

基于知识蒸馏的夜间低照度图像增强及目标检测

苗德邻 刘磊 莫涌超 胡朝龙 张益军 钱芸生

Nighttime low-light image enhancement and object detection based on knowledge distillation

MIAO Delin, LIU Lei, MO Yongchao, HU Chaolong, ZHANG Yijun, QIAN Yunsheng

引用本文:

苗德邻, 刘磊, 莫涌超, 等. 基于知识蒸馏的夜间低照度图像增强及目标检测[J]. 应用光学, 2023, 44(5): 1037–1044. DOI: 10.5768/JAO202344.0502004

MIAO Delin, LIU Lei, MO Yongchao, et al. Nighttime low-light image enhancement and object detection based on knowledge distillation[J]. *Journal of Applied Optics*, 2023, 44(5): 1037–1044. DOI: 10.5768/JAO202344.0502004

在线阅读 View online: <https://doi.org/10.5768/JAO202344.0502004>

您可能感兴趣的其他文章

Articles you may be interested in

[基于改进SSD的车辆小目标检测方法](#)

Detecting method of small vehicle targets based on improved SSD

应用光学. 2020, 41(1): 150–155 <https://doi.org/10.5768/JAO202041.0103004>

[基于深度学习的铸件X射线图像分割研究](#)

Casting X-ray image segmentation based on deep learning

应用光学. 2021, 42(6): 1025–1033 <https://doi.org/10.5768/JAO202142.0602003>

[基于深度学习的无人车夜视图像语义分割](#)

Semantic segmentation of night vision images for unmanned vehicles based on deep learning

应用光学. 2017, 38(3): 421–428 <https://doi.org/10.5768/JAO201738.0302007>

[结合深度学习的非下采样剪切波遥感图像融合](#)

Remote sensing image fusion based on deep learning non-subsampled shearlet

应用光学. 2018, 39(5): 655–666 <https://doi.org/10.5768/JAO201839.0502001>

[基于视觉SLAM和目标检测的语义地图构建](#)

Semantic SLAM based on visual SLAM and object detection

应用光学. 2021, 42(1): 57–64 <https://doi.org/10.5768/JAO202142.0102002>

[基于低照度三基色图像去噪及融合彩色图像增强方法研究](#)

Color image enhancement based on LLL tricolor image denoising and fusion

应用光学. 2018, 39(1): 57–63 <https://doi.org/10.5768/JAO201839.0102003>



关注微信公众号，获得更多资讯信息

文章编号:1002-2082(2023)05-1037-08

基于知识蒸馏的夜间低照度图像增强及目标检测

苗德邻, 刘 磊, 莫涌超, 胡朝龙, 张益军, 钱芸生

(南京理工大学 电子工程与光电技术学院, 江苏南京 210018)

摘要:为了实现夜间低照度图像的增强,提高目标检测模型在夜间低照度条件下的检测精度并减小模型的计算成本,提出了一种基于知识蒸馏和数据增强的夜间低照度图像增强以及目标检测多任务模型,基于高质量图像模型进行知识蒸馏,利用高质量图像的特征信息指导模型训练,从而使模型在夜间低照度图像中提取到与高质量图像类似的特征信息。利用这些特征信息可以实现图像的对比度增强、去噪以及目标检测。实验结果表明,提出的蒸馏方法可以提升 16.58% 的夜间低照度目标检测精度,且用该方法增强的图像可以达到主流的基于深度学习的图像增强的效果。

关键词:低照度图像增强; 目标检测; 深度学习; 知识蒸馏

中图分类号:TN206;TP399

文献标志码:A

DOI: 10.5768/JAO202344.0502004

Nighttime low-light image enhancement and object detection based on knowledge distillation

MIAO Delin, LIU Lei, MO Yongchao, HU Chaolong, ZHANG Yijun, QIAN Yunsheng

(School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210018, China)

Abstract: In order to enhance the quality of nighttime low-light image, improve the accuracy of the object detection model under the nighttime low-light condition and reduce the calculation cost of the model, a multi-task model for nighttime low-light image enhancement and object detection based on knowledge distillation and data enhancement was proposed. Knowledge distillation was performed based on the high-quality image model, and the feature information of the high-quality image was used to guide the model training, so that the model could extract the feature information similar to that of the high-quality image in the nighttime low-light images. These feature information could be used to achieve enhancement of image contrast, denoising and objects detection. The experimental results show that the proposed distillation method can improve the object detection accuracy of nighttime low-light by 16.58%, and the image enhanced by this method can achieve the effect of mainstream image enhancement based on deep learning.

