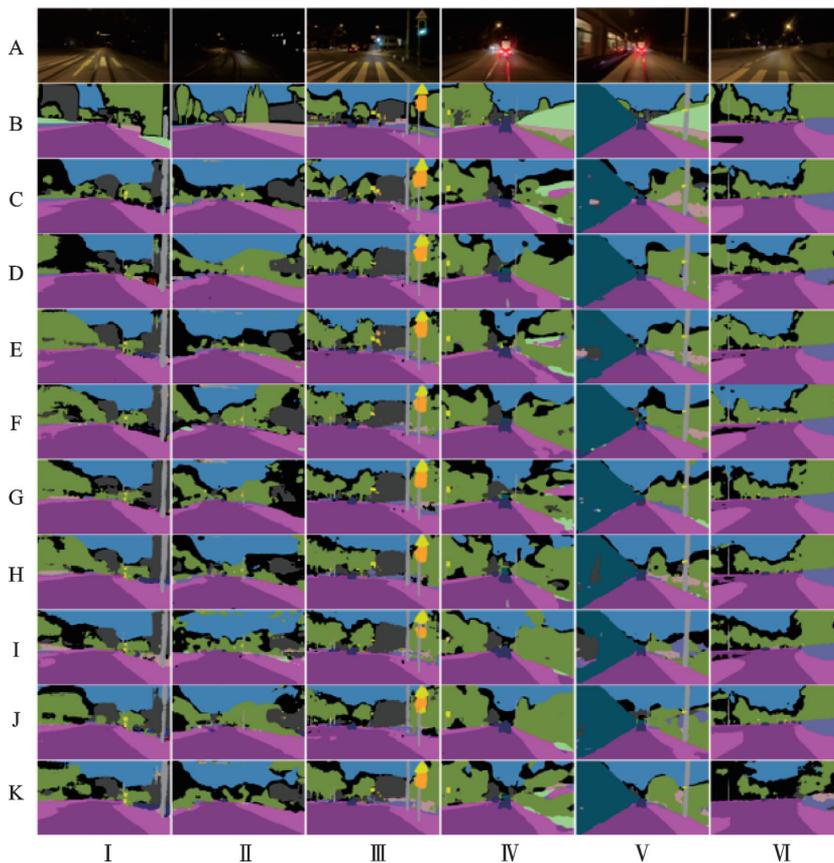


图 3 不同语义分割模型用于 ACDC 数据集的原始图像和增强图像的结果

Fig.3 Effect of different semantic segmentation models for the original and enhanced images of ACDC dataset.

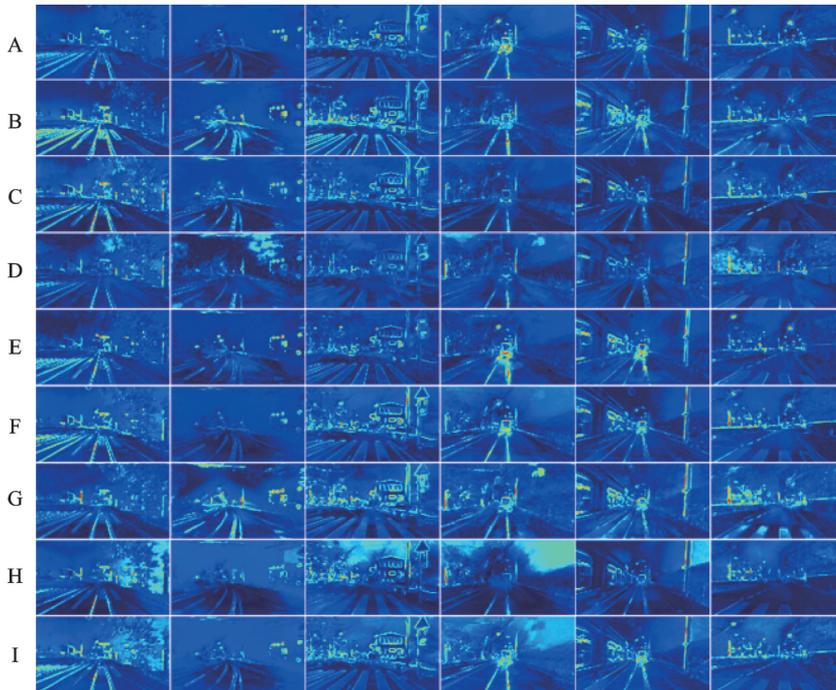


图4 经BiSeNetV1模型在ACDC数据集上提取的特征图

Fig.4 Feature maps extracted by BiSeNetV1 on the ACDC dataset

轮廓,但同时增强处理引入的噪点会影响模型的效果,导致物体分割不完整。具体来说,经过ACE、EnlightenGAN、HE和MSRCR增强后提取的特征图(图4第B、C、D、G行)中分割模型的关注点有所迁移,导致模型分割时错误地把不同物体融合在一起,使得小物体和复杂的物体如交通标志、单车和交通灯的分割效果较差(图3第B、C、D、G行)。经过KinD算法增强后提取的特征图(图4第E行)中目标边缘细节有所丢失,并且引入了大量的冗余信息,导致大物体如天空、汽车和围栏分割不完整,分割边缘也较为粗糙(图3第E行)。如图4第H行所示,经过Retinex增强后提取的特征图由于增强算法的过度曝光导致分割模型过度关注无用的曝光特征,错误地将一个物体分割为多个物体,如道路、建筑和火车等,造成模型分割物体不完整的问题(图3第H行)。经MBLLEN、Zero-DCE增强后提取的特征图(图4第F、I行)未过度引入噪点,同时细节信息保存较好,图3第F、I行也显示出该两种增强算法有利于低光照场景下的语义分割,尤其是对于大物体和边缘复杂的物体分割较为理想。

以上分析说明了增强算法虽然有利于提高如均值、标准差、信息熵,平均梯度和对比度等客观评价指标,使图像亮度有所提升,明暗变化更加明显,图像纹理特征、边缘信息都有所增加,从而提升了低光照场景中图像的人眼视觉效果。但实际上,增强算法会引入不同程度的冗余信息,干扰分割模型的判断,使模型错误地定位和分割目标物体,同时增强后的图像可能存在图像曝光、颜色失真等问题。多种因素共同作用,最终导致分割性在多数情况下有着不同程度的降低或微弱提升。

4.2 BDD数据集实验结果

4.2.1 图像增强结果

BDD数据集图像增强结果如表6和图5所示。从表6可知,大部分经图像增强处理的图像的评价指标均有了一定提升。然而,经过ACE增强的图像虽大幅提升了图像的对比度,但图像标准差有所下降;经过KinD增强的图像标准差和对比度都略微降低;经过MBLLEN增强的图像均值和信息熵略有

表 6 不同增强方法下BDD数据集图像增强指标

Table 6 Image enhancement metrics of BDD dataset by different enhancement methods

指标	均值	标准差	信息熵/(比特·像素 ⁻¹)	平均梯度	对比度	全变分
低光照图像	50.98	56.90	10.87	2.52	49.24	0.010 1
ACE	84.99	52.58	13.18	4.72	113.73	0.023 7
EnlightenGAN	119.06	63.87	13.44	4.21	72.86	0.014 5
HE	130.35	71.72	13.26	4.87	106.52	0.023 2
KinD	111.17	42.81	12.58	3.27	46.20	0.011 8
MBLLEN	119.79	52.94	11.42	1.94	41.21	0.004 7
MSRCR	126.93	58.90	14.44	7.04	198.65	0.041 4
Retinex	187.57	60.36	11.42	2.37	49.51	0.023 5
Zero-DCE	112.24	56.92	12.52	3.95	68.14	0.014 2

注:红色加粗表示增强后图像相比原图的图像评价指标有所提升。



图 5 不同图像增强算法用于BDD数据集的主观结果

Fig.5 Results of different image enhancement algorithms on BDD dataset

提升,但相比其余增强算法的提升较少,且其标准差、平均梯度和对比度相对于原图略微下降;图像的对比度在经 MSRCR 增强后有了显著的提升;经过 Retinex 增强后的图像均值有较大提升,但是平均梯度相较原图有所下降;经过 Zero-DCE 增强后图像的评价指标较原图均有提升,但是提升幅度都不大。除了 MBLLEN 算法,其余增强算法处理的图像全变分指标都变大了,说明这些增强算法处理的图像均引入了噪声,降低了图像的平滑程度。