Key words: low-light image enhancement; object detection; deep learning; knowledge distillation

引言

近年来随着深度学习理论的发展以及计算机计算性能的提升,目标检测和追踪技术已经广泛应用于智能监控、自动及辅助驾驶、信息安全系统等诸多领域;但是这些技术在现实环境下仍然面临很多问题,由于图像通常无法在最优的光条件下获取,往往会受到不可避免的环境因素如背光、非均匀照明、夜间微光等,以及技术限制如不

足的曝光时间和过低的增益等影响,这类图像的视觉质量较差,目标检测和识别的精度远低于高质量图像。考虑到大多数目标检测模型都是基于高质量图像数据进行训练和测试的,如 RCNN (region convolutional neural network)^[1]、YOLO(you only look once)^[2]、SSD(single shot multi-box detector)^[3]等,采用先对原始低照度图像进行图像增强的方式提高目标检测的精度是一种最简便的方式。

收稿日期:2022-11-24; 修回日期:2022-12-30

基金项目:微光夜视技术重点实验室基金 (J20190102);江苏省青蓝工程基金 (2017-AD41779)

作者简介:苗德邻 (1998—),男,硕士,主要从事数字图像处理研究。E-mail: 879837494@qq.com

通信作者:刘磊 (1978—),女,博士,教授,主要从事光电信息探测与处理以及半导体光电发射材料设计与制备研究。

E-mail: liu1133_cn@sina.com.cn

LLNet(low-light net)^[4]、MBLLEN(multi-branch low-light enhancement network)^[5]等端到端的低照度图像增强工作已经证明了利用神经网络提升低照度图像质量的可能性。同时,学者们也注意到 Retinex 模型在传统的低照度图像增强以及图像去雾中有着不俗的表现,由此出现了基于利用神经网络来估计 Retinex 中的照明分量以及反射分量的方法,如 Retinex-Net^[6]、LightenNet^[7]等,都能起到很好的增强效果以及自适应性。为了实现神经网络的泛用性,很多研究者采集构建了一系列低照度场景数据集,包括 SID(see-in-the-dark)^[8]、LOL(low-light)^[9]等。除此之外,基于非匹配数据和非监督学习的方法也被应用在低照度图像增强中,如 EnlightenGAN(enlighten generative adversarial networks)^[10],还有一些基于零样本学习的方法,例如 Zero-DCE(zero-reference deep curve estimation)^[11]等,这种方法具有很低的数据成本以及极高的训练和推理速度。

上述低照度图像增强模型主要以人为的标准进行优化,即使在一些图像评价指标上表现出色,也很难对下游的目标检测任务起到很好的精度提升效果。不仅如此,先增强后检测的方法还会增加在边缘设备上部署的计算力需求。受到 KRU-THVENTI S S 等人^[12]和 QIL 等人^[13]的启发,本文提出了一种新的基于知识蒸馏的目标检测网络

训练方法,不同于一些典型的知识蒸馏算法通过蒸馏大模型的知识来训练小模型^[14],该方法通过构建一组像素级对齐的高质量和低照度图像对数据集,利用高质量图像的特征信息指导网络模型学习从相应的低照度图像中提取类似的信息,可以有效提高在夜间低照度环境或其他复杂条件下的低质量图像的目标检测精度。为了便于人眼对目标检测结果进行观察,本文创新性地在原有的目标检测网络模型框架上,引入了用于图像增强的增强解码器,其主要作用是利用目标检测主干网络的特征恢复出高质量图像,最终可以在一个模型上同时实现图像的增强和目标检测。

1 基于知识蒸馏的低照度图像增强和目标检测算法

本文设计了一种结合目标检测和低照度图像增强的神经网络结构,可以巧妙地将目标检测和图像增强的算法集合在同一个卷积神经网络模型中。模型整体可以分为 3 部分:负责提取图像特征信息的主干网络;通过提取的多尺度特征进行图像重建的低照度增强解码器;识别图像特征信息并判断图像中目标位置和类别的目标检测模块。**图 1** 展示了教师网络和学生网络的网络结构以及蒸馏的位置,其中教师网络的输入为高质量图像,