图 5 所示的主观图验证了图像评价指标的变化趋势,大部分增强后的图像有了更高的亮度,更多的

信息和更丰富的色彩。经过MBLLEN、Zero-DCE增强处理后的图像(图5第F、I行)较为平滑,亮度相较原图有了一定的提升,图像细节信息保留较好,可以较清晰地识别出小物体。然而,与ACDC数据集的增强结果相似,增强算法会在计算过程中引入难以忽视的噪声,并且不同程度地造成图像失真。如经过EnlightenGAN增强的图像(图5第C行)虽有更丰富的色彩和更高的亮度,但同时也存在图像失真的问题;经过ACE(图5第B行)处理后的图像亮度虽相较原图有一定的提升,但仍然较低;经过HE和KinD增强的图像(图5第F、G行)出现了颜色失真的问题,并且引入了大量噪声和冗余特征;经过MSRCR增强的图像(图5第G行)虽然图像梯度变化更加明显,但是引入了大量噪声;经过Retinex处理后的图像(图5第H行)图像边缘较为柔和,曝光过度,导致图像丢失大量细节信息。

4.2.2 语义分割结果

以不同的图像增强方法为预处理技术,在BDD数据集上进行低光照图像的语义分割结果如表7所示。大部分增强方法会导致语义分割任务性能有不同程度的下降。其中,经过KinD增强算法处理后的图像的模型分割下降最为明显,使得CGNet、ICNet、LEDNet、ERFNet、BiseNetV1和BiseNetV2的分割性能分别降低了1.04%、0.32%、0.46%、2.41%、1.36%和1.12%。对经MBLLEN和Zero-DCE增强的图像进行分割,部分模型提升了分割性能,如CGNet、BiseNetV1和BiseNetV2模型在经MBLLEN增强处理后的图像上的分割性能分别提升了0.32%、0.52%和0.03%。LEDNet和BiseNetV1在经过Zero-DCE增强处理的图像上的分割精度也有了一定的提升,但这两种增强算法对于分割模型的选择上鲁棒性不强。

表7 BDD数据集中不同图像增强方法下低光照语义分割的mIoU性能分析

Table 7 mIoU of low-light semantic segmentation with different image enhancement methods on BDD dataset

分割方法	CGNet	ICNet	LEDNet	ERFNet	BiseNetV1	BiseNetV2	%
低光照图像	24.35	26.46	25.89	26.10	29.06	28.10	
ACE	23.54	26.03	25.10	24.32	28.28	27.48	
EnlightenGAN	23.79	25.85	25.50	24.44	28.70	27.43	
HE	23.74	25.96	25.49	24.85	28.86	28.08	
KinD	23.31	26.14	25.43	23.69	27.70	26.98	
MBLLEN	24.67	26.34	25.75	24.94	29.58	28.13	
MSRCR	24.19	26.39	25.55	24.28	28.70	27.34	
Retinex	23.98	25.77	25.26	24.66	28.13	27.42	
Zero-DCE	23.65	26.18	25.90	25.56	29.08	27.54	

注:红色加粗表示增强后图像相比原图的语义分割指标有所提升。

表8给出了BiseNetV1经过不同图像增强方法处理后的BDD数据集的不同种类物体的mIoU。未经增强处理的低光照图像在分割建筑、围栏和草地时有最佳性能,然而对于骑行者、汽车的分割效果较差,这说明在低光照场景下,较难分割边缘复杂的物体,同时受到灯光和黑暗条件的影响,很难完整分割较大的物体;经过ACE增强处理的图像在分割道路、围栏、行人和摩托车时效果较差,这是由于增强后图像亮度较低并且引入了噪点,无法准确地定位小物体和完整地分割大物体;由于经过KinD增强的图像会出现颜色失真和细节信息丢失的问题,因此在分割人行道、墙、杆子、货车、火车、单车时效果较差;经过MBLLEN增强处理的图像在分割人行道、墙、交通灯、行人、货车、火车和摩托车时均达到最好的分割效果,说明此增强算法在未引入过多的冗余信息的同时提高了特征

表8 BDD数据集中BiseNetV1在经过不同图像增强方法处理的不同种类物体mIoU性能分析

Table 8 mIoU of BiseNetV1 on BDD dataset for different object types with different image enhancement methods %

	低光照 图像	ACE	EnlightenGAN	HE	KinD	MBLLEN	MSRCR	Retinex	Zero- DCE
道路	69.52	69.50	70.67	70.44	70.17	69.70	71.60	70.59	70.71
人行道	33.12	32.85	34.70	33.16	31.15	35.07	33.25	32.89	33.05
建筑	68.64	66.95	67.23	67.21	66.00	68.09	68.35	65.45	66.82
墙	3.39	10.26	2.63	8.13	0.22	11.82	5.47	6.93	6.86
围栏	13.62	3.44	7.04	11.43	9.59	10.73	11.06	13.39	9.74
杆子	22.30	21.61	22.13	20.63	20.37	21.03	21.25	24.45	22.68
交通灯	17.22	15.20	18.02	16.38	17.92	18.32	17.98	14.45	17.87
交通标志	28.19	31.31	30.75	28.43	31.92	27.63	32.07	27.47	27.61
植被	60.46	61.06	59.89	58.25	58.66	60.26	61.23	61.04	62.48
草地	16.82	13.84	15.86	12.49	14.25	12.82	13.80	13.78	15.39
天空	69.17	66.94	70.75	63.71	67.89	68.43	70.64	67.02	69.11
行人	17.16	14.50	17.87	18.29	17.68	19.94	15.31	17.04	19.28
骑行者	1.29	4.24	5.63	4.18	5.76	3.28	2.34	4.56	6.75
汽车	77.51	80.47	79.18	80.23	79.70	79.67	81.51	79.25	80.19
货车	11.47	11.11	17.88	10.58	5.49	19.15	8.92	7.42	12.45
巴士	10.04	8.86	1.34	15.78	6.97	5.41	0.75	0.83	4.07
火车	10.42	9.56	8.45	9.23	8.02	11.58	10.85	9.09	8.89
摩托车	8.67	7.42	7.83	9.11	8.13	10.90	9.24	10.22	7.53
单车	9.13	8.33	7.61	10.82	6.57	8.25	9.77	8.43	11.37
平均交并比	29.06	28.28	28.70	28.86	27.70	29.58	28.70	28.13	29.08

注:红色加粗表示最佳分割效果,蓝色加粗表示最差分割效果。

的可识别度。

图6的分割主观图和图7的特征图进一步说明了语义分割的结果。图6中,从第I~II列分别代表采用CGNet、ICNet、LEDNet、ERFNet、BiseNetV1和BiseNetV2模型的语义分割效果。从第A~K行分别为低光照图像、真实标签、不同模型在低光照图像、经ACE、EnlightenGAN、HE、KinD、MBLLEN、MSRCR、Retinex和Zero-DCE增强后图像上的分割结果。图7中,从第A~I行分别为通过低光照图像以及经ACE、EnlightenGAN、HE、KinD、MBLLEN、MSRCR、Retinex和Zero-DCE增强后图像上提取的特征图。如图7第F、I行所示,经过MBLLEN、Zero-DCE增强的图像特征图的纹理较原图和其他增强算法处理的图像特征图更加清晰,关注到了一些不受重视的特征,并且没有明显引入冗余的特征信息。因此,从图6第H、K行可以看出,经MBLLEN和Zero-DCE增强处理的图像可以较为准确地定位小目标物体并且完整地分割大目标物体。如图7第B~H行所示,经ACE、EnlightenGAN、HE、KinD、MSRCR和Retinex增强后提取的特征图的关注点有所迁移,且存在放大不必要的纹理轮廓特征等问题,引入了大量的冗余信息,造成分割模型无法