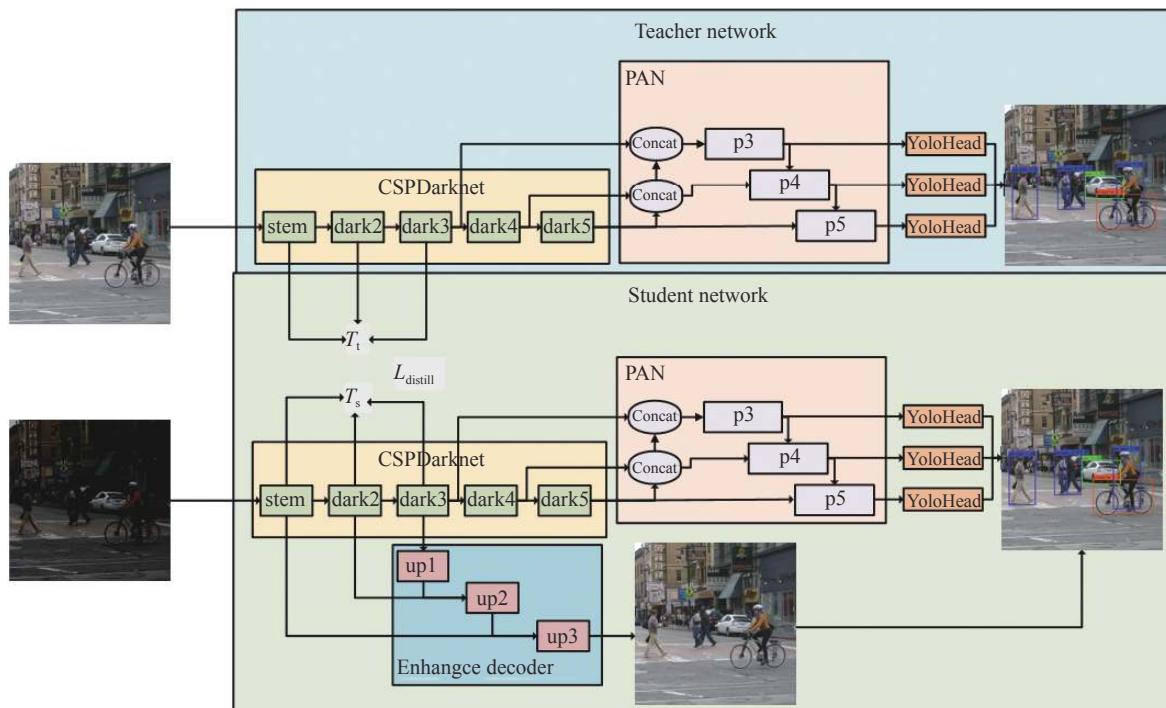


图 1 图像增强和目标检测双任务模型结构以及知识蒸馏示意图

Fig. 1 Structure diagram of dual task model of image enhancement and object detection and schematic of knowledge distillation

通过预训练使教师网络学习利用高质量图像的特征信息进行目标检测;学生网络的输入为低照度图像,通过蒸馏教师网络提取的高质量图像的特征信息进行训练,使学生网络可以在低照度图像上提取出类似于教师网络在高质量图像上提取的信息,并利用这些信息进行后续的目标检测和图像增强。由于目标检测和图像增强都需要利用主干网络所提取的特征信息,类似于文献[13]在特征金字塔上进行蒸馏的方法无法使学生网络直接学习到教师网络更浅层的特征信息,而这些浅层的信息对图像的增强有很大作用^[10],故为了通过知识蒸馏来实现多任务模型的训练,最终选择了在目标检测主干网络上进行蒸馏。

1.1 主干网络结构

在目标检测中,主干网络通常用于提取输入图像的特征,而这一部分功能和上游的图像分类网络功能相近,往往会选择成熟的图像分类模型作为主干网络。同时在底层视觉任务如图像降噪,图像超分辨率以及图像复原等网络模型中,往往会采用一种编码器-解码器的结构,即先对输入图像进行特征提取或者称为编码过程,然后根据各级特征进行恢复或者称为解码过程。其中编码器往往会选择与图像分类模型类似的结构,因此在目标检测和图像增强任务中使用同一个主干提取网络是具有可行性的。

如图1所示,主干特征提取网络框架基于CSPDarknet主干网络实现,并采用了YoloX^[15]对CSPDarknet做出的一些改进,诸如采用聚焦(Focus)来对输入图像进行降采样,通过调整CSPDarknet的卷积层深度和宽度实现更轻量化的网络模型等。