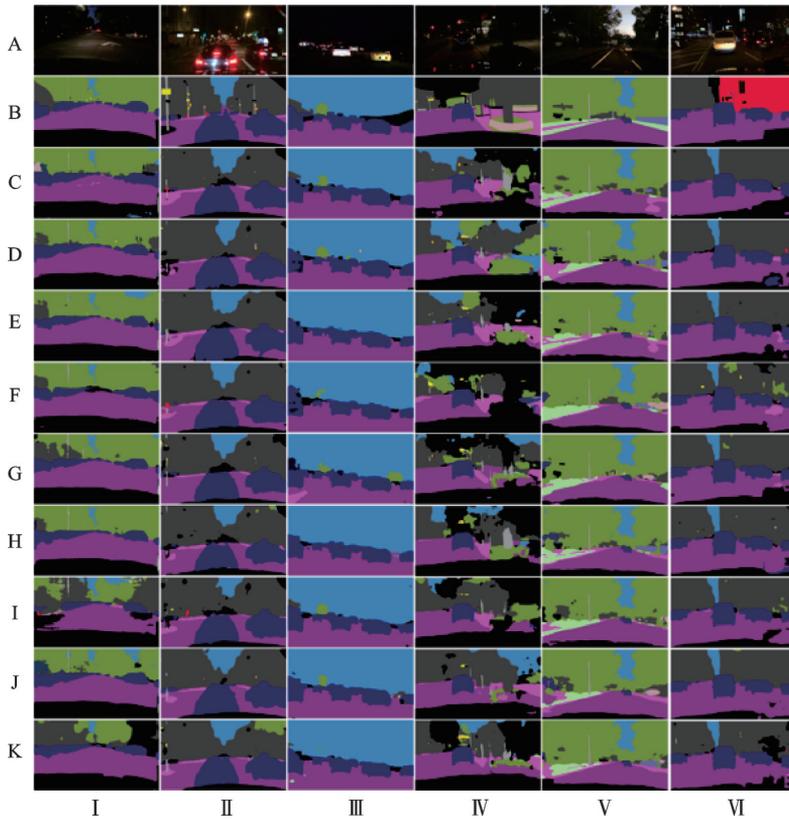


图6 不同语义分割模型用于BDD数据集的原始图像和增强图像的结果

Fig.6 Effect of different semantic segmentation models for the original and enhanced images of BDD dataset

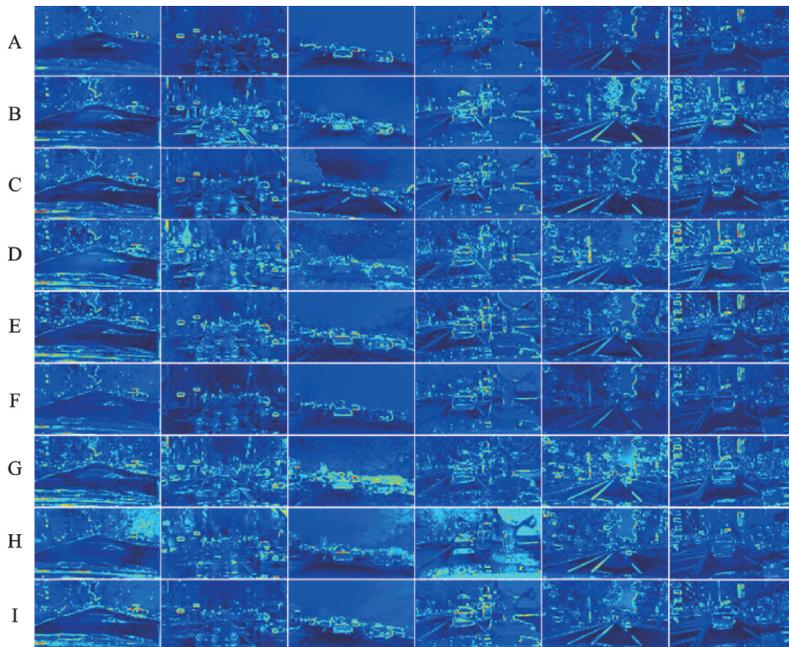


图7 经 BiseNetV1 模型在 BDD 数据集上提取的特征图

Fig.7 Feature maps extracted by BiseNetV1 on BDD dataset

专注于有用的特征,因而导致了分割性能的下降。如图6第D、F行所示,经ACE和KinD增强处理的图像,在光照较弱的区域,易错误地将不同种类的物体混合在一起,造成分割性能下降。经过Retinex增强处理的图像(图6第J行)会将曝光过度的部分当成一个整体,从而进行错误的分割。

4.3 实验分析

通过在ACDC和BDD数据集的实验,结合图像评价指标以及语义分割结果,可以发现即便增强图像的图像评价指标相对于增强前有所提高,但是仍有可能导致语义分割性能下降。这可能是由于现有的图像增强方法更多关注的是人眼视觉效果,而语义分割关注的重点为图像的纹理特征和信息,二者之间暂时还没有一致性评价指标。图像增强方法会影响到图像的多方面特征,如图像亮度、明暗程度、边缘强度、纹理特征、细节特征和平滑程度等,这些特征综合作用影响语义分割性能。然而,当信息熵太大时,图像信息杂乱,分割时可能会提取出错误的特征信息;当图像亮度太大时,图像出现曝光的情况,从而造成图像边界模糊,影响分割结果;当图像的梯度过大时,导致语义分割关注的重点集中在梯度较大的区域,忽略了重要的特征区域;当图像的对比度过大时,会造成目标物体内部的像素值差异过大,导致物体分割不完整。同时,在图像增强的过程中,不可避免地会引入噪声,即引入了语义分割任务中的冗余特征信息,会造成语义分割性能下降。因此,研究图像增强和语义分割的内在联系,从而设计出图像增强和语义分割的一致性评价指标和有效的端到端低质场景语义分割网络,对提升低质场景下的语义分割性能有重要意义。

5 结束语

本文研究了低光照图像增强对于语义分割性能的影响。首先总结和归纳了当前典型的低光照图像增强模型和语义分割模型,然后以两阶段的方式设置实验,将低光照图像和经不同增强处理的图像分别进行语义分割任务,最后通过观察语义分割指标、图像评价指标以及图像主观图、语义分割主观图、特征图,分析了增强处理在低光照图像语义分割中的作用。实验结果表明,图像增强算法虽然可以提升图像评价指标和主观视觉体验,但是对于语义分割结果的影响是不完全一致的,部分增强算法对于低光照图像的语义分割表现出促进作用,而部分增强算法带来抑制作用,这说明图像增强和语义分割关注的重点和特征不完全一致,导致增强对语义分割结果可能存在负面影响。因此,在未来的针对低光照图像语义分割研究中,可以从以下几个方面进行深入研究:

(1) 进一步研究不同种类低质图像增强对于其他机器视觉应用性能的影响。本文仅针对低光照图像和语义分割任务进行了实验研究,但是在现实生活中拍摄的图像存在不同低质种类,如雾霾图像、雨天图像、雪天图像和水下图像等,且有不同种类的机器视觉应用,如目标检测和图像分类等,因此可以针对不同种类的低质图像和不同的机器视觉应用进行研究,从而总结出低质图像增强对于机器视觉应用性能影响的规律性结论。

(2) 根据图像增强和语义分割的内在联系制定一致性评价指标。本文研究发现,图像增强算法和语义分割关注的重点不同,导致经增强处理的图像语义分割结果反而下降。因此探究出图像增强和语义分割之间的内在联系,从而设计出一致性的评价指标,有助于帮助解决低质图像语义分割性能差的问题。

(3) 设计端到端的低质图像语义分割的算法模型。低质图像存在图像信息丢失、边缘模糊和亮度对比度低等问题,造成语义分割性能的下降。然而,采用对图像先增强后分割的方式,不能确保提升语义分割性能,并且步骤复杂,不适用于现实场景。因此,设计出端到端的低质图像语义分割,有利于提升不利环境条件下的语义分割算法性能和算法的鲁棒性,适用于实际应用的部署。