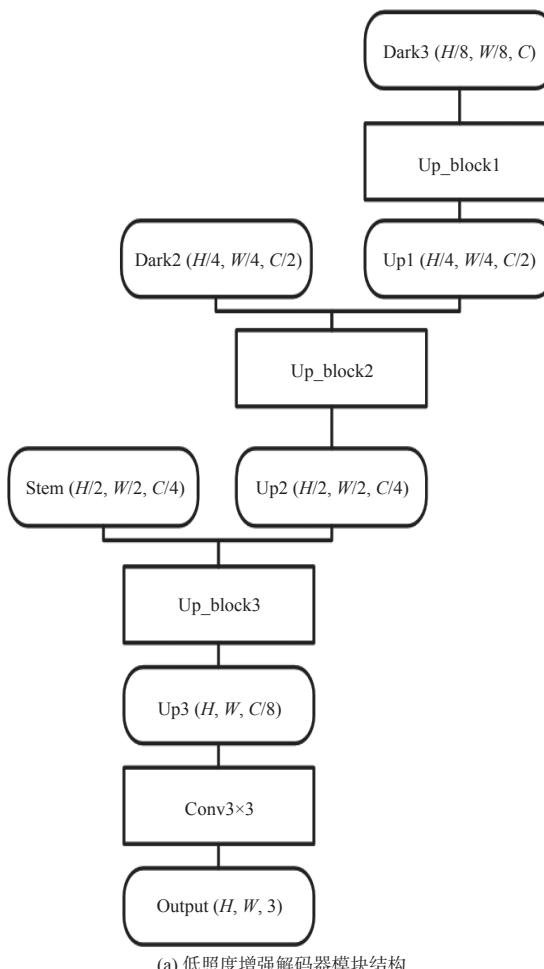
CSPDarknet可以分成5个阶段,每个阶段分别提取出不同尺度的特征信息,对于一个分辨率为 $W \times H$ 的输入图像而言,5个阶段的输出尺寸分别为 $W/2 \times H/2$ 、 $W/4 \times H/4$ 、 $W/8 \times H/8$ 、 $W/16 \times H/16$ 、 $W/32 \times H/32$ 。在目标检测网络中,最后3个阶段的输出会作为目标特征进行多尺度特征金字塔融合,而对于图像增强,我们并不需要远小于原始分辨率的图像特征,通常只会利用前3个阶段的输出结果进行解码恢复。综合而言,只需要保证学生网络第3阶段的输出与教师网络相似,就可以很大程度上实现对教师网络的蒸馏。

1.2 低照度增强解码器

与目标检测不同,图像复原模型并不是一个概率模型,所需要的特征信息相较于目标检测任务

也更多,仅通过一个尺度的特征信息生成的图像无法满足需求。基于U-net^[16]架构的图像增强模型在各个底层视觉任务中都有优秀的表现,因此基于U-net的结构设计了低照度增强解码器。

如图2所示,图像增强模块利用了主干网络中的前3个阶段的特征图,其尺寸分别为 $W/2 \times H/2$ 、 $W/4 \times H/4$ 和 $W/8 \times H/8$ 。整体采用了类似U-net的网络结构,通过跳跃连接将不同尺度的特征信息进行融合,并通过3个上采样以及卷积层将图像的重建为 $W \times H \times 3$ 的RGB(red-green-blue)图像。其中上采样操作由若干个残差块以及一个反卷积构成,具体结构如图2(b)所示,残差块的主体部分由两个 3×3 卷积层构成,残差边不做任何处理,直接将残差块的输入和输出相加得到结果。经过若干次残差块后,经过一次 3×3 的反卷积进行上采样,得到分辨率翻倍、通道数减半的输出结果。在经过上采样块处理后,输出结果与主干网络中尺度相对应的特征层进行相加,作为下一个上采样块的输入。经过3次上采样后,特征张量通过一次 3×3 卷积输出一个通道数为3的RGB图像。



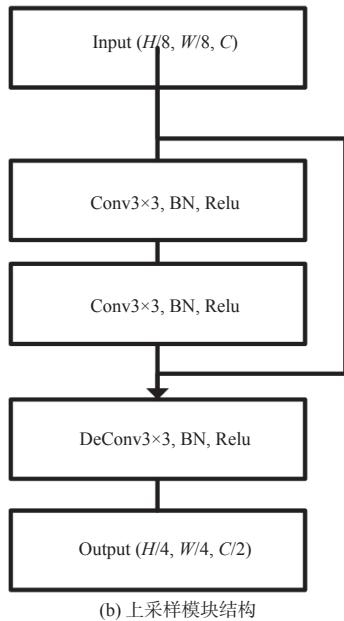


图 2 低照度图像增强模块结构图

Fig. 2 Structure of low-light image enhancement block

1.3 目标检测网络结构

目标检测结构由空间金字塔网络 (SPN)^[17]、特征金字塔网络和解耦目标检测头构成。空间金字塔用于在低尺度下生成并融合不同尺度的特征，特征金字塔主要用于融合主干特征提取网络不同尺度的特征，使目标检测头能从多尺度多语义层次的信息中进行检测，以达到更好的检测精度。不同于标准 YoloX 模型的 FPN(feature pyramid network) 特征金字塔网络，本文采用了 PAN(path aggregation network) 特征金字塔^[18]，不仅通过低尺度特征信息上采样融合到高尺度特征上，还通过下采样将高尺度信息往低尺度上融合。