参考文献:

- [1] WANG J, MACKENZIE J D, RAMACHANDRAN R, et al. A deep learning approach for semantic segmentation in histology tissue images[C]//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Greece: Springer, 2016: 176-184.
- [2] MEHTA S, MERCAN E, Bartlett J, et al. Learning to segment breast biopsy whole slide images[C]//Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2018: 663-672.
- [3] TEICHMANN M, WEBER M, ZOLLNER M, et al. Multinet: Real-time joint semantic reasoning for autonomous driving[C]//Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). [S.l.]:IEEE, 2018: 1013-1020.
- [4] HAMIAN M H, BEIKMOHAMMADI A, AHMADI A, et al. Semantic segmentation of autonomous driving images by the combination of deep learning and classical segmentation[C]//Proceedings of 2021 26th International Computer Conference, Computer Society of Iran (CSICC). Piscataway: IEEE, 2021: 1-6.
- [5] MIYAMOTO R, ADACHI M, NAKAMURA Y, et al. Accuracy improvement of semantic segmentation using appropriate datasets for robot navigation[C]//Proceedings of the 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT). Piscataway: IEEE, 2019: 1610-1615.
- [6] KIM W, SEOK J. Indoor semantic segmentation for robot navigating on mobile[C]//Proceedings of the 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN). Piscataway: IEEE, 2018: 22-25.
- [7] IRWIN F G. An isotropic 3×3 image gradient operator[J]. Presentation at Stanford AI Project, 1968, 2014(2): 271-272.
- [8] ROSENFELD A. The max Roberts operator is a Hueckel-type edge detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1981(1): 101-103.
- [9] PATODI V K. Curvature and the eigenforms of the Laplace operator[J]. Journal of Differential Geometry, 1971, 5(1/2): 233-249.
- [10] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986(6): 679-698.
- [11] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66.
- [12] PREWITT J M S. Object enhancement and extraction[J]. Picture Processing and Psychopictorics, 1970, 10(1): 15-19.
- [13] ROTHER C, KOLMOGOROV V, BLAKE A. "GrabCut" interactive foreground extraction using iterated graph cuts[J]. ACM Transactions on Graphics (TOG), 2004, 23(3): 309-314.
- [14] BOYKOV Y Y, JOLLY M P. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images[C]//Proceedings of the eighth IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2001: 105-112.
- [15] HUANG Suzhao, CHU Yihong, LAI Shanghong, et al. Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI[J]. IEEE Transactions on Medical Imaging, 2009, 28(10): 1595-1605.
- [16] POGGIO T. Vision by man and machine[J]. Scientific American, 1984, 250(4): 106-117.
- [17] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 3431-3440.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Germany: Springer, 2015: 234-241.
- [19] ZHAO Hengshuang, SHI Jianping, QI Xiaojuan, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 2881-2890.
- [20] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [21] CHENG H K, CHUNG J, TAI Y W, et al. CascadePSP: Toward class-agnostic and very high-resolution segmentation via