检测部分采用无锚点的目标检测方法，且将分类和回归的任务进行解耦，不再通过先验知识设定锚点框输出预测框相对于锚点框的位置和形状偏移，而是直接预测相对于每个网格的位置偏移以及预测框的高度和宽度。

1.4 多任务训练的损失函数

本文方法涉及到的损失函数包含 3 部分，目标检测损失、图像增强损失以及知识蒸馏损失。其中目标检测部分和 YoloX 一样由 3 个部分组成，预测框区域损失、目标置信度损失以及目标类别损失。

预测框区域损失，即获取到每个框对应的特征点后，取出该特征点的预测框，利用真实框和预测框计算 IoU(intersection over union) 损失：

$$L_{\text{IoU}} = 1 - O_{\text{IU}}^2 \quad (1)$$

$$O_{\text{IU}} = \frac{B_{\text{pre}} \cap B_{\text{gt}}}{B_{\text{pre}} \cup B_{\text{gt}}} \quad (2)$$

式中 B_{pre} 和 B_{gt} 分别代表预测框和真实框的区域。

目标置信度损失和目标类别损失的表示形式是相同的，即：

$$L_{\text{cls}} = -w_n \left[P_n^{\text{gt}} \times \lg \sigma(P_n^{\text{pre}}) + (1 - P_n^{\text{gt}}) \times \lg (1 - \sigma(P_n^{\text{pre}})) \right] \quad (3)$$

式中： P^{pre} 和 P^{gt} 分别代表预测的概率和真实的概率； w 为每个类别损失函数的权重。

对于图像增强等底层视觉任务，最常使用的是 L_1 距离或者 L_2 距离损失。 L_1 损失在对于输入和输出差距较大时不会有过大的梯度，在图像降噪任务中应用广泛，而 L_2 损失常出现在图像超分辨率以及图像复原任务中。低照度图像增强既要进行图像信息的恢复，也要想办法去除图像中的噪声，因此本文采用结合两种损失函数优点的 smooth L_1 损失作为图像增强目标的损失函数，即：

$$L_{\text{enhance}} = \begin{cases} \frac{1}{2} \|x - y\|_2, & |x - y| < 1 \\ \|x - y\|_1 - \frac{1}{2}, & |x - y| \geq 1 \end{cases} \quad (4)$$

式中： y 为参考图像； x 为模型增强后的图像。由于学生网络和教师网络采用的是完全相同的网络结构，对于两个网络而言，每一层对应的特征层输出的尺度是完全相同的。因此，无需对两个网络的特征层进行任何其他处理，只需要求两个特征层输出的距离即可。最小化教师网络和学生网络特征层差异可以表示为最小化以下函数：

$$L_{\text{distill}} = \sum_i \|F_i^T - F_i^S\|_2 \quad (5)$$

式中 F_i^T 和 F_i^S 分别代表教师网络和学生网络特征金字塔的输出特征张量。通过这种简单的损失函数可以使学生网络和教师网络在不同尺度的特征层都以相同的程度趋于一致，然而考虑到教师网络和学生网络输入图像特性的差异，本文采用了一种针对不同尺度特征进行调整的损失函数进行优化，对于靠近输入端的特征层，采用更低的优化权重，即：

$$L_{\text{distill}} = \sum_i \frac{i}{N} \times \|F_i^T - F_i^S\|_2 \quad (6)$$

最终的损失函数可以表示为

$$L = L_{\text{distill}} + L_{\text{cls}} + L_{\text{obj}} + L_{\text{IoU}} + L_{\text{enhance}} \quad (7)$$

2 实验评估

2.1 数据集的构建

若想实现高质量正常照度目标检测知识向低照度低信噪比目标检测的迁移, 就必须获得完全对齐的高质量正常照度目标图像和低信噪比低照度目标图像对。本文采用生成对抗网络以及低照度下均匀面真实噪声叠加的方式进行数据对的构造。

通常可以认为一幅图像是由图像信号和无关的噪声组成的, 即:

$$I_{i,j} = s_{i,j} + n_{i,j} \quad (8)$$

式中: $I_{i,j}$ 为图像; $s_{i,j}$ 为信号; $n_{i,j}$ 为噪声。一般情况下, 信号的大小是大于噪声的大小的。但对于低照度下的均匀面而言, 相比一个频率很低的低频信号, 图像可以看作是以信号为均值上下波动的一个随机过程。这种情况下, $s_{i,j}$ 可以通过均值滤波提取低频的方式得到, 然后将噪声的均值调整为 0, 最终提取出来的噪声可以表示为

$$n_{i,j} = I_{i,j} - s_{i,j} - \text{mean}(I_{i,j} - s_{i,j}) \quad (9)$$

最后利用非匹配的低照度图像数据集以及正常高质量图像数据集对循环生成对抗神经网络 (cycle generative adversarial networks, CycleGAN)^[19] 进行训练, 得到一个能够将正常照度图像转换为具有夜间低照度特性的图像的生成器 $G(x)$ 。将从均匀面上提取出来的噪声叠加在生成的低照度特性图像上就得到了训练所需要的低照度-高质量目标检测图像对数据集。

$$I_{i,j}^{\text{lowlight}} = G(I_{i,j}^{\text{normal}}) + n_{i,j} \quad (10)$$

2.2 训练环境

CycleGAN 具有独特的循环训练机制, 可以实现非匹配的数据集之间的相互转换。为了训练出可以将正常高质量图像转换为具有夜间低照度特性的图像, 本文采集了 1 550 张包含行人、车辆等目标的真实夜间低照度图像, 同时在 ImageNet^[20] 和 COCO 数据集中收集了大量高质量的正常照度图像 2 480 张。利用这些夜间低质量图像数据集和收集到的高质量图像组成非匹配的数据集, 对 CycleGAN 网络进行训练。利用训练得到的高质量图像-夜间低照度图像转换模型和公开的 PASCAL VOC 目标检测数据集, 生成了共 9 963 组带标签的夜间低照度-正常高质量图像对。

同时, 为了降低图像的信噪比, 对夜间的地面、墙体等均匀面场景进行采集, 构建了共 10 组不同照度分辨率为 1 920×1 080 像素的均匀面图像

视频序列, 然后从这些视频中随机提取噪声叠加到生成的 9 963 组图像对中, 将这 9 963 组数据中的三分之一即共 2 989 对图像划分为验证集, 其余 6 974 对图像划分为训练集。

训练过程分为 3 步, 首先利用原始高质量的 PASCAL VOC 数据集训练教师网络, 然后利用知识蒸馏的方式对学生网络的主干网络和目标检测网络进行 200 轮的训练, 学习率为 2×10^{-4} , 且采用余弦退火法进行学习率的下降。随后冻结主干网络, 利用图像对对低照度图像增强解码器进行训练 50 轮, 学习率为 1×10^{-4} , 然后进行学习率线性下降至 0 的方法再次训练 50 轮。本次实验基于 Nvidia RTX 2080Ti GPU 进行训练。

2.3 图像增强效果

本文通过在真实低照度数据集 LOL 数据集上的峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structural similarity, SSIM) 指标来验证网络模型在低照度图像增强任务上的效果, 同时利用自然图像质量评价器 (natural image quality evaluator, NIQE)^[21] 对比验证模型在真实低照度场景下的增强效果, 其中 NIQE 值越低代表图像质量评价越高, PSNR 和 SSIM 值越高代表图像越接近参考图像。进行对比的方法包括 EnlightenGAN、Zero-DCE 以及 CycleGAN, 结果如表 1 所示。

表 1 不同方法的图像质量评价

Table 1 Parameters of image quality assessment of different methods

Method/Index	NIQE	PSNR	SSIM
EnlightenGAN ^[18]	18.82	28.23	0.73
Zero-DCE ^[21]	23.33	28.54	0.73
CycleGAN ^[41]	27.58	28.61	0.71
Algorithm proposed	19.86	28.57	0.74

从对比结果来看, 本文的方法在 NIQE 指标上接近 EnlightenGAN 的水平, 且明显超过了 Zero-DCE 和 CycleGAN 的效果, 同时该方法在 PSNR 和 SSIM 指标上均优于 EnlightenGAN。证明了该方法在低照度图像增强上具有可行性。

本文还对 4 种方法处理的图像质量进行了对比和主观判断, 对比结果如图 3 所示。其中第 1 列展示了原始的低照度夜间图像, 第 2~5 列分别为 EnlightenGAN、Zero-DCE、CycleGAN 以及本文方法处理的结果。