- global and local refinement[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 8890-8899.
- [22] CHEN L C, PAPANDEOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(4): 834-848.
- [23] 刘峰, 居昊, 干宗良. 基于全卷积孪生神经网络的复杂监控场景下前景提取方法[J]. *南京航空航天大学学报*, 2021, 53(5): 743-750.
- LIU Feng, JU Hao, GAN Zongliang. Fully-convolutional siamese networks for foreground subtraction in complex surveillance videos[J]. *Journal of Nanjing University of Aeronautics and Astronautics*, 2021, 53(5): 743-750.
- [24] ROMERA E, ALVAREZ J M, BERGASA L M, et al. Erfnet: Efficient residual factorized convnet for real-time semantic segmentation[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 19(1): 263-272.
- [25] ZHAO Hengshuang, QI Xiaojuan, SHEN Xiaoyong, et al. Icnet for real-time semantic segmentation on high-resolution images [C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 405-420.
- [26] YU Changqian, GAO Changxin, WANG Jingbo, et al. BiseNet v2: Bilateral network with guided aggregation for real-time semantic segmentation[J]. *International Journal of Computer Vision*, 2021, 129(11): 3051-3068.
- [27] WANG Yu, ZHOU Quan, LIU Jia, et al. LEDnet: A lightweight encoder-decoder network for real-time semantic segmentation[C]// Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP). [S.l.]: IEEE, 2019: 1860-1864.
- [28] WU Tianyi, TANG Sheng, ZHANG Rui, et al. CGNet: A light-weight context guided network for semantic segmentation[J]. *IEEE Transactions on Image Processing*, 2020, 30: 1169-1179.
- [29] YU Changqian, WANG Jingbo, PENG Chao, et al. BiseNet: Bilateral segmentation network for real-time semantic segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 325-341.
- [30] 高凌飞, 王海龙, 王海涛, 等. 基于轻量级卷积神经网络的人证比对[J]. *南京航空航天大学学报*, 2021, 53(5): 751-758.
- GAO Lingfei, WANG Hailong, Wang Haitao, et al. Scene-identity face matching based on efficient convolutional neural network[J]. *Journal of Nanjing University of Aeronautics and Astronautics*, 2021, 53(5): 751-758.
- [31] PARK S, YU S, MOON B, et al. Low-light image enhancement using variational optimization-based retinex model[J]. *IEEE Transactions on Consumer Electronics*, 2017, 63(2): 178-184.
- [32] TAO Li, ZHU Chuang, XIANG Guoqing, et al. LLCNN: A convolutional neural network for low-light image enhancement [C]//Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP). [S.l.]: IEEE, 2017: 1-4.
- [33] LAND E H, MCCANN J J. Lightness and retinex theory[J]. *Josa*, 1971, 61(1): 1-11.
- [34] CASTLEMAN K R. Digital image processing[M]. Upper Saddle River: Prentice Hall Press, 1996.
- [35] JOBSON D J, RAHMAN Z, WODELL G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes[J]. *IEEE Transactions on Image Processing*, 1997, 6(7): 965-976.
- [36] GATTA C, RIZZI A, MARINI D. ACE: An automatic color equalization algorithm[C]//Proceedings of the Conference on Colour in Graphics, Imaging, and Vision. Netherlands: Society for Imaging Science and Technology, 2002: 316-320.
- [37] LV Feifan, LU Feng, WU Jianghua, et al. MBLLEN: Low-light image/video enhancement using CNNs[C]//Proceedings of the BMVC. Cham: Springer, 2018: 4.
- [38] ZHANG Yonghua, ZHANG Jiawan, GUO Xiaojie. Kindling the darkness: A practical low-light image enhancer[C]// Proceedings of the 27th ACM International Conference on Multimedia. [S.l.]: ACM, 2019: 1632-1640.
- [39] GUO Chunle, LI Chongyi, GUO Jichang, et al. Zero-reference deep curve estimation for low-light image enhancement[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 1780-1789.
- [40] JIANG Yifan, GONG Xinyu, LIU Ding, et al. Enlightengan: Deep light enhancement without paired supervision[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.

- [41] SAKARIDIS C, DAI D, VAN GOOL L. ACDC: The adverse conditions dataset with correspondences for semantic driving scene understanding[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). [S.l.]: IEEE, 2021.
- [42] YU Fisher, XIAN Wenqi, CHEN Yingying, et al. BDD100K: A diverse driving video database with scalable annotation tooling[EB/OL].(2018-05-12)[2022-02-23].<https://doi.org/10.48550/arXiv.1805.04687>.

作者简介:



艾羽丰(1998-),女,硕士研究生,研究方向:计算机视觉、语义分割等,E-mail:ai-yufeng@tju.edu.cn。



郭继昌(1966-),通信作者,男,教授,博士生导师,研究方向:智能视频、图像分析、识别及处理、滤波器理论及设计等,E-mail:jc-guo@tju.edu.cn。



安冠华(1999-),男,硕士研究生,研究方向:计算机视觉。



张怡(1997-),男,硕士研究生,研究方向:计算机视觉、显著性检测。

(编辑:刘彦东)