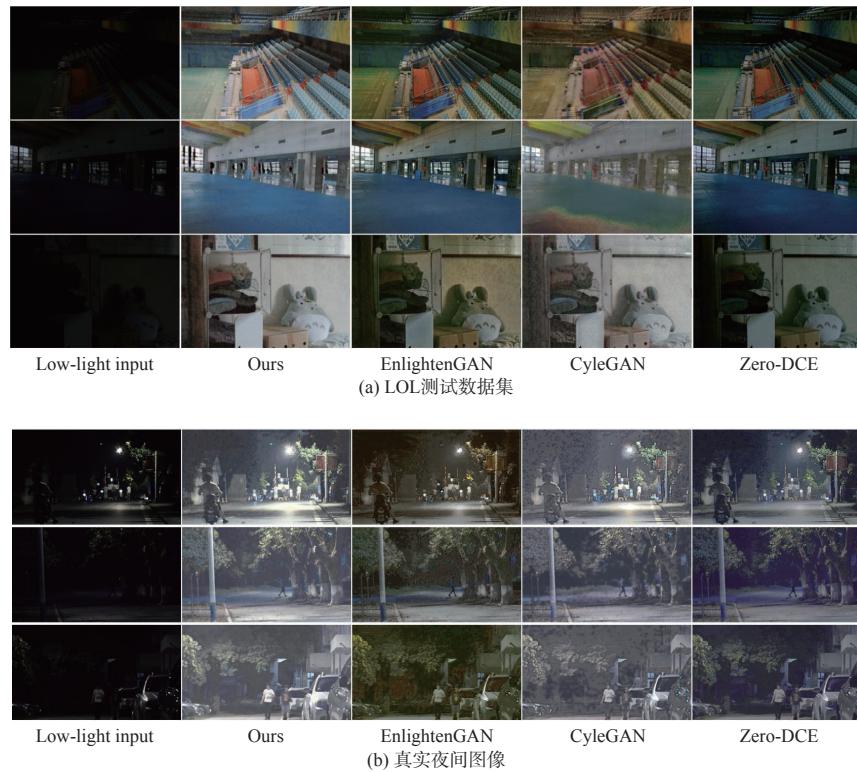


图 3 不同方法的低照度图像增强效果对比

Fig. 3 Visual comparison on low-light image enhancement effect of different methods

可以看出,本文方法对夜间低照度图像具有明显的对比度和动态范围增强以及降噪的效果。通过对比可以发现,其余3种方法处理的结果都有较为明显的噪声残留,而本文在构造的训练数据集中叠加了真实的噪声,可以使神经网络模型具有一定的降噪效果。

2.4 检测精度

为了验证知识蒸馏对夜间低照度场景下的目标检测精度提升的效果,采用本文所构建的验证集中2989张图像以及公开的夜间低照度目标检测数据集ExDark^[22]中的2194张夜间低照度图像的平均精确度(mean of average precision, mAP)作为检测精度的评估指标,并与正常照度下的精度以及未采用知识蒸馏的精度进行对比,其中mAP计算的IoU阈值为50%。结果如表2、图4~图5所示。

表2 不同方法的目标检测精度

Table 2 mAP of different methods

Method	mAP50	mAP50@ExDark
YOLOXtiny(Normal illumination)	80.47%	None
YOLOXtiny (LLL+Knowledge distillation)	54.40%	41.28%
YOLOXtiny(LL)	37.82%	23.13%

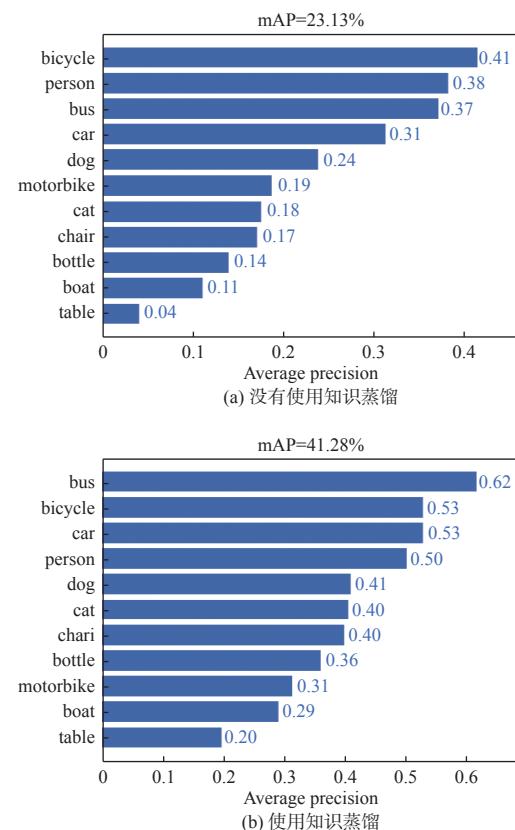


图4 在真实低照度图像数据集中的mAP指标测试结果

Fig. 4 Test results of mAP indicators on real low-light images dataset



图 5 低照度图像增强和目标检测测试结果

Fig. 5 Test results of low-light image enhancement and object detection

从测试结果整体上来看,采用知识蒸馏的方式可将在构造数据集下的夜间低照度目标的检测精度提高 16.58%,在真实数据集下的目标检测精度提高 18.15%,且对每一种目标的精度都有所提升,尤其是对车辆等交通工具的识别精度提升最大。在真实低照度图像的测试中还发现,一些小目标以及在光源附近的目标很容易出现漏检的情况,这可能是由于训练数据对中缺少真实灯光,尤其是多个光源场景下的数据所导致的。

3 结论

本文设计了一种可以同时进行夜间低照度图像增强以及目标检测任务的神经网络模型,合成构造了具有夜间低照度特性的图像数据集,利用特征知识蒸馏的方式对网络模型训练。在数据集上的实验表明,本文的方法具有可行性,能够有效提升夜间低照度下的目标检测精度,图像增强效果也能够达到甚至优于一些主流方法。此外该方法还可以有效提升高噪声低照度图像的信噪比,产生对人眼视觉友好的增强图像,多任务共享主干的结构还有效降低了计算成本。未来的工作将探索实现在利用高质量图像训练的大模型上进行低照度、小模型的蒸馏。利用对抗模型进行无监督训练也是进一步提升图像质量的探索方向。

参考文献:

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2014: 580-587.
- [2] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. New York: IEEE, 2016: 779-788.
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// European Conference on Computer Vision. Switzerland : Springer, Cham, 2016: 21-37.
- [4] LORE K G, AKINTAYO A, SARKAR S. LLNet: a deep autoencoder approach to natural low-light image enhancement[J]. *Pattern Recognition*, 2017, 61: 650-662.
- [5] LV F, LU F, WU J, et al. MBLLEN: low-light image/video enhancement using CNNs[J]. *Baghdad Monitoring and Verification Center*, 2018, 220(1): 4.
- [6] WEI C, WANG W, YANG W, et al. Deep retinex decomposition for low-light enhancement[EB/OL]. (2018-08-14) [2022-11-20]. <https://arxiv.org/abs/1808.04560v1>.
- [7] LI C, GUO J, PORIKLI F, et al. LightenNet: a convolutional neural network for weakly illuminated image enhancement[J]. *Pattern Recognition Letters*, 2018, 104: 15-22.
- [8] CHEN C, CHEN Q, XU J, et al. Learning to see in the dark[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3291-3300.
- [9] XIONG W, LIU D, SHEN X, et al. Unsupervised real-world low-light image enhancement with decoupled networks[EB/OL]. (2020-05-06) [2022-11-20]. <https://arxiv.org/abs/2005.02818v2>.
- [10] JIANG Y, GONG X, LIU D, et al. Enlightengan: deep light enhancement without paired supervision[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [11] GUO C, LI C, GUO J, 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. New York: IEEE, 2020: 1780-1789.

- [12] KRUTHIVENTI S S S, SAHAY P, BISWAL R. Low-light pedestrian detection from RGB images using multi-modal knowledge distillation[C]// 2017 IEEE International Conference on Image Processing (ICIP). New York: IEEE, 2017: 4207-4211.
- [13] QI L, KUEN J, GU J, et al. Multi-scale aligned distillation for low-resolution detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2021: 14443-14453.
- [14] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. (2015-03-09) [2022-11-20]. <https://arxiv.org/abs/1503.02531>.
- [15] GE Z, LIU S, WANG F, et al. Yolox: exceeding yolo series in 2021[EB/OL]. (2021-07-18) [2022-11-20]. <https://arxiv.org/abs/2107.08430>.
- [16] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [C]// International Conference on Medical image computing and computer-assisted intervention. Switzerland: Springer, Cham, 2015: 234-241.
- [17] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [18] WANG W, XIE E, SONG X, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2019: 8440-8449.
- [19] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of the IEEE international conference on computer vision. New York: IEEE, 2017: 2223-2232.
- [20] DENG J, DONG W, SOCHER R, et al. Imagenet: a large-scale hierarchical image database[C]// 2009 IEEE conference on computer vision and pattern recognition. New York: IEEE, 2009: 248-255.
- [21] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a 'completely blind' image quality analyzer[J]. *IEEE Signal Processing Letters*, 2012, 20(3): 209-212.
- [22] LOH Y P, CHAN C S. Getting to know low-light images with the exclusively dark dataset[J]. *Computer Vision and Image Understanding*, 2019, 178: 30-42.